A Comparative Study of Multi-Word Items in EFL Coursebooks

Paul McAleese

Abstract

Recent findings in corpus linguistics have revealed a significant proportion of authentic language to be constructed of prefabricated phrases (Biber & Conrad, 1999). Research has also shown learner knowledge of these multi-word items can promote both processing speed and pragmatic competence (Ellis, 1996; Wood, 2007). However, to date only a few studies have investigated such language in contemporary teaching materials (Hsu, 2008; Koprowski, 2005; Meunier & Gouverneur, 2007). This study used a large-scale corpus to investigate and compare multi-word items in a small number of upper-beginner-proficiency-level coursebooks used in a Japanese university context. The results showed that although a large number and variety of multi-word items are introduced, the types and representativeness of these items vary significantly among the coursebooks. The results suggested that a significant proportion of the items in question may be unrepresentative of authentic language and therefore of limited usefulness to the target learners.

Introduction

Background

This study investigates multi-word items (MWI) included in four contemporary English for general purposes (EGP) coursebooks used in a Japanese private university context with upper-beginner to pre-intermediate-level learners. Using a large-scale corpus, the extent to which the items introduced are representative of authentic language is also looked at. Based on the results, the usefulness of the MWI to learners is considered along with the implications for both teachers and coursebook writers.

Defining MWI

There has been a significant amount of study on collocation and phraseology. As a result, a wide range of terminology has been used in the literature to describe such language, including *lexical bundles* (Biber & Conrad, 1999), *lexical phrases* (Nattinger &

DeCarrico, 1992) and *formulaic language* (Wray, 2008). For the purpose of this study, the term *multi-word items* provided by Moon (1997) was used, describing them as sequences of two or more words co-occurring with a high regularity and limited structural variation. This definition includes a range of structures from multi-word compounds to semi-fixed expressions. In accordance with Moon's definition, MWI contrast to syntactically formed language; however, their degree of fixedness can allow for limited structural variations such as pluralization and verb inflections.

Benefits of MWI

Research has revealed knowledge of MWI can provide a range of benefits to language learners. Studies using large-scale corpora have revealed that between 20-50% of language is made up of varying degrees of MWI (Biber & Conrad, 1999; Erman & Warren, 2000). Studies on second language learners have also revealed knowledge of MWI increases the speed of encoding and decoding language and, as a result, speed of fluency (Ellis, 1996; Wood, 2007). A very strong case is also made by Nattinger and DeCarrico (1992) and Lewis (1994) on the benefits of functional MWI expressions in improving pragmatic competence. These types of MWI range from fixed greetings and responses to phrases for discourse management and can provide immediate communicative advantages, even to beginner-proficiency-level learners.

Evaluation of MWI

There are a number of factors that can be taken into account when evaluating the usefulness of language for leaners. Some of these include relevance to learners, learnability, frequency, and range (Mackey, 1965; White, 1988). In this study, it was decided to use the frequency and range to determine the level of usefulness to the learners in question. It was decided frequency should be used because, as it gives an estimation of how often MWI actually occur in authentic language, it can give an indication of MWI representativeness and therefore the language learners are most likely to reencounter. Range was also selected because it gives an estimation of the degree to which a given MWI occurs over different language types and genres, and therefore represents flexibility of use. It is accepted that frequency and range are not the only criteria that can be considered when evaluating MWI usefulness; however, based on the above, they are generally thought to be primary factors in terms of

language usefulness to learners (Nation & Waring, 1997; Sinclair, 1991).

Previous Coursebook Studies on MWI

Despite the increased prominence of MWI mentioned above, to date there have been relatively few studies on these items in coursebooks. The studies to date have mostly focused on MWI in intermediate to advanced-proficiency level coursebooks and have identified problems with their treatment (Hsu, 2008; Koprowski, 2005; Meunier & Gouverneur, 2007). The studies by Hsu (2008) and Meunier and Gouverneur (2007) suggested deficiencies with the number of MWI targeted for explicit attention, the order MWI were introduced, and inconsistent terminology used to identify MWI types and functions.

