

Conceptual metaphoric meaning clues in two idiom presentation methods

Sophia Skoufaki

Abstract

Cognitive linguists have been exploring how giving learners information about the motivation of L2 idioms can facilitate comprehension and retention. Results overall are encouraging, but the relative effectiveness of different proposals has not yet been examined. In this chapter I report a small-scale experiment that was set up to compare the relative effectiveness of two CL-inspired idiom presentation methods. In both conditions learners were presented with idioms that were grouped under conceptual metaphors (CMs). In one condition the meaning of each idiom was given, while in the other condition the learners were encouraged to use CM clues to guess the meaning before it was explained. Post-test results suggest that the addition of a guessing task is likely to enhance the effectiveness of presenting idioms in CM groups. The participants were relatively unsuccessful at guessing the meaning of the idioms correctly, though, and this indicates that a fair amount of guidance and feedback is required for this method to pay off in the classroom.

Keywords: idioms; metaphor; grouping; guessing; retention of form; retention of meaning.

1. Introduction

Idioms have often been considered to be an unsystematic part of the lexicon and hence difficult to teach (e.g. Sornig 1988: 281). Cognitive Linguistics (CL), however, claims that many idioms are motivated by conceptual metaphors (CMs) and metonymies (e.g. Kövecses and Szabó 1996: 337-338; Lakoff 1987: 448-449).

In recent years, cognitive linguists have been exploring the effect that giving learners information about the motivation of L2 idioms can have on comprehension and retention (e.g. Kövecses and Szabó 1996; Boers 2000). Results are encouraging, but the relative effectiveness of the various

proposals has not yet been examined. This chapter reports the results of an experiment that was set up to compare the relative effectiveness of two idiom-instruction methods, both of which involve providing learners with CM clues.

2. Two CL-inspired methods of L2 idiom presentation

The first CL proposal for L2 idiom instruction was to present idioms in CM groups. Grouping idioms appears to be a beneficial method of presentation because it seems in general to be congruent with people's preferred vocabulary learning methods. Schmitt (1997: 211-217) refers to studies which (1) show that in unprompted free recall tasks people tend to retrieve words in category clusters and (2) indicate that when words are presented in categories they are better remembered. Schmitt also cites a study showing that relatively advanced L2 learners use grouping strategies in vocabulary learning more than beginners do. However, the problem with this method of idiom instruction is that idioms differ along more than one variable, so diverse categorisation criteria have been proposed. For example, Sornig (1988: 286-288) proposes three: according to (1) the surface structure of the target idioms, (2) their motivation, and (3) the communicative purpose they serve (e.g. to describe a situation, to serve as phatics). Moon (1997) proposes a number of possible categorisations, including categorisation by speech act and by the situations in which idioms could be used.

A further teaching method is that of encouraging students to guess the meanings of idioms. Irujo (1984), for example, proposed that teachers instruct L2 learners how to guess idiom meanings from context because this instruction could lead to the development of an idiom-meaning guessing strategy and hence to the learning of more idioms than can be taught in class (Irujo 1984: 128-129). In a later publication, Irujo (1993: 217) argues that focusing students' attention on the features of idioms through urging them to guess at their meaning in context may lead to better learning since the idioms that seem to be learned best are the most transparent ones. Lennon (1998), although without presenting new empirical evidence, has also proposed that learners be encouraged to guess at the meaning of idioms. His rationale is that this will involve deeper processing and should therefore lead to better retention.

3. Cognitive Linguistic experimental data on the idiom-grouping method

Kövecses and Szabó (1996) were the first to investigate the impact of CM groupings on L2 idiom learning. Their main aim was to see whether presentation of L2 vocabulary in terms of CM groupings would lead to better vocabulary memorisation than the simple presentation of words accompanied by their L1 equivalents. Hungarian adult intermediate-level learners of English were taught phrasal verbs with *up* and *down*. Half of the participants were presented with the Hungarian equivalents of the phrasal verbs and the other half with the underlying CMs as well. In the test phase following the instruction phase, participants had to fill in gapped sentences with *up* or *down*. Half of the sentences contained the phrasal verbs from the learning phase and the rest contained phrasal verbs that had not been taught. The group that had received the CM instruction gave a higher percentage of correct responses than the other group as far as the taught phrasal verbs were concerned. The researchers did not ascertain whether this difference was statistically significant, though. Moreover, rather than being caused by this *specific* categorisation principle, the better performance under the CM condition might be the consequence of categorisation in general. Psycholinguistic experiments comparing the relative ease of vocabulary learning across different kinds of input categorisation have shown that almost any kind of grouping leads to better memorisation than no grouping at all (Higa 1963; Tinkham 1993, 1997; Waring 1997).

In his first experiment, Boers (2000) compared the effectiveness of grouping new vocabulary by CMs and by functional themes. In his experiment (Boers 2000: 554-557), Belgian native speakers of Dutch who were intermediate-level learners of English were asked to fill in a gapped text with figurative lexis used to describe anger, namely, idioms (e.g. *bite someone's head off*, *be hot under the collar*), metaphorical words and collocations (e.g. *explode*, *be fuming*, *inflammatory remark*), and phrasal verbs (e.g. *simmer down* and an expression stemming from a phrasal verb, *be steamed up*). The items had been presented to one group of students according to a CM grouping (e.g. according to whether they expressed ANGER IS A HOT FLUID IN A CONTAINER or ANGER IS MADNESS) and to another group according to functional dimensions such as *to describe acute and sudden anger*, *to describe anger as a process*. The group that was presented with these lexical items in CM groups outperformed the other group significantly.

Luciani de las Mercedes (2001: 47-49) conducted a similar study with advanced Spanish learners of English, half of whom were encouraged to think of newly-presented vocabulary (prepositions, metaphorical words, and verbal idioms) in terms of CMs and the other half in terms of functional groupings. For example, in the first condition the sentence *She needs some direction in life* appeared under the title LIFE IS A JOURNEY and *Get the most out of life* appeared under LIFE IS A CONTAINER. In the other condition these sentences appeared under the title TALKING ABOUT LIFE. The group that had received the CM instruction performed significantly better in cloze tests in the course of which they had to choose from a selection of the words which had been taught (Luciani de las Mercedes 2001: 54-55). Luciani de las Mercedes gave another group of participants exercises that alerted them to the CM basis of certain conventional metaphoric expressions. The majority of both the participants and the L2 teachers who administered the exercises considered the exercises useful, interesting and innovative (Luciani de las Mercedes 2001: 50-52).

Boers's and Luciani de las Mercedes's experimental results suggest that CM categorisation may be more effective than a functional one. However, because these researchers tested participants on more than one category of lexical items, no claims can be made about which kind(s) of vocabulary are indeed learned more effectively when CM categorisation is used. The results from Li (2002: 187-190) suggest that different categorisation principles may differentially affect the memorisation of metaphorically used words and short collocations but not that of verbal idioms. In his fourth experiment, Li compared the performance of three groups of EFL-learner, first-year Chinese university students in a cloze test where the target items, English V+NP (PP) idioms, were presented (1) without any additional context and (2) with one or two words missing from each idiom. For instance, the test for *spill the beans* was "to spill _____". These idioms had been taught to the three groups of participants in three ways:

- 1) To the first group in semantic sets¹. For example, the idioms were presented under the titles Anger, Exerting Authority, Secretiveness, Insanity, and Revelation.
- 2) To the second group in CM sets embedded in a semantic categorisation. For example, the idioms *hit the ceiling*, *foam at the mouth*, and *lose your cool* were those presented under the title Anger in both conditions, but in the second condition, the title Anger was followed by the expressions denoting the underlying CMs, THE MIND IS A CONTAINER, ANGER IS HEAT.

