Cognitive Semantic ways of teaching figurative phrases

An assessment

Frank Boers

In this article I review studies published between 1996 and 2010 in which the effectiveness of Cognitive-Semantics informed second language pedagogy was put to the test. Altogether, the published evidence is manifestly favourable, although questions remain as to the scope of application of the approach and the precise properties that produce its positive effects. It must also be recognised that Cognitive Semantic ventures into language pedagogy stand a lot to gain from a closer collaboration with 'mainstream' applied linguistics, not only with regard to general insights into the nature of second language acquisition but also with regard to this type of research methodology.

Keywords: instructed second language acquisition, idioms, phrasal verbs, motivation, metaphor awareness, comprehension, retention, receptive versus productive knowledge

1. Introduction

Lakoff and Johnson's *Metaphors We live By* (1980), which lay the foundations of Conceptual Metaphor Theory (CMT), and subsequent books that shaped the Cognitive Semantics (CS) movement (e.g., Gibbs, 1994; Lakoff, 1987), in which CMT features prominently, demonstrated that a lot of so-called idiomatic language was actually far less random and less arbitrary than had long been assumed. It soon became clear that this new way of thinking about allegedly 'dead' or 'fossilised' metaphors could be of use for second language instruction. The idea that the semantics of idioms, phrasal verbs[1] and figuratively used words in general might be explainable with reference to general conceptual metaphors held a certain promise for language pedagogy, because it suddenly looked as though great chunks of

Review of Cognitive Linguistics 9:1 (2011), **227–261**. DOI 10.1075/ml.9.1.11boe
ISSN 1877–9751 / E-ISSN 1572–0276

language which had hitherto seemed un-teachable could be made easier to learn after all. A proliferation of publications have followed which argue in favour of introducing CMT to the second language classroom (e.g., Baker, 1998; Boers & Demecheleer, 1998; Boers, 2000a/b; Danesi, 1992; Deignan et al., 1997; Dirven, 2001; Hannan, 1998; Holme, 2001, 2004; Kövecses & Szabó, 1996; Kövecses, 2001; Kurtyka, 2001; X. Li, 2010; Lindstromberg, 1991; Littlemore & Low, 2006; Littlemore, 2009, pp. 94–105; MacLennan, 1994; Ponterotto, 1994; Rundell, 2001, 2002; Scott, 1994). The rationale put forward in these publications typically incorporates the following tenets.

a. Revealing the motivation behind the meaning of L2 phrases opens up pathways for insightful learning, and insightful learning is believed to be superior to rote learning. Moreover, if learners recognise the conceptual metaphors, then we may assume they will be in a better position to work out the meaning of newly encountered L2 figurative phrases than if they were to rely on contextual clues only (e.g., Kövecses & Szabó, 1996, p. 351).
b. CMT shows how seemingly unrelated expressions are actually associated through a shared conceptual metaphor or by a shared source domain. CMT thus provides a possibility of presenting learners with L2 figurative phrases grouped in distinct sets. The phrases can then be presented in an organised fashion, and organised vocabulary is assumed to be easier to learn than a semantically random presentation of vocabulary (e.g., Boers, 2000a, p. 563).
c. Making learners conscious of the metaphorical nature of certain phrases involves imparting a heightened awareness of the literal, concrete meaning of words contained in these phrases. This concreteness stimulates mental imagery, and associating vocabulary with images is believed to make that vocabulary more memorable (e.g., Boers, 2000a, p. 563).

While these tenets seem at first sight to be perfectly commonsensical, we shall see further below that they should not be taken strictly at face value.

Apart from making pleas for the introduction of CMT to second language instruction, some authors have gone on to actually develop language learning materials, lesson plans and exercises aimed at raising students' awareness of metaphor generally and the motivations behind particular sets of L2 figurative phrases (e.g., Boers, 2000a, pp. 566–568; Boers & Lindstromberg, 2008; Lindstromberg & Boers 2008a, pp. 55–67; Juchem & Krennmayr, 2010; Rudzka-Ostyn, 2003). It is therefore reassuring that a fair number of so-called intervention studies lend support for the proposed types of instruction. In this article I will review the publications which contain reports of such quasi experimental studies, with a view to (i) evaluating the strength of the evidence available so far, (ii) fine-tuning the interpretation of early findings in the light of more recent insights, and (iii) suggesting items

for a research agenda in this domain. I will also briefly discuss the extent to which the CS proposals chime (or clash) with insights from mainstream vocabulary research. First, however, it may be helpful to illustrate more concretely what is understood by CS ways of teaching figurative phrases.

2. Cognitive Semantic ways of teaching figurative phrases: Some examples

The most direct translation of CMT for purposes of L2 vocabulary instruction is the presentation of vocabulary in sets of phrases that instantiate a shared conceptual metaphor. For example, phrasal verbs can be grouped according to the metaphors that motivate the use of the particle, as follows:

MORE IS UP; LESS IS DOWN: *cut down expenses; turn up the heating; the story was blown up*.
ACTIVE IS UP; INACTIVE IS DOWN: *they set up a business; the car broke down; the factory closed down*.
GOOD IS UP; BAD IS DOWN: *cheer up; feel up to a task; feeling down*.
KNOWING IS SEEING and VISIBLE IS UP (IN YOUR FIELD OF VISION): *the teacher turned up late; she never showed up; look it up in the dictionary*.
KNOWING IS SEEING and VISIBLE IS OUT (IN THE OPEN): *I figured it out; it turned out difficult; she found out*.

Idiomatic expressions referring to emotions have been a particularly popular subject in proposals for this type of organised input, probably because inventories of them were readily available from early CMT case studies. For example, expressions for anger can be grouped as follows (e.g., Kövecses, 1986):

ANGER IS HEAT: *Simmer down; Keep cool; She blew up at me; He's blowing off steam; She erupted; I reached my boiling point; She just exploded; I was fuming; It just added fuel to the fire*.
ANGER IS INSANITY: *He went into an insane rage; You're driving me nuts; I'm mad at you! He's fit to be tied*.
ANGRY PEOPLE ARE DANGEROUS ANIMALS: *He has a ferocious temper; Don't bite my head off! Don't snap at me! He unleashed his anger*.

Of course, this type of presentation is meant to be accompanied by classroom activities. The students may be asked to think of reasons for the existence of the given conceptual metaphors (i.e., their grounding in physical experience), to compare with expressions in their mother tongue, to detect the metaphors in a text, to invent a context in which a given expression would fit, and so on. More engagement on the part of the learners may be stimulated by asking them to categorise addi-

tional phrases under this or the other conceptual metaphor or even to propose a conceptual metaphor themselves on the basis of a sample of well-chosen instances (e.g., Beréndi et al., 2008, pp. 82–87; Boers, 2000a, pp. 564–566). For example, students may be asked to decide which of the above metaphor themes for anger each of the following expressions belongs to: *He was growling with rage; His anger welled up inside him; She got all steamed up; He began to bare his teeth; Those were inflammatory remarks; He was breathing fire; He's tearing his hair out*.

As an alternative to using conceptual metaphors as a principle for organising figurative idioms, some authors (e.g., Boers & Lindstromberg, 2008a) have found it easier to group idioms according to the source domains they can be traced back to. For example, the following expressions can all be linked to the source domain of seafaring: *take something on board; steer clear of someone; be on an even keel; the tip of the iceberg; in the doldrums; be left high and dry; a leading light*. Again, students can be asked to categorise idioms or to infer the likely source domain of a given set of idioms themselves. For instance, they may be asked to identify the common source domain (in this case, the theatre) of *behind the scenes*, *be waiting in the wings*, *take centre stage*, *in the limelight* and *the curtain comes down*.

Several pedagogy-oriented CS publications make use of pictorials to illustrate a literal reading of the figurative expressions. These pictorials may be line drawings (especially when phrasal verbs and prepositional phrases are concerned), which may be highly schematic (e.g., Tyler & Evans, 2004) or slightly more concrete (e.g., Lindstromberg, 2010). Some have also used photographic visuals (e.g., Boers et al., 2008, pp. 207–210). One of the intentions in adding these pictorials that depict source-domain scenes is to make the figurative phrases that are to be learned more memorable. The visuals add concreteness to the figurative expressions and thus make them more amenable to *dual coding* (Paivio & Desrochers, 1979; Paivio, 1986). Apart from pictures, mime and enactment techniques can often be used for the same purpose (Boers & Lindstromberg, 2008, pp. 384–385; Lindstromberg & Boers, 2005). As we will see further below, however, we may have to fine-tune our expectations about the precise mnemonic advantage of pictorials.

Insights from CS can also be put to use in the language classroom without necessarily presenting learners with organised *sets* of figurative lexis, as CMT-informed presentations are wont to. As students encounter a figurative expression during a classroom activity, the teacher may briefly point out the semantic motivation for the figurative meaning of this particular item without further interrupting the communicative activity any further at that point.[2] For example, on encountering *forge an alliance*, students may find it helpful to learn the literal sense of the word *forge* (as in *forge a new sword*). When meeting *in the wake* of, students may benefit from a line drawing on the board of a large sailing vessel with high waves in its wake. On encountering *a snap decision*, the teacher can easily inform/remind

the students of the literal sense of the action verb by snapping her fingers. To elucidate *cheer up* a teacher can briefly demonstrate what smiling does to the corners of the mouth. And so on.

Whether this more incidental approach for raising learners' metaphor awareness will in the long run also lead them to organise their L2 phrasal lexicon in the types of clusters (e.g., expressions from seafaring; up-is-good expressions) that more explicit CMT-style presentations envisage is uncertain. Neither is it certain, however, whether such an organisation is actually necessary to obtain the positive effects on learning that CS promises. Anyhow, teachers are bound to come across figuratively used words that they may find hard to link in with any conceptual-metaphor cluster. As there is evidence to suggest that knowledge of the literal sense of a word helps to make its figurative uses more transparent and memorable (Verspoor & Lowie, 2003), it is perhaps sufficient to inform learners of the literal sense or origin of a figuratively used word or expression when the opportunity for doing so presents itself. In other words, dual coding (i.e., exploiting the concreteness effect) may be the decisive and sufficient factor in making CS-style interventions effective. For instance, it has been shown that merely asking learners to hypothesise about the origin of idioms whose figurative meaning they have just been introduced to has a significant positive effect on retention, and it does not seem to matter much whether the learners' 'etymologies' are entirely correct (Boers, 2001). In short, using conceptual metaphors as headings under which figurative phrases can be grouped is not the only way of encouraging dual coding. Moreover, further below I will mention experimental studies the results of which suggest that, contrary to popular belief, a grouped presentation of vocabulary is often *not* facilitative for learning.

