2020

Using literal underpinnings to help learners remember figurative idioms: Does the connection need to be crystal-clear?

Xinqing Wang
aileen.wang@vuw.ac.nz

Frank Boers
fboers@uwo.ca

Paul Warren
paul.warren@vuw.ac.nz

Citation of this paper:

Wang, X., Boers, F., & Warren, P. (2020). Using literal underpinnings to help learners remember figurative idioms: Does the connection need to be crystal-clear? In A. Piquer-Piriz & R. Alejo-Gonzalez (Eds.), *Metaphor in foreign language instruction* (pp. 221–239). Mouton de Gruyter.

Using literal underpinnings to help learners remember figurative idioms: Does the connection need to be crystal-clear?

Xinqing Wang, Frank Boers and Paul Warren

1. Introduction

Several idiom dictionaries for language learners include information on the origins or literal underpinnings of idioms (e.g., *American Heritage Dictionary of Idioms* 2003; *Collins Cobuild Idioms Dictionary* 2012; *Oxford Idioms Dictionary for Learners of English* 2009, and https://idioms.thefreedictionary.com/). This suggests that many dictionary makers assume that this kind of information must be helpful or at least appealing to learners. Its actual benefits for language learners have not yet been thoroughly evaluated, however. One proclaimed benefit (e.g., Liontas 2017) is that information about the origins of idioms helps learners to appreciate the connection between metaphorical language and culture. Another proclaimed benefit—and the focus of the present chapter—is that awareness of their literal underpinning can aid learners' retention of the idioms (e.g., Boers, Demecheleer, and Eyckmans 2004). This expectation rests on at least three theories.

One of these theories is that knowledge of the literal underpinning of idioms renders their meaning more transparent. For example, taking a back seat in a vehicle implies that someone else will be at the steering wheel and will thus likely be in control. By analogy, then, the meaning of the idiom *take a backseat* to refer to one's non-determining role in a project or activity makes sense against the backdrop of a more generic conceptual metaphor (Lakoff and Johnson 1980), according to which projects or activities are likened to journeys. In other words, the idiomatic meaning is *motivated*, to use a term from cognitive linguistics. If this indeed results in a clearer link between the

form and the idiomatic meaning of the expression, it may be expected to facilitate acquisition (e.g., Steinel, Hulstijn, and Steinel 2007).

The second theory is Paivio's (1986) *Dual Coding Theory*, according to which concrete concepts (or lexical items with concrete meaning) are easier to remember than abstract ones owing to their imageability, i.e., their association with a mental image of the referent. Although idioms have abstract meanings, resuscitating the context in which they were originally used in a literal sense is likely to evoke images of concrete scenes, and this may thus render them memorable.

The third relevant theory is Levels of Processing theory (e.g., Cermak and Craik 1979), according to which "deep" processing creates stronger memories than "shallow" processing. Mental operations – called *elaborations* – that build rich semantic associations around lexical items are considered deep in this model. Connecting the meaning of an idiom to its literal underpinning qualifies as an example of this, and the label coined by Boers et al. (2004) for this particular type of elaboration is *etymological elaboration*. The term etymological should be interpreted here broadly as the original context in which the expression was (and sometimes still is) used literally. We shall call this the literal underpinning of an idiom.

One may justifiably wonder whether language learners truly experience information about the literal underpinning of an idiom as helpful. There is evidence, however, that many learners presented with L2 idioms spontaneously activate images related to the literal meaning of constituent words (e.g., Cieślicka 2006). This inclination seems different from most native speakers, where the experimental evidence suggests that idiomatic meaning is by default accessed directly (e.g., Siyanova-Chanturia, Conklin, and Schmitt 2011), although imagery *can* be triggered also in L1 speakers of a language (e.g., Gibbs, Nayak, and Cutting 1989). If it is true that many learners are inclined to activate imagery when they process L2 idioms, then pointing them to the "right" literal underpinning of an idiom (or, at least, a literal reading that is congruent with the figurative meaning of the idiom) may harness an inclination that is already present. That pointing learners to that underpinning will often be necessary is illustrated in the following section.

2. Obstacles to learners' autonomous recognition of the literal-figurative link

Learners face diverse challenges should they try to establish a connection between the literal and figurative meanings of idioms. For starters, the idiom may contain a key content word that is simply not yet known by the learner, as would likely be the case for

low-frequency words (e.g., *in the doldrums; get short shrift; at the end of one's tether; bury the hatchet; pass on the baton*). In addition, content words may look deceptively transparent owing to homonymy. For example, learners may assume *suit* in *follow suit* refers to clothing (while it refers to playing cards) and they may assume *shot* in *a shot in the arm* refers to a bullet wound (not an injection). In a similar vein, polysemy may cause misinterpretations. For instance, *gun* in *jump the gun* could easily be misinterpreted as a weapon (rather than a starting pistol) and *wings* in *waiting in the wings* could be mistaken for a bird's wings (rather than a theatre's wings). It is also possible for learners to mistake new words for already known ones because of formal resemblance (e.g., *reign* for *rein* in *on a tight rein*, *limp* for *limb* in *out on a limb*, and *tower* for *towel* in *throw in the towel*). Because the wording of idioms is often elliptic, learners may have insufficient clues to work out a plausible literal reading. For example, it is unclear what *thin* and *thick* refer to in *through thick and thin* (allegedly referring originally to making one's way through thick and thin bushes).