Of the studies above, Koprowski's (2005) was the only one using frequency and range to estimate the MWI usefulness. In this study, three upper-intermediate-proficiency-level coursebooks were investigated, and a total of 822 MWI were identified and classified. It was first determined that not only did the number and type of MWI vary significantly among the coursebooks but also over individual coursebook units suggesting there was no consistent MWI selection criteria. More importantly, based on MWI frequency and range data from the Bank of English (1991) corpus, Koprowski also suggested around 20% of the MWI identified were unrepresentative of authentic language. A very similar study was conducted more recently by McAleese (2013) on an upper-beginner-level coursebook; however it was limited in scope being based on only 220 MWI identified from a single coursebook. From corpus frequency and range data this study also suggested a significant proportion of the MWI identified were of limited value to the coursebook target learners. However, the author acknowledged the limited scope of the study and that further study covering a larger number of MWI and range of coursebooks at this proficiency level was required.

This study continues from McAleese's (2013) investigation above but includes three additional coursebooks to not only provide a larger sample of MWI, but also make comparisons among the coursebooks. It specifically aims to answer the following questions: (1) To what extent do the MWI numbers, types, and frequency and range scores vary among the coursesbooks? (2) How useful are the MWI to the target learners of these coursebooks?

Methodology

The Coursebooks

The coursebooks chosen for this study were used to teach EGP courses in a private Japanese university context. The learners in question were non-English majors and ranged from upper-beginner to pre-intermediate proficiency. In order to make the results of the study as relevant as possible to other teaching contexts, it was decided to use coursebooks that were both used by the author of the present study and teaching peers in the above context and also widely commercially available. It was also decided to consider only coursebooks published after 2005 in order to account for recent research and development in the field. Finally, due to the fact that most coursebooks introduced over 200 MWI, resulting in a highly time-consuming process for their analysis, only a limited number of coursebooks could be investigated within the scope of this study. Accordingly, the following four coursebooks were selected.

New Headway Pre-intermediate 3rd Edition (Soars & Soars, 2007), Oxford University Press

Touchstone 2 (McCarthy, McCarten, & Sandiford, 2005), Cambridge University Press

English Firsthand 1, 4th Edition (Helgesen, Brown, & Wiltshier, 2010), Pearson Longman

Smart Choice 2, 2nd Edition (Wilson, 2011), Oxford University Press

Of the coursebooks selected, *Touchstone 2* was of particular interest as the authors specifically stated the *Cambridge International Corpus* was used to create the coursebook and "make sure each lesson teaches (you) authentic and useful language" (McCarthy et.al, 2005, p. 4).

Identifying and Categorizing MWI

Coursebook MWI were taken from the unit vocabulary summaries or word lists. These summaries were either found in the student book or teacher book appendices or were explicitly introduced as coursebook target vocabulary in the accompanying teacher book or student CD-ROM. An item was considered a MWI when it was introduced as a complete unit or chunk of vocabulary. For example, the phrase *look forward to* would be considered a MWI if it were listed as one complete unit rather than three separate units. In some cases the coursebook introduced phrases as one unit but indicated possible substitution of intervening words such as *sprain* (*your*) *ankle*, and these were also considered MWI. Accordingly, the MWI identified also included longer

semi-fixed and fixed expressions such as *That's really kind of you* and *How do you say (that) in English?* provided they were introduced as single units as above. All MWI identified were logged verbatim and, in order to make further comparisons later in the study, categorized into four types: *compound nouns, phrasal verbs, longer expressions,* and *other two-word collocations*. As previous studies had also shown significant MWI variation between individual coursebook units (Koprowski, 2005), all coursebook MWI identified were analyzed without using sampling.

Determining Frequency and Range

A computer-based corpus was queried to determine MWI frequency and range. The corpus used was The Bank of English (1991) because it provided a very large and wide-ranging sample of language types and genres including 21 different sub-corpora. In order to collect MWI data from the corpus, the approach developed by Koprowski (2005) was adopted. The first part of the approach provided comprehensive criteria for ensuring that MWI frequency and range values also reflected structural variations and issues with polysemy. For example, with the phrasal verb *look up* it allowed for the data to also account for all verb forms of *look* such as *looked up*. Also, as *look up* has multiple meanings, it allowed the data to only reflect the coursebook meaning.

The second part of the approach allowed for frequency and range data to be incorporated into a single value for each MWI. These values were called R-scores (representativeness scores) and were calculated by averaging the MWI frequency values over the five subcorpora the item occurred most frequently in. In other words, R-scores for each MWI were determined by tallying the frequencies (words-per-million) of its five most common subcorpora and then dividing that value by five. For example, in the case of the MWI *dining room*, the five subcorpora it occurred in most were: 1) US Ephemera, 2) UK Ephemera, 3) OZ Papers, 4) UK magazines, and 5) UK books; and averaging these five frequencies gave an R-score of 24.6 (see Table 1). Accordingly, a high R-score suggests the MWI is more representative of authentic language than a low R-score.

