

Julia Miller

Research in the pipeline: where lexicography and phraseology meet

Lexicography, 2018; 5(1):23-33

© Springer-Verlag GmbH Germany, part of Springer Nature 2018

*This is a post-peer-review, pre-copyedit version of an article published in **Lexicography**, The final authenticated version is available online at: <http://dx.doi.org/10.1007/s40607-018-0044-z>*

PERMISSIONS

<https://www.springer.com/gp/open-access/publication-policies/self-archiving-policy>

Self-archiving for articles in subscription-based journals

Springer journals' [policy on preprint sharing](#).

By signing the Copyright Transfer Statement you still retain substantial rights, such as self-archiving:

*Author(s) are permitted to self-archive a pre-print and an author's **accepted manuscript** version of their Article.*

.....

b. An Author's Accepted Manuscript (AAM) is the version accepted for publication in a journal following peer review but prior to copyediting and typesetting that can be made available under the following conditions:

(i) Author(s) retain the right to make an AAM of their Article available on their own personal, self-maintained website immediately on acceptance,

(ii) Author(s) retain the right to make an AAM of their Article available for public release on any of the following 12 months after first publication ("Embargo Period"): their employer's internal website; their institutional and/or funder repositories. AAMs may also be deposited in such repositories immediately on acceptance, provided that they are not made publicly available until after the Embargo Period.

An acknowledgement in the following form should be included, together with a link to the published version on the publisher's website: "This is a post-peer-review, pre-copyedit version of an article published in [insert journal title]. The final authenticated version is available online at: [http://dx.doi.org/\[insert DOI\]](http://dx.doi.org/[insert DOI])".

When publishing an article in a subscription journal, without open access, authors sign the Copyright Transfer Statement (CTS) which also details Springer's self-archiving policy.

See Springer Nature [terms of reuse](#) for archived author accepted manuscripts (AAMs) of subscription articles.

15 April 2020



Research in the pipeline: where lexicography and phraseology meet

Julia Miller¹

Received: 1 March 2018 / Accepted: 13 June 2018
© Springer-Verlag GmbH Germany, part of Springer Nature 2018

Abstract

In the pipeline is an English idiom referring to something abstract, such as a project, that is in progress but is not yet complete. Although some research has been published on initiatives linking lexicography and phraseology, and more research may be in the pipeline, there is plenty of room for further collaboration. This paper will explore ways in which phraseology (epitomised here by the phrase known as an idiom) contributes to lexicography in diverse fields: online dictionaries; headwords; the labelling of idiom entries; example sentences; the possibility of crowdsourcing; the role of corpora; pedagogical lexicography; teachers' use of dictionaries in class; user skills; and ways in which phraseology can contribute to the portrayal of culture in a dictionary. It is hoped that this will inspire joint research projects between phraseologists and lexicographers so that future papers can report on more research in the pipeline.

Keywords Phraseology · Idiom · Lexicography · Dictionary

1 Introduction

Phraseology can present both a stumbling block to understanding for non-native speakers and, conversely, an indication of a non-native speaker's mastery of a language when she/he uses multiword expressions correctly (Li and Schmitt 2009, p. 86). Correct usage, however, involves more than just demonstrating mastery of an expression's meaning and syntax, although such mastery itself can be problematic. It also involves socio-pragmatic knowledge of when and how to use an expression, along with underlying cultural references (O'Keeffe et al. 2010).

This paper is based on a plenary session at the Asialex 2017 conference, and includes brief sections on each of the conference themes.

✉ Julia Miller
julia.miller@adelaide.edu.au

¹ School of Education, The University of Adelaide, Adelaide, SA, Australia

One common type of expression is the idiom. There are many definitions for the word ‘idiom’, but most would agree that it contains at least two words, is non-literal, has a relatively fixed form (both syntactically and lexically) and is known to many members of a speech community (Moon 1998, p. 6). The semantic opacity of idioms makes them hard for L2 speakers of a language to understand, and their structure can make it hard to search for them in a dictionary. For this reason, they are a good example of the potential pitfalls encountered when examining lexicography in relation to phraseology.

