Selecting English collocations for explicit teaching in pedagogical materials at advanced levels

Sharon Virginia Hartle

University of Verona, Italy
https://orcid.org/0000-0002-3515-2472
sharon.hartle@univr.it

Abstract

This small-scale study aims to examine the appropriacy of collocation treatment, both in global-market coursebook materials and easily accessible online sites for the study of English lexis at advanced levels. Because of the constraints of the research project, assessment is limited to studying one online site and one unit of the coursebook *Speakout Advanced*, taking into account materials design and appropriacy of collocation selection. The main question was originally how appropriate the treatment of collocation was in teaching materials for C1 level learners, prompted by the realization over the years that collocation in English is a problematic area for advanced undergraduates studying a foreign language at Verona University. The research question, however, became slightly modified during the course of the study to how appropriate the selection of collocation was in these materials, and how to develop a series of filters to aid authors when they develop pedagogical materials to teach English collocation explicitly. To study this, initially, frequency was measured across the COCA and BNC corpora and then further filters were added to examine the nature of the collocations themselves including restriction, transparency and relevance for the specific context. In trialing these filters, the secondary aim was to test their effectiveness as measures that could enhance the process of materials development.

Keywords: collocation; advanced learners; ELT digital and print materials; explicit collocation teaching; materials development and evaluation

1. Introduction

In socio-cultural contexts where digital information retrieval searches are the norm and where, thanks to the advent of Web 2.0, widespread online content publishing has proliferated, pedagogical materials for the teaching of English as a foreign language (EFL) are being produced and made available online by many educators, ranging from commercially developed materials to those published by teachers for their learners. Many learners access these materials independently and, therefore, it is crucial that the items being taught explicitly in such materials are appropriate and relevant to learners' needs as well as being selected in a principled way. The small-scale study reported in this paper aimed to address this issue by improving the process of selecting English collocations for pedagogical materials.

2. Literature review

2.1. Materials development and evaluation with a focus on lexis

A considerable amount of research has recently been done into the evaluation of English language teaching (ELT) materials, even though it is a relatively new field of research. In fact, not very long ago Tomlinson and Masuhara (2010 p. 1) cited Chapelle (2008) as pointing out how "surprisingly little research has been published on materials evaluation". This echoes similar sentiments expressed by Sheldon (1988). Much of the recent research has been carried out in particular with reference to coursebooks and materials used in traditional face-to-face classroom settings. This research has examined different stages in the learning process and many different aspects of the materials themselves (see Ellis, 1997; Sheldon, 1998; Tomlinson, 2003). The focus of these studies has often been on existing materials and how teachers adapt or evaluate them. For example, Richards (2006) focused on the development of materials according to a framework firmly grounded in principled research. In fact, many frameworks have been developed to aid practitioners with selecting materials and adapting them for use in their contexts (see e.g., Cunningsworth, 1995; Harmer, 2015; Littlejohn, 2011; McGrath, 2013; Tomlinson, 2003b).

Another important focus in the literature is developing materials with the assistance of technology. For instance, the book for teacher training by McDonough, Shaw and Hitomi (2013) has been revised to include a chapter on IT materials as well as their evaluation and adaptation. Other influential publications on technological materials include Motteram, Slaouti and Onat-Stelma's (2013) comprehensive coverage of a range of topics related to technology in ELT, including materials evaluation and adaptation for the classroom. Reinders and

Lewis (2006) provide a checklist to aid those selecting materials for independent or self-access study and they also stress the fact that the selection of such materials is not simple due to a lack of published guidelines.

This paper focuses specifically on materials that aim to provide practice of lexis and specifically collocations. The literature on the teaching of lexis is abundant. Richards (1976), for instance, discussed the acquisition of lexis as not simply a matter of knowing the meaning but, rather, as a complex process which involves developing an awareness of not only linguistic but also psycholinguistic and sociolinguistic factors. With the advent of corpus linguistics at the end of the 20th century, leading researchers (e.g., Nation, 1975, 2003; Schmidt, 2000) developed the notion of frequency as being key when selecting lexis for explicit teaching. It is a measure that is still commonly used particularly when developing materials for lower levels, following the logic that the more frequently a word occurs the more useful it is for learners.

When discussing materials development for the explicit teaching of vocabulary, Nation (2003, p. 395) emphasizes that "(…) there is a relatively small group of words (around 2000) that are much more frequent and useful in a very wide range of language uses than other words in the language". In effect, materials designed for lower levels focus on high frequency items and those for more advanced levels tend to include less frequent items. When it comes to collocation, however, this may not be such a useful principle. This will be shown in the next subsection that focuses on the nature of collocations and the challenges that learning them poses for learners.

2.2. Collocation and problems it causes for advanced L2 users

There is considerable debate as to what collocation actually is. For the purpose of this paper, it is understood in a pedagogically meaningful way which departs from the linguistic concept of co-occurrence over a range of a few words to either side of a specific item (Halliday, 1994; Sinclair, 1991), as it does not necessarily focus on semantic properties (Macis & Schmitt, 2017). Rather, the focus here is on collocation as phraseological, lexical combinations which co-occur, but which also have a reciprocal relationship, largely determined by convention. In the examined materials collocation covers a range of wordings and the items assessed in this study reflect this. Firstly, there are two item collocations of the *verb + noun* type, such as, *miss the train*, or *adjective + noun*, such as *married name*. Secondly, there exist *multi part verbs*, such as *live up to something*. Thirdly, there are idiomatic expressions such as *make a name for oneself*. These three categories of collocations were the main focus in the analysis presented below.

The reciprocal relationship between constituent parts of a collocation refers to how fixed lexical combinations are. Collocations may fall somewhere in between *weak* and *strong* (Conzett, 2001), where *friendly dog* is weak in that each part could easily be combined with other items, as in *friendly person* or *hungry dog*, and *throw in the towel* is very strong. The key to this "strength" is the expectation, as Conzett explains, "the presence of one word means you strongly expect the other to be there too" (2001, p. 74). The idea of strength is also echoed in Nesselhauf's (2003) continuum of *free* to *restricted association*. She defines *free combinations* as combinations where elements can be substituted according to their semantic properties so that the verb *buy* may be combined with *a pair of jeans* because of its meaning of "purchasing" while *drink a pair of jeans* would not make semantic sense (Nesselhauf, 2003, p. 225).

