# THE EFFECT OF TYPES OF DICTIONARY PRESENTATION ON THE RETENTION OF METAPHORICAL COLLOCATIONS: INVOLVEMENT LOAD HYPOTHESIS VS. COGNITIVE LOAD THEORY

## Abstract

Although conceptual metaphor boxes were introduced in *Macmillan English Dictionary* over a decade ago, the effect of this innovative presentation on the retention of metaphorical collocations remains unclear. To fill this gap, this paper reports on a study involving 70 Chinese EFL learners, who were randomly assigned into three groups. They were exposed to three types of collocation presentation, varying in terms of availability and salience of conceptual metaphor information. The results show that the learners who were exposed to the accessible and salient presentation of metaphor information had the best retention of metaphorical collocations and the shortest lookup time. Two competing theoretical accounts, Involvement Load Hypothesis (ILH) and Cognitive Load Theory (CLT), were compared. It was found that CLT had greater explanatory power in predicting the research results, accounting for why conceptually grouping collocations facilitated the retention of metaphorical collocations.

## 1. Introduction

Despite the importance of collocations in language acquisition (Stengers et al. 2011), relatively little attention has been paid to the presentation and treatment of collocations in dictionaries (Gouws 2015). In some dictionaries, collocations are not typographically distinguished from example sentences or the co-text. This causes difficulties for users to identify collocations and 'retrieve the information they require to use these collocations in a proper way' (Gouws 2015: 172). In some dictionaries, collocations are not embedded in the explanation or illustration of the lemma meanings, but rather presented in a non-grouped list as a separate text block of the entry. A third type of presentation is to group collocations lexico-grammatically or conceptually, such as the collocation boxes in *Longman Dictionary of Contemporary English* (LDOCE5) and the metaphor boxes in *Macmillan English Dictionary* (MED).

Although different presentation modes have been adopted by a variety of dictionaries, it was not until the 2010s that scholarly attention was drawn to the effect of dictionary presentation and dictionary use on collocation learning (Chen 2017; Dziemianko 2010, 2012, 2014; Laufer 2011). For instance, Dziemianko (2010) reported that learners were able to retain 45.6% of new collocations when using a paper dictionary, and 63.9% when using an electronic dictionary. Interestingly, in a replication study by Dziemianko (2012), a wide range of retention rates (from 26.3%

to 76.3%) was observed, depending on the form (paper vs. electronic) and the dictionary (OALD7, LDOCE5, or COBUILD7). Dziemianko's (2014: 272) study further shows that 'highlighting collocations in bold before or within examples is the best presentation mode for production purposes' and that '[c]ollocations are best retained if given twice in the entry, in bold before examples and within examples in regular font'. Contrary to these positive findings, the effect of dictionary lookup was less satisfactory in Laufer's (2011) study. On average, pre-intermediate learners with dictionary assistance retained about 0.5 collocation, and intermediate learners 1.5 collocations. Similarly, Chen (2017) found that dictionary consultation had a moderate effect, with only 2.13 (or 36.1% of the target) collocations retained.

These inconsistent results might be caused by different types of target collocations in the experiments. Some studies focus on literal collocations—'combinations where the literal meanings of the words are simply added together' (Macis and Schmitt 2017b: 51), while some others do not differentiate collocations with a high degree of semantic transparency from those with a lower degree. For example, Chen (2017) conflated semantically transparent collocations (e.g. *write a check*) and opaque ones (e.g. *keep house*). We contend that a distinction should be made between the dictionary presentation of literal collocations and metaphorical collocations for at least two reasons.

First, unlike literal collocations, metaphorical collocations cannot be understood by simply adding up the meanings of their component words (Macis and Schmitt 2017b). For instance, the meaning of the literal collocation *keen blade* is straightforward, while that of the metaphorical collocation *keen understanding* is less transparent. Indeed, recent studies show that fully transparent compounds or collocations like *keen blade* are processed much quicker than the semi-transparent ones like *keen understanding* (El-Bialy, Gagné and Spalding 2013; Gyllstad and Wolter 2016). Therefore, traditional presentation that focuses on lexico-grammatical information might not be helpful for users in decoding metaphorical collocations.

Second, corpus findings have demonstrated that the majority of collocations, particularly metaphorical collocations, are motivated (Deignan 2005; Liu 2010). Words do not co-occur in an arbitrary manner. To continue the previous example, *keen understanding* is motivated by the conceptual metaphor INTELLIGENCE IS A KNIFE. Conceptual metaphor (CM) is a comparison, in which the properties of the source domain (e.g. A KNIFE) are mapped onto the target domain (e.g. INTELLIGENCE). Studies have shown that raising CM awareness is beneficial for EFL learners in learning idioms (Berendi, Csábi and Kövecses 2008) and phrasal verbs (Yasuda 2010). Therefore, in dictionary presentation, explicating or at least implicating why words collocate in a certain pattern can enhance learners' metaphor awareness (Boers 2000) and contribute to their learning. In fact, MED has made 'a lexicographical innovation' to incorporate 'illustrative features on conceptual metaphor' (Moon 2004: 195). The following reproduces the metaphor box at *intelligence* in MED2:

Metaphor

Intelligence is like a **knife** or **blade**. The more intelligent someone is, the sharper the blade.

*He's very **sharp/sharp-witted**: he notices everything. ♦ She has a **razor-sharp** mind. ♦ I want to **cut** through the waffle and get straight to the point. ♦ We carefully **dissected** the problem. ♦ He has a **keen** intellect. ♦ They made some **pointed** remarks. ♦ He was an **incisive** critic.*

→ MIND, UNDERSTAND

As shown in this extract, the metaphor box first introduces a conceptual metaphor by explaining the mapping between source and target domains. Then, it provides a series of example sentences to illustrate the conceptual metaphor. Marked in boldface are the words and phrases that realize the mapping. Although this 'lexicographical innovation' has been around for over a decade, to the best of our knowledge, there is no empirical research examining the effect of this innovative presentation on the retention of metaphorical collocations, which by nature are most closely related to the conceptual metaphor information provided in the metaphor boxes. To fill this gap, our study attempts to determine whether three different types of dictionary presentation have a varying impact on the retention of metaphorical collocations. Based on the research results, we also compare the explanatory power of two theories to understand how dictionary presentation can be optimized to reduce the cognitive load expended in the consultation process, thereby improving metaphorical collocation retention.

## 2. Theoretical framework and research hypotheses

Two competing theoretical accounts—Involvement Load Hypothesis (ILH) and Cognitive Load Theory (CLT)—have been empirically tested to ascertain the effect of dictionary lookup on vocabulary learning, but they have not been compared in a single study. Both of them focus on the level of cognitive processing, but they differ in how cognitive loads are theorized and measured. We believe that the comparison of the two theories will throw light on how dictionary presentation attracts or distracts users' cognitive resources, eventually leading to (un)successful retention of metaphorical collocations.