One of the first places people search for the meaning and use of idioms is a dictionary. However, although idiom dictionaries specialise in such expressions, a user may be unaware that a string of words is actually an idiom. They are therefore more likely to search for individual words from the idiom in a general purpose, bilingual or learner’s dictionary. Nevertheless, finding an idiom in a dictionary, even in an online dictionary, is not always an easy task (Miller 2013). Moreover, not every dictionary will include every idiom (Miller 2013).

Previous research has looked at individual aspects of phraseology and lexicography, often in connection with particular types of dictionaries. Cowie (1998), for example, described the influence of Russian and British phraseologists in the compilation of phraseological dictionaries, while Siepmann (2008) examined the portrayal of phraseology in learner’s dictionaries and Szczepaniak and Lew (2011) explored the topic of images in idioms dictionaries. However, no recent research article has exploited a wider range of possibilities for collaboration between researchers from the two disciplines of phraseology and lexicography.

The paper that follows will address three main intersections between phraseology (exemplified here by the idiom) and lexicography—the possibilities created by the digital revolution; the place of idioms in pedagogical lexicography; and the role of idioms and dictionaries in portraying different cultures. The paper will conclude with further questions to prompt research or action.

2 Intersections between phraseology and lexicography

Lexicography is both a research-based and a practical profession, and phraseologists have great potential to contribute to the portrayal of all types of phraseological units within dictionaries of different types and for different audiences. Online dictionaries are uniquely placed to deal with problems such as which headword to choose for an idiom, how idioms should be labelled, which examples are provided, how to track emerging idioms, and how best to exploit different corpora to provide accurate information about idiom use. Such information can be particularly useful for learners of a language, and pedagogical lexicography has great potential to help learners understand and use idioms with greater confidence. The culture of a target language is also clearly displayed in the range of idioms included in a dictionary, despite frequent problems of idiom translation and equivalence. Each of these areas will now be addressed in greater detail.

2.1 The electronic and digital revolution in lexicography

Online dictionaries have great potential to simplify dictionary searches. Whereas, with a print dictionary, a user might have to look under a number of possible headwords before finding an idiom, the search should be much simpler with an online dictionary. For example, an idiom such as *a wolf in sheep's clothing* (indicating a character who appears to be innocent but has an underlying evil nature) might appear under one or more of the nouns, but an electronic or online dictionary has the capacity to provide hyperlinks between each of these so that a user can easily find the meaning of an idiom. Unfortunately, despite the capacity of online dictionaries to facilitate the search for phrases, there is often a lack of consistency in their approach. I will use a popular type of dictionary—the online monolingual English learner's dictionary—to illustrate the current situation.

There are six major online English learner's dictionaries. These will be referred to here by the names given to the popular print editions: *Cambridge Advanced Learner's Dictionary* (CALD) (<https://dictionary.cambridge.org>); *Collins COBUILD* (COBUILD) (<https://www.collinsdictionary.com/dictionary/english>); *Longman Dictionary of Contemporary English* (LDOCE) (<https://www.ldoceonline.com>); *Macmillan English Dictionary for Advanced Learners* (MEDAL) (<https://www.macmillandictionary.com>); *Merriam-Webster Learner's Dictionary* (MW) (<http://www.learnersdictionary.com>) and *Oxford Advanced Learner's Dictionary* (OALD) (<https://www.oxfordlearnersdictionaries.com>).

The expression *a wolf in sheep's clothing* provides an example of how the different dictionaries facilitate searches for what is a relatively common idiom. In COBUILD, for example, *a wolf in sheep's clothing* appears only when one types the whole idiom into the search box; it does not appear under any of the individual nouns. In CALD and MEDAL, it appears only under **wolf** or if one types the whole idiom. In LDOCE, it appears at **wolf**, **clothing** and when one types the whole idiom. In MW, it appears at **wolf**, **sheep** and **clothing**, with a direction to each of these if one types the whole idiom. In OALD, it appears at **wolf**, **sheep** and **clothing**, with a direction to **clothing** if one types the whole idiom. Only MW and OALD, therefore, cater for all the search possibilities. It is harder for the user to find the meaning in the other four dictionaries, unless they recognise that they are looking for an idiom and type in the complete expression. The nature of idioms, however, means that they are hard to identify, and a user may not know which of the component words to use in a search. Thus, while the advent of online dictionaries has made it theoretically much easier to find idioms among the entries, there is still some way to go in terms of links between potential headwords.