Even if a learner does recognize the literal meaning of an expression, this does not guarantee an accurate connection between that literal reading and the actual idiomatic meaning. For example, literally *having a lot on one's plate* may well be experienced as a good thing (especially if one is feeling hungry), and so the more negative meaning of the idiom ('being very busy; having a lot of work') seems not to follow logically from this. Besides, understanding the literal underpinning of idioms may require cultural knowledge that the learner may not yet have. This regards culture-specific experiential domains. For instance, a learner may be unfamiliar with sports such as cricket and golf, and so fail to recognize the underpinning of *hit someone for six* and *be par for the course*. It also concerns knowledge of historical events, myths, legends and fables behind expressions such as *cross the Rubicon, hang on a thread*, *a Trojan horse* and *a white elephant*. Moreover, even familiar-looking content words may prompt associations that are not shared cross-culturally. For example, the heart is conventionally referred to in western culture as the seat of emotions such as romantic love, while reason supposedly resides in the mind. This duality seems absent from Chinese, however, where the heart can symbolize both reason and emotion. It is not surprising, then, that Chinese learners of English find idioms containing the words *heart* or *mind* comparatively hard to interpret (Hu and Fong 2010).

It also needs to be acknowledged that the literal—figurative link can simply be too obscure for the literal meaning to serve as a clue for interpretation. Examples here

include idioms such as *a red herring, kick the bucket* and *break a leg*. The origins proposed in dictionaries (e.g., that the distracting scent of smoked herring was used to train hunting dogs) may consequently be felt by learners to be rather far-fetched.

The general question raised by these observations is whether all literal-figurative links can be expected to be helpful. More specifically, the question we attempt to answer in the study reported in the present chapter is whether the degree of transparency of the literal-figurative connection matters for the mnemonic effect of etymological elaboration. First, however, we need to review a small number of earlier experiments that have evaluated this instructional approach to idiom learning.

3. Some earlier experiments

A number of studies have already shown that raising learners' awareness of the literal senses of polysemous words (e.g., *hurdles* referring to obstacles for athletes to jump over in track sports; *soar* in the sense of physical upward motion) benefits their comprehension and retention of the abstract or figurative uses of said words (e.g., *hurdles* referring to problems to be overcome; *soar* to denote upward trends in economics) (see Boers 2013, for a review). Focusing more specifically on the use of literal underpinnings in teaching figurative idioms, Boers et al. (2004) reported two experiments conducted with computer-aided exercises, where each idiom was tackled in three exercise components. In one component, learners were presented with the idiom in isolation and asked to choose the most likely domain of origin. For example, when presented with *jump the gun,* they could choose between source domains such as sports, war, jurisdiction, etc. Following this, the literal underpinning was displayed for learners (e.g., an athlete who *jumps the gun* in a racing contest sets off before the starting pistol has been fired). The second component asked learners to choose the correct figurative meaning of the expressions. For *jump the gun*, for example, they were given the choice between (a) defend someone at your own risk, (b) do something before the appropriate time, and (c) be startled by an unexpected event. The correct choice was subsequently pointed out to them. The third component presented learners with a completion task, where the idiom was incorporated in a meaningful context and a missing content word needed to be supplied. For example, "Although we had agreed not to tell anyone about my pregnancy yet until we were absolutely sure, my husband jumped the _____ and told his parents straightaway." This exercise was used as a post-test to evaluate the effectiveness of the preceding steps.

The first experiment reported in Boers et al. (2004) aimed to assess the mnemonic effect of presenting the literal underpinning of the idioms. In one condition learners tackled the multiple-choice exercise on origins, and they were then asked to do the completion exercise, i.e., they skipped the exercise on the actual figurative meaning of the idioms. In the other condition, the learners did the multiple-choice exercise on the figurative meaning of the idioms and they were then asked to do the completion exercise. The former condition was found to yield the better scores in the completion exercise, which the authors took as evidence of the mnemonic effectiveness of etymological elaboration. Of particular relevance for the present chapter is the distinction which Boers et al. then made between idioms whose source domains were identified readily in the multiple-choice exercise and those which were not. In the case of idioms whose source domains were hard to guess, recall turned out *not* to be better after etymological elaboration than in the comparison condition.

In the second experiment reported in Boers et al. (2004), all participants tackled the three exercise components of the program, in the order of (1) the multiple-choice exercise on the figurative meaning of the idioms (followed by feedback), (2) the multiple-choice exercise on the idioms' origins, and (3) the completion exercise, which served as a post-test again. This time, the average post-test scores for idioms whose source domains were hard to guess were very similar to those for idioms with better guessable source domains.

Leaving aside the inconsistent nature of the findings reported in Boers et al.'s (2004) two experiments, what the findings do *not* tell us is whether it matters if learners find the proposed motivation of an idiom's meaning transparent *after* it is presented to them. For example, learners may understandably fail to guess that *jump the gun* originates from track sports, but they might nonetheless find that the latter explanation makes good sense when it is subsequently given to them. By contrast, in a case such as *follow suit* they may not only fail to guess that the expression originates from a card game, but perhaps also find the information about this origin non-illuminating if they happen to be unfamiliar with that card game. In the experiment reported below, we therefore asked learners to evaluate the connection between the proposed literal underpinning of idioms and their figurative meaning after these were both explained to them. A week later, the participants were asked to recall the meaning of the idioms, and so we could examine if there was an association between each learner's ability to recall an idiom and their appreciation of its proposed underpinning.