- 3) To the third group according to CM sets embedded in a semantic categorisation, complemented by questions aimed to evoke mental images. For example, for the idiom *hit the ceiling* the questions were (Li 2002: 158): “What image do you have in your mind when you read *hit the ceiling*? Where does this force come from? What’s the result after the ceiling was hit? Who hits the ceiling?”

The same cloze test was administered both immediately after the teaching phase and a week later. In both tests, neither of the CM conditions led to significantly more correct answers than the semantic condition. This result may be due to verbal idioms being too long to memorise correctly after just one presentation or to their being lower in transparency than other kinds of figurative expressions. These speculations seem plausible because in his first experiment, Li did find better performance in a cloze test for the CM than the semantic conditions when teaching the same kind of learners conventional metaphoric usages of words (Li 2002: 178-179).

One final point about this research is that Boers’s and Luciani de las Mercedes’s experiments tested for memorisation of both form and meaning at the same time through the use of cloze tests. This kind of test does not enable one to distinguish the effect of the instruction method on the memorisation of *form* from its effect on the memorisation of *meaning*. Idiom-instruction methods might have a differential effect on the memorisation of form but not on that of meaning or vice versa. Therefore, it could be useful to use tests that examine form and meaning retention separately.

4. Cognitive Linguistic experimental data on the idiom-meaning guessing method

Experimental testing of the viability of the idiom-meaning guessing method has been sparse, but it has recently been increasing. Irujo (1993), Kövecses and Szabó (1996), and Boers (2000) draw conclusions about the effectiveness of this method in experiments addressing other questions. Boers and Demecheleer (2001) devote an experiment to a question more directly related to this issue. Finally, recent studies by Csábi (2004), and Boers, Eyckmans and Stengers (2007) indicate that a combination of this method with the provision of unambiguous meaning clues may be an effective idiom-instruction option.

In Irujo (1993) the task was to translate paragraphs including idioms from L1 Spanish to L2 English. Irujo conducted correlation analyses between, on the one hand, the number of correct translations and, on the

other, (1) the frequency of the idiom's use by L2 native speakers, (2) the idiom's transparency, and (3) whether the L1 idiom was totally equivalent to, similar to, or different in form from the L2 idiom. The second strongest correlation was for the factor of transparency. This finding was taken as an indirect indication that learners tend to think about idiom transparency and this process helps them memorise idioms. Therefore, Irujo considered that focusing the students' attention on idiom features through urging them to guess at the meaning of idioms could lead to better learning (Irujo 1993: 217).

In addition to the issue of grouping idioms by CM, Kövecses and Szabó (1996) studied whether L2 learners, once they have been introduced to the notion of CMs, develop a strategy of guessing at the meanings of new idiomatic expressions of taught CMs. The higher performance of the CM condition (as compared to the no-categorisation condition) with respect to the untaught items led Kövecses and Szabó to conclude that it was the CM condition that enabled students to use CMs to guess at the meaning of novel phrasal verbs. This evidence is encouraging for the idiom-meaning guessing method because it seems to show that people are more likely to guess correctly the meaning of at least some idioms (i.e. phrasal verbs) once they have been encouraged to use their CM knowledge than when not.

The third experiment in Boers (2000: 559-562) replicated the design of Kövecses and Szabó (1996) but used more participants and different input. Boers used CMs to group phrasal and prepositional verbs with *up*, *down*, *in*, *out* and other prepositions. For example, *blow up* and *cut down* appeared under the title MORE IS UP; LESS IS DOWN and *come up with an idea/solution* and *find out something* appeared under VISIBLE IS OUT AND UP; INVISIBLE IS IN AND DOWN. The test phase was a cloze test where, among the words to choose from, were also words that had not been taught. Significantly better retention was found in the CM condition with regard to the taught items, but not with regard to the untaught ones. The latter finding may be explained by the fact that some of the untaught items instantiated CMs different from those used in the instruction phase (Boers (2000: 562).

In the aforementioned studies, meaning guessing was considered to occur during the solution of exercises containing unknown words. In contrast, in Boers and Demecheleer (2001) L2 learners were explicitly asked to guess at the meaning of imageable idioms² presented out of context. The finding that relates to our discussion is that at least 35% of the definitions learners gave pointed to the same general metaphorical meaning as the stimulus idiom. However, as Boers and Demecheleer admit, the number of idioms used was small. Moreover, only a small percentage of meaning guessing was absolutely successful. Nevertheless, Boers and

Demecheleer propose that students be encouraged to guess at the meanings of idioms and be given corrective feedback afterwards. They suggest that presenting idioms in contexts will facilitate correct guessing. Research on inferring the meaning of novel L2 words in context shows very low percentages of successful guesses, though (e.g. Herman et al. 1985; Nagy, Herman, and Anderson 1985a, 1985b).

In terms of retention of the form of idioms, an incorrect analysis of the link between the form and the meaning of an idiom in a guessing task may carry the risk of production errors. For example, in the first experiment reported in Skoufaki (2005)³, where Greek advanced learners of English guessed at the meaning of L2 VP idioms, many of them interpreted *chew the fat* as 'become slim' or 'lose weight' when it appeared out of context. Such an interpretation might lead to mistakes in the production of the idiom in their speech, such as *fight the fat* or *lose the fat*.

The studies reported in Csábi (2004) are also indirectly linked to the issue of the efficacy of idiom-meaning guessing. In these experiments, teenage Hungarian native speakers learning English as a foreign language were first taught meanings of the verbs *hold* and *keep*. They were then taught phrasal verbs containing these verbs (e.g. *hold back something* and *keep in something*), and, finally, VP idioms containing these verbs (e.g. *hold one's head up* and *keep something under one's hat*). In the phrasal-verb and VP idiom-teaching conditions, subjects were either given clues of their semantic motivation or just given their meaning in Hungarian. In the instruction phase of the idioms, drawings meant to evoke mental images of the literal meanings of the idioms were made on the blackboard. Learners were not asked to guess at the meaning of the expressions they were presented with, but they usually voluntarily voiced their interpretations and in many cases they were correct, so then Csábi did not have to give the L1 equivalents (Csábi, personal communication).

After each instruction phase, there was a cloze test from which the *whole* expressions were missing. The same cloze tests were administered to participants one day after the treatment in the first experiment and two days after it in the second experiment. Data were analysed not only for the correct answers to the cloze tests but also after collapsing the correct answers and those that were not absolutely correct but included the correct verb, at least. The total scores of the "motivations" condition were superior to those of the "Hungarian meaning equivalents" condition.