Another area in which the leading online learner's dictionaries lack consistency is in the labels given to phraseological units. For example, the idiom *out of the blue* has no label in CALD, LDOCE or MW. In COBUILD and MEDAL, it is labelled as a phrase; and in OALD it is called an idiom. Standardisation of terminology would help learners know how to find such terms within a dictionary, especially when idioms may be listed in a side bar rather than under one of the key component words.

The digitalization of lexicography also has great potential to increase the number of example sentences provided for each entry. Example sentences in dictionaries

should help the reader to understand how and when an idiom is used, as well as contributing to its meaning. The following instances, however, reveal that although corpus examples may succeed in supplying syntactic information, it is extremely difficult to encapsulate a true semantic portrayal of an idiom in a single example sentence. Consider which single idiom might best fit each of these sentences:

1. The decision came _____. (OALD)
2. _____ she said, 'Your name's John, isn't it?' (MEDAL)
3. Then, _____, he sold his house and left the country. (MW)
4. One day, _____, she announced that she was leaving. (CALD)
5. Even with a mysterious ex-lover who had turned up _____ after more than sixteen years' absence. (LDOCE)
6. And my winning ride came _____; you can often win when you are least expecting to. (COBUILD)

The missing idiom is *out of the blue*. Native speakers of English may be able to guess this in some cases due to their knowledge of likely collocations, such as the words *come* and *completely*. Speakers of English as an additional language (EAL), however, may understandably lack this knowledge. The syntax in the example sentences suggests that the missing word or phrase is an adverb of time, but only the COBUILD example clearly flags the idea of unexpectedness. Sentences 1, 2, 3 and 5 could equally well take the word *yesterday*, while sentences 2, 3 and 4 could take *predictably*.

Since corpus examples which truly demonstrate the meaning of an idiom are hard to find, one obvious solution is to provide more than one example each time, so that different uses and aspects of the idiom's meaning can be covered. This is easier in online dictionaries than in print dictionaries, as long as the digital page does not become cluttered (Kwary 2012). Since finding multiple example sentences is time consuming (Frankenberg-Garcia 2012), the task could potentially be crowdsourced. Rundell (2016, p. 5) defines crowdsourcing in its narrowest sense as "a distributed working method in which a large, centrally managed task is completed with the help of hundreds, even thousands, of volunteers, each of whom makes a small contribution". In this case, idioms could be distributed to volunteers, who would then note down occurrences of the idiom in audio, visual or written form, especially in online sources, and identify the examples that best encapsulate the idiom's meaning. Editorial oversight would still be necessary to decide which examples to include in a dictionary, but the initial identification of suitable examples would impinge less on the budget for paid dictionary staff while at the same time involving some of the many people who have a non-professional interest in language.

Crowdsourcing can also be used as a source of neologisms and emerging expressions. Just as 'citizen scientists' contribute to discoveries in astronomy (Marshall et al. 2015), so 'citizen lexicographers' could easily contribute to lexicography by tracking the use of new expressions. One new saying that is gaining currency, for

example, is the expression *Haters gonna hate*.¹ This expression, which means that some people will always hate, no matter what the circumstances, had 1,080,000 Google hits on 7 June 2017. This had almost quadrupled by 21 February 2018, with 3,820,000 hits. If the phrase continues to be used, it could usefully be included in a dictionary, and enthusiastic non-professionals could play a role in monitoring such new idioms for potential dictionary inclusion.

Once an idiom has been identified, an analysis of different corpora is necessary to ascertain the idioms most frequently used in certain genres. Unfortunately, this task remains extremely time consuming, as idioms can still only be identified by human reading of a text; there is as yet no means of searching for an unknown idiom electronically. For instance, it is possible to search for the idiom *out of the blue* in any digital corpus if we enter that phrase as a search term. However, if we do not know what idioms we are looking for, the entire text must be read to identify the idioms it contains. Moreover, there is a shortage of specialised corpora within which to search.