### 2.1 *Involvement Load Hypothesis*

Hulstijn and Laufer (2001) drew upon the processing depth theory (Craik and Lockhart 1972) and proposed the Involvement Load Hypothesis (ILH) to explain vocabulary learning. They contend that the cognitive involvement load of a task can be measured in three components: need, search, and evaluation (Hulstijn and Laufer 2001). Need is the non-cognitive dimension of involvement and has two values: moderate (1) and high (2). Need is moderate when a word is externally required by agents (e.g. teachers asking students to make a sentence with a given word). Need is strong when a word is internally wanted by learners themselves (e.g. learners taking initiative to look up a word when reading an article). Search and evaluation constitute the cognitive dimension of involvement. Search is measured as present (1) or absent (0), depending

on whether learners need to find out the meaning of a word. Evaluation is about making comparison and decision. Evaluation is considered to be moderate (1), if learners compare two words or several senses of a word. Evaluation is considered to be strong (2), if learners decide how additional words and the target word work in combination. The involvement load is the sum of the values for need, search and evaluation. If a learning task has a higher involvement load, the task requires a deeper level of processing and is more likely to result in vocabulary learning.

Applying ILH in this research context, we can hypothesize that the traditional presentation of mixing literal and metaphorical collocations is more beneficial to learning, because learners spend more cognitive effort going through entries in order to search for and evaluate the meanings. If, on the other hand, metaphorical collocations are grouped together and explained in a metaphor box (just as in MED), learners can directly spot the target collocation and spend less effort on searching. In this regard, ILH posits that the latter presentation design might induce a shallower level of cognitive processing, not contributive to collocation learning.

## 2.2 *Cognitive Load Theory*

An alternative perspective to explore the relationship between collocation presentation and learning is the Cognitive Load Theory (CLT). With its origin in educational psychology, CLT conceptualizes three types of cognitive load: intrinsic, extraneous, and germane (Sweller et al. 1998). Intrinsic cognitive load is the basic load of working memory required to hold, process and interact with the learning materials. It is intrinsic because it is largely determined by the nature and design of the learning task and/or material. Extraneous cognitive load is caused by the instructional design that is not directly relevant to achieving the learning goal (Sweller 1994). Therefore, extraneous cognitive load needs to be avoided to free up the finite amount of working memory for intrinsic cognitive load. Germane cognitive load is the intentional cognitive effort learners make to form and automate conceptual and knowledge schemata (Sweller et al. 1998).

Applying CLT in the current research case, we can hypothesize the relations between different types of collocation presentation and retention of collocations. In some traditional dictionaries, collocation presentation predominantly focuses on the lexico-grammatical information and does not differentiate metaphorical collocations from literal collocations. Students have to go through every individual collocation to evaluate whether it fits the context of the target collocation. This process requires many search-and-evaluate operations which are irrelevant to the learning of the target collocation. Therefore, it increases students' extraneous cognitive load and renders the presentation design suboptimal (Dang et al. 2013). On the other hand, if conceptual metaphor information is placed under the headword, students' attention will be drawn to the conceptual mapping (i.e. the relation between source and target domains). This will save time for students to search information about the target metaphorical collocation and reduces extraneous cognitive load. Additionally, this presentation design increases students' germane cognitive load, because the conceptual metaphor

information enables students to form and automate the metaphor knowledge schema that underpins the target collocations.

Given the potentially conflicting predictions of the two theories, we attempt to find out whether ILH or CLT has greater explanatory power to account for the effect of types of dictionary presentation on the retention of metaphorical collocations.

## 3. Research design

### 3.1 *Research questions*

Our study is guided by the following research questions:
1) Do different types of dictionary presentation affect the retention of metaphorical collocations?
2) Do different types of dictionary presentation affect the lookup time and the retrieval time of metaphorical collocations?

### 3.2 *Participants and treatment conditions*

The participants in this study were 70 first-year students from a top-tier university in China. They were non-English majors from three parallel classes. Judged by their grades in the National Matriculation Test, their English proficiency level was similar ($F$ (2, 67) = 1.21, $p$ > 0.05, $\eta^2$ = 0.03). The three classes were randomly assigned to three treatment conditions (see Table 1), which varied in terms of availability and salience of conceptual metaphor information. For the Control Group, conceptual metaphor information was neither available nor salient. For the Experimental Group 1, the information was both available and salient. The salience was reflected by the position in the microstructure (Dziemianko 2014) – conceptual metaphor information and between seven and ten example sentences were placed right under the headword. For the Experimental Group 2, the information was available but not salient. The conceptual metaphor information and one example sentence were placed under the metaphorical sense of the headword and that sense was placed in the middle of the entry. Examples of dictionary entries can be found in Section 3.6.

**Table 1.** Participants and treatment conditions.

| Group | N | Treatment Parameters | | Treatment Condition |
|---|---|---|---|---|
| | | Availability | Salience | |
| Control Group | 23 | No | No | No explanation about conceptual metaphor is provided, and only one example sentence is provided under the metaphorical sense of the headword. |
| Experimental Group 1 | 24 | Yes | Yes | Conceptual metaphor information is placed right under the headword, with 7-10 example sentences provided. |
| Experimental Group 2 | 23 | Yes | No | Conceptual metaphor information and one example sentence are placed |

under the metaphorical sense of the headword, which is in the middle of the entry.

### 3.3 *Selecting conceptual metaphors and metaphorical collocations*

The selection of conceptual metaphors went through three stages. First, we consulted a wide range of metaphor dictionaries (such as *Collins Cobuild English Guides 7: Metaphor*) and databases of lexicalized metaphors (such as Metaphor at Lingnan University Department of English [METALUDE]) to collect conceptual metaphors that might be unfamiliar to the students. This round of search resulted in a pool of 52 conceptual metaphors. At the second stage, we selected 20 structural metaphors, in which the structural properties of the source domain are mapped onto the target domain. Lakoff and Johnson (1980: 5) pointed out that structural metaphors are most typical in 'understanding and experiencing one kind of thing in terms of another'. At the third stage, we narrowed down to five conceptual metaphors that have the largest amount of instantiations in our databases: (a) LIGHT IS A LIQUID, (b) MONEY IS A LIQUID, (c) INTELLIGENCE IS A KNIFE, (d) INTELLIGENCE IS LIGHT, and (e) AN IDEA IS A PLANT.

The selection of the target metaphorical collocations went through two stages. First, we consulted the metaphor dictionaries and databases once again, and all the metaphorical collocations belonging to the five target conceptual metaphors were put in an item pool. This search produced a total of 162 metaphorical collocations. Second, following Laufer (2011), we selected collocations that were made up of words familiar to the students, because our test tapped the depth not the breadth of vocabulary knowledge (see Section 3.4 for detail). We defined unfamiliar words as those outside of the wordlist generated from the students' High School English textbooks and first-year College English textbooks. The elimination of collocations with unfamiliar words resulted in a total of 42 metaphorical collocations.