The results for the VP idioms seem especially interesting. Idioms were "the most difficult to remember, since the number of correct results was the lowest there" (Csábi 2004: 248). Csábi attributes this result to idioms being too long to memorise precisely in such a short period of time. Another

noteworthy result was that in her second experiment the number of entirely correct answers was higher when just the Hungarian idiom meanings were given to the learners than when their motivations were explained as well. However, when she collapsed the data of the correct and the nearly correct answers, the latter condition had a significantly higher score. These results may be taken to suggest that, whenever idioms were guessed at with the help of their motivations, guessing led to superior results in form retention as compared to when there was mere presentation of idiom meanings. However, the results cannot be totally attributed to the learners' meaning-guessing attempts. Firstly, meaning guessing was done on the students' initiative. Secondly, as Boers and Lindstromberg (2006) point out, results may be partly due to the use of pictorial input only in the experimental condition. The higher score for the traditional instruction method when totally correct answers were counted indicates that the length of VP idioms and/or other factors, such as low transparency, may block the beneficial effect that a vocabulary-instruction method involving guessing can have on memorisation.

An idiom instruction method similar to that in Csábi (2004) is reported in Boers, Eyckmans and Stengers (2007). It is a variation of the experiments on "etymological elaboration" by Boers and his colleagues (Boers 2001; Boers, Demecheleer, and Eyckmans 2004a, b). Etymological elaboration refers to the practice of making learners aware of the historical-cultural-etymological origin of idioms. This practice is believed to call up in learners' minds mental images of concrete scenes that can be stored in memory alongside the verbal form (and any propositional meaning) of the idiom – in short, lexical information can be dually coded. In the control condition, participants were asked to identify the meaning of each idiom first and then the domain of experience from which the idiom stemmed (e.g. warfare, sailing, gardening, commerce, sports). In the experimental condition these tasks were reversed, so that these students received feedback about the origins of the idioms before they were asked to guess at the figurative meaning of the expressions. Results indicated a tendency towards more correct meaning guesses in the experimental (etymological-elaboration-first) condition than in the control (etymological-elaboration-second) condition.

A secondary aim of this experiment, but one that relates to our discussion, was to see whether the condition where the etymological information preceded the meaning guessing would lead to better form retention of the idioms taught, as measured in cloze tests in which the keyword of each idiom was missing. Students in the experimental condition indeed tended to outperform their control peers. This result was interpreted

with a depth-of-processing argument: the information about the origins of the idioms was probably actively used by the experimental students to figure out the figurative meaning of the idioms. The assumption is that depth of processing of newly presented vocabulary determines retention. This is also a conclusion that Hulstijn (2003: 364) reached in his review of the literature on vocabulary learning.

There are some reasons for scepticism about the effectiveness of idiom-meaning guessing, however. For one thing, some L2 learners may prefer to use mechanical word learning strategies rather than those requiring deep processing (Schmitt 1997). For another, students' guesses are quite often wrong, and this may carry the risk of negative affect setting in. Still, Csábi (2004) and Boers, Eyckmans, and Stengers (2007) do indicate that when the idiom-meaning guessing task is properly guided, relatively high success rates can be obtained. Given that the presentation of idioms in CM groups could be part of such guidance, the following question arises: is a combination of the guessing method with the grouping method more effective in promoting retention than the grouping method alone?

5. Conceptual Metaphoric grouping of idioms versus a Hybrid idiom instruction method

Table 1 sums up the tasks for each of the conditions in the experiment.

Table 1. Experimental design overview

Condition	Instruction phase	Practice phase	Test phase
1	Metaphor groups.	“Meaning-focused”: Students see all the target items in italics in two texts. Then, they answer three questions, each including each item.	(a) Retention for re-production, i.e. of form: cloze test.
2	Metaphor groups + meaning guessing.		(b) Retention for comprehension, i.e. of meaning: answering questions containing a target expression
3	Metaphor groups.	“Form-focused”: In the texts from conditions 1 and 2, the items included in the questions in those conditions are missing.	

In condition 1, participants were presented with figurative expressions which were in CM groups and accompanied by their definitions, in Greek, along with sentences illustrating their meaning. The participants were then asked to read through all this information. In condition 2, participants were presented with the expressions in the same CM groups but without the definitions and examples. These participants were asked to write what they guessed each expression meant on the basis of the cues they were given by the (CM) group titles.

In both conditions, a practice phase was inserted between the instruction and the test phases. Participants were asked to read texts which included, in italics, all the expressions they had been taught previously and then they were given three minutes to jot down their answers to questions each of which included one of the taught expressions. They were instructed to answer each question by expressing the meaning of the taught items in their own words.

Psychological research shows that there are learning-test congruency effects. For example, DeKeyser and Sokalski (2001), in their critique of experiments by VanPatten and Cadierno (1993a, 1993b) which suggest that comprehension practice shows the same effect in production tests as production practice, point to learning-test congruency effects. They conclude that “[...] to comprehend input, language learners need practice comprehending input; to produce, they need practice producing.” (DeKeyser and Sokalski 2001: 86). In the area of L2 word learning, Barcroft (2002), for example, found higher retention scores for the L1 meaning equivalents of L2 words when the L2 words had been taught through a semantic rather than a structural elaboration task and the reverse result for the retention of L2 word forms. To control for possible congruency effects, I added one more condition (condition 3), where students were given production exercises (cloze tests) to do in the practice as well as in the test phase and compared this condition with condition 1, which had only comprehension tasks in the practice phase. Apart from the practice phase, condition 3 was identical to condition 1.

The experiment finished in all conditions with two unannounced post-tests. One was a cloze test, where parts of some of the target vocabulary were missing. After they had finished this task, participants were presented with a list of questions, each containing one of the idioms taught and requiring knowledge of the idiom’s meaning to be answered.

5.1. Participants

The participants were Greek students at the University of Athens studying various disciplines and attending classes for the CPE (Cambridge Proficiency in English) examination. Participants were randomly assigned to conditions. There were 10 participants in each condition. The experiment took place in classrooms at the Language Centre of the University of Athens, Greece.

5.2. Materials

VP idioms were the primary focus in the experiment. To construct the materials, I selected twelve VP idioms that are used to talk about two target domains, morality and comprehension. The idioms to talk about morality were related to MORAL IS UP and MORAL IS CLEAN: *take the high road*, *take the high ground*, *fall from grace*, *do the dirty on someone*, *be squeaky clean* and *dish the dirt on someone*. The idioms used to talk about comprehension were related to SEEING IS KNOWING and HOLDING IS CONTROLLING: *light dawns on someone*, *blind someone with science*, *be clear as mud*, *come to grips with something*, *get a handle on something* and *get the wrong end of the stick*. I consulted various idiom dictionaries for the selection, but the definitions chosen for use in the materials are usually those of the *Collins COBUILD Dictionary of Idioms* (Sinclair and Moon 1995) because they have the features that McKeown (1993)⁴ considers necessary for a definition to be understandable. The sentences used to illustrate the meanings of idioms were selected from the dictionaries – the criteria being lack of potentially unknown words, brevity, and use of context that helps to highlight the meaning of an idiom. If I could not find such examples, I created my own.

Although the primary aim was to compare the retention of VP idioms under different presentation methods, I also included other kinds of figurative expressions used in talk about morality and comprehension that instantiate the same CMs. The inclusion of those expressions was meant to assess whether the effectiveness of the two presentation methods under investigation would differ significantly for lexical units other than idioms. The additional expressions were six NP collocations (e.g. *dirty deed* and *high standards*), three phrasal verbs (e.g. *trace out something* and *catch on*) and two prepositional verbs (*see through someone/something*, *grope for something*)⁵.