The development of new corpora can be a slow process, but it is vital to our understanding of language in use. English, for example, has two academic spoken corpora—the British Academic Spoken English (BASE) corpus (Applied Linguistics, University of Warwick 2018) compiled between 2000 and 2005, and the Michigan Corpus of Academic Spoken English (MICASE) (The University of Michigan English Language Institute 2007)—but even these are becoming dated. For example, one BASE lecture records the introduction of new terms (emphasised here in bold font):

a diary on on your web page is c-, is now called a **blog [laugh]** whether that will ever make its way into a dictionary i don't know... have you come across **F-A-Qs** no **[laugh]** er you you probably will during the next year

There is therefore a need to develop new corpora. These could not only reflect more recent usage and include language used in other English speaking countries; they would also provide a means of comparing idiom and other language use historically. Corpus development, however, is financially costly. More sharing of existing corpora would, therefore, benefit researchers everywhere, including those seeking to provide corpus-based examples of idioms for dictionary use.

Corpora can also tell us more about the types of speech events and texts in which idioms appear. There is a perception, for example, that idioms are never used in academic speech or writing. My current research from the BASE corpus, however, indicates that this is not the case. For instance, the idiom *in the long run*, meaning 'eventually', appears with this sense 7 times in the BASE corpus (in seven different texts), 8 times in MICASE (in 5 different transcripts) and 662 times in the Oxford Corpus of Academic English (Oxford University Press 2012). Were it not for corpus evidence, the academic use of this idiom would be overlooked.

¹ I am indebted to a presentation by Damien Villers at the Interdisciplinary Colloquium on Proverbs in Tavira, 2016, for this example.

The question of which idioms are most appropriate in certain fields is also important. Let us take as examples three idioms which each appear three times in the BASE corpus: *in the pipeline*, *sit on the fence* and *out of the blue*. A 2017 search of the Springer Exemplar corpus found that *in the pipeline* appeared most in articles relating to medicine and public health (562 tokens), computer sciences (539 tokens), engineering (533 tokens) and life sciences (515 tokens).² Conversely, *sit on the fence* was rarely used, but when it did appear it was in the context of political science and international relations (six tokens), social sciences (five tokens), childhood, adolescence and society (four tokens), education (four tokens) and human geography (four tokens). *Out of the blue* appeared in a wide range of subjects, particularly medicine and public health (180 tokens), philosophy (171 tokens), psychology (99 tokens) and education (91 tokens). The relative numbers of tokens may be more indicative of the database composition, which favoured science journals over other disciplines. However, the widespread use of idioms in academic writing cannot be ignored.

Dictionaries of academic English are just one type of specialised dictionary that benefit from corpus studies. However, aside from Liu (2003, 2012), Simpson and Mendis (2003) and Grant (2005), there is little research on the use of idioms (as opposed to formulae and lexical bundles) in different corpora. Such research would indicate the frequency of idioms in different genres, disciplines and registers, and facilitate a diachronic view of idiom use that could be used in many different kinds of dictionaries.

2.2 Pedagogical lexicography

One type of dictionary that benefits particularly from information on use and register is the learner's dictionary. Use is hard to determine, but this information is vital in informing a non-native speaker if an idiom is formal or informal, and which age group might use it. For example, in Miller's (2011) survey, the expression *let off steam* was known and used by almost all the 869 English L1 survey participants of different ages in Australia and the UK. Conversely, *the milk of human kindness* was known by roughly two-thirds of participants aged over 22, but used by only 12%, while the expression was almost unknown and was never used by those aged 22 and under. Including advice on who is likely to use an idiom allows learners to gauge whether it is appropriate for them to include in a conversation with their peers. Miller's research (2011) discovered a generational rather than a regional model of idiom use, but such detailed information about user groups is hard to establish and is not normally included in learner's dictionaries.

