The figurative and polysemous nature of collocations and their place in ELT

Marijana Macis and Norbert Schmitt

There is little guidance on how to teach different meaning senses of collocations as most pedagogical materials treat collocations as word partners which co-occur together. However, if we consider meaning, collocations fall into three categories. Literal collocations are combinations where the literal meanings of the words are simply added together. Figurative collocations have idiomatic meanings which are not derivable from the component words. Duplex collocations are polysemous, having both literal and figurative meanings. This exploratory study analysed 54 collocations and found that even though the majority of the collocations appeared to be literal, a substantial percentage had both literal and figurative meanings, and relatively few seemed to be solely figurative. We discuss the teaching implications of this, depending on whether the most important collocational characteristic is a pattern of co-occurrence or of meaning. Overall, we argue that considering meaning can bring useful insights to the nature of collocations and how to teach them.

Introduction

There has been little attention to date—especially in ELT textbooks—given to the semantic qualities of collocations and the subsequent pedagogical implications of this. As a result, many teachers simply do not look at collocations as units for which meanings can be more than just literal. Instead, most definitions of collocation that teachers are familiar with focus on the pattern of co-occurrence in language, rather than the meanings that they may convey. This is exemplified by Sinclair's (1991: 170) early definition, 'Collocation is the occurrence of two or more words within a short space of each other in a text', which is reflected in typical descriptions of collocation in recent textbooks:

- 'Collocation means a natural combination of words; it refers to the way English words are closely associated with each other'. *English Collocations in Use: Intermediate* (McCarthy and O'Dell 2005: 4).
- 'It is important to know which words collocate (commonly go together)'. *Cambridge English: Objective Advanced* (O'Dell and Broadhead 2014: 13).

Advance Access publication May 17, 2016

This approach is essentially about the co-occurrence of words, i.e. that words 'partner' together. But can the consideration of their semantic properties add to our understanding of collocations? We believe that teachers should focus on more than just the literal partnering of words as, from a meaning-focused perspective, there seem to be three categories of collocation.

Literal collocations

These are combinations where the literal meanings of the words are simply added together ('powerful computer' = a computer with powerful processing abilities). The consideration of the collocations' meanings by learners is not such an issue here, as they merely need to know the meaning(s) of the component words, and then string them together in a 'Meaning A + Meaning B' manner. The real challenge is knowing which words pattern together for that particular meaning (for example not *'strong computer' or *'tough computer').

This aspect of patterning has been the focus of most collocation pedagogy and discussion to date.

Figurative collocations

Figurative collocations[1] have figurative meanings which are not derivable from their component words. For example, the collocation 'hot ticket' does not literally mean 'a ticket with a high temperature', but rather one which is 'highly sought after'. Thus, 'hot ticket' behaves like a 'chunk' rather than two words which simply partner together. However, we seldom see Figurative collocations discussed in ELT textbooks or materials.[2]

Duplex collocations

Duplex collocations[1] are polysemous,[3] having both literal and figurative meanings. For example, 'top drawer' has an A + B literal meaning ('the uppermost drawer in a cabinet'). However, it also has a separate meaning that is clearly figurative ('something that is best of its class'): He bought a *top drawer car*.

The above distinctions matter because research suggests that figurative language is stored, processed, and learnt differently than language where meaning is constructed from individual components (Siyanova-Chanturia and Martinez 2015), and so may require different teaching approaches. We know that Figurative and Duplex collocations exist from examples like the above; however, are these examples just atypical cases, or are they relatively widespread? We could not find answers to this question in the literature. Thus, our study analysed a set of collocations to obtain an initial indication of the meaning-based nature of collocations in order to provide guidance for ELT practitioners on how to integrate these into ELT syllabi.

The study
The corpus

The corpus selected for this study was the Corpus of Contemporary American English (COCA) (Davies 2008). The reason for choosing the COCA is because it is a large, freely available, and balanced corpus of American English (520-million words spread evenly across five different text types: spoken, fiction, popular magazines, newspapers, and academic texts), and as such, both teachers and textbook writers can access it.