### 3.4 *Test instrument*

To measure students' knowledge of metaphorical collocations, we designed a Metaphorical Vocabulary Knowledge Test (MVKT), adapted from the Word Associates Test (WAT: Read 1993, 1998). We regarded WAT as an advantageous tool to model on, because it 'covers both meaning and collocation' (Schmitt 2010: 226) and measures the depth of vocabulary knowledge. We followed the format of WAT: for each target word (or stimulus), eight options are provided. Four of them are correct responses (or associates) and related to the target word in three ways—paradigmatic, syntagmatic or analytic (Read 1998). Eight options are evenly divided into two boxes. The associates in the first box are paradigmatic or analytic, meaning that they are either synonyms or representing one component of the stimulus word. Those in the second box are syntagmatic, meaning that they usually collocate with the stimulus word in a sentence. The four correct associates could be located evenly in the two boxes (2-2) or split in 1-3 or 3-1. For example, as shown in Figure 1, *beautiful* is the stimulus word and the correct associates are *enjoyable*, *face*, *music* and *weather*.

beautiful

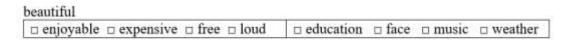| □ enjoyable  □ expensive  □ free  □ loud | □ education  □ face  □ music  □ weather |
|---|---|

**Figure 1.** An example item in WAT (*beautiful* as the stimulus word).

Following the WAT format, we found that the 42 collocations described in Section 3.3 could be divided into a list of 36 stimulus words and 42 metaphorical collocates. We then went through the original items in WAT and found that four stimulus words matched our list. One test item (*fertile*) could be used without modification and three others (*dull*, *acute* and *bright*) were modified to include three metaphorical collocates from our list. We then designed 32 test items with the remaining stimulus words and metaphorical collocates. In this process, we noticed that all the WAT stimulus words were adjectives, but in MVKT we also included nouns and verbs. This decision was informed by the following research finding: among all the target words in a self-built corpus, verbs take up nearly 50% of the words with metaphorical meanings, followed by nouns and adjectives (Cameron 2003). After this step, we had 36 MVKT items (one original WAT item, three modified WAT items, and 32 self-compiled items). Figure 2 represents one example of the self-compiled items. Based on the collocation *the germ of the idea*, we wrote a test item with *germ* as the stimulus word and *bud*, *disease*, *idea*, and *feeling* as the four correct associates. The collocation is motivated by the target conceptual metaphor AN IDEA IS A PLANT. Therefore, *bud* and *idea* are placed on the middle pane of Table 2, while *disease* and *feeling* are not the target associates in this study, so they are placed on the right pane of Table 2.

germ

| □accident  □bud  □condition  □disease | □idea  □feeling  □street  □reader |
|---|---|

**Figure 2.** An example item in MVKT (*germ* as the stimulus word).

We piloted the 36 items on 44 students from another university, who had the same English proficiency as our focal participants. Based on their responses, we made two changes. First, we decided to reduce the number of test items from 36 to 24. In the pilot study, we found that with access to dictionaries it took the students 38.3 minutes on average to cover 24 items, and 58.5 minutes to cover all of them. It would not be logistically possible to accommodate a test longer than 40 minutes during regular class time. Therefore, we used the internal consistency (measured by Cronbach's alpha) to check whether an item had a strong correlation with others, and eliminated 12 items with weaker correlation. Second, we decided to replace 12 MVKT items with 12 original WAT items. The 12 original WAT items are all words without metaphorical senses or collocations, serving as distractors. In the pilot test, according to the students' self-reports, when all the stimulus words had metaphorical senses and collocations, they could guess that the metaphorical sense was probably one of the correct answers. When the non-metaphorical and metaphorical test items were evenly distributed, the students

would have a 50% chance of getting a stimulus word with or without metaphorical collocations, preventing them from guessing the correct answers. Therefore, we kept the 12 items that had the best correlation with other items, and then added 12 original WAT items without metaphorical senses. These 24 items were piloted on another group of 38 students, and the Cronbach's alpha was 0.89, suggesting a good reliability. In this pilot test, we also asked the students to report any unfamiliar words and whether they had heard of the conceptual metaphors before the test. They confirmed that they recognized all the words but had little knowledge of the conceptual metaphors, thus validating the selection criteria of conceptual metaphors and familiar words explained in Section 3.3. Table 2 presents the 12 test items (stimulus words and associates) with metaphorical senses and collocations.[1] It looks slightly haphazard, because we do not want the students to infer the pattern of correct answers, thus ensuring test validity.

**Table 2.** Twelve target items in MVKT.

| CM | Stimulus | Associates related to target CM | | | | Associates unrelated to target CM | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| | | LS | LC | MC1 | MC2 | LS | MS | LC | MC1 | MC2 |
| a | spill | liquid | drink | light | | escape | | | | |
| a | flood | river | | light | | | | | market | memory |
| b | pump | pour | blood | savings | | press | | | | |
| a + b | flow | liquid | | moonlight | cash | | | | thought | |
| a + b | pool | lake | | light | money | | | swimming | | |
| a + b | drain | water | pond | light | income | | | | | |
| c | keen | | blade | understanding | | | eager | | reader | |
| c | acute | sharp | | mind | | | | | angle | illness |
| d | bright | shine | | idea | | | happy | | future | |
| d | dull | cloudy | lamp | pupil | | | | | knife | |
| e | fertile | plant | soil | mind | | | | egg | | |
| e | germ | bud | | idea | | disease | | | feeling | |

Notes: The letters a–e represent the five target conceptual metaphors explained in Section 3.3. LS: literal sense; LC: literal collocation; MC1: the first metaphorical collocation; MC2: the second metaphorical collocation; MS: metaphorical sense.

3.5 *Procedure*

The experiment was carried out in three sessions during regular class time. In the first session (i.e. the pretest), MVKT was administered to 70 students, who were not allowed to consult any dictionary or reference material. One week later, in the second session (i.e. the main test), MVKT was administered to those students. They were allowed to use three different purpose-built smart-phone dictionaries, corresponding to the three different treatment conditions (see Table 1). The dictionaries only contained the entries of the 24 stimulus words. To prevent the Hawthorne effect, the students were told that the test had nothing to do with their grades. It was simply a task for them to test-use a smart-phone app MDict (see Section 3.6 for detail), so they should focus on the user-experience of the smart-phone dictionary app. One week later, in the third session (i.e. the retention test), MVKT was administered to the students unexpectedly without access to any reference material. It should be noted that the items in the retention test

were randomized in a different order from the main test to mitigate the memory effect. During the three sessions, no time limit was imposed, but a stopwatch was displayed in the classroom projector. The students were asked to note down the time upon completion of the tests.