In light of the results of three norming studies⁶ I conducted before this experiment with the participation of groups of students sharing the same profile as the participants in the actual experiment, I was confident that the lexical items I intended to use would be unknown to the participants and that the example sentences, the texts and the questions in the test phase would not contain unknown words that might hinder their comprehension. The first norming study (N 12) tested only for knowledge of VP idioms (it was initially conducted as a norming study for another experiment sharing the same idioms). The second norming study (N 12) was conducted to test whether the other figurative expressions I would use in the experiment were unknown to the participants as well; it also tested for the knowledge of the idioms used in the experiment, so that the power of the first norming study would be increased. The second norming study indicated that some of the other figurative expressions were known to participants, so I replaced these expressions with others and did a third norming study (N 10) to test for the knowledge of the substitute expressions, and to increase the power of the previous norming studies as far as their results for the idioms and the other figurative expressions were concerned.

The norming studies shared the same design and procedure. Each had three phases. In the first phase, participants had to state whether they knew the input items that I intended to use in the main study and to write the meanings of those which they knew. In the second phase, they were asked to go back to all of the target expressions and underline unknown words within them. In the third phase, they saw the example sentences I intended to use in the instruction phase along with the cloze texts and questions for the test phase, the idioms and other input items being in italics. They were asked to underline any unknown words, apart from the italicised expressions. They had to circle those words if they thought not knowing them rendered the sentences that included them difficult to comprehend.

The norming studies generated very few correct definitions from the students. They ranged between 3% (for the most difficult expression) and 9% (for the easiest expression). Therefore, I felt these target expressions were extremely unlikely to be known to the participants in the actual experiment, to which we turn in the following section.

5.3. Design and Procedure

Participants were given a booklet with the input materials. On the first page were the instructions, on the second the materials for the “morality” idioms, and on the last one the materials for the “comprehension” idioms. The

instructions were adapted from Boers (2000, Experiment 1). The instructions for conditions 1 and 3 were: “English has many figurative expressions to express actions, emotions, and situations. Below are some figurative expressions referring to specific meanings. Idioms expressing each meaning are grouped according to the metaphor they express.” The instructions for condition 2 were: “English has many idiomatic expressions to express actions, emotions, and situations. Below are some idioms referring to specific meanings, namely morality and comprehension. Idioms expressing each meaning are grouped according to the metaphor they express. Please try to guess the meaning of each idiom and write your guess in Greek in the space provided. If you know what an idiom means already, please write *KNOWN* first and then write its meaning.” Appendix 1 shows what the input sheets for the expressions about morality looked like in each of the three conditions.

Before the instruction phase started in each condition, I explained to the participants that in every language there is a tendency for words with a certain first literal meaning to acquire a certain metaphorical meaning and said that the figurative expressions they were going to see were grouped according to such metaphoric associations between concepts in the English language. I then explained to them what I meant by each title in the handouts. For example, for the title “Morality as being up”, I said that the concept of height is linked to that of morality in a way that words that have a first literal meaning relevant to height tend to acquire one relevant to morality.

In conditions 1 and 3, learners were instructed to take turns in reading the idioms and the definitions in the classroom. (I acknowledge that such reading-aloud activities are controversial, but these still happen to be common practice in many Greek schools, and I did not want to deviate too much from familiar classroom practice.) I explained any unknown words. Then I asked participants to spend seven minutes reading the materials on their own and try to memorise both the form and the meaning of the expressions. In condition 2, I asked participants to underline words that were unknown to them inside the idioms (e.g. many of them asked me what *dawns* means in *The light dawns on someone*). I explained these words and told the participants they could note these meanings down.

After the guessing phase, which was timed for 20 minutes, participants in condition 2 were given the input of condition 1 as feedback. They had ten minutes to first compare their interpretations with the definitions and then read through the whole feedback booklet and ask me about any unknown words in the examples. As mentioned above, in the other

conditions participants read the input aloud in the classroom as if in a lesson and asked me for the meaning of any words unknown to them.

I conducted the experiment with two participants at a time. Given that people differed in the time they needed to finish each task, the time of the instruction phase was not the same for all participants, but the instruction phase in condition 1 never exceeded 30 minutes and, as mentioned already, I asked students in condition 2 to finish the guessing task in 20 minutes. Therefore, the time invested was more or less the same across conditions, with the exception of three participants in condition 2. They could not finish the guessing task in time and so I allowed them five more minutes, maximum. However, this extra time investment did not result in superior results in the post-tests (see below).

As mentioned in the overview of the experiment, all conditions included a practice phase. Participants had to read texts which included, in italics, all the expressions they had been taught previously and then they had three minutes to jot down their answers to questions each of which included one of the target expressions. Then we discussed their answers and, when necessary, I corrected them afterwards.

In condition 3, the practice exercise consisted in reproducing the input lexis rather than displaying comprehension only. In other words, condition 3 stimulated a stronger focus on form than the other conditions. The exercise was a cloze test constructed out of the texts used in the practice phase in the other conditions. Participants had to fill in the missing parts of the figurative expressions. The materials used in the practice phase in each condition are presented in Appendix 2.

After the end of the practice phase, I collected those papers and gave participants two cloze tests to do in ten minutes. One cloze test omitted parts of the expressions related to the concept of morality; the other target parts related to the concept of comprehension. Each cloze test tested for four of the target idioms expressing each concept. The first cloze test also tested for four NP collocations and the second for two prepositional verbs and two phrasal verbs. After this task, I took those papers away as well and asked participants to answer the questions that required retention of the idioms' meanings, giving them 20 minutes for this task. For both tasks, I advised them to ask me about the meaning of any unknown words, other than the target expressions, that they might have (see Appendix 3).

5.4. Results

5.4.1. Cloze test

Table 2 shows the mean cloze-test scores for the VP idioms under the three conditions.

Table 2. Mean cloze-test scores and standard deviations (max. = 8)

Condition	Mean	SD
1	1.60	0.70
2	4.30	1.83
3	1.90	1.79

Condition 2 (i.e. the presentation method including a meaning-guessing task) generated the highest mean cloze-test score. To test for significance I used a Kruskal-Wallis test. Mean ranks were 11.50 for condition 1; 22.70 for condition 2; and 12.30 for condition 3. The value of the test was 10.43, which yields $p < .05$. To see which conditions in particular differed, I used the multiple-comparison test described in Siegel and Castellan (1988: 213-214). The multiple-comparison test shows that the comparison between conditions 1 and 2 is significant. Since condition 2 is the one with the higher mean ranks, this means that it led to significantly higher scores than condition 1. The slightly better cloze-test score under condition 3 (i.e. the more form-focused condition) in comparison with condition 1 (i.e. the more meaning-focused condition) was not significant.

As pointed out earlier, three participants in condition 2 needed five more minutes to finish the guessing task, but they nevertheless obtained slightly lower scores than their peers. If we were to exclude their scores, the mean score under condition 2 would be even higher (4.57) (and the Kruskal-Wallis test would still yield $p < .05$). The difference between conditions 1 and 2 would again be significant and that between conditions 1 and 3 would again be non-significant.