Table 1

Calculation of R-Score for dining room

MWI	1	2	3	4	5	R-score
dining room	29.7	29.7	23.0	21.9	18.5	24.6

Prior to conducting the study, a brief pilot study was performed with another teaching colleague to ensure the method used for determining R-scores was accurate and objective, particularly with regards to structural variations and polysemy, was appropriate. Using the same BOE corpus, R-scores for the first 25 coursebook MWI (two chapters) were calculated separately by the author of the present study and a colleague. Comparing the two sets of results, 23 of the 25 R-scores gave the same R-scores and the same two polysemous MWI were identified. However, the two polysemous MWI gave slightly different R-scores with *heavy metal* giving final R-scores of 4.3 and 4.5 and *spend time* (*somewhere*) giving R-scores of 7.4 and 7.6. Accordingly, it was accepted that a certain degree of subjectivity might be involved in addressing polysemy; however, the degree was considered acceptable for the purpose of this study.

Results

MWI Numbers and Types

A total of 811 MWI were identified in the four coursebooks, and these were categorized by coursebook and MWI type. They were then tallied as a percentage of the total vocabulary covered in the each of the coursebook vocabulary summaries (see Table 2). As shown in the table, there was significant variation in MWI percentages among the four coursebooks. *Touchstone* had the highest percentage of MWI at 47.9%, and *New Headway* the lowest at 18.0%, under half the value of *Touchstone*.

Table 2

Percentage of MWI by coursebook

Coursebook	% MWI
Smart Choice 2	32.3
Touchstone 2	47.9
English Firsthand 1	31.7
New Headway Pre-int.	18.0

The percentage of total MWI allocated to each MWI type was then compared by coursebook (see Table 3). The table shows considerable variation in all four MWI types among coursebooks, and this was particularly evident with *expressions*. *Smart Choice* did not explicitly introduce any of this type, while *Touchstone* allocated 38.4%, almost four times as many as either *English Firsthand* or *Headway*. *Compounds* and *phrasal verbs* also showed significant variation among coursebooks. *Phrasal verbs* were also covered significantly less than other MWI, with *Headway* allocating the most coverage at 14.2%.

Table 3

MWI types by coursebook

Coursebook	% compounds	% phrasal verbs	% expressions	% other collocations
Smart Choice 2	35.2	4.3	0.0	60.5
Touchstone 2	13.4	3.0	38.4	45.2
English Firsthand 1	44.2	9.4	12.3	34.1
New Headway Pre-int.	30.0	14.2	11.4	44.4

R-Scores by Coursebook

R-scores for all 811 MWI identified were then calculated in order to represent MWI frequency and range. The average MWI R-score, statistical range, and standard deviation (SD) was also calculated for each coursebook (see Table 4).

Table 4

R-score average, range, and standard deviation (SD) by coursebook

Coursebook	Average R-score	Lower range	Upper range	*SD*
Smart Choice 2	4.0	0.0	97.2	11.9
Touchstone 2	23.3	0.0	494.8	66.6
English Firsthand 1	12.3	0.0	340.0	29.9
New Headway Pre-int.	16.9	0.0	367.7	41.1

The results showed significant variation in average R-scores over the four coursebooks. *Touchstone*, with a score of 23.3, had easily the highest average R-score. *Smart Choice*, with 4.0, had easily the lowest average R-score, which was almost six

times lower than the *Touchstone* score. However, average scores can often be misleading by being distorted by disproportionately high or low individual scores. For example, a single R-score that is much higher than other R-scores in the same group will produce a significantly higher average R-score value. As the consistency of coursebook R-scores also needed to be investigated, it was decided to include statistical ranges and standard deviation (SD) values for each coursebook (see Table 4).

The range and SD values showed a very wide spread of R-scores over all coursebooks (see Table 4). Examples of individual MWI that scored the highest R-scores are: *some of* (*Touchstone*, R-score = 494.8), *I don't know* (*Touchstone*, R-score = 356.4), and *move to* (*Touchstone*, R-score = 103.2). On the other hand, other R-scores were extremely low, such as *regular mail* (*Touchstone*, R-score = 0.08) and *wear* (*your*) *hair in cornrows* (*Touchstone*, R-score = 0.0).