3. Assessing the evidence

To evaluate the merits of a proposed instructional method, its effectiveness and efficiency has to be compared to that of other available methods. Arguments for its implementation are supported when the proposed method is found to result in better or faster learning than another method that teachers (or material writers) are known to resort to. The standard way of sizing up instructional methods in applied linguistics is to conduct quasi-experimental studies in which the learning outcomes obtained under the proposed method — the experimental treatment — are compared to those obtained under a comparison treatment. Several reports of this type of research concerning CS proposals for the teaching of L2 figurative phrases have appeared in peer-reviewed publications since the second half of the 1990s. The publications I will review in this article are the following (in chronolog-

ical order):[3] Kövecses & Szabó (1996) (revisited in Kövecses, 2001), Boers & Demecheleer (1998), Boers (2000a) (revisited in Boers, 2004), Boers (2000b), Boers (2001), Verspoor & Lowie (2003), Boers et al. (2004), Csábi (2004) (revisited in Beréndi et al., 2008), Lindstromberg & Boers (2005), Guo (2007), Morimoto & Loewen (2007), Beréndi et al. (2008), Condon (2008), T.F. Li (2009), Cho (2010) and Gao & Meng (2010). Several of these publications actually report more than one experiment or report multifaceted experiments. Tables 1 to 3 separate out the experiments according to whether they measured comprehension of the target phrases (Table 1), retention of the meaning of the target phrases (Table 2), or retention of the form as well as the meaning of the target phrases (Table 3).

Most of the experiments (e.g., Boers, 2000a) were run in the course of just one or two lessons, but some (e.g., Condon, 2008) ran over an extended period. Most of the studies evaluate the effect of CS-style presentations on retention in memory, either by measuring recollection of words or phrases in gap-fill tests (i.e., testing 'productive' knowledge) or by measuring how well the learners remember the meaning of the words or phrases (i.e., testing 'receptive' knowledge). With regard to in-depth comprehension, a couple of experiments were set up to test the possibility that knowledge of literal meanings of words or phrases can help

Table 1. Experiments on the influence of CS-style presentation on *comprehension* of L2 figurative phrases

Publication	Targets for learning	CS-type input	Type of post-test	N	CS scores better than comparison?
Boers & Demecheleer (1998)	Figurative use of *beyond*	Prototypical literal usage	L1 translation	73	Yes; *p* 0.007
Boers (2000b)	Figurative use of *bail out*, *weed out*, *wean off*	Literal meanings	Appreciate connotations	85	Yes, *p* 0.001
Verspoor & Lowie (2003)	Figuratively used words	Prototypical literal usage	L1 explanation	78	Yes, *p* 0.01
Lindstromberg & Boers (2005) (3 trials)	Figuratively used manner verbs (*stumble*, *hurl*, etc.)	Literal usage through enactment	Appreciate connotations	29–62	Yes, *p* 0.002
Guo (2007)	Idioms	Source domain awareness	Matching and paraphrasing	70	Yes, *p* 0.001
Beréndi et al. (2008)	Expressions for anger	Conceptual metaphors	L1 translation	43	Yes, *p* 0.03

learners infer the 'connotations' of those words or phrases when they are used metaphorically (Boers, 2000b; Lindstromberg & Boers, 2005). For example, if students know of the literal meaning of *wean off* (i.e., to gradually stop breastfeeding your infant), they may be able to infer that the figurative use of this phrase (as in *wean companies off state support*) is likely to suggest the inevitable nature of the process referred to.

Although the majority of the studies are small-scale — both in terms of number of participants and the number of targets for learning — their outcomes in favour of CS-style treatments indicate that, collectively, these studies begin to constitute a body of evidence that is hard to ignore as an incentive for implementing CS insights in mainstream language pedagogy. We do have to acknowledge, however, that a closer look reveals one or the other methodological weakness in many of the studies, and so some caution is still in order when it comes to interpreting their results. In what follows I will adopt a devil's advocate stance to scrutinise the publications for flaws that may render the reported findings less than conclusive.

The first issue is that many of the studies lack proper pre-tests. When two treatments are administered to two groups of learners with a view to assessing which treatment yields the best results in a post-test, then it should first be established that

Table 2. Experiments on the influence of CS-style presentation on *retention of the meaning of* L2 figurative phrases

Publication	Targets for learning	CS-type input	Type of post-test	N	CS scores better than comparison?
Boers (2001)	Idioms	Hypothesising about origins	Explain meaning	54	Yes, *p* 0.001
Verspoor & Lowie (2003)	Figuratively used words	Prototypical literal usage	L1 explanation	78	Yes, *p* 0.01 on delayed post-test
Li (2009), study A	Figuratively used words	Conceptual metaphors	Explain meaning	76	Yes, *p* 0.000 on immediate and delayed post-test
Li (2009), study B	Idioms	Conceptual metaphors / images	Explain meaning	127	Yes, *p* 0.000 on delayed post-test
Li (2009), study C	Proverbs	Conceptual metaphors / images	Explain meaning	90	Yes, *p* 0.000 on delayed post-test
Gao & Meng (2010)	Expressions for anger	Conceptual metaphors	L1 translation	66	Overall test scores in favour of CS, but no separate statistics given for the translation test.

Table 3. Experiments on the influence of CS-style presentation on *the form* as well as the meaning of L2 figurative phrases

Publication	Targets for learning	CS-type input	Type of post-test	N	CS scores better than comparison?
Kövecses & Szabó (1996)	Phrasal verbs with *up* and *down*	Conceptual metaphors	Fill in blanks with *up* or *down*	30	Yes, but no significance calculated
Boers (2000a), experiment 1	Expressions for anger	Conceptual metaphors	Fill in blanks with keywords	118	Yes, p 0.5
Boers (2000a), experiment 2	Up/Down lexis (*soar*, *skyrocket*, *plunge*, etc.)	Source domains	Free production in essay writing	73	Yes, p 0.001, but not anymore on a delayed test (Boers 2004)
Boers (2000a), experiment 3	Phrasal verbs with *up*, *down*, *out* and *in*	Conceptual metaphors	Fill in blanks with verb + particle	74	Yes, p 0.01
Boers (2000b)	Figurative use of *bail out*, *weed out*, *wean off*	Literal meanings	Fill in blanks with keyword	75	Yes, p 0.03
Boers (2001)	Various idioms	Hypothesising about origins	Reproduce idioms to match paraphrases	54	Yes, p 0.001
Csábi (2004), 2 trials	*hold* and *keep* (in phrasal verbs and idioms)	Literal meanings	Fill in blanks	26	Yes, but not significant for idioms
Boers et al. (2004)	Idioms from various source domains	Origins of the idioms	Fill in blanks with keywords	17–24	Yes, p 0.001
Morimoto & Loewen (2007)	Various uses of *break* and *over*	Starting from image schema	Acceptability judgement and picture descriptions	58	No, $p > 0.1$
Beréndi et al. (2008)	Expressions for anger	Conceptual metaphors	Fill in the blanks and free recall	43	Yes, p 0.007, but only borderline significance on delayed test

Table 3. (*continued*)

Publication	Targets for learning	CS-type input	Type of post-test	N	CS scores better than comparison?
Condon (2008), 4 trials	Phrasal verbs with *in*, *out*, *up* and *down*	Conceptual metaphors	Fill in blanks with verb + particle	24–71	Yes, in 2 of the trials (p 0.005), but in 1 trial the comparison group did better (p 0.005)
Li (2009), study B	Idioms	Conceptual metaphors / images	Fill in blanks with keywords	127	No.
Li (2009), study C	Proverbs	Conceptual metaphors / images	Fill in blanks with keywords	90	Yes, p 0.00 on delayed post-test
Cho (2010)	Phrases with *in*, *at* and *on*	Meaning extensions from prototypes	Fill in blanks with preposition	46	Yes, p 0.00 on delayed post-test
Gao & Meng (2010)	Expressions for anger	Conceptual metaphors	Fill in blanks and free recall	66	Overall test scores in favour of CS (except with low-proficiency learners), but no separate statistics for the productive part of the test.

the two groups of learners are on a par regarding their knowledge of the target items prior to the treatment. Some of the reports mention that groups were judged by teachers to be of similar levels of proficiency, that the groups had similar histories of language learning, or that the groups had obtained comparable exam results at the end of the previous term. However, when knowledge of a small number of discrete items is at stake in an experiment, then a more precise measurement of prior knowledge is commendable. Especially when the number of participants is small, 'accidental' differences in knowledge profiles between the two groups risk skewing the post-test results. Only eight of the publications reviewed here (Cho, 2010; Condon, 2008; Gao & Meng, 2010; Guo, 2007; Li, 2009; Lindstromberg & Boers, 2005; Morimoto & Loewen, 2007; Verspoor & Lowie, 2003) report pre-test results.[4]

The second issue regards the quantitative differences in treatment between experimental and comparison groups in some of the studies. To be able to compare the *efficiency* of two qualitatively different instructional treatments, these should involve a similar investment of time and effort on the part of the learners, and a similar amount of input to stimulate that investment. If one group of learners

spends more time studying a given set of target items because they are given more information about them, then this is likely to give them an edge over other learners who have been deprived of such incentives. Better learning outcomes ensuing from the former condition cannot in that case be taken as evidence of superior efficiency. After all, an equal amount of processing of the input material in ways compatible with the comparison treatment might have generated as good a learning outcome as the experimental treatment. One example is the experiment on phrasal verbs learning conducted by Kövecses and Szabó (1996). In that experiment, the comparison group was presented with ten English phrasal verbs and their equivalents in the students' L1, and they were asked to try and memorise them. The experimental group received the same input, but in addition, nine conceptual metaphors were explained and exemplified by several phrasal verbs. This is clearly more input than was given to the control group and thus the two treatments differ not only qualitatively but also quantitatively. In five of the other studies reviewed here (Beréndi et al., 2008; Boers, 2000a on phrasal verbs; Condon, 2008; Csábi, 2004; Li, 2009 on proverbs), the CS-style treatment also involved somewhat more input than the comparison treatment. Guo (2007) does not give any information about the comparison treatment, referred to as "traditional instruction", and so it is impossible to say whether it involved quantitatively equivalent input and engagement.