Both Conzett's (2001) notion of "weak to strong" and Nesselhauf's (2003) notion of "lesser to more restricted" are useful when considering problems that producing collocations cause for learners. Nesselhauf (2003, p. 225) defines "restricted combinations" as those where the choice of possible reciprocal constituents seems to be arbitrary and it is this "arbitrariness" which seems to be responsible for the errors that B2-C1[1] level Italian undergraduate language students commonly produce. These include *live an experience* instead of *have an experience* or *lose the train* instead of *miss the train*, where the combinations of constituents are a matter of convention and use rather than semantically determined.

Research shows that learner difficulty in the use of collocation is common in foreign language learning (Bahns & Eldaw, 1993; Durrant & Schmitt, 2009; Granger & Bestgenn, 2014) particularly at intermediate to advanced levels. Two main reasons for this emerge from the literature. Firstly, there is the issue of frequency, in that low frequency or rare collocations tend to be underused (Granger & Bestgenn, 2014). Secondly, restriction is also problematic (Durrant & Schmitt, 2009). This is reflected in Italian undergraduates' oral and written target language production in my context. What is even more noticeable, perhaps, is that learners tend to mismatch those collocations that are in the middle of the free-restricted range. This refers to constituent parts that can combine with other elements and thus so are not fixed, such as *missing trains* and *having experiences*, mentioned above. As Conzett (2001, p. 70) points out, medium-strength items, when combined, are possibly more useful for learners than extremely rare or fixed items. This raises the question whether frequency may be the best criterion to employ when selecting lexical items to be taught. For this reason, the study reported below examined the types of collocation that are targeted for explicit

---

[1] These learners were part of a group that aimed to reach a C1 level according to the *Common European Framework* (CEFR) by the end of the course, so they, in fact, represented a range of levels.

teaching in easily accessible materials and tried to determine whether they meet the needs of learners by representing "medium strength items" (Conzett, 2001). In fact, teaching materials may not include such items at higher levels if frequency of occurrence is the basis for choosing collocations. Whilst more rare items should be included at more advanced levels, this should not be to the detriment of collocations whose constituent parts are quite common. To return to the example given above, a B2-C1 level learner may well recognize both *miss* and *train* as single words but the difficulty lies in their appropriate combination, so materials need to focus on such combinations as well as more infrequently occurring lexical items. The question arises then how such items might be successfully selected for explicit teaching.

2.3. Selecting lexical items for explicit teaching

Selecting lexical items for explicit teaching involves two steps. Firstly, it needs to be decided what the item should be and, secondly, its appropriacy needs to be determined. Present-day thinking focuses on lexis as going beyond simply knowing the meaning of words and actually targeting more than merely a single word. Many feel that language teaching does not focus enough on this. Brown (2011) expresses this sentiment when he comments that there is a "widespread feeling that language teaching fails to take a rich enough view of vocabulary knowledge" (2011, p. 84). He cites Singleton (1999) who writes that "much of what has passed for vocabulary teaching (…) addresses only the tip of the lexical iceberg" (p. 272). In the past, areas such as collocation were often left to incidental acquisition rather than being taught explicitly. The thinking changed radically following the publication of such books as Lewis' (1993) *The Lexical Approach*, which grew from the ground-breaking work of Pawley and Syder (1983). Lewis (1993) focuses on high frequency lexical patterns in language rather than teaching grammatical rules and adding vocabulary later, almost as an afterthought.

When developing materials and selecting items for explicit teaching, frequency is still often considered to be the best criterion for determining appropriacy (Koprowski, 2005; Shin & Nation, 2007). Koprowski (2005), however, pointed to the limitations of using frequency as a sole measure and added the criterion of range by considering lexis across three intermediate level coursebooks. He developed a *usefulness score* based on averaging frequency across the subdivisions of the COBUILD Bank of English corpus Martinez (2013) goes even further and suggests yet another measure based on the opaque or transparent nature of the collocations, depending on learners' first languages. While frequency remains one appropriate criterion for selection of lexical items, it is certainly not the only one.

2.4. Four ways to measure the appropriacy of collocations in teaching materials

When examining the collocations included in teaching materials, the selected items can be considered in various ways. Frequency alone is limiting as a measure of appropriacy as there are infrequent items which may be useful to teach at low levels too, such as *Nice to meet you*. As Martinez (2013, p. 187) points out, this phrase is infrequent in the British National Corpus but it is of unquestionable value to learners. In addition, as mentioned above, learners at B2-C1 levels often find medium-strength collocations challenging. The question of how strong or restricted the co-occurrence of constituents is may be a useful criterion to add to frequency. It is also useful to consider how comprehensible that reciprocal relationship is, depending in particular on the learner's L1, which Martinez (2013) describes in terms of a continuum from transparent to opaque. What is meant by this is that a combination such as *take time* can be quite easily understood from the meanings of the two constituent parts, whereas the phrase *take place* may not and is therefore more opaque. However, the L1 may, among other things, determine whether or not an item is easy to understand. For example, a combination like *take place* corresponds to the Spanish *tener lugar* with a similar meaning but this is not true of Portuguese, meaning that this collocation would be more opaque for Portuguese learners (cf. Martinez, 2013, p. 188). This indicates that useful collocations to teach will vary from context to context. Thus, the filter of relevance to specific learners' L1s together with their learning contexts can be added. In fact, little credit is given to what Timmis (2008, p. 6) refers to as "professionally informed intuition". Educators working in local contexts, however, build up an awareness of what may or may not cause difficulties for their learners, which may be an added measure of appropriacy, particularly when writing for specific groups of learners that teachers know well. In effect, the criteria of restrictedness, opacity and awareness of learner relevance seem to be useful filters that can be used alongside frequency when selecting lexical items for instruction. The study reported below was an attempt to verify how effective these filters are as a framework for choosing collocations for explicit teaching to B2-C1 level learners.