## 3.6 *Dictionaries used*

In our pilot study, we asked 44 students about the preferences for paper dictionaries, PC-based dictionaries and smart-phone dictionaries. As it turned out, 91% of the students possessed smart-phone dictionaries, while 66% of them owned paper dictionaries. In terms of use frequency, they used smart-phone dictionaries most frequently (scoring 4.71 in a 5-point Likert-scale). Therefore, to accommodate the dictionary use habit of our students, we decided to use MDict, a smart-phone dictionary. Among many other standard features of smart-phone dictionary platform, MDict allows customization of dictionary database.

To correspond with the three different treatment conditions, we purposefully built mini dictionaries as follows. For the Control Group, we provided the entry information of the 24 stimulus words in MWKT (see Figure 3 for an example). The entry information was taken from OALECD4. This presentation predominantly focuses on the lexico-grammatical information of collocations.

For the Experimental Group 1, in addition to the information available to the Control Group, the conceptual metaphor was explained at the very top of the entry, followed by 7-10 example sentences, before the sense distinctions were given. The metaphorical information was taken from MECD (see Figure 4 for an example). This presentation foregrounded conceptual metaphor information in the initial position of the entry and furnished it with ample example sentences.

For the Experimental Group 2, in addition to the information available to the Control Group, the conceptual metaphor was explained under the metaphorical sense of the headword and one example sentence was provided (see Figure 5 for an example). Again, the conceptual metaphor information was taken from MECD. The presentation was less salient because of the smaller number of examples and the middle position it occupied in the entry.



**germ** / dʒəːm; *NAmE* dʒəːrm /

**noun**

1. [C, usually pl.] a very small living thing that can cause infection and disease
   微生物；细菌；病菌：
   Disinfectant kills germs.
   消毒剂可杀菌。
   Dirty hands can be a breeding ground for germs.
   脏手会滋生病菌。
2. [sing.] ~ of sth an early stage of the development of sth
   起源；发端；萌芽：
   Here was the germ of a brilliant idea.
   一个绝妙的主意就是从这里萌发的。
3. [C] (biology 生) the part of a plant or an animal that can develop into a new one
   胚芽；胚原基；芽孢；胚胎
   ⇒ see also wheatgerm

**Figure 3.** Entry for **germ** in the Control Group.

**germ** / dʒɜːm; *NAmE* dʒɜːrm /

☛ **Metaphor** 隐喻

An idea is also like a **plant**. Developing an idea is like helping a plant to grow.
一个想法也像一株植物（plant）。构思想法如同帮助植物生长一样。

- *The* **idea grew** *in her mind until she could think of nothing else.*
  这个主意在她脑子里滋生着，直至别的事儿她一点儿也想不进去。
- *I had already* **planted the idea** *in their minds.*
  我已经在他们头脑中灌输了这种思想。
- *I had* **sown the seeds of doubt.**
  我已播下了怀疑的种子。
- *The* **idea germinated** *slowly in his mind.*
  这个想法在他头脑中慢慢萌生出来。
- *This* **belief** *quickly* **took root.**
  这种信仰迅速生根了。
- *These* **beliefs** *are* **deep-rooted.**
  这些信仰根深蒂固。
- *What does this* **idea stem** *from?*
  这个想法是从哪儿来的？
- *She has a* **fertile imagination.**
  她拥有丰富的想象力。
- *It was a* **fruitful** *line of research.*
  那是成果丰硕的一系列研究。
- *The work shows evidence of* **cross-fertilization** *from many disciplines.*
  这项工作显示出多学科融合的迹象。

→ achieve , argument , mind

**noun**

1. [C, usually pl.] a very small living thing that can cause infection and disease
   微生物；细菌；病菌：
   Disinfectant kills germs.
   消毒剂可杀菌。
   Dirty hands can be a breeding ground for germs.
   脏手会滋生病菌。
2. [sing.] ~ of sth an early stage of the development of sth
   起源；发端；萌芽：
   Here was the germ of a brilliant idea.
   一个绝妙的主意就是从这里萌发的。
3. [C] (biology 生) the part of a plant or an animal that can develop into a new one
   胚芽；胚原基；芽孢；胚胎：

⇒ see also wheatgerm

**Figure 4.** Entry for **germ** in the Experimental Group 1.

10

**Figure 5.** Entry for **germ** in the Experimental Group 2.

### 3.7 *Variables and data analysis*

To answer the first research question, we focused on the accuracy rate of choosing the target metaphorical collocates in MVKT. That is, we examined the accuracy rate of the 15 words in the MC1 and MC2 columns on the middle pane of Table 2. The scoring scheme was either 1 (correct) or 0 (incorrect). Data analyses were conducted in three steps. First, as the data were normally distributed, we used one-way ANOVA to compare the pretest accuracy scores to confirm that the three groups had a comparable level of knowledge about the target metaphorical collocations. Second, we used one-way ANOVA to compare the main test accuracy scores to confirm that the three groups achieved a comparable level of accuracy with the help of the dictionaries. In the third step, we used repeated measures ANOVA to examine the differences among the three tests within the three groups. If the main test accuracy was significantly higher than the pretest accuracy *and* if the retention test accuracy was not significantly different from the main test accuracy, we were convinced that the students had shown strong signs of retention.

To answer the second research question, we treated the time the students spent in the main test as the lookup time, and in the retention test as the retrieval time. We also included the time spent in the pretest as the baseline time. We used one-way ANOVA to find out whether the three groups differed in terms of the baseline time, the lookup time, and the retrieval time. All the statistical analyses were performed using SPSS 21.

## 4. Results

### 4.1 *Accuracy rate of metaphorical collocates*

Figure 6 shows the three groups' mean accuracy scores in the three tests. One-way ANOVA revealed that the three groups displayed no statistical difference in the pretest ($F(2, 67) = 1.27$, $p > 0.05$, $\eta^2 = 0.04$), which means that they had a comparable level of knowledge about the target collocations before the treatment. One-way ANOVA also showed that the three groups displayed no statistical difference in the main test ($F(2, 67) = 1.04$, $p > 0.05$, $\eta^2 = 0.03$). This suggests that the students had a comparable level of accuracy in locating the metaphorical collocates with the help of the dictionaries.
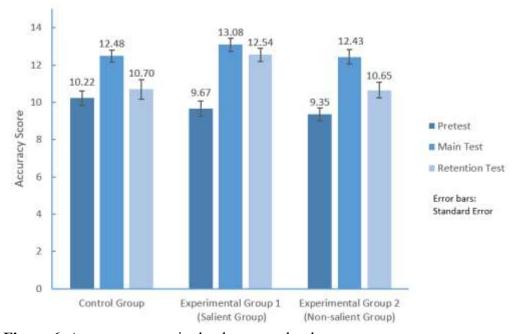


