Studies in Second Language Learning and Teaching
Department of English Studies, Faculty of Pedagogy and Fine Arts, Adam Mickiewicz University, Kalisz
SSLLT 13 (4). 2023. 883-902
https://doi.org/10.14746/ssllt.40219
http://pressto.amu.edu.pl/index.php/ssllt

The Routledge handbook of vocabulary studies: *A study in micro-bibliometrics*

Paul Meara ✉
Swansea University, UK
University of Oxford, UK
https://orcid.org/0000-0001-8863-5374
p.m.meara@gmail.com

Abstract

This paper reports a bibliometric analysis of two small data sets: a set of 34 papers that make up *The Routledge handbook of vocabulary studies* (Webb, 2020) and a set of papers dealing with second language (L2) vocabulary research taken from a single journal *Frontiers in Psychology.* Bibliometric maps based on author co-citations in these two data sets are presented and compared. Although the two data sets are comparable in terms of size, they appear to be very divergent. In particular, the significant sources identified in *The Handbook* map seem to play a relatively minor role in the *Frontiers* map. The obvious conclusion is that *The Handbook* is not as representative of L2 vocabulary research as its title might lead us to believe. The paper argues that micro-bibliometric studies like this one can sometimes highlight features that are lost in the more traditional large-scale bibliometric approach.

Keywords: micro-bibliometrics; L2 vocabulary; co-citation analysis

1. Introduction

Most of the recently published papers on bibliometrics in applied linguistics are relatively large-scale studies, which use big data bases as their information source. The deployment of databases such as the Web of Science (WoS) and Scopus

means that it is relatively easy to ask questions such as who is publishing in a given time span, where are they publishing, what keywords are they using, who are they collaborating with, who are they cited by, and so on. Once you understand how to query the data bases, questions such as these can be answered with the click of a few buttons, and work that only a few years ago would have taken many months of hard slog can now be completed within a few days. However, there is a downside to this: It encourages researchers to look at the bigger picture rather than the many smaller ones which collectively make up the bigger picture. So, for example, it becomes easy to ask questions about the contents of a whole journal, but much less easy to tease out what is happening in the many different fields that the journal may cover.

I first got interested in this issue after reading Lei and Liu's (2019) excellent account of research in the journal *System.* My own research largely involves vocabulary acquisition. Lei and Liu had noted, among other things, that vocabulary research was an emergent subfield in their data. Nation, Schmitt and Laufer, names instantly recognizable to vocabulary researchers, all get a mention in the Lei and Liu (2019) list of highly cited authors, and they note that *vocabulary learning* is an area of interest in the teaching of specific skills. They also note that Nation's two textbooks *Teaching and Learning Vocabulary* (1990) and *Learning Vocabulary in Another Language* (2001/2013), and Read's *Assessing Vocabulary* (2000) are listed among the most frequently cited references in the journal. However, what does not emerge from Lei and Liu's analysis is any sense of the dynamics within vocabulary research over the time-scale they are interested in. In contrast, Meara (2020) is a more detailed analysis of the second language (L2) vocabulary research published in *System.* It showed that there are significant changes in the vocabulary research published in this journal, and suggested that much of this research can be seen as the consolidation of a *first paradigm* in vocabulary research. This insight is lost in Lei and Liu's (2019) larger study, and only emerges when you look at the fine details of the bibliometrics.

The present paper is another example of how "micro-bibliometrics," the bibliometrics of small data sets, might be a useful tool for applied linguistics. The paper compares the research reported in a large collection of recent research papers, *The Routledge Handbook of Vocabulary Studies* (Webb, 2020) with a corpus of papers from *Frontiers in Psychology*, an open access journal that has published a surprisingly large number of high quality papers addressing L2 vocabulary acquisition. The paper argues that there are some interesting discrepancies between these two data sets.

2. Methodology

The approach used in this paper relies on an author co-citation analysis. This approach may not be familiar to readers, so it is described in detail in this section

of the paper. Readers who are already familiar with co-citation analysis can skip to the next section.

The co-citation methodology was developed by Small in an influential paper published in 1973. Essentially, we make a list of sources who are co-cited in a data set, and use these co-citations to identify "invisible colleges" in the data – sources who share common reference points and can be assumed to work with the same concepts (de Solla Price, 1965; de Solla Price & Beaver, 1966).

The method consists of a number of discrete steps. Firstly, we identify a data set, a collection of papers which define the area that we are working with. Next, we extract a list of all the authors cited by each paper in the data set. Authors who are cited together in a single paper are said to be co-cited. The next step is to draw up an ordered list of sources according to how frequently they are cited in the data set. This step usually generates a surprisingly long list of sources, but about 80% of these will only be cited in a single paper. Custom and practice is to ignore these sources, and to work with a smaller number of sources that are cited frequently in the data set.

In the second step of the analysis, then, we want to identify a manageable number of sources that are considered to be important in the data set, and to do this, we set an inclusion threshold which excludes sources cited only infrequently. This step is basically an arbitrary judgment, but normally we pick a threshold which gives us about 100 highly cited authors for further analysis. We then build a matrix which shows the number of times each of these significant sources is co-cited with another significant source. Even though this selection involves only a fraction of the full data set, it still generates colossal quantities of data. Each of the 100 authors can potentially be co-cited with 99 other authors, so our matrix contains 9,900 potential co-citation cases. In practice, of course, it is unusual to find co-citation matrices where every source is cited with every other source. Normally, patterns emerge, where some sources will frequently be cited together, while other source pairs are only rarely co-cited. Usually, the co-citation matrix is only sparsely populated, and will contain a number of empty cells. Although this makes the matrix easier to manipulate than we might expect, searching a co-citation matrix is a difficult and error-prone operation, and for that reason it is usual to present the data in the form of a co-citation map which visualizes the way authors are cited in the data set. In these maps, authors are shown as nodes and co-citations are shown as edges. The maps presented in this paper are generated by the Gephi program (Bastian et al., 2009). Gephi is able to find clusters of authors who are commonly cited together, and these clusters are considered to be indicative of common research themes in the data set.

In this paper, I will compare two data sets using this co-citation methodology. The first data set comes from *The Routledge Handbook of Vocabulary Studies*

(Webb, 2020) (henceforth, *The Handbook*). The second data set comprises a set of papers from a single journal, *Frontiers in Psychology*, published between 2016 and 2020. The rationale for this comparison is laid out in the following sections.