It must be acknowledged that the higher cloze-test scores under condition 2 might be due not to the factor of guessing *per se* but to other,

ancillary factors. For one thing, on finding out that their guesses were correct, students may have felt positive affect, which may have contributed to learning. For another, the guessing task may have stimulated participants to pay greater attention to the form (e.g. the precise lexical composition) of the idioms, as this may have been felt to contain valuable clues for the problem-solving (i.e. guessing) task.

To check for indirect support for these speculations, I did the following statistical analysis. I located the VP idioms for which each subject in condition 2 guessed the meaning correctly and then calculated the percentage of those whose form was also written correctly in the cloze test. I compared these percentages of cases of correct form retention with those in condition 1, where there was no guessing task. The rationale was that if there is a facilitating effect of correct guesses on form retention, there should be a higher mean of correct cloze-test responses for those whose meaning was inferred correctly than for those deriving from cases where there was no guessing (condition 1). I made three comparisons:

- a) between the number of correct cloze answers corresponding to correct guesses in condition 2 and the number of correct cloze answers in condition 1;
- b) between the number of correct cloze answers corresponding to the compilation of the correct and nearly correct guesses⁷ in condition 2 and the number of correct cloze answers in condition 1; and
- c) between the compiled number of the correct and nearly correct cloze answers for idioms that had been correctly guessed at in condition 2 and the compiled number of the correct and nearly correct cloze answers in condition 1.

In all three subsets of the data, condition 2 produced the highest scores. I applied a Mann-Whitney test to compare the students' performance under either condition for these subsets of data. Despite the small size of these subsets of data, the results were significant at $p < .005$ for comparisons (b) and (c), and close to significance ($p = .06$) for comparison (a).

I did the same calculations after collapsing among all kinds of figurative expressions taught. Again, condition 2 generated the best results. Comparisons (b) and (c) yielded significance at $p < .005$ again, and $p = .054$ (i.e. very close to significance) for comparison (a).

In sum, in all of the comparisons, more vocabulary was remembered correctly after the expressions' meaning had been guessed correctly (or nearly correctly) than when no guessing task had occurred. The lack of a significant difference between conditions both in the idiom data and after collapsing them with data for the other expressions when the correct answers for the items corresponding to correct guesses were compared to

those where no guessing had occurred, may be due to the low occurrence of totally correct guesses. The mean number of correctly guessed VP idioms is only 3.1 out of 12 (25.83%) (or 3.43 out of 12 [28.58%] when the data of the three apparently weaker participants who needed extra time on the guessing task are excluded). This is not very encouraging for the application of this method in contexts of learner autonomy (i.e. without explicit guidance and/or immediate feedback). Still, if one compiles these scores with those of the partly correct guesses, the mean score becomes 5.20 out of 12 (43.33 %) (or 5.57 out of 12 [46.42%] if we exclude the three weaker students).

5.4.2. Comprehension test

Participants' replies in the meaning retention test were categorised as:

- a) correct answers, that is, answers that were both to the point and unequivocally indicated that the participant remembered the meaning of the idiom;
- b) wrong answers, that is, answers that showed without any doubt that the participant did not remember the meaning of the idiom;
- c) partly correct answers, that is, answers which showed that the participant remembered only part of the idiom correctly;
- d) no answer (when the area beneath a question was left blank); and
- e) dubious answers, that is, answers which are difficult to categorise because the "Why/Why not?" part of the question was left unanswered.

Table 3 sums up the results.

Table 3. Answers falling under each category in the meaning retention task

Answer category	Condition		
	1	2	3
Correct	20 (50%)	25 (62.50%)	17 (42.50%)
Wrong	8 (20%)	8 (20%)	6 (15%)
Partly correct	2 (5%)	1 (2.50%)	0 (0%)
No answer	2 (5%)	0 (0%)	6 (15%)
Dubious	8 (20%)	5 (12.50%)	11 (27.50%)
Incomprehensible	0 (0%)	1 (2.50%)	0 (0%)

Because of the very few instantiations of the partly-correct and no-answer categories, some of the cells have expected counts under five. To be able to conduct a valid Chi-square test, I therefore compiled these categories with the wrong answers. Although condition 2 produced the highest mean score in the meaning-retention test, this better performance was not statistically significant (Pearson Chi-square = 3.65, $p > .05$).

6. Conclusions and pedagogical perspectives

This experiment led to significantly higher cloze-test scores for the instruction method combining the CM grouping with meaning guessing than for the presentation of VP idioms in CM groups only. This result indicates that the former method is a more promising method for L2 VP idiom instruction, at least as far as retention of form for purposes of (re)production is concerned. More generally, it adds support to the claim for higher effectiveness of assisting guessing by supplying information about the motivation of L2 figurative expressions relative to traditional instruction methods (see the summaries of Csábi [2004] and Boers, Eyckmans and Stengers [2007] above). Moreover, the relative success of the meaning-guessing condition in this experiment is an encouraging sign that VP idioms are not necessarily so long or so arbitrary as to obliterate evidence of the differing impact of whatever instructional methods are under investigation.

The superiority of the Hybrid method in condition 2 over the Metaphoric groups presentation method in condition 1 was enhanced by the observed connection between correct form retention responses in the cloze test and the correct idiom meaning guesses. This suggests that it is indeed the extra effort invested in the guessing task that led to superior retention.

Condition 3 was identical to condition 1 except that it contained a cloze test rather than a comprehension-questions test practice phase. Since no significant difference was found in form retention between conditions 1 and 3, this study does not show any convincing evidence of learning-test congruency effects.

As far as meaning retention is concerned, the meaning-guessing condition did not generate significantly higher scores than the other two conditions. This seems in accordance with experimental findings indicating that the retention of the meaning of new words is *not* higher when students are encouraged to guess at meaning (e.g. Mondria 2003; Mondria and Wit-de Boer 1991).

Although this experiment did not examine the effect of individual differences on the benefit of each instruction method, one finding indicates that for some people the guessing method (as in condition 2) may not be very effective. Three participants in that condition took more time to guess at the meaning of the input items, but still their scores were lower than the average score of the rest of the participants. On the other hand, we cannot tell how they would have performed under either of the other conditions. Still, given that participants had not been rated for individual differences before the experiment, we do not know what learner traits contribute (or inhibit) the effective use of the guessing method.

Given the connection between *correct* guessing and retention signalled by our data, another factor that may affect the success of the guessing method is how often participants correctly interpret an idiom. The likelihood of absolutely correct guesses seems fairly low, but if one considers also the results for *nearly* correct guesses, this method then seems to have some promise, although results are bound to depend on just how the method might actually be implemented in the classroom.

Although the results of this experiment point to the superior mnemonic effectiveness of adding a guided meaning-guessing activity to the CM-grouping method of idioms, this does not necessarily mean that these results will be replicated in all pedagogical contexts. Firstly, the participants in this study were not told they were going to be post-tested on their retention of the input lexis. In different conditions, anticipation of a test might stimulate students to apply their own mnemonics to the grouped input lexis, irrespective of any guessing stage. Secondly, in an adult-learner L2 classroom it can be assumed that learners attend classes on their own initiative and, hence, are sufficiently intrinsically motivated to try to remember as much L2 input as possible and at a relatively rapid pace. Such learners might benefit more from the instruction through the Grouping method than is indicated by the results of the above experiment simply because they are likely to invest more cognitive effort in processing the information than our participants in condition 1 may have done.