Collocation sampling procedure

Although most would agree with Sinclair's (op.cit.) general definition of collocation as being words that co-occur together, there are a range of criteria for identifying individual cases of collocation (for example Shin

and Nation 2008). Different analytical techniques identify different kinds of collocation, so it is not surprising that there is no universal list of collocations for teachers to use (even though there are some useful lists related to formulaic language in general, such as Martinez and Schmitt's PHRASE List (2012)). Given the lack of a standard collocation list, we needed to create a sample of collocations to explore in our study.

We decided to use frequency as a criterion for choosing the target collocations because learners typically acquire higher frequency words before lower frequency words. The first 2,000 word families have generally been accepted as high-frequency vocabulary, and these words are crucial to use English at a basic, yet useful level. These words also cover around 90 per cent of the total number of words in both written and spoken discourse (Nation 2006). Our rationale was that if the most frequent words are so important, then surely the collocations around these words must be important as well. So, in order to investigate collocations which would be useful for learners (but also not to replicate Shin and Nation's (op.cit.)), we focused on collocations forming around words from the 2,000 frequency level. We created a stratified sample of ten nouns (every 100th noun) from this level for a detailed analysis: 'block', 'conflict', 'drawer', 'furniture', 'knee', 'nose', 'pupil', 'shame', 'ticket', and 'zero'. As our investigation required a time-consuming manual analysis, it was impractical to include all possible collocations for each noun. We thus limited ourselves to the three most common verb + noun collocations and the three most common adjective + noun collocations, since these seem to be the most challenging combinations (Grant 2005) and are therefore the most commonly researched. It is also useful for teachers to be aware of which collocation types, and which collocations are most complex, in order that they can be targeted for instruction.

These combinations were subject to several additional criteria. First, because it is unclear how well learners recognize collocations that are non-adjacent ('confirm findings': *confirm* the very surprising *findings*), we focused mainly on adjacent collocations. Second, we wanted to analyse the collocations for meaning, and so did not include delexicalized verbs in the verb + noun sample (for example 'to have a shower'), as their meaning depends on the meaning of the word they collocate with. Finally, Mutual Information (MI) is a statistic considered to be an important indicator of the strength of the link between two words, and so we only selected collocations with an MI score of 3 or above.[4]

For each target collocation, we aimed at analysing 100 concordance lines at a time (see below). In some cases there were fewer lines, and those collocations with fewer than five lines were excluded from the study. This resulted in the six verb collocations for 'drawer' and 'pupil' being omitted. Our final set of 54 collocations displayed a range of frequencies (5–844) and MI (3.00–15.46) in the COCA corpus (see Table 1 for examples). Our sample is admittedly modest, but should be sufficient to give an initial indication of the extent of Literal, Figurative, and Duplex collocations that learners are likely to come across.

Procedure

The first author analysed 100 concordance lines for each target collocation, or as many lines as were available. Each line was tagged for

Marijana Macis and Norbert Schmitt

TABLE 1
The most common
adjective + noun and
verb + noun collocations

Sampled nouns	Adjective + noun collocations	Verb + noun collocations
Conflict	Armed conflict	To avoid conflict
	Ethnic conflict	To resolve conflict
	Potential conflict	To manage conflict
Furniture	New furniture	To buy furniture
	Antique furniture	To sell furniture
	Old furniture	To move furniture
Ticket	One-way ticket	To buy ticket
	Speeding ticket	To purchase ticket
	Hot ticket	To afford ticket

literal meaning or figurative meaning. For example, it was clear that the first of the following corpus extracts (Extract 1) was a literal meaning (= a ticket in one direction), while the second (Extract 2) was a figurative one (= an irrevocable action).

> Extract 1
>
> But what really worried him was the suspiciously low ***one-way ticket***— priced at $92.60—on an airline he had barely heard of.
>
> Extract 2
>
> Even if your locale has frigid winters or simmering summers, the weather needn't be a ***one-way ticket*** to slothdom.

In order to assure accuracy of the categorization, around 20 per cent of the target collocations were given to a second rater who was asked to rate the collocations without seeing the first rater's judgements. The two raters reached a high interrater reliability of nearly 96 per cent agreement.