Coverage of idioms also varies greatly between dictionaries, and none of the 6 major English learner's dictionaries included all 84 idioms in Miller's (2011) study. For instance, *run out of steam*, *in the land of Nod* and *spend a penny* are not in MW and *give someone an inch and they'll take a mile* and *not move/budge*

² Unfortunately, it is not possible to check exactly how many of these instances were idiomatic, as Springer Exemplar is no longer available, but the idiomatic use was in the majority, particularly in the non-engineering subjects.

an inch are not in COBUILD. Two of the ten most used idioms in the survey (*the blind leading the blind, reap what you sow*) also do not appear in COBUILD, though the other learner's dictionaries do include all the study's top ten. Coverage has improved since the study was conducted, but there are still many gaps to fill.

Even when idioms are included in a dictionary, however, students may have trouble finding them. Lexicographers are the first to understand that dictionary consultation skills need to be taught (Hartmann 2003), but unfortunately, this understanding does not seem to have been grasped by many language teachers. Although students in China are given instruction in dictionary use (Li 1997, p. 67), this seems to be the exception rather than the rule in most countries. Of 55 EAL teachers surveyed in Australia, for example, few talked about the uses that could be made of dictionaries in classroom work, and two-thirds did not refer their students to dictionaries for help with collocations or idioms (Miller 2008). Only 12 teachers in Miller's study had received even cursory training themselves in how to use dictionaries for teaching, although 38 gave their students some instruction in dictionary use.

The above study was conducted in 2005, when free online dictionaries were rare. Now that many dictionaries are freely available online, and many students have smart phones, there are more possibilities to use quality online dictionaries innovatively in classroom teaching. However, many teachers are reluctant for their students to use their phones in class for fear they may be exchanging messages or using social media rather than carrying out classroom activities. This is a pity, as online dictionaries are rich in all kinds of information for learners and present many opportunities for the teaching of idioms, especially in more specialised subject areas, where appropriate use of idioms helps learners to integrate into a speech community.

Although students do not always use online dictionaries in class, they frequently use them for private study. Research on dictionary use has examined, among other areas, microstructural features (e.g. Al-Ajmi 2002); dictionary skills (e.g. Atkins and Varantola 1998); language needs and reference skills (e.g. Béjoint 1981); the problems of dictionary workbooks (e.g. Campoy-Cubillo 2002); user needs (e.g. Hartmann 1986); EFL dictionaries (e.g. Nesi and Hail 2002); electronic dictionaries (e.g. Yuzhen Chen 2017); and online dictionaries (e.g. Dziemianko 2017). There is, however, room for research on how dictionaries are used by those wanting to know the meaning and use of idioms. Possible areas of research include the following questions:

- How do users look for idioms?
- How many lookup attempts do they make?
- Which headwords do they choose?
- How do they cope with idiom variability?
- What cross referencing/hyperlinking do they make use of?
- Do they look at the labelling material?

Although there are many studies on pedagogical dictionary use, therefore, there is still plenty of scope for studies which address how students use dictionaries in regard to idioms and other phraseological units.

2.3 Dictionaries and culture

The inclusion of some idioms in a dictionary, and the exclusion of others, reflects the fact that dictionaries are “closely related to power and politics and are important carriers of ideology” (Wenge Chen 2017, p. 871). The view of a culture portrayed in a dictionary may, therefore, be less neutral than one might expect, and, since idioms are a close reflection of culture, more studies should be undertaken to examine which dictionaries include which idioms. Even seemingly neutral idioms can provide cultural information. Take the following examples from American and British English:

- Then, *out of the blue*, he sold his house and left the country. (MW)
- And my winning ride came *out of the blue*; you can often win when you are least expecting to. (COBUILD)
- Even with a mysterious ex-lover who had turned up *out of the blue* after more than 16 years’ absence. (LDOCE)

First, we have the idea of home ownership and wealth (*he sold his house and left the country*—MW), indicating two privileges which may be impossible in other parts of the world. Then we have the idea of sports and competition (*my winning ride*—COBUILD). We do not know if this refers to a bicycle ride, a horse ride, or some other kind of sport, but its use as an example indicates that sport is valued by the speech community from which the example is taken. Finally, the fact that *a mysterious ex-lover* appears in the LDOCE example indicates a morally accepting, or at least neutral, stance on such a relationship, which may not be viewed the same way in other cultures. Idioms in dictionaries, therefore, give us an insight into a society’s ideology.