First, however, we need to point out that not all experimental evidence to date has been favourable of etymological elaboration. Szczepaniak and Lew (2011) asked upper-intermediate EFL learners to study booklets with information about 18 idioms for ten minutes under one of four presentation conditions, with differing amounts of information. The minimal condition consisted of a definition of the idiomatic meaning and an example sentence. The richest condition included, in addition, a picture of a literal interpretation of the expression and an etymological note. The learners were subsequently tested on their recall of the form of the idioms (by means of a completion task where they were required to supply a missing content word) and on their recognition of the meaning of the idioms (by means of a multiple-choice task). The test was administered again three weeks later. While both the immediate and the delayed post-tests furnished evidence in favour of adding pictures to the definitions of idiomatic meaning, neither produced evidence of the effectiveness of etymological notes. It could be argued that the ten minutes of study time allowed in Szczepaniak and Lew (2011) was perhaps insufficient for the participants who were presented with the (longish) etymological notes. It takes much less time and effort to take in a picture than a text passage, after all (Boers, Warren, Grimshaw, and Siyanova-Chanturia 2017). Because this was exclusively a pen-and-paper experiment, it is not known to what extent the participants engaged with the etymological notes that were included in the booklets. If they did not actually take in the information provided there, then it would not be surprising that it left no impact. In our experiment, we therefore used a one-on-one interview procedure to make sure the participants did consider the information given about the idioms' literal underpinnings.

A second possible account for the lack of evidence in favour of the etymological elaboration in Szczepaniak and Lew's experiment could be that a fair number of the target idioms used (e.g. *a white elephant*, *a loose cannon*, *a red herring* and *have an axe to grind*) were arguably ones where the learners found it hard to follow the explanations for the connection between the origin and the actual idiomatic meaning. Szczepaniak and Lew did not examine (through an analysis at the level of individual items) whether etymological notes perceived to be relatively transparent by the participants were the more helpful ones. In the study we report next, we did examine this.

Something else that is largely missing from the available body of research on L2 idiom learning is the potential role of individual learner traits. Exceptions are explorations of the role of (young) age on L2 metaphor comprehension (Piquer-Piriz, 2008) and explorations of so-called cognitive style differences among learners when they are taught

idioms with the etymological elaboration technique (Boers, Eyckmans, and Stengers, 2006). It is conceivable that some learners experience etymological elaborations as more useful than other learners do and are thus more inclined to appreciate the potential mnemonic potential of such elaborations even in cases where they seem far-fetched. By contrast, for learners who find it harder to appreciate the connection between an idiom's meaning and its proposed literal underpinning, the mnemonic benefits of etymological elaboration may be confined mostly to relatively transparent connections. What seems worth exploring, then, is whether such predispositions might be related to L2 student profiles more generally.

For example, one may wonder if high achievers within a given student population reap the mnemonic benefits of etymological elaboration more readily than their comparatively low-achieving peers. This possibility appears compatible with findings from research on language-learning aptitude (Robinson, 2013; Skehan, 2015), where one of the known predictors of learning success is associative memory ability. High achievers are likely to have an advantage in this regard, and, in theory, this could also apply to their remembering the meaning of an idiom in association with additional information, such as its literal underpinning.

The possibility that high achievers might benefit the most from etymological elaborations is also compatible with research on vocabulary learning strategies (Gu, 2003, 2013) which suggests that it is individuals who welcome a wide range of strategies who tend to be relatively successful learners. In addition, the high achievers within a student population are also likely to have built a larger L2 vocabulary and to have developed greater familiarity with the L2 in general. If so, they will be more likely to be familiar with the lexical constituents of new idioms, and this could reduce the learning burden. It is now relatively well established that the ability to learn new L2 lexical items tends to increase as one becomes more proficient in the language (e.g., Elgort and Warren, 2014), which is sometimes referred to as the Matthew effect (i.e., "the rich get richer faster").

In sum, there are indeed grounds for hypothesizing that, within an otherwise homogenous population of L2 students, those who have been comparatively successful may be more inclined to engage with and remember the etymological elaborations proposed by a teacher, regardless of whether some of these elaborations seem far-fetched. An evaluation of this hypothesis will therefore be part of the present study as well.

4. The present study

4.1 Method

4.1.1 Participants

The participants ($N = 25$), 22 females and 3 males aged between 19 and 22 (medium age = 20), were Chinese EFL learners majoring in English at a university in mainland China. They were all in their third year of study at the university. They shared very similar histories of EFL learning and were all considered of have intermediate-to-high proficiency in English. They thus made up a rather homogenous population. They had all passed the Test for English Majors 4 (TEM-4) before the end of their second year at the university. This criterion-referenced test is widely used in China to gauge the English proficiency of university undergraduate English majors in accordance with the National College English Teaching Syllabus for English Majors. The mean grade of the participants on this test was 75.72 (SD 6.52), and the grades ranged from 61 to 89 (and showed a normal distribution). According to the TEM descriptors, a grade of 60-69 qualifies as a pass grade, one of 70-79 is considered a good grade, and one of 80 or above is an excellent grade. So, although the sample of participants was homogenous in terms of L1 background and EFL learning history, the TEM-4 grades demonstrate different levels of EFL achievement. Whether this made a difference to their performance in the actual study will be explored further below.