Results

The results of the analysis show that the majority of our collocations, both verb + noun and adjective + noun types, are literal collocations. Out of the 54 target collocations, 42 were literal collocations (78 per cent). For example, a 'left knee' literally means 'the left joint between the thigh and the lower leg' (see Extract 3).

> Extract 3
>
> ... help him recover more quickly and effectively from arthroscopic surgery performed Sunday on his ***left knee***.

Based on our modest data set, it seems that the majority of collocations fit the typical 'word partner' view, with two literal meanings that can be essentially added together (A + B).

The results also indicate that a very small number (two collocations: 4 per cent) of our sample collocations had figurative meaning only. One must be cautious in interpreting a category with only two collocations, but our results do suggest that collocations do not generally seem to be completely figurative, although in a few cases they are, for example a 'big ticket' relates to 'something being expensive; a major expense' rather than a physically large ticket (Extract 4).

The figurative and polysemous nature of collocations

Extract 4

We're seeing the people a lot more confident making a big purchase so they're not as afraid to come in and buy something out, a **big ticket** like this ...

While few collocations seem to be wholly idiomatic, ten collocations (18 per cent) had the characteristics of Duplex collocations, i.e. they can have both literal and figurative meanings. For example, the collocation 'big nose'[5] literally means 'a nose that is big in size', but it also refers to a 'strong aroma'. Extracts 5 and 6 show the different meaning senses:

Extract 5

She was young. Maybe twenty-five. A **big nose** but otherwise pretty. (Literal)

Extract 6

... The color is wonderful—the lightest clear amber—with streams of tiny bubbles rising to the surface. The taste is dry and just pleasantly alcoholic. The fragrance is what I think winemakers would describe as a '**big nose**.' It is an unalloyed delight. (Figurative)

Overall, the two types of collocation with non-literal meanings (Figurative and Duplex collocations) combined to make up 22 per cent of the total sample. This means that over one-fifth of the sample collocations had one or more figurative meaning senses. Thus, while the majority of collocations conformed to a literal A + B pattern, a substantial percentage did not (Figure 1). It is also interesting that even though Figurative collocations appear to be rare, Duplex collocations seem to be relatively common.

Discussion

Teachers have been encouraged to teach collocations, and the most widely used activity has been noticing how two known words typically combine together (Lewis 1993). In his *Lexical Approach*, Lewis (ibid.) suggests that learners should be encouraged to notice recurring lexical chunks in the L2 input they receive. But what about collocations with figurative meanings? Most textbooks conceptualize collocations as co-occurring 'word partners' (i.e. our Literal collocations) where learners need to deal with the literal
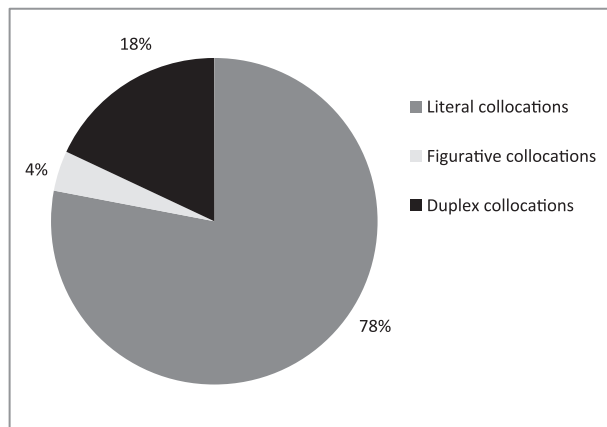

FIGURE 1
Percentages of Literal collocations, Figurative collocations, and Duplex collocations

Marijana Macis and Norbert Schmitt

meanings of the words comprising the collocation, and merely recognize the affinity between these words. If most collocations are actually Literal collocations, then the focus on co-occurrence without a further semantic analysis is a reasonable approach. However, if a substantial percentage of collocations are Figurative or Duplex, then some additional attention to their meaning seems necessary. As for collocations, to the best of our knowledge, our admittedly modest study is the first to suggest the scope of occurrence of each collocation type.