The idioms included in a dictionary can also give us an introduction to a society’s way of life, both past and present. In both English and Spanish, for example, idioms involving the word *donkey/burro* encompass meanings of hard work (*donkey work/trabajar como un burro*), but Spanish alone has frequent idioms which associate donkeys with riding and transport (e.g. *burro con dos albardas*—*saying something in two different ways to make the meaning clear*) (Korostenskiene and Tarnauskaite 2015). This may reflect the continuing use of donkeys in rural areas in Spain, which is not common in the UK. In a different example, Australian and British English include idioms associated with the game of cricket (e.g. *have a good innings*; *it’s not cricket*), which is closely linked to these cultures but not played widely in many other countries.

While idioms give us an insight into a culture, the translation of idioms is nevertheless problematic. An idiom in one language may seem to have an equivalent in another language because the words are the same; the meaning, however, may be different. For example, *bread and butter* in English refers to a means of earning a living, while the equivalent words in Polish, *chleb z masłem*, mean that something is easy (Szpila 2000, p. 80). Moreover, even when another idiom with an apparently equivalent meaning is found (e.g. *out of the blue* (English)/*à l’improviste* (French)), this does not necessarily entail that it will be used in the same way or have the same

connotations (Dobrovolskij 2000, p. 20). For instance, *out of the blue* occurs 17 times with the meaning *unexpectedly* in the Linguee parallel concordance (DeepL GmbH 2018b), but not once is the idiom *à l'improviste* given in the parallel translations. Similarly, *à l'improviste* appears 30 times in Linguee, but the English translations most commonly use *unexpectedly* or *unannounced*, rather than *out of the blue*. Such problems of equivalence in meaning and use therefore mean that register and usage notes are vital in any bilingual dictionary entry containing an idiom.

2.4 Three areas for further research

The areas of electronic and pedagogical lexicography, and dictionaries as vehicles of culture, are only the surface of a wide range of possible collaborations between the fields of phraseology and lexicography. I have taken many of my examples from online learner's dictionaries, as these are readily available to an international audience. However, there are, of course, many fruitful avenues of collaboration that may be pursued in relation to other kinds of dictionaries, and questions pertaining to some of these will be raised in the concluding section.

3 Suggestions for future research on intersections between phraseology and lexicography

Although I have referred only to idioms in this paper, and not to other types of phraseological units, there are, as we have seen, many connections between phraseology and lexicography. I would like to end this paper with an exhortation to readers to consider the following questions in relation to many different types of dictionaries:

- Which idioms do you include in L1 dictionaries in your country?
- Which idioms do you include in learner's dictionaries for your language or language variety?
- How do you choose these idioms?
- Do the idioms in your dictionaries reflect current usage?
- In what situations do you use idioms in your language (and do you have corpora to prove this)?
- Would a learner of your language find enough information in your dictionary to know how and when to use an idiom?
- Do the idioms in your dictionaries have helpful example sentences?
- Do idioms in your bilingual dictionaries have information that helps users to understand any differences between apparent idiom equivalents?
- Are idioms consistently presented in your dictionaries, or do they appear under a variety of different labels (e.g. *phrase*, *expression* and *idiom*)?

Phraseology and lexicography should have many links, but too often, as with many other disciplines, the insights between researchers are not shared. I hope that this paper may prove to be a starting point for much future collaboration.