**Figure 6.** Accuracy scores in the three tests by three groups.

As evident in Figure 6, the Control Group scored 1.78 points lower on the retention test than on the main test, and 0.48 point higher than on the pretest. Repeated-measure ANOVAs showed that the three test results differed significantly ($F(2, 44) = 10.78$, $p < 0.01$, $\eta^2 = 0.33$). Tukey's post-hoc tests showed that the Control Group scored significantly lower on the retention test than on the main test, but the difference between the retention test and the pre-test scores was not significant. This suggests that when exposed to the traditional dictionary presentation, the students had little retention of the target metaphorical collocations. As regards the Experimental Group 1, the retention test score was 2.87 points higher than the pretest score and was only 0.54 point lower than the main test score. Repeated-measure ANOVAs showed that the three test results differed significantly ($F(2, 46) = 35.98$, $p < 0.01$, $\eta^2 = 0.61$). Tukey's post-hoc tests showed that the Experiment Group 1 scored significantly higher on the retention test than on the pre-test and that the difference between the retention test score and the main test score was not significant. This suggests that the accessible and salient metaphor information in the dictionary facilitated the retention of the metaphorical collocations. As regards the Experimental Group 2, the retention test score was 1.3 points higher than

the pre-test score but 1.78 points lower than the main test score. Repeated-measure ANOVAs showed that the three test results differed significantly ($F(2, 44) = 20.44$, $p < 0.01$, $\eta^2 = 0.48$). Tukey's post-hoc tests showed that the retention test score of Experiment Group 2 was significantly lower than the main test score but the difference between the retention test and the pretest scores was not significant. These results suggest that the students showed moderate signs of retention of the target metaphorical collocations.

4.2 *Processing time*

Figure 7 presents the baseline time, lookup time and retrieval time across the three groups. One-way ANOVA revealed that the groups did not differ significantly in terms of the baseline time ($F(2,67) = 2.62$, $p > 0.05$, $\eta^2 = 0.07$) or the retrieval time ($F(2,67) = 2.12$, $p > 0.05$, $\eta^2 = 0.05$). However, they differed significantly in terms of the lookup time ($F(2,67) = 7.79$, $p < 0.05$, $\eta^2 = 0.23$). Tukey's post-hoc tests showed that the Experimental Group 1 took significantly shorter lookup time than the Experimental Group 2 and the Control Group, while the Experimental Group 2 and the Control Group did not differ from each other. Taken together, the results suggest that dictionary presentation of metaphorical information, if made available and salient, can significantly reduce the collocation lookup time in the main test, but does not reduce the retrieval time in the retention test.
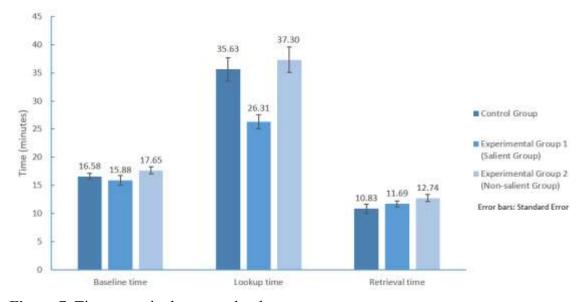


**Figure 7.** Time spent in three tests by three groups.

5. Discussion

Our study found that the Control Group only retained 0.48 metaphorical collocation, a figure virtually identical to the 0.5 collocation retained by the pre-intermediate learners in Laufer's (2011) study, showing little sign of collocation retention after dictionary consultation. On the contrary, the Experimental Group 1 on average retained 2.87 metaphorical collocations (or 84.2% of the correctly found collocations). This figure is even better than the 70.3% of collocations retained by the learners using the electronic

version of *Collins Cobuild Advanced English Dictionary* in Dziemianko's (2010) study. The Experimental Group 2 retained 1.3 metaphorical collocations (or 42.2% of the correctly found collocations). The figure is similar to the 1.5 new collocations retained by the intermediate students in Laufer's (2011) study. Percentage-wise, it is also consistent with the 45.6% of collocations retained by the learners using a paper dictionary in Dziemianko's (2010) study. In the following paragraphs, we will explain and compare the divergent results from the theoretical accounts of ILH and CLT.

### 5.1 *Predictions of ILH*

As described in Section 2.1, ILH operationalizes involvement load into three components: need, search, and evaluation. While keeping the score ranges for the need and the evaluation components, we changed the range for the search component. Originally in ILH, search was scored 1 or 0, either to consult dictionary/teacher or not. However, in this experiment, different dictionary presentations incurred different lengths of lookup time (see Figure 7). Therefore, following Folse (2006), we changed the range from 0 to 2, with 1 meaning moderate cognitive effort on search and 2 high cognitive effort. With this new rating scheme, we found that the Experimental Group 1 had the lowest involvement load, while the Control Group and the Experimental Group 2 shared the same load (see Table 3). According to ILH, it might be hypothesized that the Experimental Group 1 would have the lowest retention rate, and that the Control Group and the Experimental Group 2 would have the same outcome. The results in Sections 4.1, however, did not confirm these hypotheses, because the Experimental Group 1 had the best outcome, and the Experimental Group 2 showed better signs of retention than the Control Group.

**Table 3.** Involvement loads in three treatment conditions.

| Group | Need | Search | Evaluation | Involvement load |
|---|---|---|---|---|
| Control Group | 1 | 2 | 1 | 4 |
| Experimental Group 1 | 1 | 1 | 1 | 3 |
| Experimental Group 2 | 1 | 2 | 1 | 4 |

ILH posits that more time spent on searching presupposes a deeper level of involvement. As the Experimental Group 1 spent significantly less time on searching for metaphorical collocates, they should have had a shallower level of cognitive involvement, leading to a less satisfactory retention rate. However, the quantitative findings proved otherwise. Why did a lower lookup time result in a better retention rate? Why did ILH fail to predict the results? We entertain two possibilities as follows.

First, the involvement load in ILH is a simple sum of three unweighted components. However, some studies have found that the three components had different contribution to the cognitive involvement load (Yaqubi et al. 2010). In this study, the dictionary given to the Experimental Group 1 required the least cognitive effort on search, because the conceptual metaphor information was placed right under the headword. Relative to other treatment conditions, the Experimental Group 1 had the best retention rate of the target collocates. This seems to reject the linear, causal understanding that more effort

on search leads to a deeper level of cognitive involvement, which in turn improves learning.