Another consideration for pedagogical practice is that complementing the CM-grouping presentation with a guessing stage seems rather time-consuming. Moreover, some learners may need a lot of guidance about the guessing task. Continual failure to guess the meaning of the idioms might even decrease the face-validity of the task and also increase negative affect among the students. Therefore, a more straightforward method such as the presentation of the target lexis in CM groups may prove more practical – especially if, for example, there is insufficient time for corrective feedback during or after the guessing stage.

If circumstances seem right for opting for the guessing method, then the question remains how it should be introduced to the learners. This experiment has shown that some learners find it difficult to understand and/or do the guessing task. We have seen that guessing the meaning of unknown words in context has been promoted as a vocabulary-learning strategy and that Irujo (1984) has proposed that guessing the meanings of idioms should be taught as a strategy so that learners will learn as many idioms as possible. However, encouraging learners to hypothesise about the CM basis of an idiom when they encounter it in context may be risky. The limited ability of at least some learners to discern the metaphorical basis of idioms poses threats to the success of meaning guessing if it is attempted in contexts of learner autonomy. Basically, guessing only on the basis of contextual clues leads to very few correct guesses and to limited meaning retention. Accordingly, the success rate of the guessing method seems to be very dependent on guidance (e.g. cues) and feedback.

Finally, it should be noted that this study examined only ways of *presenting* sets of idioms to L2 learners and thus only concerns ways of assisting *semantisation*, the initial word learning phase whereby the formal characteristics of a new word are matched with semantic content (Beheydt 1987). Therefore, any pedagogical applications of these findings should be seen as the *beginning* of a series of tasks with this vocabulary.

Appendix 1: Sample materials in the instruction phase

Input in conditions 1 and 3 for the expressions about Morality (first page)

Morality	
<i>Being moral as being up</i>	
1. Take the high road	To follow the course of action which is the most moral or most correct and which is least likely to harm or upset others e.g. The American President has become unpopular since the war. So, he has decided to <i>take the high road</i> about world peace.
2. A low-down trick	A very dishonest or unfair action (used informally to heavily criticize someone) e.g. Mafia members don't behave on the basis of any kind of moral principles. They will do any <i>low-down trick</i> to become richer and more powerful.
3. Fall from grace	To have made a mistake or done something wrong or immoral, and as a result to have lost one's power or influence and spoiled one's good reputation e.g. He <i>fell from grace</i> with his parents after his extra-marital affair.
4. High standards	Very moral principles e.g. When I was a child, my parents used to punish me frequently. They had very high standards!
5. Low behaviour	Not honest or fair behaviour e.g. Ruining her ex-husband's reputation by telling all those lies was low behaviour.
6. Take the high ground	To express your negative opinion of other people's, especially your rivals', actions as far as morality is concerned e.g. The opposition-party leader <i>took the high ground</i> on the Prime Minister's sex scandals in order to gain popularity.

Input in conditions 1 and 3 for the expressions about Morality (second page)

<i>Being moral as being clean</i>	
1. Be squeaky clean	To live a very moral life and not appear to have any vices (often used to suggest that this way of life is unnatural or uninteresting, or that someone is not as virtuous as they seem). e.g. In the 1980s this young singer made a great success because she <i>was</i> pretty and <i>squeaky clean</i> . That's why people were shocked when she turned into a sex bomb in the 1990s.
2. Do the dirty on someone	To behave unfairly or very badly towards someone, often without them knowing e.g. His boss hated Mark so much that he just waited for Mark to make a tiny mistake and then <i>did the dirty</i> on him and fired him.
3. Clean fun	Entertainment that is not sexually immoral or offensive e.g. Today, most people think that flirting is just a harmless way of communication and that it's <i>clean fun</i> .
4. Dish the dirt on someone	To gossip about someone, especially by saying things that may embarrass or upset that person, or damage their reputation. e.g. In his autobiography, the singer <i>dishes the dirt on his band</i> . Many people were surprised because they thought he was friends with the other band members.
5. Dirty deed	An action/lie that is unfair, dishonest, or unkind e.g. Ladies and gentlemen of the jury, it is not difficult to decide whether this man is guilty or not! Robbing an old lady is certainly a <i>dirty deed</i> .
6. Clean fight	Honest or fair competition e.g. He called for a <i>clean fight</i> in the election. He wanted an end to 'negative campaigning'.

Guessing task in condition 2 (expressions about Morality)

Morality

Being moral as being up

1. Take the high road:
2. A low-down trick:
3. Fall from grace:
4. High standards:
5. Low behaviour:
6. Take the high ground:

Being moral as being clean

1. Be squeaky clean:
2. Do the dirty on someone:
3. Clean fun:
4. Dish the dirt on someone:
5. Dirty deed:
6. Clean fight:

Appendix 2: Materials in the practice phase⁸

Materials for the practice phase in conditions 1 and 2 (first page)

Morality

Quentin Tarantino's films have been criticized as advertising immoral behaviour by some film critics and members of the audience. These people think that in Tarantino's films violence is presented as *clean fun*, because the gangsters in them do not have any remorse after killing someone or after doing some other kind of *dirty deed*, such as *dishing the dirt on* a fellow gangster.

Fans of Tarantino's films, however, say that the violence in his films is like that in comic books: It does not affect people's opinions about the world or their behaviour, because the combat scenes are exaggerated and therefore unrealistic. For Tarantino's supporters, people who criticize Tarantino are exaggerating and just *taking the high ground* to show that they *are squeaky clean*. They find these films very entertaining, so, no matter how violent they are, he will never *fall from grace* for these people. They also think that Tarantino does a *clean fight* with other film directors, because his films have gained the public's love not because of their violence but because of their clever dialogue and great story-telling.

In a way, I agree with both opinions. Although Tarantino's heroes certainly don't have *high standards*, they are not just killing machines. They do *low-down tricks* and have other kinds of *low behaviour*, but they do not always *do the dirty on* people. For example, in 'Pulp Fiction', the character played by Samuel Jackson decides to stop being a gangster in the end of the story and *take the high road*.

Please answer the questions below in a way that shows that you know the meaning of the figurative expressions in the questions.

1. Why does the text call killing a "dirty deed"?

.....

2. Why do Tarantino's supporters think that people who criticize Tarantino are taking the high ground to show that they are squeaky clean?

.....

3. Could you mention examples of low behaviour in Tarantino's films?

.....

Materials for the practice phase in conditions 1 and 2 (second page)

Comprehension

What can go wrong when you don't make a plan before you start writing an essay? First, writing might be much more tiring and time-consuming when you have not *come to grips with* what you want to write about. If you don't write a plan, every now and then you'll be *groping for* what to write next. Second, because you are paying attention mainly to the content of your essay, you may neglect its form. In other words, a reader may not be able to *take in* all the meaning of your essay because of long sentences and complex expressions you may have used. If a teacher just *looks over* your essay, he or she may be unable to *see through* the text to the main points of the essay. Often a poorly organized essay is *clear as mud* to the teacher. If, however, you have made a plan of the essay you will be able to *trace out* complicated ideas.