The participants were informed the study was about strategies for learning English idioms, but the precise purpose was only explained to them after data collection was completed. They all gave written consent for their data to be used for this research project.

4.1.2 Instruments and procedure

The idioms presented to these participants were semi-randomly selected from dictionaries such as *Collins Cobuild Idioms Dictionary*, *Oxford Idioms Dictionary for Learners of English,* and *American Heritage Dictionary of idioms*. These were all expressions whose dictionary entries included notes about their origins (not all entries do) and/or for which origins are proposed in online resources (e.g., *the Phrase Finder* on https://www.phrases.org.uk/ and *the Free Dictionary* on https://idioms.thefreedictionary.com/). As we were interested in the role of item properties, we needed a substantial number of idioms. Altogether, 80 idioms without close equivalents in the learners' L1 (Mandarin) were selected. Because interviewing each participant about so many idioms was deemed unrealistic, the collection of idioms

was divided into four equal sets of 20 idioms each. Each set was then used in interviews with separate groups of six (and in one case, seven) participants. For motivational reasons, we also added a small number of idioms that do have close equivalents in Mandarin and that were thus easier to guess the meaning of, but these were excluded from the analysis.

The two interviews were conducted a week apart. During the first interview (which took about an hour on average), the learner was informed about the meaning and the literal underpinning of the idioms and was asked to evaluate the clarity of the link between the two. The second, shorter, interview served to determine how well the participants remembered the meaning of the idioms, and if they remembered this in conjunction with the literal underpinnings they had been told about. The interaction in both interviews was in the participants' L1 whenever they preferred so. All the interviews were audio-recorded and transcribed. The Mandarin excerpts were translated into English by the first author (who is a native speaker of Mandarin) to facilitate inter-rater reliability procedures (see below).

The first interview proceeded along the following steps per idiom:

1. The idiom was presented in isolation (e.g., *pull one's weight*) and the participant was asked if he or she was familiar with it, and, if so, to give its meaning. The purpose of this step was merely to ascertain that the participant was not yet familiar with the idiom. If a participant did give evidence that he or she already knew the idiom, the interview moved on to the next idiom on the list.

2. The idiom was presented in a short sentential context without revelatory semantic clues (e.g., *He needs to pull his weight*), and the interviewee was asked to hazard a guess at the meaning of the expression. Embedding the idiom in a sentence served to illustrate at least the word class of its constituents and thus made the interpretation task more realistic than guessing the meaning of a completely decontextualized expression. At the same time, keeping the sentential context minimal was a way of avoiding variability in the interpretability of the idioms owing to a factor (i.e., the availability of contextual clues) extraneous to the semantic transparency of the idioms as such.

3. The origin of the expression was explained and then the interviewee was asked to make another guess at the figurative meaning of the idiom (or if the new information confirmed their earlier interpretation). The information about the origin of the idioms was taken from dictionaries or online resources. For example, according to *American Heritage Dictionary of Idioms*, *pull one's weight* comes from rowing, where all crew members are expected to pull on their oar(s) to give momentum to the boat. In case more than one

origin was proposed by different sources, the one that seemed the most plausible was chosen.

4. The researcher then explained the meaning of the idiom given in the dictionary (or confirmed the interpretation proposed by the participant in the previous step, if that happened to be correct). The idiom was then also presented in a slightly more elaborate context to illustrate its meaning and use (e.g., *If he doesn't start pulling his weight, he'll lose his job*). These example contexts were borrowed mostly from the dictionaries. After both the literal underpinning and the actual idiomatic meaning of the expression were clarified, the interview moved on to the next idiom in the set.

After their set of 20 idioms had been tackled following the above steps, the participants were given a list of them, preceded by the following instruction:

You may feel that the figurative (idiomatic) meaning of some idioms follows in a straightforward manner from the literal (source) meaning of the expressions. For other idioms, you may find it much less obvious how the figurative (idiomatic) meaning is derived from the literal (source) meaning. You may even feel that there is no clear link at all. Now that you know the literal meaning as well as the figurative meaning of the idioms, please circle the number on the scale below that best represents how you feel about the relationship between the literal meaning and the figurative meaning of the idiom.

5= The link between the literal and the figurative meaning of the expression is very clear to me. Given the literal use of the expression, it is easy for me to see how the figurative use is derived from it.

0 = The link between the literal and the figurative meaning of the expression is not at all clear to me. Although I (now) know both the literal (source) meaning of the expression and its idiomatic meaning, I cannot see how the literal use could have given rise to the figurative use of the expression.

      non-transparent   *transparent*

⟵⟶

0	*1*	*2*	*3*	*4*	*5*
The link between the literal and figurative meaning is not transparent at all.	*The link between the literal and figurative meaning is very vague.*	*The link between the literal and figurative meaning is vague.*	*The link between the literal and figurative meaning is somewhat clear.*	*The link between the literal and figurative meaning is clear.*	*The link between the literal and figurative meaning is very clear.*

The participants were encouraged to express their thoughts (in L1 if they wished) and explain their reasoning as they carried out this rating task per idiom. At the end of the interview, the participants were asked not to discuss individual items encountered in the interview with other students. They were not told they would be asked to recall the meaning of the idioms the following week.