The third issue concerns the degree of resemblance between the test-format and the nature of the tuition or the task the learners engaged in, (unknowingly) as preparation for the test. In Cho (2010) the post-test items were accompanied by pictorial prompts, which may have been to the advantage of the experimental students, because visuals were part of the explanations they had received, whereas the comparison group had received only verbal explanations. In Csábi (2004), the post-test required students to produce L2 phrases, but the preparatory task for the comparison group had consisted in producing L1 translations, i.e., a task that may have drawn the students' attention away from the form of the L2 phrases. On the other hand, it must be mentioned that CS-driven learning is not the ideal preparation for *productive* tests. CS draws students' attention to the motivation behind the *meaning* of the figurative phrases rather than the precise words that make up the phrases. It is well-known that tasks which draw students' attention to the meaning of words will primarily foster retention of meaning, whereas tasks that draw students' attention to the form of words will primarily foster retention of form (Barcroft, 2002). It is therefore not surprising that the comparative success obtained through CS-style instruction tends to become less pronounced as the post-test requires more productive knowledge, as is the case with gap-fill tests that are to be completed with full phrases rather than single words (Csábi, 2004; Beréndi et al., 2008). After CS-style instruction, students tend to remember the

meaning of idioms, but often fail to remember their exact lexical composition, and so they are prone to substituting words (e.g., *add oil to the fire* instead of *add fuel to the fire* — Beréndi et al., 2008, p. 78; Boers, 2000a, p. 557). The mnemonic benefit of CS-inspired instruction is thus likely to be greater in terms of learners' receptive knowledge than their productive knowledge of the taught figurative phrases.[5] Still, the evidence from the studies under review shows that the CS treatment — despite its semantic orientation — often fosters better retention than the comparison treatment also at the level of productive knowledge, albeit mostly when the phrases to be learned are short and consist of familiar words — as tends to be the case with most phrasal verbs, for example.

A fourth issue has to do with the 'ecological validity' of the choice of comparison treatments. Few of the authors of the studies reviewed here justify their choice of comparison treatment. It is often taken to be representative of what is offhandedly — and perhaps even dismissively — referred to as "the traditional method", but little evidence is given that the chosen treatment for the comparison group actually mimics a type of instruction that is established practice in language education. Of course, what is established practice differs from one educational setting to the next. Still, one cannot help wondering if the differential results reported in some of the studies are not due to the weakness of the comparison treatment rather than the proclaimed strength of the experimental treatment. In this context it is also worth mentioning that statistical significance is sometimes reached not because the experimental treatment is so effective, but rather because the comparison treatment is so *in*effective. In Cho's (2010) study on prepositions, for example, the comparison group's score actually *fell* between the pre-test and the delayed post-test. This was probably the negative side-effect of massed input (the students dealt with over 40 instances of three confusable prepositions in one go) and the use of 'correct-the-wrong-preposition' exercises during instruction. The general question this raises is whether the comparison treatments in the CL intervention studies are always 'ecologically valid' — that is, reflecting the way the given targets are known to be taught in real pedagogical practice.

The fifth issue is the absence of delayed post tests in the majority of the studies. The results that are reported typically concern tests that were administered shortly after the instructional treatments. Such immediate post tests are not always predictive of long-term learning effects, and should ideally be complemented by delayed post tests, if possible administered weeks after the treatment stage (Schmitt, 2008). The couple of studies that did include a delayed post-test show that the better outcome under the CS treatment is maintained (e.g., Li, 2009; Verspoor & Lowie, 2003), but we must bear in mind that "better" is used in relation to the sometimes bitterly poor results in the comparison groups. In Condon's (2008) study on phrasal verbs, the longer-term learning gains were in fact negligible even

in the experimental condition. In her experiment, the mean scores on a 15-item delayed post test were just around 1.5 points higher than the scores the students obtained in a pre-test targeting precisely the same items.

The sixth issue concerns how big a part chance may have played in the tests used in some of the studies. In a gap-fill test where students are to choose between just two or three options (e.g., between *up* and *down* in Kövecses & Szabó, 1996, between *in*, *on* and *at* in Cho, 2010, and between *hold* and *keep* in Csábi, 2004), the chance factor is substantial if the test consists of relatively few items. This also holds for "agree-disagree" or "true-false" type questions (used, e.g., in Boers, 2000b, and Lindstromberg & Boers, 2005). But even in some of the experiments which required students to make a choice among a wider range of options, this choice was in actual fact more restricted than first impressions lead one to assume. An example is Boers (2000a), where students were asked to match phrasal verbs with blanks in a text. The phrasal verbs were listed in the appropriate morphological form (e.g., with past tense ending) to fit the blank where they belonged. After filling in test items one feels confident about and test items one can tackle through a process of elimination thanks to morphological clues, the greater the chances of making lucky guesses in the matching task subsequently, thanks to the reduced set of remaining possibilities. A small difference in knowledge between two groups can thus influence test results beyond what students have truly learned during the treatment. In other words, the nature of the test format can sometimes enlarge differences between groups' abilities, and so levels of statistical significance can be reached which might not have been reached had a different test format been used.

The seventh issue regards the way the participants' responses were processed in some of the studies. For example, none of the studies reported in Li (2009), where participants were asked to "explain the metaphorical meaning" of the target items, detail how the tests were marked. The same holds for tests meant to measure productive knowledge. In some cases, the researchers explain that they tallied the correct responses in gap-fill tests twice: once counting only completely correct responses and a second time including partially correct responses (Csábi, 2004; Beréndi et al., 2008). The latter meant accepting spelling mistakes and incomplete phrases, and this more lenient marking tended to yield more convincing *p*-values than the former. No details are given about how such minor mistakes were dealt with in the data analyses of the other studies that tested productive knowledge. We have already conceded that the proposed CS-style presentations of vocabulary are not ideally suited to engender retention of the precise form of L2 phrases. It is therefore not surprising that a scoring procedure that is lenient about spelling errors will yield a higher score than a stricter procedure.[6]

The eighth issue is the way alternative explanations for favourable findings are sometimes given insufficient consideration. For example, Boers et al.'s (2004)

study, in which it is reported that informing learners of the literal origins of figurative idioms helps these learners to work out the idiomatic meaning usage for these idioms, the participants tackled the tests on-line, without supervision. Given the lack of control over the way the students went about completing the tests (e.g., whether they might have consulted additional sources of information), a cautious interpretation of the results is in order. Another example is the study by Lindstromberg & Boers (2005), where it was hypothesised that a profound understanding of the concrete meaning of manner verbs (e.g., that *hurl* is more forceful than *throw*) would help learners appreciate nuances of the metaphorical uses of these verbs (e.g., that *hurl accusations at someone* expresses a higher degree of intensity than *accuse someone*). To help experimental students acquire that profound knowledge of the literal meaning of over 20 manner verbs, mime and enactment techniques were used. In the comparison groups, the verbs' meanings were explained only verbally (the ecological validity of which is, again, not beyond doubt). In post-tests, the students' appreciation of the 'connotations' of the metaphorically used manner verbs was gauged, and the experimental students were found to perform best. This seems to corroborate the hypothesis. However, it is also possible that students in the comparison group failed to capture the nuance of the metaphorical usage of some verbs not because the knowledge they had acquired of the literal meaning was too shallow, but because by the time they were given the post-test they were not even sure anymore of the 'basic' meaning of some of the verbs (a possibility rightly pointed out by one of that paper's anonymous reviewers).

The ninth and last issue concerns the insufficient detailing of the quantitative data in the early studies. Just a *p*-value is mentioned, but no information on the type of statistic that was used to obtain that *p*-value is given in Boers & Demecheleer (1998), Boers (2000a), and Boers (2001). The reader then has to take it for granted that the choice of statistic was appropriate for the data at hand. It is regrettable also that some of the reports mention *p*-values, but not the mean test scores of the two groups (and the standard deviations) (Boers, 2000b; Boers & Demecheleer, 1998). Guo (2007) does give mean scores, but it is not clear how many test items there were.

Despite the weaknesses noted in experimental designs and the sometimes unorthodox ways of reporting data, there is still no denying the obvious trend in the collection of studies reviewed here: in virtually all of the comparisons, the CS-informed presentation of L2 figurative lexis comes out on top. Each of the above criticisms — which are all by themselves relatively minor — applies to only some of the studies, but never to the collection as a whole. Surely, then, the collective evidence outweighs the imperfections of individual studies.

At the same time, the review does suggest that larger-scale and more rigorously designed experiments would be welcome to solidify the case for CS-style

pedagogy. In addition, new studies would be welcome to compare the effectiveness of different versions of CS-inspired instruction, and to pinpoint exactly what benefits can realistically be expected from them. For, example, Li's (2009) studies on idioms and proverbs suggests that stimulating students to attach a specific mental image to individual target phrases leads to better retention than grouping sets of target phrases under conceptual metaphors. Other researchers have started investigating the potential of including extra stimuli for cognitive engagement in CS-style approaches to figurative lexis. It is generally agreed among researchers of vocabulary acquisition that the rate at which words are acquired is positively influenced by the degree of cognitive engagement on the part of the learners (Schmitt, 2008). The motivated nature of figuratively used lexis opens avenues for cognitive engagement which have so far been left pretty much unexplored in L2 pedagogy. More particularly, giving students clues about the motivation can help them make 'educated guesses' at the meaning of figurative phrases, such as idioms, before seeking confirmation or rectification of those guesses. The cognitive effort this guided guessing task incites is predicted to positively influence retention, a prediction which has been borne out by at least three CS experiments. Skoufaki (2008) presented idioms under the heading of the conceptual metaphors the idioms instantiate and asked one group of participants to use these clues to try and infer the meaning of the expressions, after which the meanings were clarified. A comparison group was given the same input but with the meanings of the idioms given from the start. The inclusion of the meaning-guessing task led to significantly better retention as measured in a post-test. In a similar vein, Boers et al. (2007) report that information about the origin of idioms was a helpful clue for students to identify the right paraphrase of their figurative meaning in a multiple-choice task. It also generated better retention than giving that information after the figurative meaning had already been explained. Boers et al. (2008) confirm this finding in connection with the use of visuals as clues for guided meaning-guessing (although the mnemonic effect seemed most pronounced when it came to post-tests measuring receptive rather than productive knowledge — see further below).

It is worth mentioning, however, that each of these studies also shows that quite often the students' guesses at the idiomatic meanings require corrective feedback. Information about conceptual metaphors or the origins of the expressions may contain useful clues, but they do certainly not guarantee correct interpretations of the figurative meanings. In Skoufaki's (2008) study, for example, 75% of the experimental students' guesses would have required some form of corrective feedback. Hu & Fong (2010), who tested the usefulness of conceptual-metaphor clues for the interpretation of idioms with Chinese learners of English, report think-aloud data that show how easily such clues can be overridden by mistaken cognates and by cross-cultural differences.

4. Hopes and expectations

Signs of at least three controversies are strewn across the dozens of publications that argue in favour of CS-informed pedagogy. These concern the question of learner independence (i.e., learners' autonomous application of CS insights to newly encountered lexis), the use of visuals as a mnemonic aid, and the question what student profiles CS caters to.