Special problems occur when you write an essay as part of an examination. If you have not made a plan, at some point during writing, you may not know which other argument to make in the essay. So you may just write nonsense and try to *blind people with science* using formal and complex expressions. Moreover, if you just keep on writing without having a plan in mind, maybe you will need to change the order of the paragraphs or rewrite whole parts of the essay. If *the light* only *dawns on* you later on in the writing process, it may be too late to make these changes. Finally, you may *get the wrong end of the stick* with the essay topic, because you didn't examine it properly before writing. It may be too late to start writing all over again, when you *have got a handle on* the topic.

Please answer the questions below in a way that shows that you know the meaning of the figurative expressions in the questions.

1. The text mentions many problems a student may face when he or she writes an essay without making a plan. Write the problem which is mentioned first in your own words, that is, without using the figurative expressions in the text.
.....
2. What does a teacher do if he looks over an essay?
.....
3. According to the text above, why do some students try to blind the teacher with science in an essay?
.....

Excerpts of the practice phase exercises in condition 3

Quentin Tarantino's films have been criticized as advertising immoral behaviour by some film critics and members of the audience. These people think that in Tarantino's films violence is presented as *clean fun*, because the gangsters in them do not have any remorse after killing someone or after doing some other kind of _____ *deed*, such as *dishing the dirt on* a fellow gangster.

[...] For Tarantino's supporters, people who criticize Tarantino are exaggerating and just *taking* _____ to show that they *are squeaky clean*. [...]

In a way, I agree with both opinions. Although Tarantino's heroes certainly don't have *high standards*, they are not just killing machines. They do *low-down tricks* and have other kinds of *low behaviour*, but they do not always *do* _____ *on* people. For example, in 'Pulp Fiction', the character played by Samuel Jackson decides to stop being a gangster in the end of the story and *take the high road*.

What can go wrong when you don't make a plan before you start writing an essay? First, writing might be much more tiring and time-consuming when you have not *come to grips with* what you want to write about. [...] If a teacher just looks over your essay, he or she may be unable to _____ *through* the text to the main points of the essay. Often a poorly organized essay is *clear* _____ to the teacher. If, however, you have made a plan of the essay you will be able to *trace out* complicated ideas.

Special problems occur when you write an essay as part of an examination. [...] If *the* _____ *only* _____ *on* you later on in the writing process, it may be too late to make these changes. Finally, you may *get the wrong end of the stick* with the essay topic, because you didn't examine it properly before writing. It may be too late to start writing all over again, when you *have got a handle on* the topic.

Appendix 3: Materials in the test phase

Fill in the blanks to complete the figurative expressions in the text below. The answers can be anything between one and four words.

Andrew's family was one of the richest families in the village. Everyone admired and respected them. Andrew was also the mayor of the village. So, he thought he had to set a good example for the rest of the people. He and his wife often went to church and gave money to those in need. Everyone in the village thought they were squeaky _____ (clean). However, after many years of happily married life, Andrew cheated on his wife. He did it mainly because he was bored. He was still in love with her. So, he was feeling awful for having done this. In the end, he could not go on hiding this secret from his wife and told her. His wife was very angry with him and couldn't forgive him because of her _____ (high) standards. It had never occurred to her that he would do _____ (the dirty) on her! She didn't speak to him for a week to punish him for his _____ (dirty) deed and took _____ (the high ground). Andrew tried to be as nice to her as possible. He bought her flowers and asked her to forgive him for his _____ (low) behaviour. He promised that he would have only _____ (clean) fun in the future but there was nothing he could do to change her mind. He had _____ (fallen) from grace.

At school, Mary was one of the best students in the class but found mathematics very difficult. Even when the teacher _____ (traced) out the theory to the students, everything he said was clear _____ (as mud) to her. She was angry with herself, because her worst enemy, John, was incredibly good at maths. What annoyed Mary most was the teacher's impression of her. Although the teacher told her parents that she was a good student, Mary could _____ (see) through him. She could feel that the teacher admired John and thought that Mary was stupid. The teacher usually gave the class a problem to solve after he had explained the theory. John immediately came _____ (to grips) with the problem, but Mary needed a lot of time to catch _____ (on). First, she got the wrong _____ (end of the stick). When she realised her mistake, she could not _____ (take) it in and wasted a lot of time panicking. By the time the light _____ (had dawned on) her, John raised his hand and told the teacher that he had solved the problem.

Please answer the questions below.

1. Would you be angry if a friend of yours dished the dirt on you? Why/Why not?
2. Do you think you would grope for the solution of a problem in astrophysics? Why/why not?
3. Do you usually take the high road or not? Why/Why not?
4. Is it good to get the handle on a problem fast or not? Why?
5. Mention three actions that you would call "low-down tricks". Explain why.
6. Is it right or wrong to blind people with science? Why?
7. Describe what a "clean fight" means to you.

Acknowledgements

This chapter is based on my PhD research. I would, therefore, like to thank my supervising committee, Richard Breheny, Gillian Brown, and John Williams for their valuable feedback and advice.

Notes

1. A categorisation is called semantic if the words share a superordinate term (e.g. *pencil*, *envelope* and *pen* are all kinds of *stationery*). I compare the results of Li's experiment with those of Boers (2000) and Luciani de las Mercedes (2001) because, in essence, semantic categorisation does not differ from functional. For example, in Boers's experiment, the group title *to describe acute and sudden anger* could be rephrased as *acute and sudden anger*, that is, a kind of anger, the superordinate semantic category for these idioms.
2. Following Lakoff (1987: 447), Boers and Demecheleer (2001: 255) I define imageable idioms as "idioms that have associated conventional images".
3. Skoufaki (2006) reports the pilot study of that experiment. It is summarised in Boers and Lindstromberg (2006) and in the introduction to the present volume.
4. In cases where the definition was disjointed or focused attention on one of its parts and so caused the others to be neglected, I modified the definition so that it would not be misunderstood. My translations of the definitions into Greek were proofread by a native Greek speaker who does not know English. I asked her to judge whether my translations constituted natural-sounding Greek and discussed with her how I should improve them where she found fault.
5. Initially, I intended to use three phrasal verbs (*make out something*, *look over something*, *take in something*) and three prepositional verbs (*catch on*, *see through someone/something*, *grope for something*). However, on the basis of the results of the second norming study, I excluded from the analysis *make out something* and *look over something*. Unfortunately, I could find only one alternative for the phrasal verbs that were deleted, so the final materials had one expression less in the group "Understanding something as seeing it".
6. For a detailed description of the norming studies see Skoufaki (2006).
7. For example, a participant had defined *the light dawns on someone* as "someone understands something gradually". Although this idiom actually means "understand suddenly something I should have understood earlier", I considered this answer as nearly correct because it contained at least the "understand" component of the definition. A wrong definition was, for example, "People start to suspect someone".
8. I acknowledge that these texts are so packed with idioms that they may come across as contrived and unnatural. They served the purpose of the research experiment, but I make no claims about their suitability as pedagogical material.