According to the results of our study, therefore, by far the most numerous types were Literal collocations, making up about three-quarters of our small sample. This would suggest that the vast majority of collocations can be taught as literal A + B combinations. There were very few Figurative collocations and this finding is similar to Grant's (op. cit.) study in which she looked at the extent of 'core idioms' in various sources and found that only 103 idioms were truly 'core'. In contrast, Duplex collocations, with both literal and figurative meanings, made up nearly one-fifth of our target collocations. So, is this enough to necessitate teachers considering figurative meanings when they teach collocations? We would argue that it is: the percentage of Figurative and Duplex collocations appears to be substantial, and cannot be written off and ignored as peripheral behaviour.

An important property of formulaic sequences is their figurativeness. The three categories of collocations clearly differ in terms of their meaning characteristics, and this implies that they might need to be addressed differently in pedagogy. One way of thinking about this is to consider whether receptive or productive knowledge of a collocation is required. Let us first take the case of receptive knowledge where the collocation needs to be recognized and understood in reading or listening. There is probably little need to worry about Literal collocations, or the literal meaning senses of Duplex collocations, as learners should be able to understand the meaning(s) of the two component words and how they come together. We might speculate that knowing the collocations could add to reading fluency, but probably would not aid comprehension to any great degree. However, for Figurative collocations and the figurative meanings of Duplex collocations, the meaning is opaque and unlikely to be understood.

If productive knowledge is required, then collocation knowledge becomes more crucial, as learners are required to produce language of their own volition. While meaning can be expressed in many ways, it is well known that skilful use of collocation is one of the factors that distinguishes proficient language use. So, knowledge of Literal collocations and the literal meaning senses of Duplex collocations is especially important for appropriacy. However, for the expression of certain meanings, knowledge of Figurative collocations and the figurative meanings of Duplex collocations is highly advantageous. This is because they express idiosyncratic and opaque meanings that are tied to specific situations. For example, the collocation 'big nose' does not only figuratively mean a 'strong aroma', but it also belongs almost exclusively to the world of wine tasting.

This discussion leads to the conclusion that collocations need to be approached in two different ways depending on their category. Literal

collocations and the literal meaning senses of Duplex collocations require knowledge of their co-occurrence patterns. Most collocation exercises and activities to date have attempted to enhance learner awareness of this patterning, and so this need is already beginning to be addressed. On the other hand, Figurative collocations and the figurative meanings of Duplex collocations require knowledge of their idiosyncratic meanings. This leads to the question of how these might best be taught. Martinez (2013) proposed a framework (the Frequency-Transparency Framework or FTF) that would help teachers choose which multi-word expressions should be targeted for explicit instruction. According to this framework, if there are two highly frequent phrases of similar frequency but different degrees of semantic transparency, the more opaque expression should be prioritized. If two phrases are infrequent but one of them is transparent and the other one is opaque, the opaque one should be explicitly taught as such phrases may be a barrier to successful comprehension. Martinez (ibid.) emphasizes, though, that this framework operates successfully only if frequency is measured in a way which reflects the needs of the learners. For example, if learners are enrolled on a flying course, then the vocabulary related to that specific area should be regarded as 'frequent' for that particular group of students.

Furthermore, there is some limited evidence that teaching verb-noun collocations holistically has advantages over teaching them with matching exercises (Boers, Demelcheleer, Coxhead, and Webb 2014). This is mainly because in a holistic approach, learners are less prone to make incorrect verb-noun associations (as is the case with 'underline the correct verb' exercises).

These approaches have been discussed in this article to help ELT practitioners visualize what seems to be an effective way of teaching figurative meanings. Instead of breaking down collocations into component parts, a more sensible approach is to present them as intact wholes since there seems to be more to learn. How to present the literal and figurative meanings of Duplex collocations, and whether they should be taught in conjunction with each other, are also interesting questions. Boers' (2000) work on metaphors suggests that there might be value in teaching figurative collocation meanings through some form of linguistic analysis and analogy because enhancing learners' metaphor awareness may facilitate retention of unknown figurative language. For example, the 'excellent quality' meaning of 'top drawer' might be explained by an analogy that things in the top drawer of a dresser are usually good, and/or the metaphorical theme that HIGHER→BETTER. Therefore, activities in which learners are asked to group figurative expressions under general metaphoric themes (for example 'to tighten the screws' under MACHINERY) or explain the figurative meaning of a collocation by using metaphors (for example 'to hit the roof' may be explained in terms of ANGER) may be a good start. However, teachers must be aware that some collocations are opaque and hence not easy to imagine. In this case, a more direct, explicit approach seems a sensible option.