The first empirical investigation of the effectiveness of CMT instruction, by Kövecses & Szabó (1996; revisited in Kövecses, 2001), generated data suggesting that learners could successfully transfer their awareness of conceptual metaphors to the processing of phrasal verbs they had not yet been taught. Experimental students were taught English phrasal verbs organised in sets that illustrated various conceptual metaphors to motivate the use of the particle (either *up* or *down*) in each of the phrasal verbs. In a gap-fill post-test targeting ten of the taught phrasal verbs (in which students were required to fill in only the particle), the experimental students did not significantly outperform the comparison group who had not received information about conceptual metaphors.[7] However, they did when they were presented with ten gap-fill items targeting phrasal verbs with *up* or *down* which they had *not* been taught during the treatment. The comparison group did not score above chance on the latter ten items, whereas the experimental group did by a large margin. Not all of the 'new' phrasal verbs instantiated the conceptual metaphors explained during the instruction stage, according to the authors' analysis. To the authors, this could only mean that the experimental students had somehow managed to successfully transfer their enhanced metaphor awareness to their processing of as yet unfamiliar phrasal verbs — at least at the level that they could make educated rather than blind guesses as to the right choice between the two particles. Similarly encouraging were data presented by Boers (2000a, pp. 564–566) which suggested that students were quite capable of linking figurative expressions to one of three source domains they were given the choice between.

More recent studies, however, give much less reason to be hopeful about the feasibility of learner-autonomous applications of metaphor awareness. Boers' (2000a, pp. 560–562) experiment on phrasal verbs did not confirm the aforementioned findings of Kövecses & Szabó's (1996) study. The students who had been introduced to conceptual metaphor groupings outperformed the comparison group in a gap-fill post-test only with regard to the explicitly taught phrasal verbs, not with regard to novel ones. It may be worth mentioning though that the gap-fill test was slightly more challenging in Boers' experiment, since students were required to supply the verbs as well as the particles; and more particles than just *up* and *down* were included. When Beréndi et al. (2008, pp. 83–84) gave a group of students the task of identifying the conceptual metaphors behind a set

of figurative expressions, this task turned out to be well beyond these students' abilities — even though unfamiliar words in the expressions were explained to the students. Identifying conceptual metaphors (or source domains) behind figurative phrases is obviously harder than linking the phrases to a narrow set of pre-identified metaphors. Also when conceptual metaphors or 'etymologies' are used as clues for meaning-guessing — a classroom activity which, as we saw above, fosters retention —, students' hypotheses will often require corrective feedback (Boers et al., 2007; Hu & Fong, 2010; Skoufaki, 2008).

In this light, it is worth mentioning that the classroom conditions created for participants to make inferences about the meaning of given phrases are quite different from opportunities for learning *outside* the classroom, where students encounter unfamiliar figurative phrases more incidentally (rather than in sets of phrases which all happen to illustrate the same metaphor), and where lack of comprehension of individual words is likely to short-circuit potential attempts at discerning an underlying conceptual metaphor (or source domain) and thus also potential attempts at guessing the meaning of the phrase through awareness of its literal origin. Interestingly, research by Cieslicka (2006) suggests that learners tend to spontaneously resort to imagery when they learn figurative idioms: they will tend to activate a literal reading of the expressions alongside the figurative reading. However, their ability to do so will evidently be dependent on their knowledge of the literal meaning of the words contained in the idiom. If they are not familiar with the word *keel* in *be on an even keel*, then it is hard to imagine that they will spontaneously resort to seafaring imagery to make sense of the expression, even if they have been given CS training that has enhanced their metaphor awareness. It is only on finding out about the meaning of the word that they might possibly relate the literal meaning of the word to its figurative usage in the idiom. However, a great many words that are used in idioms are homonymous or polysemous even at the level of their literal usage, and the appropriate interpretation of the idioms by metaphor-conscious learners will hinge on the right identification of the basic word meaning. For example, *take something on board* and *across the board* exploit different literal senses of the word *board* (the board of a ship and a board to write on, respectively). A metaphor-conscious learner coming across the idiom *show someone the ropes* may mistakenly try to interpret the expression by relating the word *ropes* to a boxing ring or even to the gallows rather than to a sailing vessel. It therefore stands to reason that enhancing language learners' metaphor awareness (be it through explanations about conceptual metaphors, groupings along source domains, or information about the literal meanings or origins of incidentally encountered phrases) is not likely to help learners 'proactively' interpret — let alone produce — newly met figurative expressions when they are left to their own devices. The need for guidance and feedback seems inevitable. Guidance may be

especially welcome in cases where associations of seemingly equivalent L1 and L2 words are profoundly divergent due to cultural differences. Hu & Fong (2010), for instance, report how hard it is for Chinese learners to interpret the meaning of English idioms that revolve around words like *heart* and *mind*, due to contrasting cultural models.

Whether phrasal verbs might be more amenable than idioms to learner-independent metaphor processing is doubtful. As is demonstrated by Kövecses & Szabó (1996), particles such as *up* and *down* can be used in ways that instantiate about a dozen different conceptual metaphors. The question this raises is how a learner could possibly pick the right one to interpret a newly encountered phrasal verb, unless the context gave sufficient clues. The learner should choose not to associate *give up* with the metaphor HAPPY IS UP, not to associate *set down* with LESS IS DOWN, and so on. If knowledge of conceptual metaphors *per se* is not an entirely reliable source for the interpretation of phrasal verbs (i.e., receptively), it is perhaps even less reliable when it comes to gauging what particle should be used (i.e., productively). Given COMPLETION IS UP as well as MORE IS UP (and half a dozen other up-down metaphors), it must be hard for learners (metaphor-conscious or otherwise) to predict the particle *up* in *use up* (as in *we've used up all our ammunition*). After all, the quantity of what is being used up goes *down*. In a similar vein, knowledge of the core meaning of a polysemous word does not enable one to predict its extended uses with any certainty. Morimoto & Loewen (2007), for example, report only minimal improvements in their participants' acceptability judgements of uses of the words *break* and *over* after teaching them the core meanings. The general point is, of course, that linguistic motivations provide *retrospective* explanations, but they do not have great predictive power (and Cognitive Semanticists have never claimed otherwise). It follows that the pedagogical exploitation of a linguistic motivation will be useful first and foremost as a way of *deepening* learners' understanding of figuratively used words and phrases rather than equipping learners with a tool for working out meanings and usages independently.

If learner-independent proactive applications of metaphor awareness are not likely to be very fruitful, then the question arises how we can account for the aforementioned intriguing finding by Kövecses & Szabó (1996) that students who had been given tuition about conceptual metaphors scored above chance on a gap-fill test targeting phrasal verbs that were *not* covered during the tuition stage. I speculate that several factors besides the influence of metaphor awareness may have conspired to cause this effect. Firstly, given the absence of a pre-test in this study, we cannot rule out the possibility that the experimental group enjoyed a bit of a head start over the comparison group. Secondly, the chance factor (due to the binary decision task and the small number (10) of test items) may have played a part. Thirdly, it is conceivable that the better performance of the experimental group

was fuelled by the greater number of exemplars of *up/down* phrasal verbs these students became acquainted with as part of that tuition — input that was denied to the comparison group. Although the phrasal verbs targeted in the gap-fill post test were not actually taught, they did have 'analogues' among the ones presented to the experimental group (but not presented to the comparison group). For example, in tackling the test items *cast down*, *break down* and *turn up* the experimental students may have benefited from having become acquainted with *knock down*, *run down* and *pop up*, respectively. This does not mean that metaphorical thinking was absent from the experimental students' problem-solving strategy (just as we cannot rule out the possibility that some metaphorical thinking was also present in the way the comparison group went about the task — and this consideration holds for all the intervention studies reviewed here), but it may have been much less 'top-down' than hypothesised by the authors.

Let's now turn to the second bone of contention mentioned at the start of this section — the use of visuals. Pictorial elucidation of the concrete source domain behind figurative phrases is believed to provide an extra stimulus for dual coding and thus for retention. Given the importance that is attached to imagery and dual coding by CS-inspired pedagogues, it is not surprising that several of the learning and teaching resources they propose feature visuals (e.g., Kövecses et al., 1996–1998; Lindstromberg & Boers, 2008). In the discussion that follows, it will be assumed that we are dealing with pictorials that are well-chosen and that are congruent with the meaning of the phrases they are intended to elucidate. In actual fact, this is not so self-evident. MacArthur & Boers (Forthcoming) give examples from popular text books and self-study vocabulary books where idioms are accompanied by line-drawings that do not at all elucidate the meaning of those idioms. Instead the drawings are multimodal puns, the humour of which one can appreciate only if one is already familiar with the idiom. Among the examples are a drawing of a bird perched on a door handle to illustrate *fly off the handle* (McCarthy & O'Dell, 2002, p. 91) and a drawing of a woman holding a mirror to inspect the inside of her own nose to illustrate *look down your nose [at someone]* (Gude & Duckworth, 1999, p. 44). It is doubtful whether such pictorials help learners comprehend the idiomatic meanings. At least, we may presume that CS-informed pedagogues choose pictorials that are faithful to the source domains of the figurative expressions.

In assessing the merits of such pictorial elucidation as an aid for learners to remember L2 figurative phrases, it is relevant to distinguish between the contribution pictorials make to the retention of the meaning of the phrases and their contribution to the retention of the precise lexical makeup of the phrases. In other words, it is relevant to distinguish between the uses of pictorials to foster receptive and productive knowledge. One reason for this is that pictures do not directly

elucidate words; they elucidate concepts. A picture may call to mind a concept, and may thus indirectly prime a word or expression associated with that concept, *if* that word or expression is readily retrievable from memory. There is of course no guarantee that a learner will be able to recollect the precise L2 word or phrase for any concept that comes to mind, whether or not the picture associated with the concept is also stored in memory.

Boers et al. (2008) report three experiments the results of which confirm that pictorial elucidation is beneficial at least for retention of meaning. In the first experiment, students were presented with sentences containing figuratively used words (e.g., *embrace* in *I wonder if our college is going to embrace the minister's guidelines*) and pictures illustrating the literal usage of these words (e.g., a picture of two people in an embrace). A post-test measuring the students' retention of the meaning of the figurative uses of the words showed they outperformed a comparison group who learned about the literal meanings of the words through verbal explanations instead of pictures. In the second experiment, students were shown a drawing illustrating the prototypical spatial sense of *beyond*. This elucidated that *beyond* expresses a certain distance (and through inference a lack of accessibility), as in *the ball went beyond the neighbour's hedge*. In a post-test, they were asked to interpret figurative uses (e.g. *his stories are getting beyond belief*). They outperformed a comparison group that was given a complete dictionary entry on *beyond* to study, but no pictorial elucidation. In the third experiment, which focused on figurative idioms, students were first given explanations about the origins of a set of idioms and only afterwards were they asked to identify the figurative meaning of those idioms in a multiple-choice test. For one cohort of students, the 'etymologies' were accompanied by pictures (e.g., of a circus artist performing a balancing act to illustrate *a balancing act*). Another cohort had received only the verbal explanations about the 'etymologies'. The former did significantly better than the latter in the subsequent multiple-choice test on the idiomatic meaning of the expressions.