3. Results

3.1. Data: *The Routledge Handbook of Vocabulary Studies*

Let us begin with an analysis of *The Handbook* in its own terms. This volume consists of 34 research chapters, plus an editorial introduction. A complete list of the chapters, providing a good overview of the contents, will be found in Appendix A. Most vocabulary researchers would recognize this list as a reasonable coverage of current topics in the field. They would also recognize in the author list a large number of researchers in L2 vocabulary studies who have published extensively, and would be considered as authoritative sources.

Even with a small data set, such as *The Handbook* data, generating a co-citation matrix is a bigger job than it sounds, mainly because at the time of writing it was not possible to extract the citation data automatically, and the lists of cited sources needed to be extracted (painstakingly) by hand. I collected citation data from the 34 research chapters; Webb's introductory editorial was not included in the analysis mainly because it made a point of citing all the authors of the other 34 chapters, and this skews the data in a way we want to avoid. The resulting lists identified each source cited in each paper; each source is counted only once, irrespective of how many times it is cited in a single paper, and where a paper has many authors, each one of them is fully counted. The citation lists from the 34 research chapters were extensive. I identified a total of 1,537 unique sources in this data set, which means that each of the chapters is citing about 45 unique sources. Table 1 summarizes this data.

Table 1 The number of sources cited N times in *The Handbook* data set. The table shows that only one source is cited 30 times, while 1087 sources are cited only once

# citations	30	29	28	27	26	25	24	23	22	21	20	19	18	17	16
# sources	1	1	1						1	1	1				1
# citations	15	14	13	12	11	10	9	8	7	6	5	4	3	2	1
# sources	2	1	1	5	5	8	5	12	12	9	35	48	77	223	1087

71% of these sources are cited only once in the data set, but some sources are cited in multiple chapters. However, the table clearly shows that there is a lot of variation in the way sources are cited. A handful of sources are cited in almost every

paper, another handful are cited in about two thirds of the chapters. A total of 102 sources are cited in at least five chapters. This is a somewhat unusual distribution.

The ten most cited sources in this data set are: Schmitt (30), Nation (29), Webb (28), N. C. Ellis (22), Laufer (21), Meara (20), Cobb and Coxhead (16), Hulstijn (15) and Pellicer-Sánchez (15). These names will be instantly recognizable to anyone working in L2 vocabulary research. It is worth pointing out here that most of these highly cited sources are based in CANZUK (Canada, Australia, New Zealand and the United Kingdom) countries; the exceptions are Laufer and Hulstijn, based in Israel and the Netherlands, respectively. Only infrequently does *The Handbook* cite sources who do not publish in English.

The second step in our analysis requires us to identify a set of about 100 sources who are frequently cited in the data set, and discard the less frequently cited sources. This process discards some data, but it greatly simplifies the reporting process. The figure of 100 sources seems to be an arbitrary convention rather than a theoretically motivated choice. Its main purpose is to make the visualizations of the data easier to handle. The nearest we can get to this conventional figure in *The Handbook* data set are the 102 sources who are cited in at least five chapters. The analysis that follows is based on this figure.

Figure 1 shows Gephi's analysis of *The Handbook* data set. This map requires some explanation. The map is based on the 99 sources who are cited at least five times in the data set. These sources are identified as labeled nodes in the map, and the lines between these nodes (sometimes called edges) identify co-citation links between the nodes. Thicker edges indicate more frequent co-citation links. Three additional sources (Schmitt, Nation, and Webb) are co-cited with almost every other source in the map, and they completely pervade the network of co-citations in the data set. However, the extent of this dominance means that they do not actually make a distinctive contribution to the structure of the network. Consequently, we can greatly simplify our analysis by eliminating these three super sources. This leaves us with a "donut" map made up of the remaining 99 sources. We can further simplify our map by eliminating the weakest edges linking the nodes.

Figure 1 shows Gephi's interpretation of the data when we take into account only edges which appear at least five times in the data. The nodes are sized according to their betweenness centrality as larger nodes play a more important role in structuring the network. The majority of these sources make up a single connected structure in Figure 1, but there are a number of "outliers." These are sources who are frequently cited in the data set, but are not consistently co-cited with the same other sources. This means that their links to other sources in the network are not strong enough for them to meet our threshold of five co-citations.

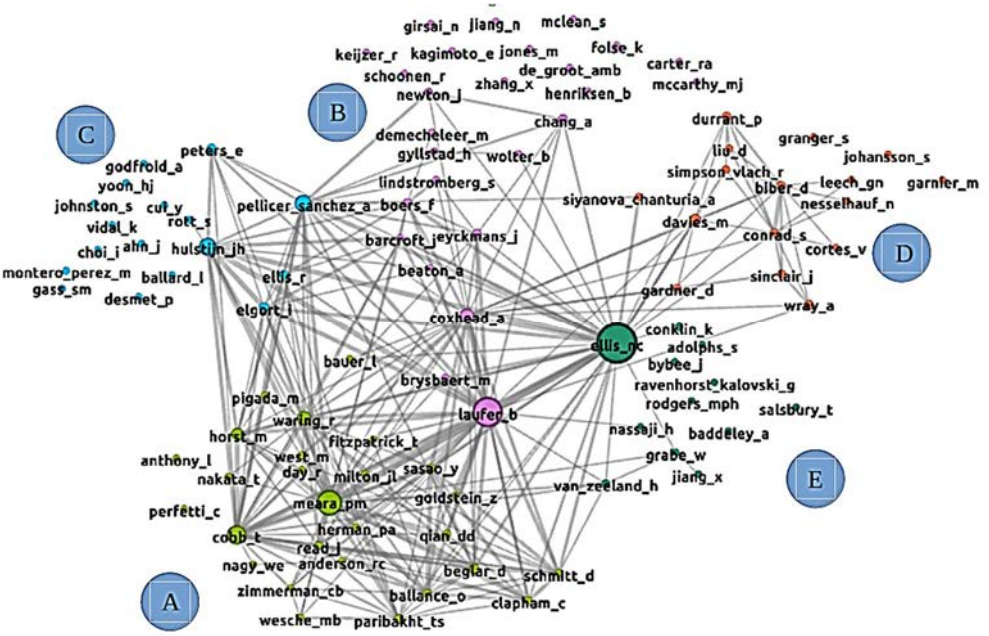

Figure 1 A simplified version of *The Handbook* data (Thresholds for inclusion: 5 citations in the data set; edges with strength < 5 removed. Three nodes, Schmitt, Nation, and Webb, are co-cited with almost all the other sources, and they have been removed from the map in order to make its distinctive structure more obvious)

3.2. Analysis: *The Handbook* data set

Gephi identifies five clusters in the data. The largest and most densely connected cluster in the map is Cluster A, with 27 sources. The two main sources in this cluster are Meara and Cobb, but all the interconnections within this cluster are very strong, and there are only two detached outliers, Anthony and Perfetti. There seem to be two main foci in this cluster: Waring, Horst and Day are all concerned with reading in an L2, and frequently co-cited with Nagy, Anderson and Herman, all major sources in L1 reading research. The second sub-cluster is concerned with vocabulary testing. Three sub-themes can be identified here: Beglar, Diane Schmitt, Clapham and Ballance focus on vocabulary size; Paribakht and Wesche are an important testing methodology for partial vocabulary knowledge, while Read and Zimmerman are best known for their important work on vocabulary testing theory.