Furthermore, regardless of their type (either Duplex or Figurative), some collocations, when used figuratively, form part of bigger lexical chunks or act as modifiers. For example, in the majority of cases, the collocation 'big ticket'

Marijana Macis and Norbert Schmitt

is used with the noun 'item' (for example 'The report says that NASA's new budget could hurt the future of space science by cutting spending on small programs to pay for big ticket items'). Despite this, the figurative meaning of 'big ticket' and other similar collocations as chunks of language does not change, no matter whether or not they are part of bigger phrases. Of course, this can be seen only if the collocations are embedded in the context. The upshot is that teachers should not become too 'narrowly focused' on the two-word limit. They should be aware that two-word collocations might not necessarily be the end of the phrase, and that the meaning of collocations can sometimes be connected to the meaning of a longer phrase.

Unfortunately, it is unclear whether current textbooks have a principled approach to teaching different types of collocations; hence teachers will still need to think about their own teaching approaches. The first step is realizing that collocations are not all the same in terms of meaning. In addition to the possible teaching approaches discussed above, some additional guidelines based on the results of this study may prove useful for teachers.

1 Make learners aware of the existence of the three types of collocations.
2 Make use of corpora and show learners how figurative meanings are used in context and what longer phrases (if any) they form part of. For example, by reading the concordance line below and paying attention to the surrounding context, learners can clearly see that 'top drawer' is used figuratively:

> ... KOTB: What about Angelina Jolie? She's always in the top drawer, isn't she?

> Mr KARGER: Of course. She played Mariane Pearl in 'The Mighty Heart' this summer and did a great job in that, and now she's in 'Beowulf', ...

Another option would be to present both literal and figurative meaning senses of the same collocation in context and ask learners to decide when the collocation in question is used literally or figuratively.
3 Teach learners to successfully guess from context, as context will determine whether a collocation is used literally or figuratively (this is of course only possible when learners know the majority of words in a text). For beginners, teachers giving clues might be a better approach.
4 Encourage learners to make use of specialized collocation dictionaries, for example the *Oxford Collocations Dictionary for Students of English* or the *Longman Collocations Dictionary and Thesaurus for Intermediate-Advanced Learners*.

Cautions and caveats   It is important to be careful when interpreting the results of this exploratory study. First of all, our sample is small. However, it is an initial indication of the existence of Figurative and Duplex collocations, and teachers should become familiar with the different types of collocations. The best way to interpret the reported percentages is as being indicative in nature only. Only further research with larger and more varied samples of collocations (also using different identification criteria) will be able to show how these figures need to be adjusted.

Moreover, teachers have to be careful when working with these collocation categories as they are not always clear-cut. The reality is fuzzy because so many different factors come into play: the learner's L1, how rich the context is, metaphor awareness, etc. Despite this, collocation meanings can differ and collocations are not only 'literal'.

Caution is also necessary in the case of Literal collocations because some of them may have more than one meaning sense. Even though they follow the A + B pattern, their accurate interpretation, for example in the case of the collocation 'take a break', will depend on the learner's existing understanding that 'take' can also mean 'have' and that 'break' also refers to 'rest'.

Finally, the reported 18 per cent figure is representative of Duplex collocations across the whole corpus. However, teachers should not assume that these collocations will be equally distributed across the different sub-corpora, but instead understand that certain collocations will be found more often in specific genres.

Conclusion

Collocations have been widely acknowledged as important in language learning and use. The results of this study suggest that a substantial percentage of them have figurative meanings. This indicates that collocations cannot be seen as merely the co-occurrence of words. With collocations, just as with individual words, meaning matters.