Whether pictorial elucidation is equally beneficial for retention of form (i.e., for learners' recollection of the precise makeup of the phrases) is doubtful, however. The aforementioned experiment on the use of pictorials to elucidate the origins of idioms (Boers et al., 2008, pp. 201–205) also revealed that the comparison group, who had received only verbal explanations about the idioms' origins, actually outperformed the experimental group in a gap-fill post-test, i.e., a test of (partial) productive knowledge. Intrigued by this finding, Boers et al. (2009) set up a new experiment in which students were again given information about the origins of English idioms prior to a gap-fill test. For half of the idioms, the explanations were accompanied by pictorials, while the other half were presented without. Results on the gap-fill test showed that the presence of pictorial elucidation made no

difference when it came to students' recollection of words they were already familiar with. When it came to recollecting words that were new to them, however, the presence of pictures in the input affected performance *negatively*. Interestingly, the students seemed to remember the concepts elucidated by the pictures, but could not remember the precise words used in the idiomatic expressions. For instance, they would fill in *violin* instead of *fiddle* in *play second* _____, *rope* instead of *rein* in *keep a tight* _____ *on someone*, and *coin* instead of *toss* in *argue the* _____. In fact, this finding is not surprising given the phenomenon known as 'picture superiority' (Nelson et al., 1976): when words are presented alongside pictures, it is the pictures that stand the best chance of being remembered. The pictures are also likely to draw learners' attention away from the words, and so picture superiority may come at the expense of remembering the precise verbal input. To (re-) produce words or phrases accurately, one obviously needs to be able to recall their form or composition. In case the expression to be learned is made up of words that are already familiar, the distraction by pictures may not matter, as the target words are already well entrenched in memory. It is then 'just' a matter of linking the words whose form one is already familiar with the concept elucidated by the picture. The use of pictorial elucidation is thus likely to have a positive influence on learners' recollection of morphologically simple, high-frequency words, such as prepositions and phrasal-verb particles. When as yet unfamiliar words are concerned, however, the picture may well help entrench the concept in memory, but due to its distracting effect, it may reduce the chances of remembering the exact form of the words that are to be associated to that concept. To foster retention of the precise lexical composition of idiomatic expressions, a certain form-focus ('structural elaboration') may usefully complement the meaning-focus ('semantic elaboration') that is stimulated by the use of pictorials. With a view to stimulating retention of form, some authors have started investigating the scope of *phonological* motivation, evident in the conventionalisation of alliterative and assonant phrases (e.g., *time will tell; from pillar to post; rule the roost; through thick and thin; steer clear*) (Boers & Lindstromberg, 2009, Ch. 6), and how exploiting this type of motivation can help language learners remember the wording of phrases in addition to remembering their meaning (Lindstromberg & Boers, 2008bc).

The last point to ponder in connection with the experimental data reviewed here is the profile of the learner who is most likely to reap the rewards of CS-style instruction. One dimension along which learners differ is their cognitive style or their learning style. It is evident that CS-style instruction encourages mental imagery as a way of comprehending and remembering (the meaning of) figuratively used words and phrases. Not all learners share an equally strong inclination towards imagery processing, however. According to cognitive-style analysis (e.g., Riding, 1991), some people have a preference for thinking in words rather than

images. In several experiments on the learning of figurative phrases through CS-style instruction, Boers et al. (2008) consistently found positive correlations between the students' learning gains and their relative inclination towards imagery processing as gauged by means of cognitive-style questionnaires. This does not mean that learners who would not by themselves resort to imagery processing so easily cannot benefit from CS-style instruction — perhaps they just need more input to that effect — but it does suggest that complementary pedagogical approaches are welcome to accommodate diverging learning styles.

Another learner-profile variable that may enhance or dampen the effectiveness of CS-style instruction is the type of motivation for language learning. Language majors may be more appreciative of activities that make them reflect on the properties of language, such as the pervasiveness of metaphor and metonymy, than students for whom the learning objective is purely instrumental, for example being able to function in English in an multinational business environment. It is perhaps not surprising that experimental results are more favourable to CS in the studies by Boers (2001) and Boers et al. (2004), where the participants were language majors, than in those by Csábi (2004) and Condon (2008), where the participants were secondary school pupils and students of economics, respectively. On the other hand, Juchem & Krennmayr (2010) describe how they integrated insights from conceptual metaphor in a course of business English for students of commerce and economics, and they report that the students evaluated this approach quite positively. More longitudinal research would be welcome in order to find out if the observed enthusiasm on the part of instrumentally-motivated students might not be due first and foremost to the novelty of the instruction, and may gradually diminish.

Also meriting more systematic research is the question whether CS-style instruction is equally suitable for learners at different levels of proficiency. Altogether, the comparative effectiveness of the interventions put to the test in the various experiments that we have reviewed appears least conspicuous when student populations at the lower end of the proficiency scale were concerned. The clearest example is the study by Gao & Meng (2010) on teaching idiomatic expressions for anger. In their study, the grouping of the expressions according to conceptual metaphors gave an edge to relatively advanced students, but not to low-proficiency ones. At least three reasons may help explain the lower effectiveness of the CS treatments in experiments conducted with the participation of low-proficiency learners. Firstly, lower-proficiency learners may not just need to get to grips with the expressions that are targeted for learning and testing, but they may also find it harder than higher-proficiency learners to cope with the explanations used during tuition (if these are put in the L2) and to make good use of the L2 verbal contexts intended to elicit the use of the target words or phrases in the tests. Secondly, lower-proficiency

learners are less likely to be familiar with the 'basic' senses of the L2 words that are used figuratively in the target expressions. While higher-proficiency learners benefit from the CS technique of linking a novel, figurative use of a word to an already familiar use of the same word, lower-proficiency learners may need to get acquainted with both to catch the point of the CS intervention. Thirdly, the target phrases that were selected for some of the experiments may simply be perceived by lower-proficiency learners to be lacking in utility at the stage of learning they are at. While none of these explanations necessarily mean that CS-style instruction is only suitable at upper-intermediate and advanced levels, they do suggest the need for a better match between the choice of targets for learning and the proficiency level of the learners. We shall return to this issue in the next section.

5. Comments from mainstream second language vocabulary research

At a general level, CS approaches can be expected to receive broad support from 'mainstream' researchers of second language vocabulary acquisition, for at least three reasons. Firstly, the CS proposals involve a fair degree of intentional vocabulary learning, and many a prominent second language vocabulary researcher endorses the view that explicit instruction is a useful supplement to opportunities for incidental learning (Laufer, 2005; Nation, 2001). Secondly, the CS proposals tend to stimulate cognitive engagement on the part of the learners, and it is generally agreed that engagement helps create durable memory traces (Laufer & Hulstijn, 2001; Schmitt, 2008). Thirdly, CS gives centre stage to multiword units, and this chimes well with applied linguists' growing interest in formulaic language generally (e.g., Schmitt, 2004; Wray, 2002). At the same time, certain characteristics of some of the CS proposals are bound to raise eyebrows in mainstream applied linguistics circles. I will discuss three such characteristics that have been pointed out to me in a critical vein at applied linguistics conferences where I have had the privilege to report CS work.

The first is the tendency to present target phrases in sets, grouped according to the conceptual metaphors they are believed to instantiate or according to the source domain they are believed to derive from. While presenting vocabulary in an organised fashion may at first sight seem pedagogically sound — and a great many popular text books and self-study materials display this method —, various studies have, unfortunately, revealed that it is not likely to facilitate learning (Erten & Tekin, 2008; Finkbeiner & Nicol, 2003; Tinkham, 1993, 1997; Waring, 1997). Especially if the words that are grouped together belong to the same 'semantic' set, in the sense that they are near-synonyms, antonyms, hyponyms, or, more generally, words belonging to the same word class and referring to the same semantic

field, this form of presentation is actually likely to *add* to the learning burden. In effect, it gives the learner not only the task of trying to remember the new words, but also that of avoiding confusion among them. Presenting all these too similar those words together may lead to cross-associations which could be avoided if they were not co-presented in the first place. It is understandable that teachers and textbook writers feel inclined to present words together that have something in common (e.g. nouns for professions, manner-of-motion verbs, and adjectives for personality traits), but it is precisely this commonality of the co-presented words that makes the learning task harder. If, in addition, some of the semantically clustered words show formal resemblances, then the risk of erroneous memory traces becomes even greater.

In this light, we may have to reassess the pedagogical soundness of asking learners to study *sets* of phrasal verbs, for example, no matter what principle of organisation is adopted. The CS approach will favour clustering phrasal verbs according to the semantics of the particle, but sceptics may ask whether this is fundamentally different from an organisation based on the verbs themselves. It is not clear, after all, why the risk of erroneous cross-associations should be lower when studying the set *show up*, *turn up*, *pop up*, *bring up*, *look up* and *call up* than when studying the set *turn up*, *turn in*, *turn out*, *turn down*, *turn over* and *turn into*. The question raised by the aforementioned research on interference effects is if learners should be asked to study such sets in the first place, or if a more distributed learning process might not be more effective and efficient.

Of course, there are ways of reducing the risk of cross-associations brought about by the grouped presentation of lexis. One way is to present learners with a set in which only few items are as yet unfamiliar, and to make sure these few novel items show as little formal resemblance as possible. Not only does this minimise the risk of erroneous form-meaning mappings, the presence of familiar items helps the learners integrate new input with old knowledge, and it can help instil self-confidence, too. Given the purpose for which they were selected, the set of target phrases used in many of the experimental studies we have reviewed in this article should not be taken as exemplary of pedagogical practice. After all, these experiments deliberately targeted phrases that were deemed to be new to the participants. A fair amount of 'mental crowding' was thus to be expected. In actual practice, well-informed teachers would undoubtedly rely much more on distributed learning.[8]

A second way of reducing the risk of interference is to opt for a clustering that is 'episodic' rather than 'semantic'. I borrow the term 'episodic' from Tulving (2002), but use it loosely to refer to story lines, scenarios or frames in which words and expressions are linked syntagmatically rather than paradigmatically. An example of an episodic set of words is *castle*, *haunted*, *night*, *ghost*, *afraid*, and

scream. Not only can the learner 'hook' each word to a moment in the 'frame', the set contains few words that belong to the same word class — unlike semantic sets where words of the same class are co-presented (nouns for body parts, adjectives for hot/cold, etc.). Although the evidence is inconclusive (Erten & Tekin, 2008), there is some agreement among second language vocabulary researchers that a presentation of vocabulary along episodic lines is more helpful than the use of semantic sets — although the spaced presentation of vocabulary remains advisable (Nation, 2000). If we then reconsider the sets of figurative idioms targeted in some of the CS studies, it is unmistakable that some of these actually lend themselves well to an episodic take. For example, it has been suggested that idioms for anger that instantiate the HOT FLUID IN A CONTAINER metaphor such as *fuming*, *exploding* and *simmering down*, can be presented in an order that corresponds to the expected sequence of events in the source domain (and which is mapped onto the target domain) (Beréndi et al., 2008, pp. 89–90). Idioms originating from boxing can be presented in an order that corresponds to the expected sequence of events in a boxing match: *flexing your muscles*, *not pull your punches*, *lower your guard*, *take it on the chin*, *be on the ropes* and *throw in the towel* (Lindstromberg & Boers, 2008a). And so on. Notice that the latter is certainly not an exhaustive set of boxing idioms. In keeping with the recommendation for distributed learning, it is advisable to delay the presentation of 'same-moment-in-the-frame' idioms until an initial repertoire is already in place. For instance, *stick your neck out* might otherwise at first compete for memory space with *lower your guard*, *in a tight corner* might compete with *on the ropes*, *be down for the count* might compete with *throw in the towel*, etc.