Second, ILH is predominantly about acquisition of word meanings, and overlooks other aspects of vocabulary knowledge (e.g. collocation). Existing studies primarily focus on the predictability of ILH in learning new words (Folse 2006; Kim 2011) or nonsense words (Keating 2008). This study, however, examines the metaphorical collocations of known words. The disparity suggests that ILH might work better in predicting the learning of unknown words than new knowledge of known words. That is, ILH has better predicative power for the breadth than the depth of vocabulary knowledge.

## 5.2 *Predictions of CLT*

From the CLT perspective, we can rate the cognitive load in three aspects: intrinsic, extraneous, and germane cognitive loads. As Sweller (2010: 124) noted, 'for a given task and given learner knowledge levels, [intrinsic cognitive load] is fixed'. Therefore, in this study, the intrinsic cognitive load was regarded as being constant across the three treatment conditions, because the test items were the same for all three groups and their knowledge of the metaphorical collocations was comparable in the pretest and the main test. We then rated the extraneous cognitive load in three levels. -2 indicated a large amount of search-and-match operations irrelevant to learning (Kalyuga 2011), -1 a moderate amount, and 0 no amount. Germane cognitive load was also rated in three levels: 2 indicated a high potential of effortful construction of structured knowledge schemata, 1 a moderate potential, and 0 no potential. Table 4 presents the ratings across the three treatment conditions. It should be noted that the total cognitive load here is taken to mean the 'good load' beneficial to learning (Sweller 2010). Therefore, 0 and -1 do not mean that students have not spent any cognitive effort, but that the effort is not productive or even counterproductive. As can be seen, the Experimental Group 1 has the highest total cognitive load, followed by the Experimental Group 2 and the Control Group. This order is consistent with the quantitative findings presented in Sections 4.1, which suggests that CLT makes better predictions than ILH.

**Table 4.** Cognitive load in three treatment conditions.

| Group | Intrinsic | Extraneous | Germane | Total cognitive load |
|---|---|---|---|---|
| Control Group | Constant | -2 | 0 | -2 |
| Experimental Group 1 | Constant | 0 | 2 | 2 |
| Experimental Group 2 | Constant | -1 | 1 | 0 |

CLT has better explanatory power because it captures two important aspects of dictionary lookup: search efforts that are counterproductive to vocabulary learning, and cognitive efforts that are facilitative in building knowledge schemata. In ILH, search efforts are simply regarded as signs of deep-level processing. In CLT, search efforts irrelevant to learning increase the extraneous cognitive load and reduce the germane cognitive load (Sweller 2010). In the Control Group, the students had to go through the entry and perform multiple search-and-evaluate operations to decide which piece of the

entry information fitted into the target collocation, thus increasing their extraneous cognitive load (Dang et al. 2013). In the Experimental Group 1, as conceptual metaphor information was provided under the headword, the students did not have to waste their cognitive efforts on searching. As such, the reduced extraneous cognitive load freed up the germane cognitive load to build and internalize the metaphor knowledge schemata.

In addition, CLT can explain why dictionary presentation with conceptual metaphor information led to better retention. In the Experimental Group 1, the students were exposed to detailed explanation of conceptual metaphor and ample example sentences. Previous studies have found that studying examples with varied features could increase germane cognitive load (Paas and van Merrienboer 1994). Therefore, exposing students to different example sentences, as instantiations of the conceptual metaphor, optimally increases germane cognitive load to build the metaphor knowledge schema, which in turn benefits the retention of metaphorical collocations. On the contrary, the traditional dictionary presentation does not provide metaphor information and has little potential to activate the germane cognitive load, leaving students unable to establish the metaphor knowledge schema. In other words, the students in the Control Group got the answers right in the main test because they simply searched and found the correct collocations in the dictionary. They did not learn the conceptual mapping that binds the source and target domains in the collocations. That is why they achieved poorer results on the retention test, even though they had a comparable level of accuracy on the main test.

## 6. Implications

### 6.1 *Presentation of metaphorical collocations*

Although metaphor boxes in MED were not designed for collocation presentation, the conceptual metaphor information in these boxes proved effective in reducing the lookup time and increasing the retention rate of metaphorical collocations. These findings suggest that metaphorical collocations should not be presented in the same way as literal collocations. While the lexico-grammatical approach has its place in glossing semantically transparent collocations, there is a need to adopt a cognitive approach (Boers 2000; Liu 2010) to explicate the motivation for metaphorical collocations in dictionaries. Conceptual metaphors can activate the cognitive process of understanding the new knowledge by associating with the old knowledge. That is why in this study metaphor information was presented immediately after the headword. It could set the schematic background for learners and set their cognitive processes in motion. When they looked up a word, for instance **drain**, they were immediately introduced to two conceptual metaphors: MONEY IS A LIQUID and LIGHT IS A LIQUID. Their prior knowledge about **money**, **liquid** and **light** was activated. The subsequent example sentences further furnished the connections that made up a complex conceptual network. Previous research has pointed out that 'more, and more varied, feature categories of the target words' in dictionary consultation resulted in better vocabulary learning (Zou 2016: 384). In a similar vein, conceptual metaphor information can expose learners to

a connected conceptual network, through which learners understand the motivation for metaphorical collocations and achieve better retention.

## 6.2 *Findability of multiword units in digital dictionaries*

The research findings show that more search time increases extraneous cognitive load, which is counterproductive to learning. Therefore, we suggest that multiword query should be incorporated in digital dictionaries to reduce lookup time. In paper dictionaries, it is difficult to look up multiword units, because entries are usually organized orthographically (Lew 2012). The difficulty is magnified when learners try to look up complex collocations, namely combinations of single collocations with mutual elements (Gouws 2015). For instance, when learners come across the following sentence 'I want to cut through the waffle and get straight to the point' (MED2) and try to understand **cut through the waffle**, they have to decide whether to consult the **cut** entry or the **waffle** entry. Actually, *Macmillan English Dictionary Online* supports partial matching, but it is still not powerful enough to make online search effective. For instance, when we queried **cut through the waffle**, 10 partially matched phrases were returned (see Figure 8). Although the example sentence containing the collocation is located in the metaphor box at the **intelligence** entry, it is apparently unsearchable.
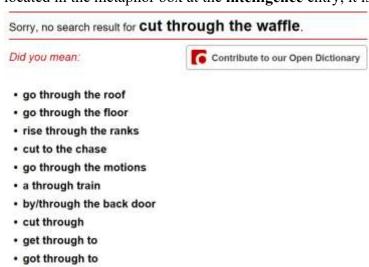
Sorry, no search result for **cut through the waffle**.