Cluster B, the second largest cluster, consists of 13 core sources, together with 13 detached sources who are frequently cited in the data set, but have co-citations which are varied, and not strong enough for them to reach the threshold

of five appearances. This cluster is dominated by Laufer, who emerges as a key figure in this map. This will not surprise anybody who is familiar with the main lines of L2 vocabulary research. The main focus of the sources in this cluster seems to be vocabulary pedagogy, though some of the sources do not obviously fit into this theme. It is possible to identify some sub-groupings within the cluster: the main sub-group is based around Boers and his colleagues in Belgium and New Zealand.

Cluster C identifies 17 sources, but only five of these have strong connections with the rest of the map. I think the main focus of this cluster is a methodological one in that several members of this cluster (Pellicer-Sánchez, Elgort and Godfroid) make use of eye-tracking methodology to investigate L2 vocabulary. Godfroid, Ahn, Choi, Ballard, Cui, Johnston, Lee, Sakar and Yoon co-authored a particularly influential paper that appeared in 2018 (Godfroid et al., 2018). Montero-Pérez, Peters and Desmet also work with eye-tracking, but mainly in the context of captioned TV input for language learning.

Cluster D, with 16 members, again seems to be a methodological cluster. Here, however, the methodological resource is large scale corpora. The key theoretical figures here are Biber and Sinclair. Wray is an important source on formulaic sequences.

The final cluster, Cluster E, is odd in that it consists of only five connected nodes: N. C. Ellis, Conklin, Grabe, Jiang, and van Zeeland. Five other disconnected sources sit on the edge of this cluster, while Nassaji has a strong co-citation link only with Laufer in Cluster B. I think the main theme that ties this cluster together concerns input in vocabulary learning tasks. However, it is notable that some of the sources in this cluster are not themselves actively involved in L2 vocabulary research, though their work does inform some of the other sources. The best example of this is Baddeley, best known for his work on memory; Bybee is mostly cited in connection with her work on morphologically complex words; Adolphs is a corpus linguist. N. C. Ellis is by a long way the key figure in the overall graph, largely, I think, because he has very many co-citation links with the sources in Cluster D, which is otherwise linked only loosely with the rest of the map. His work is mostly cited in connection with implicit/explicit learning methodologies. This cluster is also notable in that it includes a high proportion of North American sources.

Most vocabulary researchers would easily recognize the picture that emerges from this map. The practical pedagogical concerns that we find in Cluster B are a continuation of a research strand that we can trace back at least to the early 20th century (cf. for instance Henmon, 1917). The themes that make up Cluster A are more recent, though they can be traced back at least as far as Richards' (1976) paper and early attempts to be specific about what it means to "know" a word. Work on formulaic sequences and multiword vocabulary has been a feature of the field at least since the early 1980s (cf. Pawley & Syder,

1983), and is becoming increasingly important. The one new methodological theme we can identify in this map is the work that uses eye-tracking as its basic method of data collection.

Readers should bear in mind that the map in Figure 1 has deliberately been simplified to make the clustering structure more obvious. In reality, the work reported in *The Handbook* is entirely dominated by three widely co-cited sources: Schmitt, Nation, and Webb. The appearance of Nation's (1990) book was a defining moment in L2 vocabulary research, and the research agenda that this work outlined has largely defined the way the modern research has developed. Nation's influence is also apparent in the large number of his former students who appear in the map in Figure 1. Webb, who is one of these former students is, of course, the editor of the current volume, and this may partly explain why he is also a key source in *The Handbook* data set. Schmitt is important because he is the focus of vocabulary research that developed at the University of Nottingham in the UK. Much of Schmitt's own work built on earlier work by colleagues at Nottingham, particularly Carter and McCarthy, Adolphs and Conklin, who all appear as significant sources in the map. Their work made extensive use of corpus research in English. Schmitt, like Nation, was also fortunate in that he was able to build an extraordinarily impressive group of graduate students, whose research is characterized by a willingness to work with sophisticated new research techniques. Some of these former students appear in Figure 1 in their own right: Pellicer Sanchez in Cluster C, van Zeeland in Cluster E, and Pigada in Cluster A.

Without the simplifications, the map of the research reported in *The Handbook* is actually much more homogeneous and more tightly connected by co-citation than Figure 1 implies, and this can be seen in Figure 2. The figure shows an unsimplified map in which I have restored the co-citations of Schmitt, Nation, and Webb.

The analysis in Figure 2 clearly shows that *The Handbook* data set is basically a single dense network, rather than a structure with a looser composition. The analysis identifies three clusters, rather than five. The interconnections between these clusters are fairly uniform, but there is perhaps a suggestion that the cluster at the right hand edge of the map is more closely associated with Schmitt than with the other major sources. This cluster is mainly a set of sources that deal with corpus linguistics – the Nottingham sources identified earlier. The important thing here is the homogeneity exhibited by the map. All the sources are closely interconnected.