In short, the concern that CS-style instruction may have given insufficient consideration to the potential negative effects of learning vocabulary in semantic sets is mitigated by the realisation that the types of sets used in CS are in fact not necessarily of the kind known to negatively affect learning. Nevertheless, the words of caution about presenting vocabulary in sets are an incentive to conduct more longitudinal studies concerning the alternative, that is, the effects of pointing out motivations to students in a more incidental fashion, as suitable targets are encountered during meaning-focused, communicative activities. One can point out the hot-liquid-in-a-container image when students come across the phrase *simmer down* without introducing a whole list of same-metaphor phrases. Similarly, one can point out the boxing origin of *lower you guard* when students hit on this expression without necessarily trying to teach them more boxing idioms at that point in time. On the one hand, this may feel like abandoning one of the major strengths of CS, deemed to hold such great promise for second language learning, namely its ability to find system and coherence in idiomatic language. On the other hand, we have seen on several occasions in this article that, of the

different strengths of CS as a paradigm in descriptive linguistics, it is probably this one which is the hardest to implement in ways that chime with insights from mainstream applied linguistics.

The second characteristic of CS-style instruction that meets scepticism on the part of second language researchers is the type of target lexis that often features in CS studies — idioms. While formulaic sequences generally are now recognised to play a crucial role in language and language learning (e.g., Nattinger & DeCarrico, 1992; Ellis, 2008), figurative idioms as a subset of formulaic language are still considered 'the icing on the cake'. They are considered low-frequency and low-utility items, and thus certainly no priority for learning. If advocates of CS-style instruction want to make a case for the teaching of idioms, then they will have to show that they are of sufficient utility to merit teaching and intentional learning time. Boers & Lindstromberg (2009, pp. 63–67) sum up the arguments in favour of familiarizing students with a broad range of idioms at least for receptive purposes. Firstly, the frequency of occurrence of idioms as a class is seriously underestimated. Some text genres (e.g., popular novels) contain many figurative idioms, more perhaps than strong collocations — and the latter *are* deemed worthy of attention by applied linguists outside CS. Also, hand counts reveal a higher occurrence of idioms generally than computer-aided corpus counts, because idioms are much less lexically and formally fixed than is often assumed. This is attested, for example, in newspaper headlines (Herrera & White, 2010). Secondly, learners find idioms — unlike collocations consisting of known words — hard to comprehend, even when they are encountered in highly transparent contexts (Boers et al., 2007). Given the higher-than-assumed frequency of occurrence of idioms, this can have wider repercussions for text comprehension than one would initially expect. A reason for the latter is that idioms fulfil an important evaluative function at crucial points in discourse (McCarthy, 1998, pp. 129–149). There thus seem to be sufficient arguments to give broad attention to figurative idioms in second language instruction.

Whether idiom teaching should aim at productive knowledge as well as receptive knowledge is a different matter, however. For one thing, it tends to take a very high level of confidence for learners to start using L2 idioms actively (Irujo, 1993). This caution on the part of learners is probably justified, too. MacArthur & Boers (Forthcoming) lament how much instruction is required to help learners come to grips with the evaluative stance and subtle 'usage restrictions' of an idiom. CS-style presentations offer no quick fix for this problem, as students often ignore clues about the usage of an idiom that are present in the information they are given about its source domain.[9] So, as far as *productive* mastery of many figurative idioms is concerned, it looks as though we will have to agree with applied linguists who maintain that this is too ambitious a goal for non-advanced learners. In that

sense, some of the experiments we have reviewed and which gauge productive knowledge of idioms in a post-test involve a mismatch of student profiles and learning targets. Furthermore, *if* productive mastery is strived for, then priority should clearly be given to idioms that are comparatively frequent. Again, the selection of idioms for some of the experiments may give a wrong impression about the CS proposals, since few of the idioms targeted in those studies (e.g., Beréndi et al, 2008; Skoufaki, 2008) are signalled, for instance, in the corpus-based *Collins Cobuild Dictionary of Idioms* as frequent enough to merit prioritisation. But again, it must be borne in mind that these selections were made for experimental purposes, not for actual pedagogical implementation.

The third and final concern we will address here is also to do with multiword units, or formulaic sequences, as they are more generally called in applied linguistics circles. The principal function of formulaic sequences, according to many a psycholinguist and applied linguist is their facilitation of fluency (see Boers & Lindstromberg, 2009, Ch. 2, for a review). Formulaic sequences are recognised and produced by native speakers as ready-made chunks. Wray (2002) argues that (young) native speakers acquire formulaic sequences (i.e., a vast array of standardised multiword strings) 'holistically', and that herein lies an important advantage they have over second language learners. The word strings have been acquired as unanalysed chunks and can be retrieved as prefabricated wholes, which speeds up processing. According to Wray (2002), the holistic mode of acquiring formulaic sequences has its roots in pre-literacy, when the child does not yet discern word boundaries. Most adult second language learners, by contrast, have already gone through literacy training before they embark on learning the second language and will consequently tend to process formulaic sequences through an analytic mode, i.e. they will break the sequences down in words during the acquisition stage and will subsequently have to reassemble them if they want to produce them. This analytic processing compromises fluent use of formulaic sequences, so the theory goes, as it cannot compete with the speed at which holistically stored strings are retrieved. From a perspective where holistic processing of chunks is presented as ideal for purposes of fluency, it must seem paradoxical that CS-style instruction actually encourages analysis of phrases which native speakers normally process as unanalysed chunks. By having learners contemplate the non-arbitrary features of phrases (e.g., the motivated choice of particle in a phrasal verb or the polysemous nature of the keyword in an idiom), distinct parts of the phrases are attended to as contributors to meaning, and this goes against the grain of holistic processing. As Skoufaki (2008) points out, CS pedagogy stimulates the 'semantisation' stage of vocabulary learning, i.e. the initial forging of solid meaning-form connections (Beheydt, 1987). After that a lot remains to be done to make the acquired knowledge easily accessible for usage when appropriate. Given the relative depth

of analysis that CS stimulates during the semantisation stage, it is indeed likely that CS-instructed learners will devote rather more processing time to the figurative phrases they have been taught when they meet or need these on later occasions.[10] For one thing, they may still give pause to the figuration of the phrase, the nature of which they might not have been conscious of had they not received CS instruction. Whether this level of consciousness will prevent them also in the long term from processing idiomatic expressions 'without further ado', i.e., as native-speakers do most of the time, is hard to predict. Wray (2002, pp. 189–90) does concede that second language learners can also reach a good level of fluency thanks to well-entrenched — or automatised — 'procedural' knowledge, whereby assembling the word string becomes faster through practice (on automaticity, see, e.g., Segalowitz, 2007). According to Boers & Lindstromberg (2009, p. 78) the pragmatic implication of Wray's model for instructed second language acquisition is as follows: "If it is really so difficult for teenage and adult learners to switch off their analytical mode, then let us explore and assess ways of harnessing this mode rather than merely noticing its robustness. If students are not likely to process newly encountered chunks truly holistically, and if they must be taught to glue the constituent words of them together anyhow, then what could be so wrong with elaboration techniques that help the students appreciate the meaning and lexical makeup of these chunks in the first place?"

Still, it is true that CS proponents are generally silent on the challenge of helping learners develop fluency.[11] This, however, just illustrates the general truth that any proposed pedagogical intervention will occupy only a modest place within a language curriculum, due to the complex and multifaceted nature of language learning. Paul Nation's (2001, 2007) *Four Strands* framework is helpful to delimit the action radius of any language learning method, including CS-informed instruction. According to Nation, effective second language learning is fuelled by four types of stimuli, which he calls the four strands. The first strand is meaning-focused input. This refers to ample exposure to comprehensible input, the importance of which was underscored by Krashen's (1985) *Input Hypothesis*. The second strand is meaning-focused output. This relates to the desirability that learners experience gaps in their L2 communicative abilities which they subsequently try to fill. The importance of this strand is emphasised by Swain's (1993) *(Pushed) Output Hypothesis*. The third strand is fluency development. This refers to activities that foster the smooth retrieval and effortless use of the language elements the students have been learning. It is a strand the importance of which has been highlighted by DeKeyser (2001, 2007), among others. The fourth strand is language-focused instruction. This is the strand where teachers or materials writers try to draw learners' attention to specific language features. It may be 'proactive' in the sense that target features are taught and subsequently put into practice, or

it may be 'reactive' in the sense that (corrective) feedback is given during or after a communicative activity (Ellis et al., 2002). It is this fourth strand which people associate most closely with the term "teaching". It is evident that the CS-types of instruction we have discussed here fit this fourth strand, alongside many other types of language-focused instruction. To put things into perspective even more, Nation argues that a well-balanced language curriculum should devote approximately equal amounts of time to each of the four strands. This means that CS-style interventions have to cohabit with other kinds of interventions within a relatively confined space of the curriculum.

6. Conclusion

From the collection of studies reviewed here it seems undeniable that Cognitive Semantic ways of teaching figurative phrases constitute an effective channel for fostering comprehension and retention of those phrases. Both the available experimental evidence and models of memory suggest that the approach can be expected to be especially beneficial for students' retention of the meaning of the phrases, and to a lesser extent for their retention of the form of the phrases (i.e., the precise lexical makeup, spelling, etc.). When productive mastery is sought — which need not be the case for all phrases at all levels of proficiency — the semantic elaboration techniques employed under the Cognitive Semantic treatment are usefully supplemented by (a) structural elaboration techniques intended to help students remember the formal features of the given phrases and (b) practice intended to help students develop fluency with the learned phrases.

Our review has also laid bare a number of controversies within the circle of pedagogy-oriented Cognitive Semanticists, and these should incite further empirical research. Among the items on the research agenda is the utility of pictorials to help elucidate and remember the meaning of figurative phrases and their potentially negative side-effects on learners' recollection of linguistic form. In itself, the fact that research has turned to comparing the relative effectiveness of different *versions* of Cognitive Semantic instruction is an indication of how much the movement has matured.