References

- Barcroft, Joe
2002 Semantic and structural elaboration in L2 lexical acquisition. *Language Learning* 52: 323-363.
- Beheydt, Ludo
1987 The semantisation of vocabulary in foreign language learning. *System* 15: 55-67.
- Boers, Frank
2000 Metaphor awareness and vocabulary retention. *Applied Linguistics* 21: 553-571.
2001 Remembering figurative idioms by hypothesising about their origins. *Prospect* 16: 35-43.
- Boers, Frank, and Murielle Demecheleer
2001 Measuring the impact of cross-cultural differences on learners' comprehension of imageable idioms. *ELT Journal* 55: 255-262.
- Boers, Frank, Murielle Demecheleer, and June Eyckmans
2004a Etymological elaboration as a strategy for learning idioms. In *Vocabulary in a Second Language: Selection, acquisition and testing*, Paul Bogaards and Batia Laufer (eds.), 53-78. Amsterdam/Philadelphia: John Benjamins.
2004b Cultural variation as a variable in comprehending and remembering figurative idioms. *European Journal of English Studies* 8: 375-388.
- Boers, Frank, June Eyckmans, and H el ene Stengers
2007 Presenting figurative idioms with a touch of etymology: More than mere mnemonics? *Language Teaching Research* 11: 43-62.
- Boers, Frank, and Seth Lindstromberg
2006 Cognitive linguistic applications in second or foreign language instruction: Rationale, proposals, and evaluation. In *Cognitive Linguistics: Current Applications and Future Perspectives*, Gitte Kristiansen, Michel Achard, Ren e Dirven and Francisco J. Ruiz de Mendoza Ib a nez (eds.), 305-358. (Applications of Cognitive Linguistics 1). Berlin/New York: Mouton de Gruyter.
- Bogaards, Paul, and Batia Laufer (eds.)
2004 *Vocabulary in a Second Language: Selection, acquisition and testing*. Amsterdam/Philadelphia: John Benjamins.
- Cohen, Andrew D.
1987 The use of verbal and imagery mnemonics in second-language vocabulary learning. *Studies in Second Language Acquisition* 9: 43-62.

- Csábi, Szilvia
 2004 A cognitive linguistic view of polysemy in English and its implications for teaching. In *Cognitive Linguistics, Second Language Acquisition, and Foreign Language Teaching*, Michel Achard and Susanne Niemeier (eds.), 233-256. (Studies on Language Acquisition 18.) Berlin/New York: Mouton de Gruyter.
- DeKeyser, Robert M., and Karl J. Sokalski
 2001 The differential role of comprehension and production practice. *Language Learning* 51: 81-112.
- Herman, Patricia A., Richard C. Anderson, P. David Pearson, and William E. Nagy
 1985 Incidental acquisition of word meanings from expositions that systematically vary text features (Tech. Report No. 364). Urbana: University of Illinois Center for the Study of Reading.
- Higa, Masanori
 1963 Interference effects of intralist word relationships in verbal learning. *Journal of Verbal Learning and Verbal Behavior* 2: 170-175.
- Hüllen, Werner, and Rainer Schulze (eds.)
 1988 *Understanding the Lexicon: Meaning, sense and world knowledge in lexical semantics*. Tübingen: Max Niemeyer.
- Hulstijn, Jan
 2003 Incidental and Intentional Learning. In *The Handbook of Second Language Acquisition*, Catherine J. Doughty, and Michael H. Long (eds.), 349-381. Oxford: Blackwell.
- Irujo, Susanne
 1984 *The Effects of Transfer on the Acquisition of Idioms in a Second Language*. Doctor of Education thesis. Boston University. UMI.
 1993 Steering clear: avoidance in the production of idioms. *IRAL* 31: 205-219.
- Kövecses, Zoltán, and Peter Szabó
 1996 Idioms: A view from Cognitive Semantics. *Applied Linguistics* 17: 326-355.
- Lakoff, George
 1987 *Women, Fire, and Dangerous things: What categories reveal about the mind*. Chicago: University of Chicago Press.
- Lennon, Paul
 1998 Approaches to the teaching of idiomatic language. *IRAL* 36 : 11-30.
- Li, F. Thomas
 2002 The Acquisition of Metaphorical Expressions, Idioms and Proverbs by Chinese Learners of English: A conceptual metaphor and image schema approach. Ph. D. diss., Chinese University of Hong Kong.

- Luciani de las Mercedes, María
2001 Metáfora y Adquisición de una Segunda Lengua. Ph. D. diss., Universidad Nacional del Litoral, Santa Fe, Argentina.
- McKeown, Margaret G.
1993 Creating effective definitions for young word learners. *Reading Research Quarterly* 28: 17-31.
- Mondria, Jan-Arjen
2003 The effects of inferring, verifying, and memorizing on the retention of L2 word meanings. An experimental comparison of the "Meaning-Inferred Method" and the "Meaning-Given Method". *Studies in Second Language Acquisition* 25: 473-499.
- Mondria, Jan-Arjen, and Marijke Wit-De Boer
1991 The effects of contextual richness on the guessability and retention of words in a foreign language. *Applied Linguistics* 12: 249-267.
- Moon, Rosamund
1997 Vocabulary connections: multi-word items in English. In *Vocabulary: Description, Acquisition, and Pedagogy*, Norbert Schmitt and Michael McCarthy (eds.), 40-63. Cambridge: Cambridge University Press.
- Nagy, William E., Patricia A. Herman, and Richard C. Anderson
1985a Learning words from context. *Reading Research Quarterly* 20: 233-253.
1985b Learning word meanings from context: How broadly generalisable? (Tech. Report No, 347). Urbana: University of Illinois Center for the Study of Reading.
- Schmitt, Norbert
1997 Vocabulary learning strategies. In *Vocabulary: Description, Acquisition, and Pedagogy*, Norbert Schmitt and Michael McCarthy (eds.), 199-227. Cambridge: Cambridge University Press.
- Siegel, Sidney, and N. John Castellan
1988 *Nonparametric Statistics for the Behavioral Sciences*. New York and London: McGraw-Hill.
- Sinclair, John, and Rosamund Moon (eds.)
1995 *Collins COBUILD dictionary of idioms*. London: HarperCollins.
- Skoufaki, Sophia
2005 Use of conceptual metaphors: A strategy for the guessing of an idiom's meaning? In *Selected Papers on Theoretical and Applied Linguistics from the 16th International Symposium (April 11-13, 2003)*, M. Mattheoudakis and A. Psaltou-Joycey (eds.), 542-556. Thessaloniki: Aristotle University of Thessaloniki.
2006 Investigating Idiom Instruction Methods. Ph. D. diss., University of Cambridge, UK.

Sornig, Karl

- 1988 Idioms in language teaching. In *Understanding the Lexicon: Meaning, Sense and World knowledge in Lexical Semantics*, Werner Hüllen, and Rainer Schulze (eds.), 280-290. Tübingen: Max Niemeyer.

Tinkham, Thomas

- 1993 The effect of semantic clustering on the learning of second language vocabulary. *System* 21: 371-380.
- 1997 The effects of semantic and thematic clustering on the learning of second language vocabulary. *Second Language Research* 13: 138-163.

VanPatten, Bill, and Teresa Cadierno

- 1993a Input processing and second-language acquisition – A Role for Instruction. *Modern Language Journal* 77: 45-57.
- 1993b Explicit instruction and input processing. *Studies in Second Language Acquisition* 15: 225-243.

Waring, Robert

- 1997 The negative effects of learning words in semantic sets: a replication. *System* 25: 261-274.
-