Comparing coursebooks, there were noticeable differences among upper-range scores, with *Touchstone* at 494.8 and *Smart Choice* at 97.2. Looking at SD scores, *Touchstone* also had the highest score of 66.6, around six times higher than the lowest score which was *Smart Choice* at 11.9. These noticeable variations in R-scores show not only very inconsistent MWI frequency and range values among the different coursesbooks but also within individual coursebooks. Furthermore, the overall proximity of the average R-scores to the lower range values suggests a significant proportion of the coursebook MWI to have very low R-scores.

Disproportionately Low R-Scores

With the above results in mind, it was decided to do further study on the MWI with very low R-scores. At present, there appear to be no clear criteria in the literature that can be used to determine what an appropriate minimum frequency or range for different learner proficiency levels is. Consequently, for the purpose of this study, it was decided to investigate the percentage of R-scores under 0.5. This value was chosen because it would be equivalent to a single 0.5-word-per-million word value, and accordingly, represent the least common 225 words (0.00005%) in the 450-million-word BOE corpus. To put this in perspective, single words with equivalent R-scores are *rejectionist* (64 occurrences, R-score = 0.5) and *microflora* (22 occurrences, R-score = 0.3). Similar MWI would likely have limited usefulness to a learner at this proficiency level, even for receptive use. Table 5 shows the percentage of MWI with R-scores under

0.5 by coursebook.

Table 5

R-scores under 0.5 by coursebook

Coursebook	% MWI
Smart Choice 2	28.5
Touchstone 2	26.2
English Firsthand 1	25.9
New Headway Pre-int.	10.4

Higher values in the table indicate higher percentages of MWI R-scores under 0.5 and therefore greater proportions of the MWI being less representative of authentic language. From the table, it is clear that *Headway* did comparatively well with 10.4% of its MWI having R-scores under 0.5. The three remaining coursebooks, *Smart Choice*, *Touchstone*, and *English Firsthand*, all had over 25% of their MWI under 0.5. On average, over 20% (185 MWI) of the MWI from all the coursebooks investigated had R-scores under 0.5.

Further Investigation of Disproportionately Low R-Scores

In order to obtain specific examples of MWI with extremely low frequency and range values it was then decided to further identify MWI with even lower R-scores. Investigating R-scores under 0.05, a total of 45 individual MWI were then identified. Within this total, 22 MWI did not occur even a single time in the corpus. For a list of examples of the MWI identified see Table 6.

Table 6

Examples of R-scores under 0.05

Coursebook	Example
Smart Choice 2	*storm chasing* (R-score = 0.04) *dancing ability* (R-score = 0.04) *drive (a) racecar* (R-score = 0.04)
Touchstone 2	*Turkish rug* (R-score = 0.04) *I kind of like cold weather* (R-score = !) *wear (your) hair in cornrows* (R-score = !)
English Firsthand 1	*computer table* (R-Score = 0.04) *poetry slam* (R-score = 0.02) *mini-notebook computer* (R-score = 0.0) *How do you say (that) in English?* (R-score = 0)
New Headway Pre-int.	*in somebody's footsteps* (R-score = !) *That's really kind of you* (R-score = !)

Note. ! = Does not occur in corpus

Discussion

Regarding the extent that MWI are introduced in the coursebooks, the results from this study appear to reflect the general importance of MWI outlined in the literature. On average, the proportion of MWI introduced reflects their level of occurrence in authentic language estimated in previous corpus studies (Biber & Conrad, 1999; Erman & Warren, 2000). However, as with Koprowski's (2005) study, this study found inconsistencies among the coursebooks regarding proportions and types of MWI introduced. *Touchstone 2* allocated 47.9% for its target vocabulary to MWI, while *New Headway* allocated a significantly lower 18.0%. Additionally, while 40% of the *Touchstone 2* MWI were *expressions*, this type was not introduced at all in *Smart Choice 2*.

The MWI R-score results from this study also concur with findings from previous studies (Koprowski, 2005; McAleese, 2013), indicating that not only are these scores inconsistent among the coursebooks, but around 20% of the items identified have very low R-scores including a number of MWI not appearing even once in a 450 million word corpus.

While coursebook writers may have other considerations when selecting MWI, it is difficult to think of a justification for including items such as *Turkish rug* (*Touchstone*

2, R-score = 0.04) or *storm chasing* (*Smart Choice 2*, R-score = 0.04) at this proficiency level. Furthermore, a number of the low R-score MWI could be easily substituted by alternative or shorter items, producing much higher R-scores. For example, in *English Firsthand*, the compound *mini-notebook computer* gave an R-score of 0.0, but simply substituting the item with *laptop computer* would increase the R-score to 2.5. In other cases, it appeared coursebook writers had tried to use MWI when a single-word item might have been a more appropriate choice. For example, in *English Firsthand*, *flower vase* is introduced with an R-score of 0.18, when simply using *vase* would have produced an R-score of 7.2.