*Did you mean:*    Contribute to our Open Dictionary

- go through the roof
- go through the floor
- rise through the ranks
- cut to the chase
- go through the motions
- a through train
- by/through the back door
- cut through
- get through to
- got through to

**Figure 8**. Search result in *Macmillan English Dictionary Online*.

A mature fuzzy matching technology would need to make the entire database searchable (Lew and Mitton 2012). When users type in the multiword unit, it can be parsed by semantic and ranking algorithms. Then, the snippet of **cut through**, **waffle** and INTELLIGENCE IS A KNIFE are simultaneously displayed as search results. Users can click on the entries for more detailed and elaborate information. To achieve this function, we need to make single collocation (e.g. **cut through**) and conceptual metaphor (e.g. INTELLIGENCE IS A KNIFE) as both standalone and embedded entries. Although this presentation might seem redundant in paper dictionaries that are limited by a finite storage space, it is actually less of a problem in digital dictionaries that have more malleable space. Multiword query and fuzzy matching technology have great potentials to significantly reduce lookup time and extraneous cognitive load,

because users do not have to think about which word to look up or expend a large amount of effort browsing and evaluating senses in an entry. This frees up users' cognitive resources for germane cognitive load to analyze and internalize the motivation for metaphorical collocations.

## 7. Conclusion

Collocations are difficult to learn, because many of them are semantically semi-transparent or even non-transparent (Macis and Schmitt 2017a). As the results in the Control Group suggest, the traditional way of presenting metaphorical collocations is counterproductive for retention. Students spend an unnecessary amount of effort searching for collocations, and their finite cognitive resources are not optimally expended on building conceptual schemata that motivate the collocations. In this regard, Cognitive Load Theory is a good heuristic to guide the design of collocation presentation. Therefore, we propose a cognitive approach to presenting metaphorical collocations, whereby motivation for collocations is explicated in dictionary entries. We also argue that better findability can be achieved by incorporating multiword query and fuzzy matching technology in digital dictionaries. This function will save students the trouble of deciding which component word to look up or skimming through a large amount of texts, thereby reducing users' extraneous cognitive load and increasing their germane cognitive load for processing collocations.

One limitation of this study is that the students' lookup behaviors were not fully captured. In our design, we inferred the lookup behaviors from the time-on-task (Hill and Laufer 2003), reported by the students using a stopwatch displayed in the classroom projector. In this way, we examined the time-on-task but not the time-on-entry. It would be interesting to know whether the students simply gave up in the middle of the search process, because the collocation information was hard to find in traditional presentation. Two recent studies have dealt with the 'bathtub effect' in dictionary consultation (i.e. senses located in the middle position of an entry are the most difficult to be identified), but the results were mixed (Dziemianko 2017; Nesi and Tan 2011). Future research can determine whether the bathtub effect is borne out in the search for collocations, particularly metaphorical collocations. With screen-recording technology, we would be able to know how much time and effort students are willing to invest in lookup. This information could be used to improve the availability and salience of dictionary presentation. Actually, in the pilot study, we tried to build a screen-recording app to capture the whole lookup process, but about one third of our participants used iPhones, running an operating system incompatible with our app. We are aware that eye-tracking technology has been used to understand users' lookup behaviors (Lew et al. 2013; Tono 2011), but these studies tend to focus on computer-based dictionaries. As reported in Section 3.6, a majority of our participants, like many L2 users, are habitual users of smart-phone dictionaries, with significantly smaller presentation space compared to computer dictionaries. As such, it would be interesting for future research to find out how screen size influences users' lookup behavior and how the presentation of metaphorical collocations should vary or remain the same across screen sizes.

## Note

1 As congruency might be an intervening variable in this study, we checked whether word-for-word translation of the 15 target colocations could be found in *Xinhua Cidian* (*New Chinese Dictionary*) or *Xiandai Hanyu Cidian* (*Contemporary Chinese Dictionary*), the two most popular and influential dictionaries in China. Our search revealed that no literal translation of the target collocations was found in either of the dictionaries.

## References

**A. Dictionaries**

**Chinese Academy of Social Sciences. (ed.). 2016.** *Xiandai Hanyu Cidian* [Contemporary Chinese Dictionary] (Seventh edition). Beijing: The Commercial Press.

**Deignan, A. (ed.). 1995.** *Collins Cobuild English Guides 7: Metaphor*. London: Harper Collins.

**Hornby, A. S. (ed). 1997.** *Oxford Advanced Learner's English-Chinese Dictionary* (fourth edition). Beijing: The Commercial Press. (OALECD4)

**Lexicography Research Centre of The Commercial Press. (ed.). 2013.** *Xinhua Cidian* [New Chinese Dictionary] (fourth edition). Beijing: The Commercial press.

**Mayor, M. (ed.). 2009.** *Longman Dictionary of Contemporary English* (fifth edition). Harlow: Pearson Education Ltd. (LDOCE5)

**Rundell, M. (ed.). 2002.** *Macmillan English Dictionary for Advanced Learners*. Oxford: Macmillan Education. (MED)

**Rundell, M. (ed.). 2005.** *Macmillan English-Chinese Dictionary for Advanced Learners*. Beijing: Foreign Language Teaching and Research Press. (MECD)

**Rundell, M. (ed.). 2007.** *Macmillan English Dictionary for Advanced Learners* (Second edition). Oxford: Macmillan Education. (MED2)

**B. Other literature**

**Boers, F. 2000.** 'Metaphor Awareness and Vocabulary Retention.' *Applied Linguistics* 21.4: 553–571.

**Berendi, M., S. Csábi and Z. Kövecses. 2008.** 'Using Conceptual Metaphors and Metonymies in Vocabulary Teaching' In Boers, F. and S. Lindstromberg (eds.), *Cognitive Linguistic Approaches to Teaching Vocabulary and Phraseology*. Berlin/New York: Mouton de Gruyter, 101–132.

**Cameron, L. 2003.** *Metaphor in Educational Discourse*. London and New York: Continuum.

**Chen, Y. 2017**. 'Dictionary Use for Collocation Production and Retention: A CALL-based Study.' *International Journal of Lexicography* 30.2: 225–251.

**Chun, D. M. and J. L. Plass. 1997.** 'Research on Text Comprehension in Multimedia Environments.' *Language Learning & Technology* 1.1: 60–81.

**Craik, F. and R. S. Lockhart. 1972.** 'Levels of Processing: A Framework for Memory Research.' *Journal of Verbal Learning and Verbal Behavior* 11: 671–684.

**Dang, T.-D., G.-D. Chen, G. Dang and L.-Y. Li. 2013.** 'Rolo: A Dictionary Interface that Minimizes Extraneous Cognitive Load of Lookup and Supports Incidental and Incremental Learning of Vocabulary.' *Computers & Education* 61: 251–260.

**Deignan, A. 2005.** *Metaphor and Corpus Linguistics*. Amsterdam: John Benjamins Publishing.

**Dziemianko, A. 2010.** 'Paper or Electronic? The Role of Dictionary Form in Language Reception, Production and the Retention of Meaning and Collocations.' *International Journal of Lexicography* 23.3: 257–273.