References

- Al-Ajmi, Hashan. 2002. Which microstructural features of bilingual dictionaries affect users' look-up performance? *International Journal of Lexicography* 15 (2): 119–131.
- Applied Linguistics, University of Warwick. 2018. British Academic Spoken English <https://warwick.ac.uk/fac/soc/al/research/collections/base/>. Accessed 21 Feb 2018 (BASE).
- Atkins, B. T. Sue, and Krista Varantola. 1998. Monitoring dictionary use. In *Using Dictionaries: Studies of Dictionary Use by Language Learners and Translators*, ed. B. T. Sue Atkins, pp. 84–122. *Lexicographica. Series Maior*. Tübingen: Max Niemeyer Verlag.
- Béjoint, Henri. 1981. The foreign student's use of monolingual English dictionaries: A study of language needs and reference skills. *Applied Linguistics* 2 (3): 207–222.
- Cambridge University Press. 2018. *Cambridge dictionary*. <https://dictionary.cambridge.org/>. Accessed 21 Feb 2018 (CALD).
- Campoy-Cubillo, Mari Carmen. 2002. Dictionary use and dictionary needs of ESP students: An experimental approach. *International Journal of Lexicography* 15 (3): 206–228.
- Chen, Wenge. 2017. A discourse approach to critical lexicography: Methods, principles and applications. In *Asialex 2017. Lexicography in Asia: challenges, innovations and prospects. Proceedings of the 11th International Conference of the Asian Association for Lexicography*, ed. Hai Xu. Guangdong University of Foreign Studies: Center for Linguistics and Applied Linguistics, Guangdong University of Foreign Studies.
- Chen, Yuzhen. 2017. Dictionary use for collocation production and retention: A CALL-based study. *International Journal of Lexicography* 30 (2): 225–251. <https://doi.org/10.1093/ijl/ecw005>.
- Collins. 2018. *Collins English dictionary*. <https://www.collinsdictionary.com/dictionary/english/>. Accessed 21 Feb 2018 (COBUILD).
- Cowie, Anthony Paul. 1998. Phraseological dictionaries: Some East-West comparisons. In *Phraseology: theory, analysis, and applications*, ed. Anthony P. Cowie, 209–228. Oxford: Clarendon Press.
- DeepL GmbH. 2018. *Linguee*. <https://www.linguee.com/> Accessed 21 February 2018 (Linguee).
- Dobrovolskij, Dmitrij. 2000. Contrastive idiom analysis: Russian and German idioms in theory and in the bilingual dictionary. *International Journal of Lexicography* 13 (3): 169–186.
- Dziemianko, Anna. 2017. Dictionary form in decoding, encoding and retention: Further insights. *RECALL* 29 (3): 335–356. <https://doi.org/10.1017/S0958344017000131>.
- Frankenberg-Garcia, Ana. 2012. Learners' use of corpus examples. *International Journal of Lexicography* 25 (3): 273–296. <https://doi.org/10.1093/ijl/ecs011>.
- Grant, Lynn E. 2005. Frequency of 'core idioms' in the British National Corpus. *International Journal of Corpus Linguistics* 10 (4): 429–451.
- Hartmann, R.R.K. 1986. Dictionaries of English: The user's perspective. In *Dictionaries of English: Prospects for the record of our language*, ed. Richard W. Bailey, 121–135. Ann Arbor: The University of Michigan Press.
- Hartmann, R.R.K. 2003. Why lexicography needs a strong academic foundation. In *Dictionaries and Language learning: How can dictionaries help human and machine learning? Papers submitted to the Third ASIALEX Biennial International Conference*, ed. Minoru Murata, Shigeru Yamada, and Yukio Tono. Tokyo: Asian Association of Lexicography.
- Korostenskiene, Julija, and Greta Tarnauskaite. 2015. A contrastive analysis of equine idioms in English, Spanish and Lithuanian. *Verbum* 6: 123–134. <https://doi.org/10.15388/Verb.2015.6.8812>.
- Kwary, Deny Arnos. 2012. Adaptive hypermedia and user-oriented data for online dictionaries: A case study on an English dictionary of finance for Indonesian students. *International Journal of Lexicography* 25 (1): 30–49. <https://doi.org/10.1093/ijl/ecr008>.
- Li, Jie, and Norbert Schmitt. 2009. The acquisition of lexical phrases in academic writing: A longitudinal case study. *Journal of Second Language Writing* 18 (2): 85–102. <https://doi.org/10.1016/j.jslw.2009.02.001>.
- Li, Lan. 1997. Dictionaries and their users at Chinese universities: With special reference to ESP learners. In *Lexicography in Asia*, ed. Tom McArthur and Ilan Kernerman. Hong Kong: Password Publishers Ltd.
- Liu, Dilin. 2003. The most frequently used spoken American English idioms: A corpus analysis and its implications. *TESOL Quarterly* 37 (4): 671–700.