3. The study

My interest in materials for explicit teaching of collocations to advanced learners is closely linked to my teaching context. Verona University provides English instruction to a range of levels, but my work involves teaching advanced undergraduate students, representing B2 to C1 levels according to the *Common European Framework* (Council of Europe, 2001). One of the most important areas

that need to be developed is lexis and in particular collocations. This led me to compare the items selected for teaching collocation in one module from an existing advanced level coursebook and the collocation exercises available on one online site. While the study is limited in terms of resources, it offers important insights into the usefulness of the four criteria for selection discussed earlier, thus providing a useful point of reference for teachers.

3.1. Aims and research questions

The study aimed to apply insights gained from previous research into the teaching of lexis to identify a set of filters that can be combined to create a framework for selecting collocations for explicit teaching included in teaching materials. The research question was originally how appropriate the treatment of collocation was in teaching materials for advanced learners, prompted by the realization over the years that collocation in English is a problematic area for students of foreign languages at Verona University. However, in the course of the study, this was slightly modified and two specific research questions were formulated:
1) How appropriate is the selection of collocation in the EFL materials in question?
2) Does the framework developed in the study determine the appropriacy of collocations effectively?

3.2. Methodology

The small-scale study was based on a micro-examination of materials for teaching lexis, in particular collocation, comparing an online site that was produced privately and a unit from an advanced-level coursebook. More specifically, assessment focused on items included in one exercise from the online site *Better English Lessons* (Brown & Brown, n.d.), which is easily accessible when doing a Google search for "studying English collocations". The website choice was somewhat problematic because Google searches of "English collocation" yielded results which fell into four categories, that is, collocation dictionaries, sites aimed mainly at teaching collocations, sites that provide definitions or examples of collocations, and publisher sites which feature exercises complementing existing coursebooks or promoting specific teaching materials. Using simple Google searches is the most common method used by my learners to find practice exercises, so, putting myself in their shoes, I changed the search to "collocation exercises". The site *Better English Lessons* was the number one hit and the website offered a range of practice activities. As many of my undergraduate learners are specializing in business studies as well, the section on business collocations

!

seemed to be useful for them. As for the coursebook, Unit One of *Speakout Advanced* (Clare & Wilson, 2012) was chosen. The book was used at the time by C1 learners in the university language canter were using at that time. Since the number of basic two-word collocations was limited, I expanded my sample to include multiword items and idioms which can also be considered types of collocations because of the restricted nature of lexical combinations that co-occur in such items.

The methodology was a mixed methods approach combining quantitative and qualitative procedures. The former involved normalized (per million words)[2] frequency analysis of the data by lemma. This was done across the Corpus of American English (COCA)[3] (Davies, 2008) and British National Corpus (BNC). The latter drew on qualitative interpretation based on my own pedagogical intuition of what is relevant to my learners, combined with the notion of transparency (Martinez, 2013) and the free-restricted continuum (Nesselhauf, 2003). This choice was determined by the realization that selecting lexical items according to frequency alone is possibly not as useful as a combining several criteria. This is because, as Dörnyei (2007) points out, "(…) in most cases a mixed methods approach can offer additional benefits for the understanding of the phenomenon in question" (p. 72).

The results produced by corpus analysis are only as good as the corpus being used and the issue of idiomaticity, for instance, was one such problematic area. Some items from my samples were not actually present in the corpora, such as *tax your brain* but a simple Boolean Google search for revealed 285,000 hits. On the other hand, a simple collocate search for *bated + noun* in the first position on the right-hand side of the key word in COCA returned *bated breath* as being the most frequent collocation with 0.12 occurrences per million words. Such instances then needed to be examined and interpreted qualitatively.

Looking at the examples of these occurrences to investigate genre frequency in COCA's word and phrase interface showed that the usage and meaning were appropriate with all instances referring to "waiting expectantly and holding your breath anxiously or excitedly in anticipation of something". A collocate search for *verb + bated breath* in up to three left-hand positions of the

---

[2] Frequency searches across different corpora are often normalized to a one occurrence per million words to avoid the results being skewed in the raw data. This may happen, for instance, because corpora are of different sizes, or because of a concentration of occurrences in certain texts. The generally adopted formula is: Frequency per million words = (frequency ÷ text no. words) x 1,000,000.

[3] The American Corpus (COCA) enables mutual information searches to determine how frequently items co-occur when measured against the total number of occurrences in the database. Although mutual information scores are sometimes problematic with less frequently occurring items, such as 1-3 tokens, a minimum occurrence rate can be set when doing such searches to avoid this problem.

key term showed the most common collocations to be different forms of *wait.* Semantically, then, these results were reliable but when it came to genre, the item did not prove to be suitable for Business English materials, as these collocations were described as Business Collocations in the materials under investigation. This is because the most common occurrences came from fiction and an examination of seven instances from academic genres returned zero occurrence in texts were related to business issues.

Particular care also needs to be taken when dealing with idioms in corpora since there is a danger that the literal meaning of items such as *black sheep*, for instance, is not the idiomatic one required. For this reason, idiom analysis was also done manually. When limiting the search for *[art*] black sheep*, ten results returned *The Black Sheep* as the name of a wool shop in a fictional text and two examples returned this item as the name of a brewery, *The Black Sheep* brewery in Masham, Yorkshire. This is further illustration of the fact that corpus data needs to be handled with care.