At the same time, the review has illustrated at multiple points that Cognitive Semantic ventures into language pedagogy stand to gain a lot from a closer collaboration with 'mainstream' applied linguistics, and this at three levels:

a. at the level of research methodology, when it comes to the rigorous design of experimental studies, the ecological validity of the comparison treatments, and the analyses and reporting of quantitative data,

b. at the level of pedagogical implementation, more particularly the selection of targets for learning and the desirability of spaced, distributed learning, and
c. at the level of curricular integration, i.e. the judicious insertion of Cognitive Semantic interventions within the larger scheme of language learning and language teaching.

Acknowledgement

I would like to thank Francisco J. Ruiz de Mendoza Ibanez and his team for inviting me to the 2009 CRAL conference at the University of La Rioja, Spain. This most stimulating conference, which purposefully brought together second language vocabulary researchers and pedagogy-oriented Cognitive Semanticists, provided an excellent opportunity for the kind of cross-pollination I have pleaded for in this article.

Also a great many thanks to Seth Lindstromberg for his help — again — on an earlier draft of this paper.

Notes

1. For simplicity's sake, I will use the term *phrasal verb* as shorthand for multiword verbs generally, including prepositional verbs.

2. A distinction is made in applied linguistics circles between language instruction where preselected elements of the target language are taught as the focus of the lesson, and lessons that revolve primarily around a communicative activity and where the teacher intervenes occasionally to draw the students' attention to language elements when this is deemed necessary or helpful. The former type has been termed forms-focused instruction; the latter form-focused instruction (Long, 1991). In most of the learning experiments reported by Cognitive Semanticists, the instruction of which the effectiveness is put to the test is of the former type: participants are presented with a number of words or phrases to be learned, they are led to process this input in a particular way, and subsequently the amount of learning that has taken place is evaluated.

3. I cannot be certain that this is an exhaustive list of published studies that put the effectiveness of CS-style instruction to the test and that include a comparison treatment. If I have overlooked any, I would be very grateful to be informed about them. Also, there must be many *un*published studies that I have no knowledge of. I have deliberately excluded from this review studies such as those reported by Tyler (2008) which did not involve any comparison group or in which the experimental treatment was compared to a *no-treatment* condition.

4. We have to concede that the lack of proper pre-testing may shed some doubt on the outcomes of some of the experiments. At the same time, it should be pointed out that any difference in knowledge profiles between two treatment groups that a pre-test might reveal may be to the advantage of either group. If there were any such differences in the experiments which did not include a proper pre-test, then it would be a very unlikely coincidence if the difference were in all cases to the advantage of the experimental group.

5. It is a pity that in Gao and Meng's (2010) study, in which the post-test contained both a component that measured receptive knowledge and one that measured productive knowledge, no breakdown of the post-test scores per component is given. This information could corroborate the thesis that semantic elaboration techniques, such as CS-inspired interventions, are beneficial first and foremost for retention of meaning.

6. At a more general methodological level, we may wonder whether productive tests are an appropriate choice to evaluate the effectiveness of semantic elaboration techniques in the first place (DeKeyser & Sokalski, 2001).

7. Kövecses and Szabó (1996) refrained from calculating statistical significance, and were wisely cautious about drawing conclusions where the difference in scores between the two groups was not very pronounced. Given the time lapse since the experiment was conducted, it is unrealistic to expect that the participants' test sheets are still retrievable. However, a crude retrospective statistical analysis (by means of a *Wilcoxon Signed Ranks Test*) of the raw data as reported in the article suggests non-significance for the taught phrasal verbs (p 0.1) and significance (at p 0.01) for the new ones. This actually confirms the way the authors evaluated their findings.

8. The fact that the learning conditions in such experiments are sometimes quite different from what can be expected in genuine pedagogical practice evidently sheds some doubt on the 'ecological validity' of the experiments. This, however, is a methodological issue that stretches far beyond CS studies alone.

9. Boers et al. (2007) do report some evidence that awareness of particular source domains (e.g. games and sports) behind some idioms can help learners estimate whether these idioms might be typical of informal rather than formal registers.

10. Skehan (1998) argues that people use two different modes of language processing, to meet the demands of two distinct conditions of communication, that is, 'planned' versus 'real-time' communication. CS-instruction probably caters first and foremost for the former.

11. I am referring here to linguistic fluency (e.g. speech without awkward hesitations), not to what has become known in Cognitive Semantics circles as *conceptual* fluency (Danesi, 1995).

References

Baker, J. (1998). Metaphor. *English Teaching Professional, 7,* 13–14.

Barcroft, J. (2002) Semantic and structural elaboration in L2 lexical acquisition. *Language Learning*, *52*, 323–363.

Beheydt, L. (1987). The semantisation of vocabulary in foreign language learning. *System, 15,* 55–67.

Beréndi, M., Csábi, S. & Kövecses, Z. (2008). Using conceptual metaphors and metonymies in vocabulary teaching. In F. Boers & S. Lindstromberg (Eds.), *Cognitive Linguistic Approaches to Teaching Vocabulary and Phraseology* (pp. 65–99). Berlin: Mouton de Gruyter.

Boers, F. (2000a). Metaphor awareness and vocabulary retention. *Applied Linguistics, 21*, 553–571.

Boers, F. (2000b). Enhancing metaphoric awareness in specialised reading. *English for Specific Purposes, 19*, 137–147.

Boers, F. (2001). Remembering figurative idioms by hypothesising about their origin. *Prospect, 16*, 35–43.

Boers, F. (2004). Expanding learners' vocabulary through metaphor awareness: What expansion, what learners, what vocabulary? In M. Achard & S. Niemeier (Eds.), *Cognitive Linguistics, Second Language Acquisition, and Foreign Language Teaching* (pp. 211–234). Berlin: Mouton de Gruyter.

Boers, F. & Demecheleer, M. (1998). A cognitive semantic approach to teaching prepositions. *English Language Teaching Journal, 53*, 197–204.

Boers, F. & Lindstromberg, S. (2008). From empirical findings to pedagogical practice. In F. Boers & S. Lindstromberg (Eds.), *Cognitive Linguistic Approaches to Teaching Vocabulary and Phraseology* (pp. 375–393). Berlin: Mouton de Gruyter.

Boers, F. & Lindstromberg, S. (2009). *Optimizing a Lexical Approach to Instructed Second Language Acquisition*. Basingstoke, UK: Palgrave Macmillan.

Boers, F., Demecheleer, M. & Eyckmans, J. (2004). Etymological elaboration as a strategy for learning figurative idioms. In P. Bogaards & B. Laufer (Eds.), *Vocabulary in a Second Language: Selection, Acquisition and Testing* (pp. 53–78). Amsterdam: John Benjamins.

Boers, F., Eyckmans, J. & Stengers, H. (2007). Presenting figurative idioms with a touch of etymology: More than mere mnemonics? *Language Teaching Research*, *11*, 43–62.

Boers, F., Lindstromberg, S., Littlemore, J., Stengers, H. & Eyckmans, J. (2008). Variables in the mnemonic effectiveness of pictorial elucidation. In F. Boers & S. Lindstromberg (Eds.), *Cognitive Linguistic Approaches to Teaching Vocabulary and Phraseology* (pp. 189–216). Berlin: Mouton de Gruyter.

Boers, F., Piquer Píriz, A.M, Stengers, H. & Eyckmans, J. (2009). Does pictorial elucidation foster recollection of idioms? *Language Teaching Research, 13*, 367–382.

Cho, K. (2010). Fostering the acquisition of English prepositions by Japanese learners with networks and prototypes. In S. De Knop, F. Boers & T. De Rycker (Eds.), *Fostering Language Teaching Efficiency through Cognitive Linguistics* (pp. 259–276). Berlin: Mouton de Gruyter.

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

Condon, N. (2008). How Cognitive Linguistic motivations influence the learning of phrasal verbs. In F. Boers & S. Lindstromberg (Eds.), *Cognitive Linguistic Approaches to Teaching Vocabulary and Phraseology* (pp. 133–158). Berlin: Mouton de Gruyter.

Csábi, S. (2004). A cognitive linguistic view of polysemy in English and its implications for teaching. In M. Achard & S. Niemeier (Eds), *Cognitive Linguistics, Second Language Acquisition, and Foreign Language Teaching* (pp. 233–256). Berlin: Mouton de Gruyter.

Danesi, M. (1992). Metaphorical competence in second language acquisition and second language teaching: The neglected dimension. In J.A. Alatis (Ed.), *Language, Communication and Social Meaning* (pp. 489–500). Washington, DC: Georgetown University Round Tables on Languages and Linguistics.

Danesi, M. (1995). Learning and teaching languages: The role of 'conceptual fluency'. *International Journal of Applied Linguistics, 5*, 3–20.

Deignan, A., Gabrys, D. & Solska, A. (1997). Teaching English metaphors using cross-linguistic awareness-raising activities. *ELT Journal, 51*, 352–360.

DeKeyser, R. (2001). Automaticity and automatization. In P. Robinson (Ed.), *Cognition and Second Language Instruction* (pp. 125–151). Cambridge: Cambridge University Press.

DeKeyser, R. (2007). Situating the concept of practice. In R. DeKeyser (Ed.), *Practicing in a Second Language: Perspectives from Applied Linguistics and Cognitive Psychology* (pp. 1–18). Cambridge: Cambridge University Press.

DeKeyser, R. & Sokalski, K. (2001). The differential role of comprehension and production practice. *Language Learning, 51*, 81–112.

Dirven, R. (2001). English phrasal verbs: Theory and didactic application. In M. Pütz, S. Niemeier & R. Dirven (Eds.), *Applied Cognitive Linguistics II: Language Pedagogy* (pp. 3–27). Berlin: Mouton de Gruyter.

Ellis, N. (2008). Phraseology: The periphery and the heart of language. In F. Meunier and S. Granger (Eds.), *Phraseology in Foreign Language Learning and Teaching* (pp. 1–13). Amsterdam: John Benjamins.

Ellis, R., Basturkmen, H. & Loewen, S. (2002). Doing focus on form. *System, 30*, 419–435.

Erten, İ.H. & Tekin, M. (2008). Effects on vocabulary acquisition of presenting new words in semantic sets versus semantically unrelated sets. *System*, *36*, 407–422.

Finkbeiner, M. & Nicol, J. (2003) Semantic category effects in second language word learning. *Applied Psycholinguistics, 24*, 369–383.

Gao, L-Q. & G-H. Meng (2010). A study of the effect of metaphor awareness raising on Chinese EFL learners' vocabulary acquisition and retention. *Canadian Social Science, 6*(2), 110–124.

Guo, S-F (2007). Is idiom comprehension influenced by metaphor awareness of the learners? *The Linguistics Journal, 3*, 148–166.

Gibbs, R.W. (1994). *The Poetics of Mind: Figurative Thought, Language and Understanding*. Cambridge: Cambridge University Press

Gude, K. & Duckworth, M. (1999). *Proficiency Masterclass: Workbook*. Oxford: Oxford University Press.