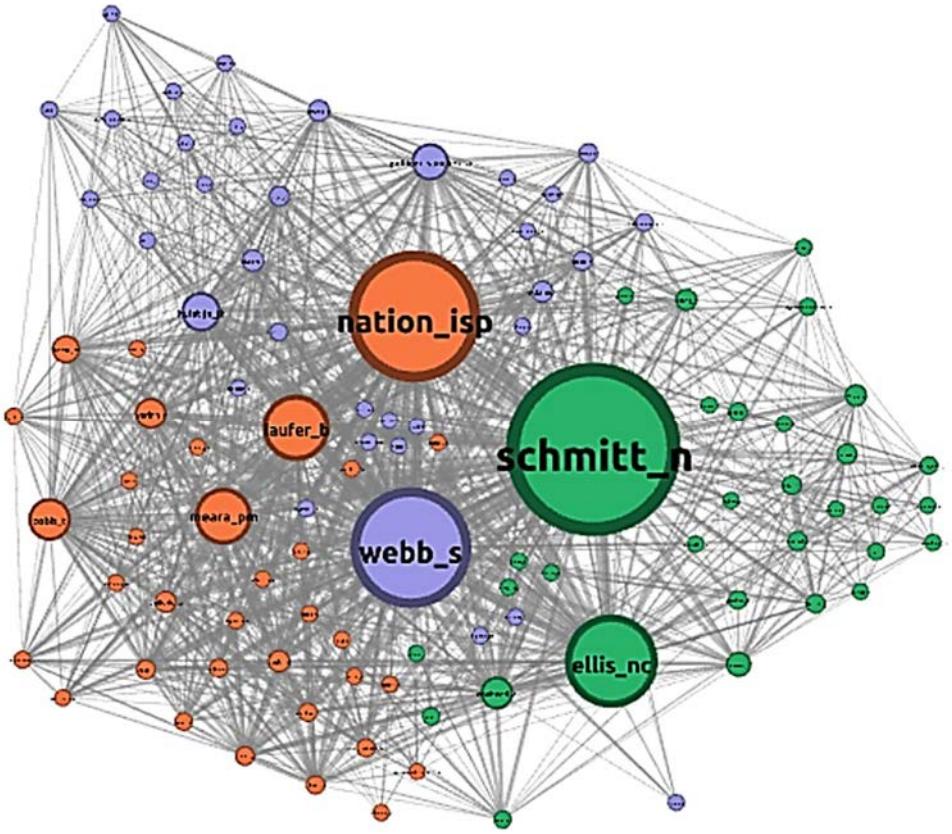

Figure 2 *The Handbook* data including Schmitt, Nation, and Webb (Thresholds for inclusion: 5 citations in the data set, edge weight 4+)

3.3. Data: *Frontiers in Psychology*

The question which now arises is whether the picture which emerges from the co-citations in *The Handbook* is an accurate and faithful representation of what is really going on in current vocabulary research. It is not immediately obvious how we can answer this question, or what other data set can be used as a comparator. The obvious solution would be to take a random selection of papers published in 2020, and test whether an author co-citation analysis of this new data set closely resembles the structure that we find in Figure 1. I rejected this approach on the grounds that random sampling would not select high-quality research such as we find in *The Handbook*. A depressingly large amount of the research recorded in the VARGA database, for example, Meara (n.d.) is of rather poor quality. Because of this, a random selection of papers would be biased towards work

that is not theoretically or methodologically innovative, and it would lack editorial consistency. After some thought, I decided that the best approach would be to take a set of thirty-four research papers from a single journal.

This left me with the problem of which journal to choose as the source for the data set. Few applied linguistics journals publish many L2 vocabulary-related papers in a relatively short space of time. *Applied Linguistics* published only three relevant papers in 2020; *System*, usually a good source for this type of material, published only two, and most of the major research journals in applied linguistics were ruled out of consideration by this criterion. The best source might have been *Vocabulary Learning and Instruction,* edited by Raymond Stubbe, but this is a relatively new journal that is still finding its feet. Eventually, after more considerable thought, and several false starts, I decided to work with a set of 34 papers from the journal *Frontiers in Psychology.* This journal may seem an odd choice, but it has published a surprisingly large number of papers that report innovative research on L2 vocabulary acquisition, and I have been consistently impressed by the range of topics it has covered and the quality of the writing. *Frontiers in Psychology* has been described as a *predatory journal,* one of a number of new journals that pursue a policy of open access publishing, where authors bear the cost of publication in return for readers being granted free access to the paper once it is published. Many such journals appear to have rather low standards, but, in my view, this criticism does not apply to *Frontiers in Psychology.* It has consistently published high-quality, challenging research, much of it authored by big names in applied linguistics. This is a journal in which I would be very happy to publish my own research.

I would have liked to have a set of 34 papers that were all published in a single year, preferably 2020. This would have given me a data set that exactly matched the 34 papers published in *The Handbook,* but this proved to be impossible. In this event, I was forced to adopt a more lenient selection criterion that included papers published between 2016-2020. The 34 papers I identified for the *Frontiers in Psychology* data set are listed in Appendix B. Like *The Handbook* chapters, the list of topics identified by these titles is extensive and it includes a wide range of L2 vocabulary-related topics.

3.4. Analysis: The *Frontiers in Psychology* data set

The first thing to note about this data set is that even at the most superficial level it looks very different from *The Handbook* data set. Table 2 reports the number of sources that are cited N times in this data. We can see immediately that this distribution is very different from the distribution reported in Table 1.

The *Frontiers* data set identifies 3617 unique sources, more than twice the number of sources that were identified in the *Handbook* data set. This is partly because the papers in the *Frontiers* data set frequently cite papers with multiple authors, a feature which is missing from the *Handbook* data set. Furthermore, none of these sources are cited in the majority of the papers in the data set.

Table 2 The number of sources cited N times in the *Frontiers in Psychology* data set

# citations	14	13	12	11	10	9	8	7	6	5	4	3	2	1
# sources				1		2	4	10	10	27	56	124	472	2911

The table shows that only one source is cited 11 times in the data set, while 2911 sources are cited only once. The most cited source in the *Frontiers in Psychology* data set is cited only 11 times, and the data contains no sources who are cited universally. 80% of the sources are cited only once, and a greater proportion of the sources are cited only two or three times, a much larger share than we find in *The Handbook* data. The most frequently cited sources are Kroll (11 citations), Bates and Bialystok (9), Baayen, Brysbaert, Perfetti and Rodríguez-Fornells (8). Equally striking is that none of these sources appears in the list of highly cited sources that accompanies Table 1. In fact, if we extend this comparison to include the 58 most frequently cited sources in *The Handbook* and the 54 most frequently cited sources in the *Frontiers in Psychology* data set, there are only *five* sources that appear in both lists. Surprisingly, only R. C. Anderson, Brysbaert, Hulstijn, Meara, and Schmitt share this distinction. The obvious interpretation here is that the two data sets are referencing very different sources, and perhaps addressing very different questions. And indeed, we can see this very clearly in Gephi's analysis of the *Frontiers in Psychology* data set shown in Figure 3.