One week later, in the second interview, the participants were presented with the same set of idioms and asked to try and recall their figurative meaning. They were also asked if they could recall anything else they had learned about the idioms in the first interview. The latter question was intended to gauge if successful recall of idiomatic meaning often coincided with recall of the literal underpinnings.

4.1.3 Analysis

The participants' meaning guesses in the first interview and their meaning recall responses in the second interview were assessed by the first author and two L1 speakers of English (PhD students in Applied Linguistics). A three-point scoring protocol was adopted, distinguishing between no or incorrect responses, partially correct responses, and fully correct responses. It was agreed that responses were counted as correct if they contained all the meaning aspects contained in the definition from the dictionary that was used in the interview. For example, *pull one's weight* was defined as doing one's share in a common task—or, in other words, to work as hard as other people in a group. If the response was "to do one's share in teamwork", it was rated as 1. If it was just "to make an effort" or "to do one's best", it was rated as 0.5. In case of disagreement among the three assessors, a verdict was made by majority vote.

As explained, step one in the interview sequence helped us to identify which idioms were already known by any given participant. For example, some participants remembered learning idioms such as *let the cat out of the bag, a rule of thumb* and *make ends meet* from their EFL textbooks. Step two revealed which (if any) idioms a participant correctly guessed the meaning of even before any hint about its origin was given. New idioms that at least some participants managed to guess the meaning of in step 2 included *a drop in a bucket, have a lot on one's plate* and *in the driver's seat*. If a learner already knew or understood a given idiom, such an item would likely still be known and understood one week later also in the absence of any intervention, and so these instances were discarded from further analysis. Some initial meaning responses generated in step 2 were not fully correct and thus given a score of 0.5. These instances were retained for

analysis because they did entail room for further learning. This pruning procedure left us with 360 initial responses indicating no knowledge or comprehension of the given idioms and an additional 54 initial responses demonstrating partial comprehension of the given idioms, that is, 414 out of the total 500 initial responses (25 participants x 20 idioms). The idioms, with exclusion of those which all members of the group already knew or interpreted correctly at the start, are listed in the appendix.

4.2 Results and Discussion

Informing the participants about the origin of the expressions was found to considerably increase the likelihood of correct inferences in the first interview. In 49.7% of the 360 instances where participants had initially failed entirely to propose a correct interpretation, they now produced correct interpretations, and in an additional 19.5% they produced partially correct ones. For the 54 instances where guessing had initially already been partially successful, the information about the origin of the idiom also occasionally (11.1%) helped learners to arrive at a fuller understanding.

One week later, no fewer than 65% of the items which participants had demonstrated no comprehension of at the start of the first interview were recalled correctly, and an additional 10.5% of the recalls were partially correct. In addition, 44.5% of the items that had already elicited partially correct guesses at the very beginning were now accurately recalled. In sum, learning gains were thus attested for close to 71.5% of the instances where learners' prior comprehension was either nil or incomplete. When the participants recalled the precise meaning of an idiom in the second interview, this typically (91.25%) coincided with their recollection of the idiom's suggested origin, as displayed in Table 1. If they failed to recall an idiom's meaning altogether, they would often (65.59%) also fail to recall its origin.

Table 1. *Meaning recall and recall of literal underpinnings (total N = 414)*

Meaning recall	Recall of origin			No recall of origin	
	n	%		n	%
Correct	240	91.25	Versus	23	8.75
Partially correct	45	77.59		13	22.41
Failed	32	34.41		61	65.59

The next question is whether the proposed literal underpinnings of the idioms were perceived by the learners as relatively transparent motivations of the idioms' meanings. The participants' ratings of the degree of transparency (at the end of the first interview) were very unevenly spread across the six-point scale, with the lowest points, 0 to 2, ticked very seldom (together only about 8%). Idioms and their motivations that attracted such occasional low ratings included *on the back burner, let the cat out of the bag*, *take a backseat, hit the roof, a loose cannon, a wet blanket, hit the roof*, *jump the gun, have cold feet*, *be on the same page, in the wake of, not be up to scratch, hold your horses, bark up the wrong tree* and *hand over fist*. The two highest points on the scale, 4 and 5, were selected the most often (together about 73.5%), which suggests that, by and large, the participants thought the origin of the idioms that they had been presented with offered a relatively clear motivation for the idioms' meanings. It is worth mentioning, however, that there was considerable disparity among participants' judgements. It was not uncommon for idioms to receive a rating of 1 or 2 from one learner but ratings of 4 or 5 from others. Making predictions about which "etymological notes" will strike individual learners as clear vs. far-fetched thus appears problematic, even within a relatively homogenous group of learners. This is also one of the reasons why we felt it was important to analyse the data by individual responses instead of using averaged ratings (see below).

It is somewhat ironic that the ratings for some of the idioms which we mentioned earlier as examples of likely transparent versus non-transparent cases did not confirm our expectations. For instance, *take a backseat* elicited an average rating of only 3.50, whereas *red herring* elicited an average rating of 4.33. We need to be cautious about these comparisons, though, because they cross over from one idiom set to another and do thus not reflect the same learners' judgements. Still, the data do suggest that it may be very hard for teachers, lexicographers and researchers to make reliable predictions about which motivations for idioms' meanings will be experienced as "making good sense" by an individual learner or group of learners.