Significantly, there were also a number of longer-word-length MWI introduced with very low R-scores. Longer-length MWI, by their nature, can be expected to have much lower frequencies (Biber & Conrad, 1999); and, as with a number of the *expressions* identified in this study, very low R-scores have resulted. A strong argument could be made for the inclusion of a few of these longer MWI due to their potential usefulness to the target learner. For example, *How do you say* (*that*) *in English?* (*English Firsthand*, R-score = 0) would most likely have pragmatic usefulness in the classroom. However, in many cases, these *expressions* could be broken into shorter sentence stems that would result in higher R-scores. For example, the MWI *I kind of like cold weather* (*Touchstone*, R-score = !) could be shortened to the sentence stem *I kind of like . . .*, giving an R-score of 0.7.

Frequency and range are widely considered to be primary factors for vocabulary selection in contemporary teaching materials as they provide language that the learners are most likely to reencounter in different contexts (Nation & Waring, 1997; Sinclair, 1991). The large proportion of low R-scores found in this study are cause for concern, particularly for lower proficiency-level learners, whose materials would generally be expected to start with the most commonly used MWI. Learners and teachers will be assuming that there is a clear reason for any material to be included in coursebooks, particularly those from well-established publishers. In the case of the coursebooks investigated, it appears that a significant number of MWI included have been selected in an unsystematic way. As Lewis (1997) points out regarding coursebooks, "the printed word has the power to authenticate itself" (p. 182), and accordingly, failure to at least remove suspect material is a disservice to learners. The use of corpus frequency and range data, while not the only consideration to be made, should provide an empirically

based and objective method for MWI selection.

Implications for Teaching and Coursebook Development

The scope of this study was limited to MWI in four coursebooks, so further research is needed before more general recommendations can be made. However, the findings in this study suggest that MWI frequency and range data need to be considered more in vocabulary selection. Large corpora and corpus-derived teaching materials are now commonplace and can be easily checked by coursebook writers, teachers, and even learners, to estimate language representativeness. Even large-scale corpora such as the COCA (Corpus of Contemporary American English) and BNC (British National Corpus) provide limited free online access to the public. Corpus-derived teaching resources, such as the *Phrasal Expressions List* (Martinez & Schmitt, 2012) and lists of frequent spoken collocations (Shin & Nation, 2008), can also provide proficiency-level-specific starting points for MWI selection. Although many teachers may not be involved in the coursebook selection process, such resources could also easily be used to supplement existing coursebook MWI. As for coursebook writers, they have a responsibility to at least check for and remove material that is of questionable usefulness to target learners. Accordingly, it stands to reason that very unrepresentative MWI should also be removed or at least annotated in some manner.

Limitations of this Study

It is important to recognize that there are a number of limitations to this study. The method used in this study to identify MWI takes the items in the forms the coursebook writers have chosen to represent them in on the target vocabulary lists. Writers may consciously choose to represent MWI as longer strings of words in order to provide more context for the learners. For example, extending the MWI *wear* (*your*) *hair in* to *wear* (*your*) *hair in cornrows* may give a learner more context but result in much lower frequency values. However, it can also be argued that writers should be including MWI in their most representative form on the vocabulary lists in the first place. Writers may also have reasons other than frequency or range when drawing attention to MWI. For example, MWI such as *allergic to* or *use by date* may have very low frequency or range values but might be of practical use for a language learner studying abroad. It is also recognized that the coursebooks will likely include many

useful MWI that are not explicitly identified in vocabulary lists, and these were therefore not addressed in this study.

Another limitation is in using a corpus to estimate MWI representativeness. Even very large computer-based corpora can only provide an estimate of language authenticity. Although the The Bank of English (1991) used in this study includes over 450 million words and 20 subcorpora covering a range of language types and genres, it includes relatively more written language and British English, which could distort the results to a certain degree. Finally, some small methodological limitations were revealed using Koprowski's (2005) approach to calculating R-scores. For instance, the pilot study used in this study revealed that the small number of MWI exhibiting polysemy gave slightly inconsistent R-score values.