There are some problems that make it difficult to generate a map which is directly comparable to the map shown in Figure 1. In order to get close to the conventional figure of 100 nodes, it was necessary to include all sources in the *Frontiers in Psychology* data set that were cited at least four times (the equivalent threshold applied in Figure 1 was five citations). It was also necessary to apply a generous criterion for the inclusion of edges linking the nodes. All edges that appear only once in the data set were discarded, and the analysis shown in Figure 3 is based on the remaining edges that appear at least twice. In Figure 3, the edges of weight 2 have been eliminated in the interests of simplicity, so each edge in Figure 3 represents at least three co-citations (the equivalent threshold in Figure 1 was five or more co-citations). One implication of these adjustments is that it is much easier for a source to be included in the *Frontiers in Psychology* co-citation map than it is for a source to be included in *The Handbook* map. A

map of the *Frontiers in Psychology* data that used the same inclusion criteria as Figure 1 would contain only seven nodes joined by 20 edges.

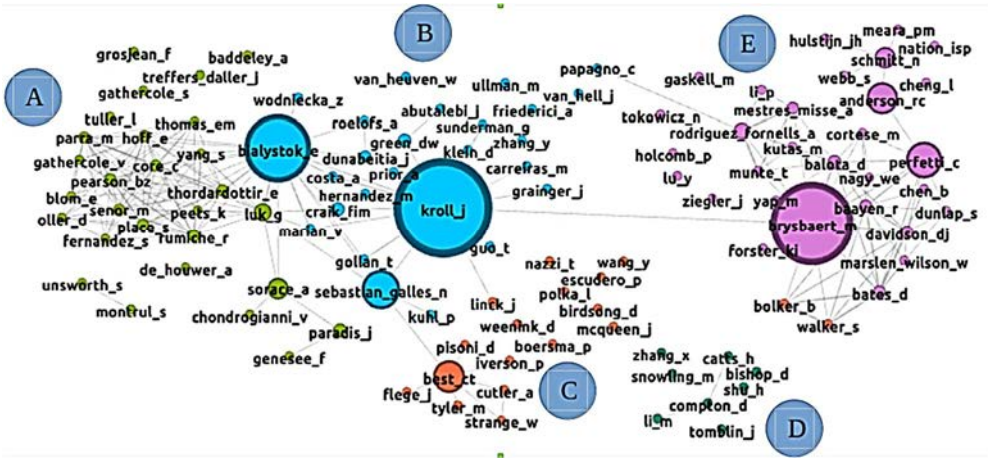

Figure 3 A simplified version of the *Frontiers in Psychology* data (Thresholds for inclusion: 4 citations in the data set; edges with strength < 3 removed. Nodes are sized according to their betweenness centrality)

Gephi identifies five clusters in the data set that is displayed in Figure 3. Cluster A on the left hand edge of the map is clearly a cluster that deals mainly with vocabulary acquisition in bilingual children, rather than ordinary L2 learners. This cluster is made up of a core of 17 sources who are very frequently co-cited, a semi-detached group centered on Sorace, a detached group consisting of Unsworth and Montrul, and five fully detached sources. Sources in this cluster have very strong co-citation links with Bialystok, a significant figure in Cluster B. I think this alignment mainly grows from a long-term collaboration between Bialystok, Craik and Luk.

Many of the sub-themes that can be identified in Cluster A also make an appearance in Cluster B, the central cluster in the map. The cluster consists of 20 linked sources plus six singletons with ties too weak to reach the inclusion threshold. The cluster has close ties with Cluster A, as already noted. The main differences between the two clusters are that Cluster B is more concerned with vocabulary skills of older bilinguals, and the sources tend to be more interested in formal models than are the sources that appear in Cluster A.

Cluster C, dominated by Best, appears to be a set of sources whose main concern is spoken word recognition in an L2, and the role that phonology plays in these processes.

Cluster D is the smallest cluster identified by Gephi in this data set. It consists of two loosely linked sources (Compton and Catts) and a group of six disconnected

singletons. The sources in this cluster are mainly cited in connection with L2 reading disorders, with a particular focus on reading Chinese characters.

Finally, Cluster E, the largest cluster, at the right hand edge of the map, is massively dominated by Brysbaert, who provides an important connecting link with Cluster B. The sources in this cluster are mainly concerned with experimental studies of written word recognition, and this work is informed by a series of methodological studies published by Brysbaert in the journal *Behavior Research Methods*. Also important in this cluster are some sources who figure because they identify statistical methods used to process the behavioral data (Bates, Bolker, Walker, Baayen, for example). The remaining sources in this cluster fall into four main subgroups. We have a large sub-cluster who work on the neurolinguistics of word recognition and vocabulary learning, a small sub-cluster who are concerned with reading in an L2, and a small group whose main focus is the effect of morphology on vocabulary acquisition. We also have a small sub-group consisting of Hulstijn, Meara, Nation, Schmitt, and Webb, loosely related to the rest of the cluster only via a co-citation link that connects R. C. Anderson and Perfetti, both cited here for their work on L1 reading development.

4. Discussion

Three main points arise out of a comparison between the two maps shown in Figure 1 and Figure 3. Firstly, the two maps do indeed look very different. Figure 1, and more specifically Figure 2, *The Handbook* maps, show a densely connected set of sources, almost always co-cited with three dominant super-sources: Schmitt, Nation, and Webb. There are distinguishable clusters within this data set, but we have to work quite hard to tease them out. Only the sources that work on multi-word, formulaic vocabulary stand out from the pack, and even that distinction is largely lost if we include the co-citations with Schmitt, Nation, and Webb. I have suggested elsewhere (Meara, 2020) that a structure of this sort might be typical of a research agenda that is heavily influenced by a single paradigm, which encourages researchers to focus on a relatively small set of problems, and investigate them using a limited range of methodologies. In some contexts, this might appear to be a strength, but it is also possible to see it as a problem if you think that the focus is becoming increasingly narrow and impervious to relevant developments in related fields. In contrast, the *Frontiers in Psychology* map, Figure 3, looks very different. Here we have five clearly distinguishable clusters that identify different but related themes, and it is obvious that each cluster is aware of the work that informs the other clusters. This figure suggests that the map is showing us a “small world” of research, where most of

the co-citations lie within each cluster, and there are relatively few long-range co-citation links between the clusters. Long-range, between-cluster co-citations can be found, but they are generally weak. In short, the research mapped in Figure 3 seems to be much more varied and less homogeneous than the work that is described by *The Handbook* maps. It is difficult to say whether this is a good thing or not. My gut feeling is that it is not. *The Handbook* data set strikes me as relatively inward looking, and it concerns me that this might reflect an unwillingness to engage with problems outside the dominant paradigm.