The collocations taken from the website and the coursebook were then analyzed by lemma searches across the COCA) and the BNC for normalized frequency (per million words). In some cases, however, manual analysis had to be carried out, as discussed above. As has already been mentioned, using frequency alone as a criterion for the selection of items for explicit teaching of lexis may be problematic, particularly at more advanced levels. For this reason Koprowski's (2005) *usefulness score* was initially used in this study to assess the frequency and range of selection of collocations. However, since the results were still related to frequency as such, it was felt that they did not add much to the initial frequency analysis. For this reason, the three qualitative measures outlined in section 2.4. were applied as well. More specifically, each item was given a 1-5 rating along the free-restricted, transparent-opaque and learner relevance continua, with the last of these based on my "professionally informed intuition" (Timmis, 2008, p. 6). The frequency results and the scores on the other measures were then averaged and I established a value of 2 out of 5 as being an appropriate measure of usefulness, which reflected the medium-strength, medium-frequency nature of items that I felt were suitable for my learners.

4. Results

4.1. Nature of the resources under investigation

These online lessons were originally developed by the authors of *Better English Lessons* to help their students. The claim made on the website states that these exercises have become very popular and are used by almost 20,000 people

every day. The authors have made some attempt to classify the lessons into *easy* or *business* categories. While the materials may have been initially designed to help learners in a specific context, they are now widely available and the users are informed that no fee is charged.

The items from one exercise were analyzed in depth which was based on filling gaps in sentences, thus requiring language production (see Appendix A). The focus was on "business collocations" but a quick glance at some of the items showed that the link to the business English domain was quite tenuous. This was the case, for example, with *strong bond*, *heavy burden*, or *tax my brains*, which can be widely used. This in itself may raise concerns about the selection approach adopted by the developers of this site.

Once the items for analysis had been determined, they were investigated across the COCA and the BNC and assessed according to the four filters mentioned above. This was done by searching for lemmas rather than words to account for different possibilities and allowed a minimum occurrence of ten items. As explained above, manual analysis had to be carried out in some cases, particularly when dealing with restricted items. The data were interpreted according to my own professional intuition, informed by collocate searches in the COCA for frequency and relevance for certain items. The frequency results and the scores on the other measures were then averaged and an assessment was made as to which items would be the most useful for my advanced learners. As will be shown below, in order for lexical items like these to be of real value to learners, materials writers must adopt a principled approach when authoring them.

In private e-mail communication with one of the authors of *Speakout Advanced*, the procedure for selecting lexical items for explicit teaching in the coursebook was elucidated. It was as follows: (1) the items were selected according both to level and with reference to the authors' shared teaching experience, (2) they were then looked up in the *Longman Dictionary of Contemporary English* (n.d.), the *LTP Dictionary of Selected Collocations* (Hill & Lewis, 1997) and on other websites such as the *Compleat Lexical Tutor* (Cobb, n.d.), (3) the BNC corpus was used as well but to a limited extent because of time constraints, (4) more recently, the *English Vocabulary Profile* (Cambridge University Press, n.d.) has also been employed to inform the decisions made, (5) the final stage in the process is the selection of items fitting in with the topics of the units. The authors emphasize their belief in the importance of lexical sets and idiomatic English in the introduction to the *Teacher's Book* and organize their materials into the main units, the *Lookback* sections where language is tested and recycled, the workbook tests and extra resources. Figure 1 shows the breakdown of the different lexical items, and it becomes clear that single words, rather than multi-word items are given high priority.
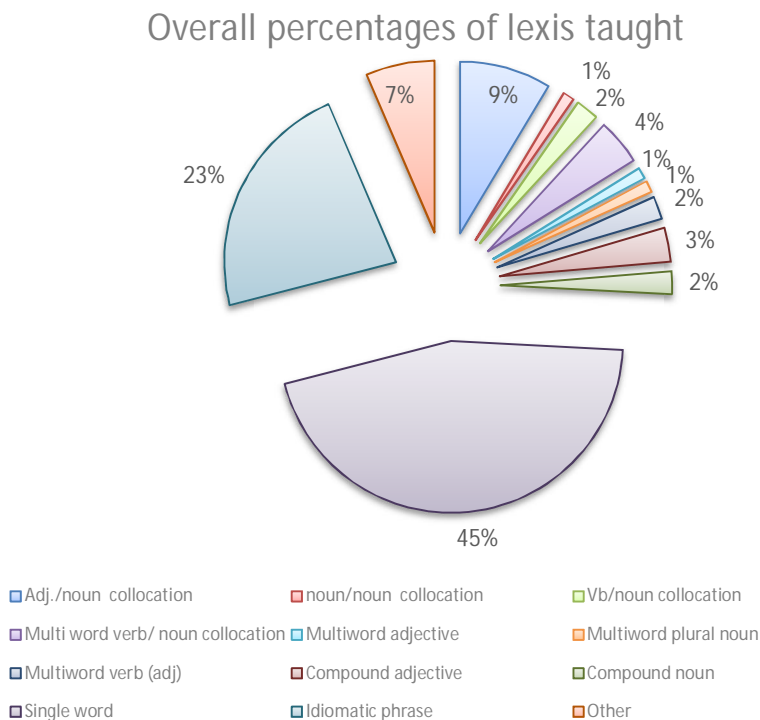
Overall percentages of lexis taught

Figure 1 Breakdown of the percentages of different types of lexical item taught in Unit One of *Speakout Advanced*
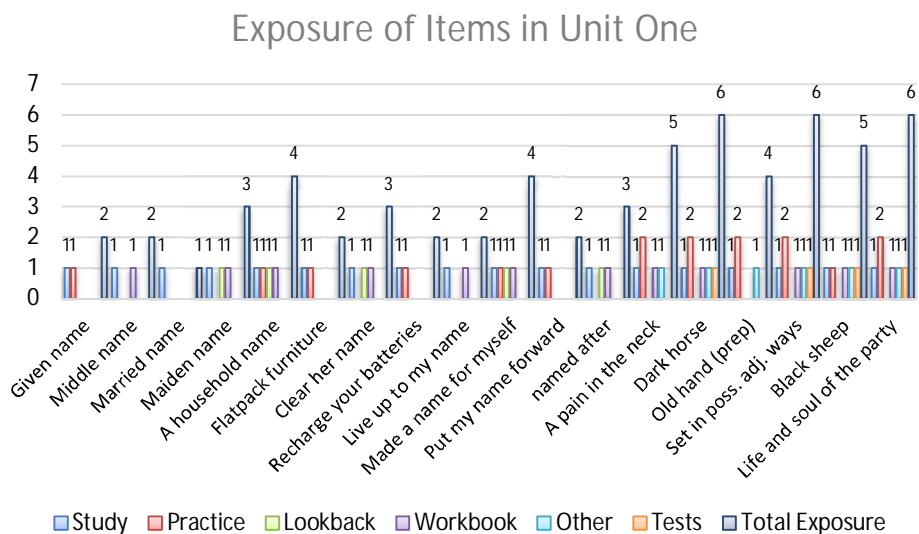
Exposure of Items in Unit One