Conclusions

This study has investigated MWI in a small number of upper-beginner proficiency-level coursebooks. It has showed that while coursebook writers have endeavored to include MWI in the materials, there are large inconsistencies on the proportions and types of MWI addressed. Using a large scale corpus, this study has also shown a significant proportion of the MWI to have disproportionally low frequency and range scores suggesting that they are unrepresentative of authentic English and therefore of limited use to the target learners. However, further study is required before wider-reaching conclusions can be made. Further study checking for MWI frequency and range with other or multiple large-scale corpora, or even using an alternative method for determining representativeness, could provide more support to these results. Importantly, investigating coursebook writers' reasons for introducing low frequency and range MWI may also help clarify any alternative justifications used.

References

Biber, D., & Conrad, S. (1999). Lexical bundles in conversation and academic prose. In H. Carter, R., & McCarthy, M. (1988). *Vocabulary and language teaching*. London, England: Longman.

Ellis, N. C. (1996). Sequencing in SLA: Phonological memory, chunking, and points of

order. *Studies in Second Language Acquisition, 18*(1), 91-126. From http://dx.doi.org/10.1017/S0272263100014698

Erman, B., & Warran, B. (2000). The idiom principle and the open choice principle. *Text, 20*(1), 29-62. From http://dx.doi.org/10.1515/text.1.2000.20.1.29

Helgesen, M., Brown, S., & Wiltshier, J. (2010). *English firsthand 1 Student's Book*. Hong Kong: Pearson Longman.

Hsu, J. T. (2008). Role of the multi-word lexical units in current EFL/ESL coursebooks. *US-China Foreign Language, 6*, 27-39.

Koprowski, M. (2005). Investigating the usefulness of lexical phrases in contemporary coursebooks, *ELT Journal, 54*(9), 322-332. From http://dx.doi.org/10.1093/elt/cci061

Lewis, M. (1994). *The lexical approach: The state of ELT and a way forward*. London, England: Language Teaching Publications.

Lewis, M. (1997). *Implementing the lexical approach*. Hove, England: Language Teaching Publications.

Mackey, W. F. (1965). *Language teaching analysis*. London: Longman

Martinez, R., & Schmitt, N. (2012). A phrasal expression list. *Applied Linguistics, 33*(3), 299-320. From http://dx.doi.org/10.1093/applin/ams010

McAleese, P. (2013). Investigating multi-word items in a contemporary ELT coursebook. In N. Sonda & A. Krause (Eds.), *JALT2012 Conference Proceedings* (pp. 321-329). Tokyo: JALT.

McCarthy, M., McCarten, J., & Sandiford, H. (2005). *Touchstone 2 student's book*. New York, NY: Cambridge.

Meunier, F., & Gouverneur, C. (2007). The treatment of phraseology in ELT Coursebooks. In L. Hildalgo, L. Querera, & J. Santana (Eds.), *Language and computers: Corpora in the foreign language classroom61* (pp. 119-139). Amsterdam, Netherlands: Rodopi.

Moon, R. (1997). Vocabulary connections: multi-word items in English. In N. Schmitt & M. McCarthy (Eds.), *Vocabulary: Description, acquisition and pedagogy* (pp. 40-63). Cambridge, England: Cambridge University Press.

Nation, P., & Waring, R. (1997). Vocabulary size, text coverage and word lists. In N. Schmitt & M. McCarthy (Eds.), *Vocabulary: Description, acquisition and pedagogy* (pp. 6-19). Cambridge, England: Cambridge University Press.

Nattinger, J. R., & DeCarrico, J. S. (1992). *Lexical phrases and language teaching*. New York, NY: Oxford.

Shin, D., & and Nation, P. (2008). Beyond single words: The most frequent collocations in spoken English. *ELT Journal, 62*(4), 339-348. From http://dx.doi.org/10.1093/elt/ccm091

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

Soars, L. & Soars, J. (2007). *New headway pre-intermediate* (3rd ed.). *Student's book*. New York, NY: Oxford.

The Bank of English (1991). An English language corpus developed by COBUILD (a division of HarperCollins Publishers) and The University of Birmingham. For more information: http://www.titania.bham.ac.uk/docs/about.htm

White, R.V. (1988). *The ELT curriculum: Design, innovation and management*. Oxford, England: Basil Blackwell.

Wilson, K. (2011). *Smart choice 2 student's book* (2nd ed.). New York, NY: Oxford.

Wood, D. (2007). Mastering the English formula: Fluency development of Japanese learners in a study abroad context. *JALT Journal, 29*(2), 209-230.

Wray, A. (2008). *Formulaic language: Pushing the boundaries*. New York, NY: Oxford.