The second point to emerge from comparing the maps is that they highlight very different concerns. The most strongly linked cluster in Figure 3, for example, is Cluster A, a set of sources concerned with the vocabulary development of bilingual children. This work is barely cited at all in *The Handbook* data set, despite the obvious relevance of the topic to L2 vocabulary acquisition in general. Similarly, the focus on formal models that we find in Cluster B in Figure 3 is also not referenced in *The Handbook* data set. Both data sets flag up innovative research methods, eye-tracking studies in Figure 1, and brain scanning methods in Figure 3, but it is difficult to avoid the conclusion that the work mapped in Figure 3 is much better founded, and theoretically better motivated than the smaller-scale eye-tracking studies reported in *The Handbook* map. On the other hand, there is not a great deal of linguistic research being cited in the *Frontiers in Psychology* map. Figure 3 does identify some sources whose work is concerned with phonology, and there is a specific focus on learning and using vocabulary in tone languages, but the important sources that deal with corpus linguistics identified in Cluster D in the *Handbook* map do not figure in the *Frontiers in Psychology* map.

The third feature that needs comment is that there is indeed some overlap between the two maps, but the extent of this overlap is rather limited. We have already noted that only a handful of sources have a significant presence in both maps. Importantly, the five core sources cited in *The Handbook*, Schmitt, Nation, Webb, Meara, and Hulstijn, do not have anything like the same status in the *Frontiers in Psychology* map. These five sources together form a nearly isolated sub-cluster in the *Frontiers in Psychology* map, and it is difficult to avoid the conclusion that these sources, whom many readers would identify as the leading figures in current vocabulary research, play a relatively small role in the much larger L2 vocabulary research enterprise that we have identified in *Frontiers in Psychology*. Is this something that we should be worried by? Again, it is difficult to tell, but Figure 3 suggests that it would not take much to sever the links between psychological research on L2 vocabulary learning and work on L2 vocabulary learning that takes place within an applied linguistics tradition. This is surely something that should set alarm bells ringing.

Of course, we could argue that comparing a collection of *The Handbook* papers to a collection of papers from a specific journal is a bit like comparing apples with oranges. Handbooks play a very distinctive role in the research literature, stressing consensus rather than innovation, reviewing the status quo rather than breaking new ground. For that reason, we might expect a bibliometric map based on a handbook collection to look rather different from a map based on a set of up-to-the-minute research papers. Nonetheless, the extent of the differences I have sketched out here is strikingly stark. It might be that my choice of *Frontiers in Psychology* was an unfortunate one, and that other journals might have produced equally diverse outcomes. This is an empirical question, but unfortunately one which is beyond the scope of this paper. My hope is that the data presented here might encourage vocabulary researchers who work in the tradition of *The Handbook* to become more aware of the many rich strands of vocabulary research that are to be found in other research traditions.

5. Conclusion

This paper is a rather unusual bibliometric study in that it is not based on a huge data set. It uses only two sources, each made up of only 34 research studies, and it has used only one of the many bibliometric tools available to researchers, co-citation analysis as described in Small's influential 1973 paper. It has not asked any of the standard questions that have become commonplace in bibliometric studies: who published the most papers, where they are based, who they collaborate with, which journals they publish in, and so on. Nor does it focus on how the pattern of publications changes over time. Nevertheless, I hope that this paper illustrates how a micro-bibliometric approach can highlight some interesting features of the current research. I hope too that readers will agree that engaging with small-scale bibliometrics in this way, interpreting the maps in light of a detailed knowledge of a field, can sometimes be just as illuminating and thought-provoking as the more traditional massive overview.

References

Bastian, M., Heymann, S., & Jacomy, M. (2009). Gephi: An open source software for exploring and manipulating networks. *Proceedings of the International AAAI Conference on Web and Social Media, 3*(1), 361-362. https://doi.org/10.1609/icwsm.v3i1.13937

Carter, R., & McCarthy, M. (Eds.) (1988). *Vocabulary and language teaching.* Longman.

De Solla Price, D. (1965). Networks of scientific papers: The pattern of bibliographic references indicates the nature of the scientific research front. *Science, 149*(3683), 510-515. https://doi.org/10.1126/science.149.3683.510

De Solla Price, D., & Beaver, D. (1966). Collaboration in an invisible college. *American Psychologist, 21*(11), 1011-1018. https://doi.org/10.1037/h0024051

Godfroid, A., Ahn, J., Choi, I., Ballard, L., Cui, Y., Johnston, S., Lee, S., Sarkar, H., & Yoon. H. (2018). Incidental vocabulary learning in a natural reading context: An eye-tracking study. *Bilingualism: Language and Cognition, 21*(3), 563-584. https://doi.org/10.1017/s1366728917000219

Henmon, V. A. C. (1917). The measurement of ability in Latin. Part 1. Vocabulary. *Journal of Experimental Psychology, 8*(9), 515-538. https://doi.org/10.1037/h0075900

Lei, L., & Liu, D. (2019). The research trends and contributions of *System's* publications over the past four decades (1973-2017): A bibliometric analysis. *System, 80,* 1-13. https://doi.org/10.1016/j.system.2018.10.003

Meara, P. M. (n.d.). *Varga ~ bibliographical resources in vocabulary acquisition*. https://www.lognostics.co.uk/varga/

Meara, P. M. (2020). The emergence of a *first paradigm* in vocabulary research: The bibliometrics of *System*. *Vocabulary Learning and Instruction, 9*(1), 1-2. https://doi.org/10.7820/vli.v09.1.meara

Nation, I. S. P. (1990). *Teaching and learning vocabulary.* Newbury House.

Nation, I. S. P. (2001/2014). *Learning vocabulary in another language.* Cambridge University Press.

Read, J. (2000). *Assessing vocabulary*. Cambridge University Press.

Pawley, A., & Syder, F. (1983). Two puzzles for linguistic theory: Native-like selection and native-like fluency. In J. C. Richards & R. Schmidt (Eds.), *Language and communication* (pp.191-225). Routledge.