Figure 2 Exposure to sample lexical items in Unit One of *Speakout Advanced*

48% of the items for explicit reaching, however, were introduced for receptive skills work and were therefore not included in the analysis. For the remaining 52%, differences in lexical categories dictated that collocations needed to be combined with multiword items and idioms. The items were recycled in various sections providing learners with different degrees of exposure, which was another factor that had to be taken into consideration. Thus the lexis taught in Unit One was initially examined to measure the type of exposure that was offered to learners (see Figure 3). Qualitative analysis was then carried out for these items in the same way as for the online items.

4.2. Frequency analysis

The analysis of the items included in the online materials revealed that in 13 out of 15 activities in the form of gap-filling the collocations were selected alphabetically from A-C. Various collocations were extrapolated from the actual sentences as the target collocations were not always clear. In the item *Economists are forecasting an economic* _____, for instance, the target could be a range of semantically determined options. The same could be said about *I'm afraid that thought-provoking* _____ *is out of print*, where it was only the fact that all the collocations began with the letter *b* that provided the necessary contextual information that led us to make an informed guess that the answer was probably *book*. When the online materials were analyzed for normalized frequency across the COCA and the BNC, the results proved to be variable, ranging from high frequency for collocations like *board meeting* and *annual budget* to very low frequency for *taxed my brain* or a *thought-provoking book*. Even a quick glance at the results shows that the selection of collocations was neither principled nor systematic (see Figure 3).

When examining the lexical items chosen from Unit One of *Speakout Advanced*, the normalized frequency scores in Figure 4 indicate that the high priority items receiving the most exposure tended to reflect low frequency or range. For example, *The life and soul of the party* occurred 6 times in the materials but is not present at all in the BNC. High priority was also given to *set in poss. adj. ways* and *a dark horse* which are rare in the two corpora. Therefore, choices regarding items that should be highlighted and to which learners should be frequently exposed seem to reflect the belief of the authors that idioms and low frequency items are appropriate for advanced learners.
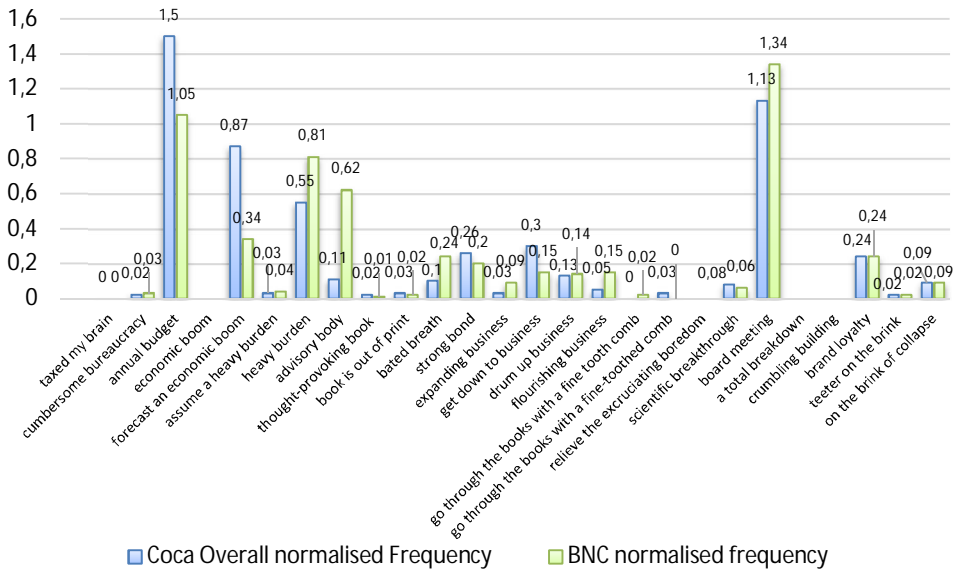
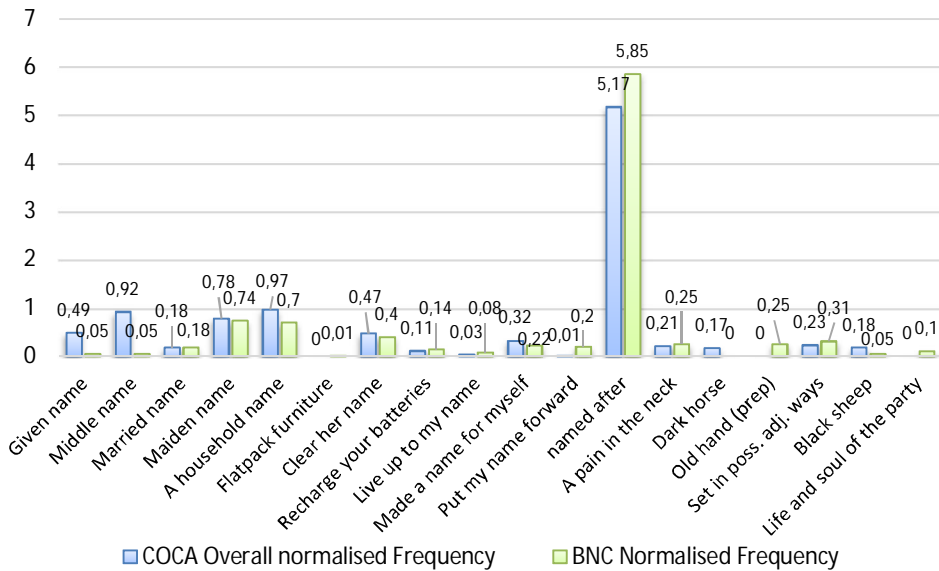Figure 3 Normalized frequency patterns for collocations in *Better English Lessons.*