The literal underpinning that elicited the lowest average rating (the only average below 3.00) concerned *a wet blanket.* One participant explained his low rating (0) as follows: "If it [the idiom] comes from putting out fire, I see it as something useful in a difficult or dangerous situation. This is in contrast with the negative meaning of the expression as stopping other people's enthusiasm." (translated from Mandarin by the first author). It seems this learner found the proposed source-target domain mapping

improbable, because of his association of fire with danger. His low transparency rating did not, however, prevent him from accurately recalling both the idiom's meaning and its proposed origin one week later.

This brings us to the key question we set out to address here, notably whether the degree of perceived transparency of the literal-figurative link influences the mnemonic effect of etymological elaboration. Table 2 gives a first impression of a trend that appears to be in favour of idioms receiving relatively high transparency ratings. Because the participants' ratings were so unevenly distributed, we divide them here into just two categories for preliminary, descriptive purposes: ratings 0 to 3 ("low") versus ratings 4 and 5 ("high").

Table 2. Meaning recall: transparent vs. non-transparent underpinnings (N = 414)

	Correct		partially correct		failed	
Transparency	n	%	n	%	n	%
0-3	66	60.55	14	12.84	29	26.60
4-5	192	62.95	49	13.11	64	20.98

What Table 2 fails to capture, of course, is variation in recall successes due to many other potentially influential characteristics of individual idioms as well as characteristics of the individual participants. Regarding the latter (and as discussed previously), the students' comparative success as EFL learners may be particularly relevant.

We therefore turned to mixed-effects regression models for ordinal data, using *clmm* (cumulative link mixed modelling) in the *ordinal* package in R (Christensen, 2018). The dependent variable was the Recall score, with the scores of 0, 0.5 and 1, representing ordered categories of wrong/null, partially correct and fully correct. The fixed effect predictors were Transparency (i.e., the learner's rating of the transparency of the proposed connection between the literal underpinning and the meaning of the idiom) and the student's TEM-4 grade (i.e., the measure of the learner's EFL achievement or English proficiency). Since we hypothesized (see above) that the effect of perceived transparency may vary with the level of EFL achievement or proficiency, we included the interaction between Transparency and TEM-4 grade as another predictor. The random effects were participants and items (i.e., the different idioms).

Including these random effects allowed the models to take into account variation both between participants and between idioms. Comparison of models with and without the interaction of Transparency and EFL achievement showed that the interaction made a significant contribution to explaining the variance in recall scores (χ^2 (1) = 4.72, p < .05). The full model, with the two simple fixed effects and their interaction, was therefore retained. A Type-III ANOVA test showed that in addition to the interaction, the simple fixed effects of Transparency (χ^2 (1) = 4.01, p < .05) and TEM-4 grade (χ^2 (1) = 7.13, p < .01) also made significant contributions to the full model. The interaction effect is illustrated in Figure 1[1].
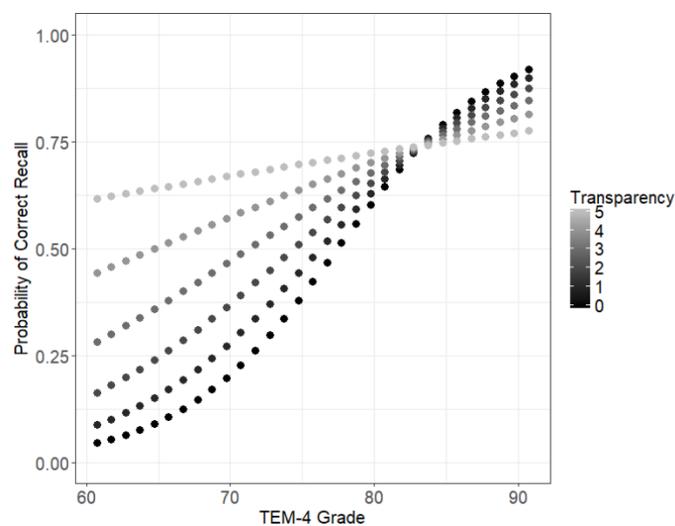

Figure 1. Probability of correct recall based on TEM-4 grade and Transparency

These effects reflect the following patterns. Overall, fully correct recall was more likely for idioms perceived to have transparent literal-figurative connections, while unsuccessful recall was more likely for those whose proposed underpinnings were perceived to be rather obscure. This supports an affirmative answer to our principal research question: Transparency matters. Fully correct recall was also more likely for the learners with comparatively high TEM-4 scores. This is in line with our prediction that the high-achieving students in the EFL programme would also be the high achievers in our experiment.

---

[1] Because the *predict* function has not been developed for *clmm*, this plot shows values derived using the *predict* function in *clmm2*. However, because *clmm2* only allows one random effect, we chose to keep items as the random effect since items explain more variance than participants.