Richards, J. C. (1976). The role of vocabulary teaching. *TESOL Quarterly, 10*(1), 77-89. https://doi.org/10.2307/3585941

Small, H. (1973). Co-citation in the scientific literature: A new measure of the relationship between two documents. *Journal of the American Society for Information Science, 24*(4), 265-269. https://doi.org/10.1002/asi.4630240406

Webb, S. (2020). *The Routledge handbook of vocabulary studies.* Routledge.

APPENDIX A

The 34 papers that make up *The Handbook* data set. The papers are listed in order of their appearance.

Nation, I. S. P. The different aspects of vocabulary knowledge (pp. 15-29).
Wood, D. Classifying and identifying formulaic language (pp. 30-45).
Doczi, B. An overview of conceptual models and theories of lexical representation in the mental lexicon (pp. 46-65).
Qian, D. D., & Lin, L. H. F. The relationship between vocabulary knowledge and language proficiency (pp. 66-80).
Vilkaite-Lozdiene, L., & Schmitt, N. Frequency as a guide for vocabulary usefulness (pp. 81-96).
Coxhead, A. Academic vocabulary (pp. 97-110).
Liu, D. L., & Lei, L. Technical vocabulary (pp. 111-124).
Peters, E. Factors affecting the learning of single words (pp. 125-142).
Boers, F. Factors affecting the learning of multiword sequences (pp. 143-157).
Pellicer-Sánchez, A. Learning single words vs. multiword items (pp. 158-173).
Conklin, K. Processing single-word and multiword items (pp. 174-188).
Miralpeix, I. L1 and L2 vocabulary growth (pp. 189-206).
Kormos, J. How does vocabulary fit into theories of second language learning (pp. 207-222).
Webb, S. Incidental vocabulary learning (pp. 225-239).
Lindstromberg, S. Intentional L2 vocabulary learning (pp. 240-254).
Newton, J. Approaches to learning vocabulary inside the classroom (pp. 255-270).
Gu, P. Y. Strategies for learning vocabulary (pp. 271-287).
Dang, T. N. Y. Corpus-based word lists in second language vocabulary research, learning and teaching (pp. 288-303).
Nakata, T. Learning words with flash cards and word cards (pp. 304-319).
Ballance, O., & Cobb, T. Resources for learning single-word items (pp. 320-335).
Meunier, F. Resources for learning multiword items (pp. 336-350).
Laufer, B. Evaluating exercises for vocabulary learning (pp. 351-368).
Yanigasawa, A., & Webb, S. Measuring depth of vocabulary knowledge (pp. 371-386).
Gyllstad, H. Measuring knowledge of multiword units (pp. 387-405).
Kremmel, B. Measuring vocabulary learning progress (pp. 406-418).
Sasao, Y. Measuring the ability to learn words (pp. 419-432).
Godfroid, A. Sensitive measures of vocabulary knowledge and processing (pp. 433-453).
Kyle, K . Measuring lexical richness (pp. 454-476).
Barcroft, J. Key issues in teaching single words (pp. 477-492).
Wolter, B. Key issues in teaching multiword units (pp. 493-510).
Spatgens, T., & Schoonen, R. Single, but not unrelated: Key issues in researching single-word items (pp. 511-528).
Siyanova-Chanturia, A., & Omidian, T. Key issues in researching multiword items (pp. 529-544).
Read, J. Key issues in measuring vocabulary knowledge (pp. 545-560).
Anthony, L. Resources for researching vocabulary (pp. 561-590).

APPENDIX B

The 34 papers that make up the *Frontiers in Psychology* data set. The papers are listed in order of appearance in the journal.

Cook, S., Pandža, N., Lancaster, A., & Gor, K. (2016). Fuzzy nonnative phonolexical representations lead to fuzzy form-to-meaning mappings. *Frontiers in Psychology,* 2016.1345. https://doi.org/10.3389/fpsyg.2016.01345

Escudero, P., Mulak, K., Fu, C. S. L., & Singh, L. (2016). More limitations to monolingualism: Bilinguals outperform monolinguals in implicit word learning. *Frontiers in Psychology,* 2016.01218. https://doi.org/10.3389/fpsyg.2016.0128

Hayes-Harb, R., & Cheng, H. W. B. (2016). The influence of the pinyin and zhuyin writing systems on the acquisition of Mandarin word forms by native English speakers. *Frontiers in Psychology,* 2016. 00785. https://doi.org/10.3389/fpsyg.2016.00785

Ma, T., Chen, R., Dunlap, S., & Chen, B. (2016). The effect of number and presentation order of high-constraint sentences on second language word learning. *Frontiers in Psychology,* 2016.1396. https://doi.org/10.3389/fpsyg.2016.01396

Boerma, T., Leseman, P., Wijnen, F., & Blom E. (2017). Language proficiency and sustained attention in monolingual and bilingual children with and without language impairment. *Frontiers in Psychology,* 2017.01241. https://doi.org/10.3389/fpsyg.2017.01241

Chen B. G., Ma, T. F., Liang, L. J., & Liu, H. H. (2017). Rapid L2 word learning through high constraint sentence context: An event-related potential study. *Frontiers in Psychology,* 2017.2285. https://doi.org/10.3389/fpsyg.2017.2285

Haman, E., Wodniecka, Z., Marecka, M., Szewczyk, J., Białecka-Pikul, M., Otwinowska, A., Mieszkowska, K., Łuniewska, M., Kołak, J., Miekisz, A., Kacprzak, A., Banasik, N., & Forys-Nogala, M. (2017). How does L1 and L2 exposure impact L1 performance in bilingual children? Evidence from Polish-English migrants to the United Kingdom. *Frontiers in Psychology,* 2017.01444. https://doi.org/10.3389/fpsyg.2017.01444

Heim, S., Klann, J., Schattka, K., Bauhoff, S., Borcherding, G., Nosbüsch, N., Struth, L., Binkofski, F., & Werner, C. J. (2017). A nap but not rest or activity consolidates language learning. *Frontiers in Psychology,* 2017.00665. https://doi.org/10.3389/fpsyg.2017.00665

Mieszkowska, K., Łuniewska, M., Kołak, J., Kacprzak, A., Wodniecka, Z., & Haman, E. (2017). Home language will not take care of itself: Vocabulary knowledge in trilingual children in the United Kingdom. *Frontiers in Psychology,* 2017.01358. https://doi.org/10.3389/fpsyg.2017.01358

Repetto, C., Pedroli, E., & Macedonia, M. (2017). Enrichment effects of gestures and pictures on abstract words in a second language. *Frontiers in Psychology,* 2017.02136. https://doi.org/10.3389/fpsyg.2017.02136

Rossi, E, Cheng, H., Kroll, J. F., Diaz, M. T., & Newman, S. D. (2017). Changes in white-matter connectivity in late second language learners: Evidence from diffusion tensor imaging. *Frontiers in Psychology,* 2017.02040. https://doi.org/10.3389/fpsyg.2017.02040

Buccola, B., Dautriche, I., & Chemla, E. (2018). Competition and symmetry in an artificial word learning task. *Frontiers in Psychology,* 2018.02176. https://doi.org/10.3389/fpsyg.2018.02176

Guo, J., Zou, T. & Peng, D. (2018). Dynamic influence of emotional states on novel word learning. *Frontiers in Psychology* 2018.00537. https://doi.org/10.3389/fpsyg.2018.00537.