Hannan, P. (1998). Particles and gravity: Phrasal verbs with *up* and *down*. *Modern English Teacher,* 7, 21–27.

Herrera, H. & White, M. (2010). Canonicity and variation in idiomatic expressions: Evidence from business press headlines. In S. De Knop, F. Boers &T. De Rycker (Eds.), *Fostering Language Teaching Efficiency through Cognitive Linguistics* (pp. 167–187). Berlin: Mouton de Gruyter.

Holme, R. (2001). Metaphor, language learning, and affect. *Humanising Language Teaching,* 3. www.hltmag.co.uk. Pilgrims.

Holme, R. (2004). *Mind, Metaphor and Language Teaching*. Basingstoke, UK: Palgrave Macmillan.

Hu, Y-H. & Fong, Y-Y. (2010). Obstacles to Conceptual-Metaphor guided L2 idiom interpretation. In S. De Knop, F. Boers, and T. De Rycker (Eds.), *Fostering Language Teaching Efficiency through Cognitive Linguistics* (pp. 293–317). Berlin: Mouton de Gruyter.

Irujo, S. (1993). Steering clear: Avoidance in the production of idioms, *International Review of Applied Linguistics in Language Teaching*, *31*, 205–219.

Juchem, C. & Krennmayr, T. (2010). Corpus-informed integration of metaphor in materials for the business English classroom. In S. De Knop, F. Boers, and T. De Rycker (Eds.), *Fostering Language Teaching Efficiency through Cognitive Linguistics* (pp. 319–337). Berlin: Mouton de Gruyter.

Kövecses, Z. (1986). *Metaphors of Anger, Pride and Love: A Lexical Approach to the Study of Concepts*. Amsterdam: John Benjamins.

Kövecses, Z. (2001). A cognitive linguistic view of learning idioms in an FLT context. In M. Pütz, S. Niemeier & R. Dirven (Eds.), *Applied Cognitive Linguistics II: Language Pedagogy* (pp. 87–115). Berlin: Mouton de Gruyter.

Kövecses, Z. & Szabó, P. (1996). Idioms: A view from Cognitive Semantics. *Applied Linguistics, 17*, 326–355.

Kövecses, Z, Tóth, M. & Babarci, B. (1996–1998). *A Picture Dictionary of English Idioms* (four volumes). Budapest: Eötvös University Press.

Krashen, S.D. (1985). *The Input Hypothesis: Issues and Implications*. London: Longman.

Kurtyka, A. (2001). Teaching phrasal verbs: A cognitive approach. In M. Pütz, S. Niemeier & R. Dirven (Eds.), *Applied Cognitive Linguistics II: Language Pedagogy*, 29–54. Berlin: Mouton de Gruyter.

Lakoff, G. (1987). *Women, Fire and Dangerous Things: What Categories Reveal about the Mind*. Chicago: University of Chicago Press.

Lakoff, G. & Johnson, M. (1980). *Metaphors We Live By*. Chicago: University of Chicago Press.

Laufer, B. (2005). Focus on Form in second language vocabulary acquisition. In Susan Foster-Cohen (Ed.), *EUROSLA Yearbook 5* (pp. 223–250). Amsterdam: John Benjamins.

Laufer, B. & Hulstijn, J. (2001). Incidental vocabulary acquisition in a second language: The construct of task-induced involvement, *Applied Linguistics, 22*, 1–26.

Lazar, G. (1996). Using figurative language to expand students' vocabulary. *ELT Journal, 50*, 43–51.

Li, T. F. (2009) *Metaphor, Image, and Image Schemas in Second Language Pedagogy*. Koln: Lambert Academic Publishing.

Li, X. (2010). Conceptual Metaphor Theory and teaching of English and Chinese idioms. *Journal of Language Teaching and Research, 1*(3), 206–210.

Lindstromberg, S. (1991). Metaphor and ESP: A ghost in the machine? *English for Specific Purposes, 10*, 207–225.

Lindstromberg, S. (2010). *English Prepositions Explained*. 2nd ed. Amsterdam: John Benjamins.

Lindstromberg, S. & Boers, F. (2005). From movement to metaphor with manner-of-movement verbs. *Applied Linguistics, 26*, 241–261.

Lindstromberg, S. & Boers, F. (2008a). *Teaching Chunks of Language*. Rum, Austria: Helbling Languages.

Lindstromberg, S. & Boers, F. (2008b). The mnemonic effect of noticing alliteration in lexical chunks. *Applied Linguistics*, *29*, 200–222.

Lindstromberg, S. & Boers, F. (2008c). Phonemic repetition and the learning of lexical chunks: The mnemonic power of assonance. *System, 36*, 423–436.

Littlemore, J. (2009). *Applying Cognitive Linguistics to Second Language Learning and Teaching*. Basingstoke, UK: Palgrave Macmillan.

Littlemore, J. & Low, G. (2006). *Figurative Thinking and Foreign Language Learning*. Basingstoke, UK: Palgrave MacMillan.

Long, M.H. (1991). Focus on form: A design feature in language teaching methodology. In K. de Bot, R.B. Ginsberg & C. Kramsch (Eds.), *Foreign Language Research in Cross-Cultural Perspective* (pp. 39–52). Amsterdam: John Benjamins.

MacArthur, F. & Boers, F. (Forthcoming). Using visuals to illustrate the source domains of idioms: Can they help learners appreciate usage restrictions too? In S. Niemeier & C. Juchem (Eds), *Metaphors and Language Teaching*. Berlin: Mouton de Gruyter.

MacLennan, C.H.G. (1994). Metaphors and prototypes in the teaching and learning of grammar and vocabulary. *International Review of Applied Linguistics, 32*, 97–110.

McCarthy, M. (1998). *Spoken Language and Applied Linguistics*. Cambridge: Cambridge University Press.

McCarthy, M. & O'Dell, F. (2002). *English Idioms in Use*. Cambridge: Cambridge University Press.

Morimoto, S. & Loewen, S. (2007). A comparison of the effects of image-schema-based instruction and translation-based instruction on the acquisition of L2 polysemous words. *Language Teaching Research, 11*, 347–372.

Nation, P. (2000). Learning vocabulary in lexical sets: Dangers and guidelines', *TESOL Journal*, *9*, 6–10.

Nation, P. (2001). *Learning Vocabulary in another Language*. Cambridge: Cambridge University Press.

Nation, P. (2007). The four strands. *Innovation in Language Learning and Teaching, 1*, 1–12.

Nattinger, J.R. & DeCarrico, J.S. (1992). *Lexical Phrases and Language Teaching*. Oxford: Oxford University Press.

Nelson, D.L., Reed, V.S. & J.R. Walling (1976). Pictorial superiority effect. *Journal of Experimental Psychology: Human Learning and Memory, 2*, 523–528.

Paivio, A. (1986). *Mental Representations: A Dual Coding Approach*. New York: Oxford University Press.

Paivio, A. & Desrochers, A. (1979). Effects of an imagery mnemonic on second language recall and comprehension. *Canadian Journal of Psychology, 33*, 17–28.

Ponterotto, D. (1994). Metaphors we can learn by: How insights from cognitive linguistic research can improve the teaching/learning of figurative language. *Forum, 32*, 2–7.

Riding, R.J. (1991). *Cognitive Styles Analysis*. Birmingham, UK: Learning and Training Technology.

Rudzka-Ostyn, B. (2003). *Word Power: Phrasal Verbs and Compounds. A Cognitive Approach*. Berlin: Mouton de Gruyter.

Rundell, M. (2001). Cats, conversations, and metaphor. *Humanising Language Teaching, 3*. www.hltmag.co.uk. Pilgrims.

Rundell, M. (2002). Metaphorically speaking. *English Teaching Professional, 23*, 21.

Schmitt, N. (Ed.). (2004). *Formulaic Sequences: Acquisition, Processing and Use*. Amsterdam: John Benjamins.

Schmitt, N. (2008). Instructed second language vocabulary learning. *Language Teaching Research, 12*, 329–363.

Scott, M. (1994). Metaphors and language awareness. In L. Barbara & M. Scott (Eds.), *Reflections on Language Learning* (pp. 89–104). Clevedon, UK: Multilingual Matters.

Segalowitz, N. (2007). Access fluidity, attention control, and the acquisition of fluency in a second language. *TESOL Quarterly*, *41*, 181–186.

Skoufaki, S. (2008). Conceptual metaphoric meaning clues in two idiom presentation methods. In F. Boers & S. Lindstromberg (Eds.), *Cognitive Linguistic Approaches to Teaching Vocabulary and Phraseology* (pp. 101–132). Berlin: Mouton de Gruyter.

Swain, M. (1993). The output hypothesis: Just speaking and writing aren't enough. *The Canadian Modern Language Review*, *50*, 158–164.

Tinkham, T. (1993). The effect of semantic clustering on the learning of second language vocabulary. *System, 21*, 371–380.

Tinkham (1997). The effects of semantic and thematic clustering on the learning of second language vocabulary. *Second Language Research, 13*, 138–163.

Tulving, E. (2002). Episodic memory: From mind to brain. *Annual Review of Psychology, 53*, 1–25.

Tyler, A. (2008). Cognitive Linguistics and second language instruction. In P. Robinson & N.C. Ellis (Eds.), *Handbook of Cognitive Linguistics and Second Language Acquisition* (pp. 456–488). New York/London: Routledge.

Tyler, A. & Evans, V. (2004). Applying cognitive linguistics to pedagogical grammar: The case of *over*. In M. Achard & S. Niemeier (Eds.), *Cognitive Linguistics, Second Language Acquisition, and Foreign Language Teaching* (pp. 257–280). Berlin: Mouton de Gruyter.

Verspoor, M. & Lowie, W. (2003). Making sense of polysemous words. *Language Learning, 53*, 547–586.

Waring, R. (1997). The negative effects of learning words in semantic sets: A replication. *System*, *25*, 261–274.

Wray, A. (2002). *Formulaic Language and the Lexicon.* Cambridge: Cambridge University Press.

Author's address

Frank Boers
School of Linguistics and Applied Language Studies
Victoria University of Wellington
Von Zedlitz building, Kelburn Parade, PO box 600
Wellington 6140,. New Zealand

Frank.Boers@vuw.ac.nz

About the author

Frank Boers has over 20 years of experience teaching EFL, ESP, EAP, TEFL, linguistics and applied linguistics in Belgium. His early publications were Cognitive Linguistic takes on polysemy and metaphor, but he has gradually shifted his research interest to instructed second language acquisition. He is author (with Seth Lindstromberg) of *Optimizing a Lexical Approach to Instructed Second Language Acquisition* (Palgrave Macmillan) and co-editor of two volumes on pedagogy-oriented Cognitive Linguistics (Mouton de Gruyter). Frank is now an Associate Professor at the School of Linguistics and Applied Language Studies of the Victoria University of Wellington, New Zealand.