- Liu, Dilin. 2012. The most frequently-used multi-word constructions in academic written English: A multi-corpus study. *English for Specific Purposes* 31 (1): 25–35. <https://doi.org/10.1016/j.esp.2011.07.002>.
- Macmillan Publishers Limited. 2009–2018. *Macmillan dictionary*. <https://www.macmillandictionary.com/>. Accessed 21 Feb 2018 (MEDAL).
- Marshall, Philip J., Chris J. Lintott, and Leigh N. Fletcher. 2015. Ideas for citizen science in astronomy. *Annual Review of Astronomy and Astrophysics* 53 (1): 247–278. <https://doi.org/10.1146/annurev-astro-081913-035959>.
- Merriam-Webster, Incorporated. 2018. *Merriam-Webster learner's dictionary*. <http://www.learnersdictionary.com/>. Accessed 21 Feb 2018 (MW).
- Miller, Julia. 2008. Teachers and dictionaries in Australia: Is there a need to train the trainers? *TESOL in Context* 17 (2): 11–19.
- Miller, Julia. 2011. *Never judge a wolf by its cover: An investigation into the relevance of phrasemes included in advanced learners' dictionaries for learners of English as an additional language in Australia*. Flinders University, Adelaide, Australia. [Unpublished PhD thesis].
- Miller, Julia. 2013. Birds of a feather don't always flock together: User problems in identifying headwords in online English learner's dictionaries. *Lexikos* 23: 273–285. <https://doi.org/10.5788/23-1-1215>.
- Moon, Rosamund. 1998. *Fixed expressions and idioms in English*. Oxford: Clarendon Press.
- Nesi, Hilary, and Richard Haill. 2002. A study of dictionary use by international students at a British university. *International Journal of Lexicography* 15 (4): 277–305.
- O'Keeffe, Anne, Michael McCarthy, and Ronald Carter. 2010. Idioms in everyday use and in language teaching. In *Applied linguistics in action: A reader*, ed. Guy Cook and Sarah North, 212–230. New York: Routledge.
- Oxford University Press. 2012. Oxford corpus of academic English. Accessed via Sketch Engine 29 May 2018 (OCAE).
- Oxford University Press. 2018. *Oxford learner's dictionaries*. <https://www.oxfordlearnersdictionaries.com/>. Accessed 21 Feb 2018 (OALD).
- Pearson. 2018. *Longman dictionary of contemporary English online*. <https://www.ldoceonline.com/>. Accessed 21 Feb 2018 (LDOCE).
- Rundell, Michael. 2016. Dictionaries and crowdsourcing, wikis and user-generated content. In *International handbook of modern lexis and lexicography*, ed. Patrick Hanks and Gilles-Maurice de Schryver. Berlin: Springer.
- Siepmann, Dirk. 2008. Phraseology in learners' dictionaries: What, where and how? In *Phraseology in foreign language learning and teaching*, ed. Sylviane Granger and Fanny Meunier, 185–202. Philadelphia/Amsterdam: John Benjamins Publishing Company.
- Simpson, Rita, and Dushyanthi Mendis. 2003. A corpus-based study of idioms in academic speech. *TESOL Quarterly* 37 (3): 419–441.
- Szczepaniak, Renata, and Robert Lew. 2011. The role of imagery in dictionaries of idioms. *Applied Linguistics* 32 (3): 323–347. <https://doi.org/10.1093/applin/amr001>.
- Szpila, Grzegorz. 2000. False friends in phraseology: An English Polish contrastive study. In *International Conference on English and American Literature and Language: Tradition and postmodernity*, pp. 77–82. Krakow: Jagiellonian University Press.
- The University of Michigan English Language Institute. 2007. Michigan corpus of academic spoken English. Accessed 21 Feb 2018 (MICASE).