Köpke, B., & Genevska-Hanke, G. (2018). First language attrition and dominance: Same or different? *Frontiers in Psychology,* 2018.01963. https://doi.org/10.3389/fpsyg.2018.01963

Liu, L. Q., &Kager, R. (2018). Monolingual and bilingual infants' ability to use non-native tone for word learning deteriorates by the second year after birth. *Frontiers in Psychology,* 2018.00117. https://doi.org/10.3389/fpsyg.2018.00117

Meade, G., Midgley, K. J., & Holcomb, P. J. (2018). An ERP investigation of L2-L1 translation priming in adult learners. *Frontiers in Psychology,* 2018.00986. https://doi.org/10.3389/fpsyg.2018.00986

Poltrock, S., Chen, H., Kwok, C., Cheung, H. T., & Nazzi, T. (2018). Adult learning of novel words in a non-native language: Consonants, vowels, and tones. *Frontiers in Psychology,* 2018.01211. https://doi.org/10.3389/fpsyg.2018.01211

Unsworth, S., Chondrogianni, V., & Skarabela, B. (2018). Experiential measures can be used as a proxy for language dominance in bilingual language acquisition research. *Frontiers in Psychology,* 2018.01809. https://doi.org/10.3389/fpsyg.2018.01809

Zhang, Y., Chen, B., Tang, Y., Yao, P., & Lu, Y. (2018). Semantic similarity to known second language words impacts learning of new meanings. *Frontiers in Psychology,* 2018.02048. https://doi.org/10.3389/fpsyg.2018.02048

Zhang, K., Peng, G., Li, Y. H., Minett, J. W., & Yang, W. S. Y. (2018). The effect of speech variability on tonal language speakers' second language lexical tone learning. *Frontiers in Psychology,* 2018.01982. https://doi.org/10.3389/fpsyg.2018.01982

D'Anselmo, A., Prete, G., Zdybek, P., Tommasi, L., & Brancucci A. (2019). Guessing meaning from word sounds of unfamiliar languages: A cross-cultural sound symbolism study. *Frontiers in Psychology,* 2019.00593. https://doi.org/10.3389/fpsyg.2019.00593

Macedonia, M., Repetto, C., Ischebeck, A., & Mueller, K. (2019). Depth of encoding through observed gestures in foreign language word learning. *Frontiers in Psychology,* 2019.00033. https://doi.org/10.3389/fpsyg.2019.00033

Marsh, J. E., Hansson, P., Eriksson Sörman, D., & Körning Ljungberg, J. (2019). Executive processes underpin the bilingual advantage on phonemic fluency: Evidence from analyses of switching and clustering. *Frontiers in Psychology,* 2019.01355. https://doi.org/10.3389/fpsyg.2019.01335

D'Angelo, N. D., Krenca, K., & Chen, X. (2020). The overlap of poor reading comprehension in English and French. *Frontiers in Psychology,* 2020.00120. https://doi.org/10.3389/fpsyg.2020.00120

Dong, Y., Tang, Y., Wu, S. X. Y., Dong, W. Y., & Li, Z. (2020). Longitudinal effects of mediums of word explanation on L2 vocabulary learning strategies among Chinese grade-7 students. *Frontiers in Psychology,* 2020.00702. https://doi.org/10.3389/fpsyg.2020.00702

Fan, N. (2020). Strategy use in second language vocabulary learning and its relationships with the breadth and depth of vocabulary knowledge: A structural equation modeling study. *Frontiers in Psychology,* 2020.00752. https://doi.org/10.3389/fpsyg.2020.00752

Guan, C. Q., & Fraundorf, S. H. (2020). Cross-linguistic word recognition development among Chinese children: a multilevel linear mixed-effects modeling approach. *Frontiers in Psychology,* 2020.00544. https://doi.org/10.3389/fpsyg.2020.00544

Koizumi, R., & In'nami, Y. (2020). Structural equation modeling of vocabulary size and depth using conventional and Bayesian methods. *Frontiers in Psychology,* 2020.00618. https://doi.org/10.3389/fpsyg.2020.00618

Sengottuvel, K., Vasudevamurthy, A., Ullman, M. T., & Earle, F. S. (2020). Learning and consolidation of declarative memory in good and poor readers of English as a second language. *Frontiers in Psychology,* 2020.00715. https://doi.org/10.3389/fpsyg.2020.00715

Suarez-Coalla, P., Martinez-Garcia, C., & Carnota, A. (2020). Reading in English as a foreign language by Spanish children with dyslexia. *Frontiers in Psychology,* 2020.00019. https://doi.org/10.3389/fpsyg.2020.00019

Wallace, M. P., & Lee, K. (2020). Examining second language listening, vocabulary, and executive functioning. *Frontiers in Psychology,* 2020.01122. https://doi.org/10.3389/fpsyg.2020.01122.

Wiener, S., & Lee, C. Y. (2020). Multi-talker speech promotes greater knowledge-based spoken Mandarin word recognition in first and second language listeners. *Frontiers in Psychology,* 2020.00214. https://doi.org/10.3389/fpsyg.2020.00214

Zhang, Y., Lu, Y., Liang, L. J., & Chen, B. G. (2020). The effect of semantic similarity on learning ambiguous words in a second language: An event-related potential study. *Frontiers in Psychology,* 2020.01633. https://doi.org/10.3389/fpsyg.2020.01633

Zhang, Y., Cao, N. N., Yue, C., Dai L., & Wu, Y. J. (2020). The interplay between language form and concept during language switching: A behavioral investigation. *Frontiers in Psychology*, 2020.00792. https://doi.org/10.3389/fpsyg.2020.00791