Figure 4 Normalized frequency patterns for collocations in *Speakout Advanced*

4.3. Qualitative measures

Table 1 shows the ratings for the free-restricted and transparent-opaque criteria for the items included in the online materials. The rating of 3 on the free-restricted

continuum, for instance, was given to *taxed my brain*, although this is not an item which is confined to the world of business. The value of 5 was awarded to *book* with *out of print* but not to *book* with *thought-provoking*, which was rated as 1 as the collocation is not restricted at all. The transparent- opaque filter returned similar results, with the most opaque results, with the rating of 5 being given to *taxed my brain*, *bated breath*, *get down to business* and *drum up business*. However, the link to business issues was tenuous to say the least for many of these collocations such as a *strong bond* or *excruciating boredom*.

Table 1 Scores on the free-restricted and transparent-opaque filters in *Better English Lessons*

Item	Free-restricted criterion (a score of 1-5)	Transparent-opaque criterion (a score of 1-5)
taxed my brain	3	5
cumbersome bureaucracy	1	1
annual budget	1	1
economic boom	2	2
forecast an economic boom	3	1
assume a heavy burden	4	3
heavy burden	1	1
advisory body	3	1
thought-provoking book	1	1
book is out of print	5	1
bated breath	5	5
strong bond	2	2
expanding business	1	1
get down to business	5	5
drum up business	5	5
flourishing business	3	1
go through the books with a fine tooth comb	5	3
go through the books with a fine-toothed comb	5	3
relieve the boredom	5	3
excruciating boredom	1	1
scientific breakthrough	4	1
board meeting	2	1
breakdown in negotiations	3	3
a total breakdown	1	1
crumbling building	1	1
brand loyalty	4	1
teeter on the brink	5	3
on the brink of collapse	4	3

Table 2 shows the free-restricted/transparent-opaque ratings for the items in Unit One of *Speakout Advanced* items. In this case, only three items, that is *middle name, married name* and *maiden name* received lower scores.

These three items constituted a lexical set, which suggests that the inclusion of items purely because they have one constituent part in common may not be the most appropriate basis for selection. A collocate search of the COCA provided more appropriate collocations for these items such as *full/good/bad/assumed + name*, the last of these being particularly appropriate for Italian learners since it is a false cognate with a relatively high level of opaqueness. The organizing principle of choosing items according to lexical sets is also evident in the greater exposure given to items reflecting Unit One's main theme of *Names*.

Table 2 Scores on the free-restricted and transparent-opaque filters in *Speaking Advanced*

Item	Free-restricted criterion (a score of 1-5)	Transparent-opaque criterion (a score of 1-5)
given name	4	4
middle name	1	1
married name	1	3
maiden name	3	3
a household name	3	5
flatpack furniture	5	5
clear her name	4	5
recharge your batteries	3	3
live up to my name	5	5
made a name for myself	5	5
put my name forward	5	3
named after	3	3
a pain in the neck	3	5
dark horse	5	5
old hand (prep.)	5	5
set in poss. adj. ways	5	5
black sheep	4	4
life and soul of the party	5	4

The final step was to decide which of these items would be appropriate for my learners. Given the difficulties Italian undergraduates have with mid-strength collocations, the learner relevance (usefulness) filter was added to identify these items, which took into consideration the fact that these learners specialized in international commerce. These items were medium-frequency, medium-strength and received medium transparency-opacity scores, taking into account the possible impact of the first language (i.e., Italian). It was clear that the only tangible criterion for the selection of items taught in the online materials was alphabetical order. There was no indication of register and out of the 27 items (one is duplicated, because of various spellings, not to mention the misspelling *excruciating* as *excrutiating**) only 9 scored more than 2 on the

learner relevance measure (see Figure 5). These items were *bated breath, get down to business, drum up business, relieve the boredom, breakdown in negotiations, brand loyalty, teeter on the brink and on the brink of collapse*, with all of them representing a challenge to my learners.
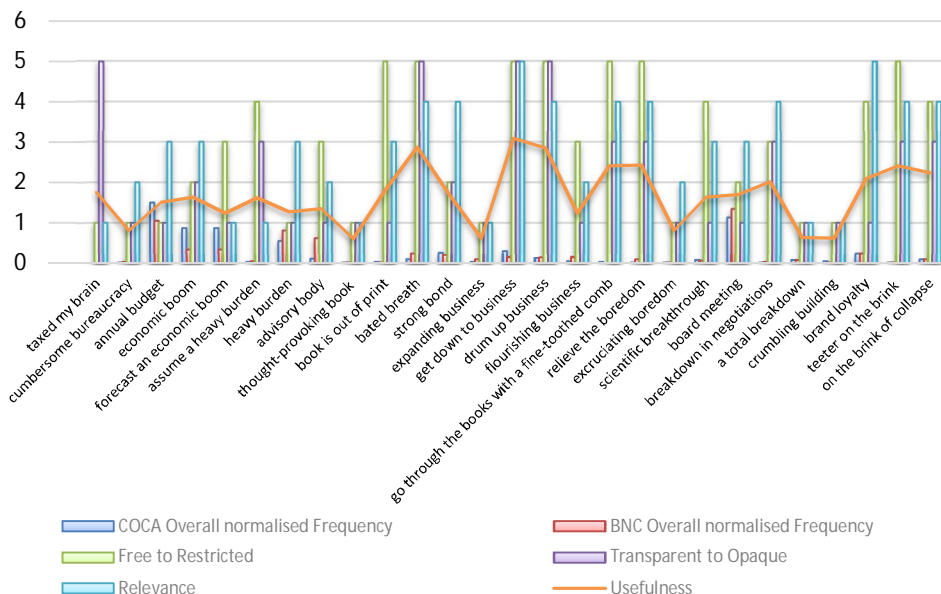