**Dziemianko, A. 2012.** 'Why One and Two Do not Make Three – Dictionary Form Revisited.' *Lexikos* 22: 195–216.

**Dziemianko, A. 2014.** 'On the Presentation and Placement of Collocations in Monolingual English Learners' Dictionaries: Insights into Encoding and Retention.' *International Journal of Lexicography* 27.3: 259–279.

**Dziemianko, A. 2017.** 'Dictionary Entries and Bathtubs: Does it Make Sense?' *International Journal of Lexicography* 30.3: 263–284.

**El-Bialy, R., C. L. Gagné, and T. L Spalding. 2013**. 'Processing of English Compounds is Sensitive to the Constituents' Semantic Transparency.' *The Mental Lexicon* 8.1, 75–95.

**Folse, K. 2006.** 'The Effect of Type of Written Exercise on L2 Vocabulary Retention.' *TESOL Quarterly* 40.2: 273–293.

**Gouws, R. H. 2015.** 'The Presentation and Treatment of Collocations as Secondary Guiding Elements in Dictionaries.' *Lexikos* 25: 170-190.

**Gyllstad, H. and B. Wolter. 2016. '**Collocational Processing in Light of the Phraseological Continuum Model: Does Semantic Transparency Matter?' *Language Learning*, 66.2: 296–323.

**Hill, M. and B. Laufer. 2003.** 'Type of Task, Time-on-task and Electronic Dictionaries in Incidental Vocabulary Acquisition.' *International Review of Applied Linguistics* 41.2: 87–106.

**Hulstijn, J. H. 2003.** 'Incidental and Intentional Learning' In Doughty, C. J. and M. H. Long (eds), *The Handbook of Second Language Acquisition*. Oxford: Blackwell, 349–381.

**Hulstijn, J. H. and B. Laufer. 2001.** 'Some Empirical Evidence for the Involvement Load Hypothesis in Vocabulary Acquisition.' *Language Learning* 51.3: 539–558.

**Kalyuga, S. 2011.** 'Cognitive load theory: How many types of load does it really need?' *Educational Psychology Review* 23.1: 1–19.

**Keating, G. 2008.** 'Task Effectiveness and Word Learning in a Second Language: The Involvement Load Hypothesis on Trial.' *Language Teaching Research* 12.3: 365–386.

**Kim, Y. 2008.** 'The Role of Task-induced Involvement and Learner Proficiency in L2 Vocabulary Acquisition.' *Language Learning* 58.2: 285–325.

**Knight, S. 1994.** 'Dictionary Use While Reading: The Effects on Comprehension and Vocabulary Acquisition for Students of Different Verbal Abilities.' *The Modern Language Journal* 78.3: 285–299.

**Lakoff, G. and M. Johnson. 1980.** *Metaphors We Live by*. Chicago: University of Chicago Press.

**Laufer, B. 2011.** 'The Contribution of Dictionary Use to the Production and Retention of Collocations in a Second Language.' *International Journal of Lexicography*, 24.1: 29–49.

**Lew, R. 2012.** 'How Can We Make Electronic Dictionaries More Effective?' In Granger, S. and M. Paquot (Eds.), 2012. *Electronic Lexicography*. Oxford: Oxford University Press, 343–361.

**Lew, R., M. Grzelak and M. Leszkowicz. 2013.** 'How Dictionary Users Choose Senses in Bilingual Dictionary Entries: An Eye-tracking Study.' *Lexikos* 23.1: 228–254.

**Lew, R. and R. Mitton. 2012.** 'Online English Learners's Dictionaries and Misspellings: One Year On.' *International Journal of Lexicography* 26.2: 219-233.

**Liu, D. 2010.** 'Going Beyond Patterns: Involving Cognitive Analysis in the Learning of Collocations.' *TESOL Quarterly* 44.1: 4–30.

**Macis, M. and N. Schmitt. 2017a.** 'Not Just "Small Potatoes": Knowledge of the Idiomatic Meanings of Collocations.' *Language Teaching Research* 21.3: 321–340.

**Macis, M. and N. Schmitt. 2017b**. 'The Figurative and Polysemous Nature of Collocations and Their Place in ELT.' *ELT Journal* 71.1: 50–59.

**Moon, R. 2004.** 'On Specifying Metaphor: An Idea and Its Implementation.' *International Journal of Lexicography* 17.2: 195–222.

**Nesi, H. and K. H. Tan. 2011.** 'The Effect of Menus and Signposting on the Speed and Accuracy of Sense Selection.' *International Journal of Lexicography* 24.1: 79–96.

**Paas, F. G. and J. J. Van Merriënboer. 1994.** 'Variability of Worked Examples and Transfer of Geometrical Problem-Solving Skills: A Cognitive-Load Approach.' *Journal of Educational Psychology* 86.1: 122–133.

**Read, J. 1993.** 'The Development of a New Measure of L2 Vocabulary Knowledge.' *Language Testing* 10.3: 355–371.

**Read, J. 1998.** 'Validating a Test to Measure Depth of Vocabulary Knowledge' In Kunnan, A. (ed.), *Validation in Language Assessment*. Mahwah, NJ: Lawrence Erlbaum, 41–60.

**Schmitt, N. 2010.** *Researching Vocabulary: A Vocabulary Research Manual*. Houndmills: Palgrave Macmillan.

**Stengers, H., F. Boers, A. Housen and J. Eyckmans. 2011.** 'Formulaic Sequences and L2 Oral Proficiency: Does the Type of Target Language Influence the Association.' *International Review of Applied Linguistics* 49.4: 321–343.

**Sweller, J. 1994.** 'Cognitive Load Theory, Learning Difficulty, and Instructional Design.' *Learning and Instruction* 4.4: 295–312.

**Sweller, J. 2010.** 'Element Interactivity and Intrinsic, Extraneous, and Germane Cognitive Load.' *Educational Psychology Review* 22.2: 123–138.

**Sweller, J., J. V. Merrienboer and F. Paas. 1998.** 'Cognitive Architecture and Instructional Design.' *Educational Psychology Review* 10.3: 251–296.

**Tono, Y. 2011.** 'Application of Eye-tracking in EFL Learners' Dictionary Look-up Process Research.' *International Journal of Lexicography* 24.1: 124–153.

**Yaqubi, B., R. A. Rayati and N. A. Gorgi. 2010.** 'The Involvement Load Hypothesis and Vocabulary Learning: The Effect of Task Types and Involvement Index on L2 Vocabulary Acquisition.' *Journal of Teaching Language Skills* 29.1: 145–163.

**Yasuda, S. 2010.** 'Learning phrasal verbs through conceptual metaphors: A case of Japanese EFL learners.' *TESOL Quarterly* 44: 250–273.

**Zou, D. 2016.** 'Comparing Dictionary-induced Vocabulary Learning and Inferencing in the Context of Reading.' *Lexikos* 26.1: 372–390.