The interaction effect revealed by the regression model is as follows. While the students with low TEM-4 grades performed generally more poorly than their high-achieving peers on the recall test, this was especially acute for idioms whose transparency they had rated as low. Idioms whose literal-figurative connection they deemed highly transparent were recalled the best at this lower end of the TEM-4 grades. However, as TEM-4 grades increased, the impact of transparency gradually diminished and disappeared by grade 80 (see Figure 1). So, if the transparency factor was found to be a predictor of recall, this is attributable to the low achievers' data. Returning to the anecdote of the student who had given *a wet blanket* the lowest transparency rating but nonetheless successfully recalled it, it is perhaps no coincidence that this was a student with a relatively high TEM-4 grade (78). According to the interview data, this student was clearly willing to reflect on the literal underpinning of the idiom proposed by the researcher, perhaps precisely because its motivation for the idiomatic meaning was puzzling to him. It is possibly this willingness to put a certain effort into evaluating a proposed literal-figurative connection that helped him to entrench this association in memory (recall that the student remembered both the idiom's meaning and its proposed underpinning).

5. Conclusion

The findings presented here suggest that, if etymological elaboration aids learners' retention of the meaning of idioms, its effect is not confined to idioms whose meaning learners find to be straightforwardly derived from the proposed origins. Rather, it appears that explanations about origins that are experienced as somewhat "far-fetched" can also serve this mnemonic purpose. However, the findings also indicate that this mnemonic purpose is generally served more easily in the case of idioms where the learner finds the proposed literal underpinning comparatively straightforward or plausible. Interestingly, the interaction effect that emerged from our mixed effects regression model suggests that this influence of the perceived transparency of the literal-figurative connection was the greatest for students who obtained comparatively low TEM-4 exam grades, and who could by that proxy be considered the less successful or less advanced EFL students in the sample. For the participants with the higher TEM-4 grades, who could be considered the high achievers in this student population, there was no noticeable impact of the transparency variable on idiom recall.

Despite the role of said variables, the fact remains that most of the students in this study gained knowledge of a considerable number of new idioms through an instructional procedure which engaged them with the literal underpinnings of the expressions. We need to be cautious not to oversell the proclaimed benefits of this etymological elaboration approach, however. As mentioned, not all evidence to date has been favourable of its implementation (Szczepaniak and Lew 2011). We also need to bear in mind that the present study did not include a comparison treatment. As such, it provides no evidence that learning idioms through etymological elaboration should be given precedence over other learning procedures. The jury is still out, so to speak. What we can say with a degree of confidence is that the learning gains attested here were quite substantial, at least according to a one-week delayed post-test. Of course, a further delayed post-test would be required to check if the observed learning gains are truly durable.

At the same time, it is possible that idiom knowledge was underestimated in our experiment. That is because we only gave full credit to learners' paraphrases of the idioms if these paraphrases included all the meaning components mentioned in the dictionary entry that was used as benchmark. It cannot be ruled out that one or the other meaning component was implied rather than explicitly verbalized in a respondent's paraphrase. Even in an interview procedure that takes recourse to the respondent's L1, an interviewer's efforts to solicit nuanced knowledge may fail. In future (conceptual) replications, it might be worth adding a different type of test, such as a multiple-choice meaning-recognition test where the respondent is required to select the meaning paraphrase that best captures the idiom's meaning.

If degree of transparency of the literal-figurative connection was found to play a role in learners' recall of the idioms' meanings in this study, this raises the question what other variables might play a part in this regard. Further analyses of the interview and test data and inclusion of alternative idiom-related variables in new mixed effects models will be required to identify these. In a similar vein, the potential influence of other learner traits than proficiency merits further investigation in future research on L2 idiom learning and teaching.

Finally, we need to be aware that the learning aim in the present experiment may in one respect be considered modest—that is, remembering the meaning of the idioms. For one thing, the post-test presented the learners with the English idioms again, and thus did not require them to recall the lexical makeup of the expressions. For another, it is well documented that idioms serve subtle pragmatic and evaluative functions in discourse

(e.g., O'Keeffe, McCarthy and Carter 2007: 80–99) which are often not captured by dictionary definitions. Developing expert productive knowledge of idioms would take extensive exposure to L2 natural discourse (MacArthur 2010). It would nonetheless be interesting to explore, by means of a longitudinal investigation, if the intervention we have described here stimulates learners' long-term engagement with the idiomatic dimension of their target language.

Acknowledgements

References

Boers, Frank. 2013. Cognitive Semantic ways of teaching figurative phrases: An assessment. In Francisco Gonzalvez-Garcia, Maria S. Pena Cervel & Lorena Perez Hernandez (eds.), *Metaphor and metonymy revisited beyond the contemporary theory of metaphor,* 229–263. Amsterdam: John Benjamins.

Boers, Frank, Murielle Demecheleer & June Eyckmans. 2004. Etymological elaboration as a strategy for learning figurative idioms. In Paul Bogaards & Batia Laufer. (eds.), *Vocabulary in a second language: Selection, acquisition and testing*, 53–78. Amsterdam: John Benjamins.

Boers, Frank, June Eyckmans & Helene Stengers. 2006. Motivating multiword units: Rationale, mnemonic benefits, and cognitive style variables. In Susan. H. Foster-Cohen, Marta M. Krajnovic, & Jelena M. Djigunovic (eds.), *EUROSLA Yearbook Vol. 6*, 169–190. Amsterdam: John Benjamins.

Boers, Frank, Paul Warren, Gina Grimshaw & Anna Siyanova-Chanturia. 2017. On the benefits of multimodal annotations for vocabulary uptake from reading. *Computer Assisted Language Learning* 30. 709–725.