Figure 5 Final analysis including the learner relevance (usefulness) filter for *Better English Lessons*

As Figure 6 shows, in the case of the collocations in the unit of *Speakout Advanced*, the learner relevance scores were much more homogeneous, indicating that most of the items were useful to the specific group of learners. Apart from low-frequency and restricted items, priority is also given to idiomatic expressions such as *set in poss. adj. ways* and *a dark horse* which are rare in the corpora used. However, despite their infrequency, such idioms are useful for advanced learners. Two of these lexical items, that is, *flatpack furniture* and *recharge your batteries* were taught receptively in a listening task and then recycled in the production phase in which learners had to use them in a questionnaire. *Flatpack furniture* does not appear at all in the corpora while *recharge your batteries* has a low frequency score. However, both items proved to be very popular with in a class that I taught recently. When the students were asked to reflect on what they had learned and select the most interesting lexis, these items were repeatedly mentioned, perhaps because they are culturally interesting and highly relevant. This shows that issues related to relevance and personal choice

cannot always be measured by corpora results and this is where the pedagogical intuition of teachers working in a specific context can be a useful resource.
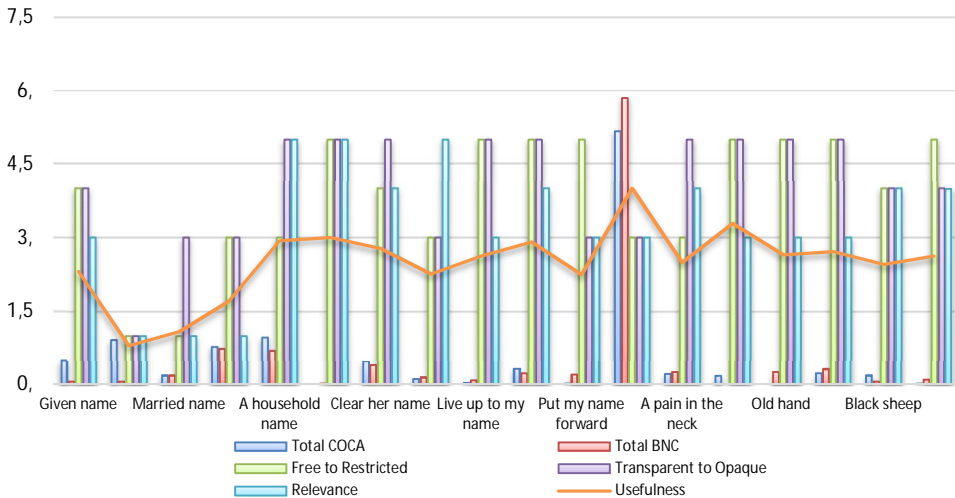

Figure 6 Final analysis including the learner relevance (usefulness) filter for *Speaking Advanced*

5. Conclusion

Although this was a small-scale study and its results cannot be generalized, its findings suggest that combining information about frequency and range with measures of restriction and opaqueness may lead to a principled selection of lexical items for instruction. In effect, these filters could be a useful starting point for those wanting to develop materials for the explicit teaching of collocations at advanced levels. In fact, Martinez (2013) argues that his frequency transparency framework probably needs to be applied in different ways depending on different contexts and should only be taken as a point of departure, which is exactly how the measures applied in this study should be considered. Thus, educators aiming to develop materials for their local contexts can draw upon their pedagogical intuition, taking into account the native language and preferences of their learners. The learner relevance framework introduced here is subjective but it provides a starting point for discussion and can be further experimented with. It has helped identify items of medium strength and medium frequency which are often problematic for advanced learners. This is certainly true of my own context, but replication in other learner populations and contexts would be extremely useful.

Although it must be remembered that the analysis was based only on one sample, it is also worth underlining that the selection of items for the online materials proved to be largely inappropriate, being primarily determined by alphabetical order. There was no real attention to productive appropriacy or learner relevance of the items selected and at times there was no attempt to ensure their suitability for the Business English domain. Given the easy access to such online materials and the relatively high numbers of users of this website, there are certainly reasons for concern. Thus, there is a need for awareness-raising among those writing online materials regarding the criteria they can apply when choosing lexical items to include. The four-stage framework used in the study proved to measure the usefulness of the lexical items effectively, which suggests that it should be trialed further. Hopefully, it has a chance to constitute a useful tool for educators and materials writers who embark on the increasingly widespread practice of developing online materials as part of e-learning or blended learning programs. It is hoped that measures like the ones proposed here may inform the criteria they develop when focusing on selecting collocations for explicit teaching to advanced learners.

References

Bahns, J., & Eldaw, M. (1993). Should we teach EFL students collocation? *System*, *21*(1), 101-114.

Brown, D. (2011). What aspects of vocabulary knowledge do textbooks give attention to? *Language Teaching Research*, *15*(1), 83-97.

Brown, P., & Brown, C. (n.d.). *Business English grammar lessons.* http://www.better-english.com/

Cambridge University Press, Cambridge English Language Assessment, British Council, Cambridge University, University of Bedfordshire, & English UK (n.d.). *English vocabulary Profile.* https://www.englishprofile.org/wordlists

Clare, A., & Wilson, J. (2012). *Speakout adanced.* Harlow: Pearson Education.

Cobb, T. (n.d.). *The compleat lexical tutor.* www.lextutor.ca

Conzett, J. (2001). Integrating collocation into a reading and writing course. In M. Lewis (Ed.), *Teaching collocation* (2nd ed, pp. 70-86). Hove: LTP.

Council of Europe. (2001). *The Common European framework of reference for languages: Learning, teaching, assessment.* http://doi.org/10.1017/S026 7190514000221

Cunningsworth, A. (1995). *Choosing your coursebook.* London: Heinemann.