Final version received March 2016

Notes

1 Collocations are defined differently depending on research purpose and theoretical orientation. Some authors would refer to Figurative collocations and Duplex collocations as idioms (for example Moon 1998). Also, the relative ease with which a collocation is interpretable does not solely depend on the semantic properties of the item itself. If the speaker, for example, already has awareness that 'top' metaphorically maps on to 'best' or 'first' (for example 'top choice'), or that 'hot' maps on to 'popular' (for example 'hot topic'), then these items should be more easily understood.

2 A number of scholars have explored the interaction between patterning and meaning (for example Granger and Paquot 2008), but most textbooks and pedagogical discussion seem to focus mainly on the patterning aspect of collocation.

3 The term 'polysemous' can denote different types of polysemy. For example, polysemous in the case of collocations could mean 'more than one literal meaning' or 'more than one figurative meaning'. In this study, however, we focus on 'polysemous'

as referring to 'one literal and one figurative meaning'.

4 Mutual information (MI) is a statistic which shows how strongly one word collocates with another. For example, 'nice' collocates with many words ('nice day', 'nice smile', 'nice car') and so would have a low MI score. 'Tectonic' only occurs with a few words (mainly 'plates'), so has a high MI score. An MI of 3 is often given as the minimum score to be considered a collocation.

5 MI score: 3.84.

References

Boers, F. 2000. 'Metaphor awareness and vocabulary retention'. *Applied Linguistics* 21/4: 553–71.

Boers, F., **M. Demecheleer**, **A. Coxhead**, and **S. Webb.** 2014. 'Gauging the effects of exercises on verb-noun collocations'. *Language Teaching Research* 18/1: 54–74.

Davies, M. 2008. The Corpus of Contemporary American English (COCA): 520 Million Words, 1990–2015. Available at http://corpus.byu.edu/coca/ (accessed on 31 October 2014).

Granger, S. and **M. Paquot.** 2008. 'Disentangling the phraseological web' in S. Granger and F. Meunier (eds.). *Phraseology: An Interdisciplinary Perspective.* Amsterdam: Benjamins.

Grant, L. E. 2005. 'Frequency of "core idioms" in the British National Corpus (BNC)'. *International Journal of Corpus Linguistics* 10/4: 429–51.

Lewis, M. 1993. *The Lexical Approach: The State of ELT and a Way Forward.* Hove: Language Teaching Publications.

Martinez, R. 2013. 'A framework for the inclusion of multi-word expressions in ELT'. *ELT Journal* 67/2: 184–98.

Martinez, R. and **N. Schmitt.** 2012. 'A phrasal expressions list'. *Applied Linguistics* 33/3: 299–320.

McCarthy, M. and **F. O'Dell.** 2005. *English Collocations in Use: Intermediate.* Cambridge: Cambridge University Press.

Moon, R. 1998. *Fixed Expressions and Idioms in English.* Oxford: Oxford University Press.

Nation, I. S. P. 2006. 'How large a vocabulary is needed for reading and listening?'. *The Canadian Modern Language Review* 63/1: 59–81.

O'Dell, F. and **A. Broadhead.** 2014. *Cambridge English: Objective Advanced* (fourth edition). Cambridge: Cambridge University Press.

Shin, D. and **P. Nation.** 2008. 'Beyond single words: the most frequent collocations in spoken English'. *ELT Journal* 62/4: 339–48.

Sinclair, J. M. 1991. *Corpus, Concordance, Collocation.* Oxford: Oxford University Press.

Siyanova-Chanturia, A. and **R. Martinez.** 2015. 'The Idiom principle revisited'. *Applied Linguistics* 36/5: 549–69.

The authors

Marijana Macis is a PhD student at the University of Nottingham. Her research interests include vocabulary acquisition, knowledge of formulaic sequences, and incidental learning from reading.
Email: marijana.macis@nottingham.ac.uk

Norbert Schmitt is a Professor of Applied Linguistics at the University of Nottingham. He is interested in various aspects of vocabulary acquisition, knowledge, and testing.
Email: norbert.schmitt@nottingham.ac.uk