Cermak, Laird S. & Fergus I. M. Craik. 1979. *Levels of processing in human memory*. Hillsdale, NJ: Lawrence Erlbaum.

Cieślicka, Anna B. 2006. Literal salience in on-line processing of idiomatic expressions by second language learners. *Second Language Research* 22. 115-144.

Christensen, Rune H. B. 2018. Ordinal - regression models for ordinal data. R package version 2018.4-19. http://www.cran.r-project.org/package=ordinal/.

Elgort, Irina & Paul Warren. 2014. L2 vocabulary learning from reading: Explicit and tacit lexical knowledge and the role of learner and item variables. *Language Learning* 64. 365-414.

Gibbs, Raymond W., Nandini P. Nayak & Cooper Cutting. 1989. How to kick the bucket and not decompose: Analyzability and idiom processing. *Journal of Memory and Language* 28. 576-593.

Gu, Yongqi. 2003. Fine brush and freehand: The vocabulary learning art of two successful Chinese EFL learners. *TESOL Quarterly* 37. 73-104.

Gu, Yongqi. 2013. Vocabulary learning strategies. In Carole A. Chapelle (ed.), *The encyclopedia of applied linguistics*. Oxford: Wiley-Blackwell.

Hu, Ying H. & Yu Y. Fong. 2010. Obstacles to conceptual-metaphor guided L2 idiom interpretation. In Sabine De Knop, Frank Boers & Antoon De Rycker (eds.), *Fostering language teaching efficiency through Cognitive Linguistics*, 293-317. Berlin: Mouton de Gruyter.

Lakoff, George & Mark Johnson. 1980. *Metaphors we live by*. Chicago, IL: University of Chicago Press.

Liontas, John I. 2017. Why teach idioms? A challenge to the profession. *Iranian Journal of Language Teaching Research* 5(3): 5-25.

MacArthur, Fiona. 2010. Metaphorical competence in EFL: Where are we and where should we be going? A view from the language classroom. *AILA Review* 23. 155-173.

O'Keeffe, Anne, Michael McCarthy & Ron Carter. 2007. *From corpus to classroom: Language use and language teaching*. Cambridge, UK: Cambridge University Press.

Paivio, Allan. 1986. *Mental representations: A dual coding approach*. New York, NY: Oxford University Press.

Piquer-Píriz, Ana M. (2008). Reasoning figuratively in early EFL: some implications for the development of vocabulary. In Frank Boers & Seth. Lindstromberg (Eds.), *Cognitive Linguistic Approaches to Teaching Vocabulary and Phraseology* (pp. 233–257). Berlin: Mouton de Gruyter.

Robinson, Peter. 2013. Aptitude in second language acquisition. In Carole A. Chapelle (ed.), *The encyclopedia of applied linguistics*. Oxford: Wiley-Blackwell.

Siyanova-Chanturia, Anna, Kathy Conklin & Norbert Schmitt. 2011. Adding more fuel to the fire: An eye-tracking study of idiom processing by native and non-native speakers. *Second Language Research*, 27. 251−272.

Skehan, Peter. 2015. Foreign language aptitude and its relationship with grammar: A critical overview. *Applied Linguistics* 36. 367-384.

Steinel, Margarita. P., Jan, H. Hulstijn & Wolfgang Steinel. 2007. Second language idiom learning in a paired associate paradigm: Effects of direction of learning, direction of testing, idiom imageability, and idiom transparency. *Studies in Second Language Acquisition* 29. 449-484.

Szczepaniak, Renata & Robert Lew. 2011. The role of imagery in dictionaries of idioms. *Applied Linguistics* 32. 323-347.

Appendix:

Target idioms (excluding ones known/understood by all respondents in the group already before the instructional procedure)

<u>Set A</u>: go belly-up; a drop in the bucket; on the same wavelength; get into gear; red tape; follow suit; pull one's weight; your bread and butter; throw in the towel; a sitting duck; give the green light; spill the beans; have/get cold feet; jump ship; play into someone's hand; par for the course; on the back burner; take the bull by the horns; take it on the chin.

<u>Set B</u>: the ball is in your court; win hands down; on the ropes; weigh someone down; pass the baton; give someone the cold shoulder; leave someone high and dry; in the doldrums; beat around the bush; waiting in the wings; a shot in the arm; throw your hand in; a loose cannon; on the same page; let the cat out of the bag; (not) up to scratch; take a back seat; hit the roof/ceiling; a feeding frenzy.

<u>Set C</u>: in the driving/driver's seat; play your cards close to your chest; down and out; rub someone the wrong way; turn over a new leaf; (hit) below the belt; make ends meet; ring a bell; a red herring; bury the hatchet; have a lot on your plate; sit on the fence; pass the buck; jump the gun; stick your neck out; take the plunge; teething problems; hold your horses; in the wake of something; a rule of thumb; show someone the ropes.

<u>Set D</u>: a hot potato; have a green thumb; go with the flow; (start) from scratch; a can of worms; flex your muscles; give someone a leg up; pull a rabbit out of the hat; bite the bullet; get your second wind; cut corners; cut corners; come out of your shell; get something off your chest; hand over fist; throw your hat/cap into the ring; play it by ear; a wet blanket.