Davies, M. (2008). *The Corpus of Contemporary American English (COCA): 560 million words, 1990-present.* https://corpus.byu.edu/coca

Dörnyei, Z. (2007). *Research methods in applied linguistics.* Oxford: Oxford University Press.

Durrant, P., & Schmitt, N. (2009). To what extent do native and non-native writers make use of collocations? *IRAL, 74*(2), 157-177.

Ellis, R. (1997). The empirical evaluation of teaching materials. *ELT Journal*, *51*(1), 36-42.

Granger, S., & Bestgenn, Y. (2014). The use of collocations by intermediate vs. advanced non-native writers: A bigram-based study. *IRAL*, *52*(3), 229-252.

Halliday, M. A. K. (1994). *An introduction to functional grammar* (2nd ed.). London: Arnold.

Harmer, J. (2015). *The practice of English language teaching* (5th ed.). Harlow: Pearson Education.

Hill, J., & Lewis, M. (1997). *The LTP dictionary of selected collocations.* Hove: LTP.

Koprowski, M. (2005). Investigating the usefulness of lexical phrases in contemporary coursebooks. *ELT Journal*, *59*(4), 322-332.

Lewis, M. (1993). *The lexical approach.* Hove: LTP.

Littlejohn, A. (2011). The analysis of language teaching materials: Inside the Trojan Horse. In B. Tomlinson (Ed.), *Materials development in language teaching* (pp. 179-211). Cambridge: Cambridge University Press.

Macis, M., & Schmitt, N. (2017). The figurative and polysemous nature of collocations and their place in ELT. *ELT Journal*, *71*(1), 50-59.

Martinez, R. (2013). A framework for the inclusion of multi-word expressions in ELT. *ELT Journal*, *67*(2), 184-198.

McDonough, J., Shaw, C., & Hitomi, M. (2013). *Materials and methods in ELT: A teachers' guide* (3rd erd.). London: Wiley-Blackwell.

McGrath, I. (2013). *Teaching materials and the roles of EFL/ESL teachers practice and theory*. London and New York: Bloomsbury.

Nation, I. S. P. (1975). Teaching vocabulary in difficult circumstances. *ELT Journal*, *30*, 21-24.

Nation, P. (2003). Materials for teaching vocabulary. In B. Tomlinson (Ed.), *Developing materials for language teaching* (pp. 394-405). London: Continuum.

Nesselhauf, N. (2003). The use of collocations by advanced learners of English and some implications for teaching. *Applied Linguistics*, *24*(2), 223-242.

Pawley, A., & Syder, F. H. (1983). Two puzzles for linguistic theory: Nativelike selection and nativelike fluency. In J. C. Richards & R. Schmidt (Eds.), *Language and communication* (pp. 191-225). London: Longman.

Reinders, H., & Lewis, M. (2006). An evaluative checklist for self-access materials. *ELT Journal*, 60, 272-278.

Richards, J. C. (1976). The role of vocabulary teaching. *TESOL Quarterly*, *10*(1), 77-89.

Richards, J. C. (2006). Materials development and research: Making the connection. *RELC Journal*, *37*(1), 5-26.

Schmidt, N. (2000). *Vocabulary in language teaching*. Cambridge: Cambridge University Press.

Sheldon, L. E. (1988). Evaluating ELT textbooks and materials. *ELT Journal*, *42,* 237-246.

Shin, D., & Nation, P. (2007). Beyond single words: The most frequent collocations in spoken English. *ELT Journal*, *62*(4), 339-348.

Sinclair, J. M. (n.d.). *COBUILD Collins Birmingham University International Language Database*. Birmingjam: Birmingham University.

Sinclair, J. M. (1991). *Corpus, concordance, collocation.* Oxford: Oxford University Press.

The Longman dictionary of contemporary English. (n.d.). London: Longman.

Motteram, G., Slaouti, D, & Onat-Stelma, Z. (2013). Second language teacher education for CALL: An alignment of practice and theory. In M. Thomas, O. Reinders, & M. Warschauer(Eds.), *Contemporary Computer Assisted Language Learning: The role of digital media and incremental change* (pp. 56-71). London: Bloomsbury.

Timmis, I. (2008). The lexical approach is dead: long live the lexical dimension! *Modern English Teacher*, *17*(3), 5-10.

Tomlinson, B. (Ed.). (2003a). *Developing materials for language teaching.* London: Continuum.

Tomlinson, B. (2003b). Developing principled frameworks for materials development. In B. Tomlinson (Ed.), *Developing materials for Language Teaching* (pp. 107-129). London: Continuum Publishing.

Tomlinson, B., & Masuhara, H. (2010). Research for materials development in language learning: Evidence for best practice. In B. Tomlinson & H. Masuhara (Eds.), *Research for materials development in language learning: Evidence for best practice* (pp. 1-20). London: Continuum.

Appendix A

Exercise 7 from the Better English Lessons (better.english.com)

I have to go to Head Office to attend a _____ meeting. It's very important.

The matter is very serious. We need to refer it to our advisory _____ for consideration

Working together under such pressure created a strong _____ between the team. They became very close.

I'm afraid that that thought-provoking _____ is out of print.

The tax authorities are going through the _____ with a fine tooth comb.

Economists are forecasting an economic _____ .

There must be something I can do to relieve the excrutiating _____ of this job

I've been taxing my _____ about this but haven't come up with a solution.

The customers have great _____ loyalty. They won't buy anything else.

The argument led to a total _____ in the negotiations.

We have achieved a significant scientific _____

We await the news with bated _____ .

The negotiations are teetering on the _____ of total collapse.

I'm afraid our expenditure has gone over the annual _____ .

I don't know how they can work in such a crumbling _____ . They need to spend more money on the up-keep or knock it down.

If you take on this job, you'll be assuming a heavy _____ .

We should try to cut down some of this cumbersome _____ .

This is an expanding _____ . It is really flourishing.

It's time we got down to _____ .

We need to drum up some new _____ . Sales are going down