

Have the Modal Verb Phrase Structures Been Well Presented in Malaysian English Language Textbooks?

Laleh Khojasteh (Corresponding author)

Department of Foreign Languages, Shiraz Branch, Islamic Azad University, Shiraz, Iran

Tel: 60-123-804-248 E-mail: khojastehlaleh@yahoo.com

Reza Kafipour

Department of Foreign Languages, Shiraz Branch, Islamic Azad University, Shiraz, Iran

Tel: 98-917-315-8394 E-mail: rezakafipour@gmail.com

Received: October 21, 2011

Accepted: November 9, 2011

Published: March 1, 2012

doi:10.5539/ells.v2n1p35

URL: <http://dx.doi.org/10.5539/ells.v2n1p35>

Abstract

To date, many studies focused on the behavior of English grammatical structures in various national and international textbooks and questioned the authenticity of the language and the grammar presented in these textbooks. To see if Malaysian English language textbooks are free from these issues, this study compared the ways in which modal verb phrase structures are presented in these textbooks and real language use. This was addressed in one research question and the design applied was qualitative content corpus analysis. The findings reveal that evidently, there are very great differences in the relative frequency of use of the nine verb phrase structures in which modals can occur in real language and the ones in Malaysian English language textbooks (Form 1-5). The survey result also reveals that secondary learners are not really exposed to other verb structures, particularly structures with passive, progressive and perfect aspects.

Keywords: Modal verb phrase structures, English Grammatical structures, Real language use, Grammar teaching

1. Introduction

It has been noted by many grammarians and applied linguistic researchers that modal auxiliary verbs count among one of the most grammatical trouble makers for second and foreign language learners and teachers. For instance, Palmer (1974) notes that the complexity of modal auxiliary verb forms and their semantic functions cannot be compared with any other grammatical structures. Several other researchers highlighted the difficulties ESL and EFL learners usually have in handling this grammar feature (Celce-Murcia & Larsen Freeman, 1999; Decapua, 2008; Manaf, 2007). However, it has been argued that although lack of grammatical equivalence between learners' target language and first might cause a great challenge for them to produce a particular language structure, lack of fit between descriptions of language phenomena in textbooks and real communication situations may play a greater role in this deficiency (Romer, 2005).

To date, many studies questioned the authenticity of the language, grammar, pragmatics and vocabulary and phraseology presented in various national and international textbooks, and strongly noted that if learners were presented with appropriate grammatical structures in line with real language use, they would have encountered fewer difficulties handling relevant structures in communicative situations (Biber, Conrad, Reppen, Byrd and Helt, 2002; Gilmore, 2004; Meunier and Gouverneur, 2009; Romer, 2005, 2004a, 2004b; Nordberg, 2010, Mukundan and Khojasteh, 2011; Vellenga, 2004). Comparing the authenticity, grammar and vocabulary in textbooks and reference corpora such as Longman Spoken and Written English (LSWE) Corpus and British National Corpus (BNC), these studies indeed demonstrate that by ignoring frequent features of the language spoken or written by real language users, many textbooks implicitly portray these linguistic features as monolithic phenomena, which behave in the same way regardless of different contexts and situations of use. As regards to the grammar one good example is Barbieri & Eckhardt's (2007) study which was based on the comparison of reported speech in seven ESL/EFL grammar textbooks used in Germany and Longman Spoken and Written English (LSWE) Corpus. The findings of their study show that textbooks neglect important information on the use of this structure in real language in general and direct reporting verbs in specific. For example they reported that although in real language use *say* and *tell* are

the most used reporting verbs, EFL textbooks in Germany tend to under represent most additional reported speech other than *say* and *tell*. Furthermore, there is a general neglect of information on the tense backshifting rule, register and context-dependent variation. Finally they conclude that the books were not written based on empirical studies because it is not clear which principles informed textbooks authors' decisions about which reporting verbs to present. Likewise, Mukundan and khojasteh (2010) investigated 9 central modal auxiliary verb forms and their verb phrase structures in a pedagogic Corpus in Malaysia and BNC and found a mismatch between the two. The finding reveals that for certain modal auxiliaries; there is a discrepancy between modal frequency order in the textbook corpus and the BNC. For example, when modal *could* is considered the fourth most common modal in BNC, this modal is standing in the seventh position in the textbook corpus and surprisingly enough *could* is not even taught implicitly in Malaysian secondary school level.

If such is the case, present-day textbooks might lack a broad empirical foundation which leads us to the first reason for carrying out such a study; because non-empirically-based teaching materials can be increasingly misleading. More often than not, these textbooks may be based on a limited collection of example sentences which are not attested instances of language use but rather intuitive examples which have been invented for the purpose of illustrating a particular point, or which have been taken from earlier linguistic studies and grammar books. In this case, material developers unwittingly miss the quantitative data on the distribution of different modal auxiliary verb forms or functions and on co-occurrences with context features such as common modal verb phrase structures.

By comparing a textbook corpus with reference corpora or real-language corpora, indeed, it is imperative to examine the language to which learners are exposed to in order to develop more effective pedagogical materials for EFL/ESL learners. Having mentioned that, this study does not suggest for EFL/ESL learners to be members of native speaker communities. From corpus linguistics point of view, native speaker corpus is not intended to serve as a language model to be exactly imitated by L2 learners (Gavioli and Aston, 2001). They suggest, however that, reference corpora such as BNC can guide the researchers to identify the mismatch between EFL/ESL materials and native speaker corpus, and can cater for usefulness, difficulty and learnability of certain grammatical structures in these materials. Insights such as high frequent forms from native speaker corpus, from Nation's (2010) view point "give [the language learner] a better return for learning efforts" (p. 1).

Accordingly, since in this study, it is argued that "the verb phrase is one of the most important elements of the sentence" (Mindt, 2000, p. 84), the question was asked in the contextual analysis of modal auxiliary verbs was:

"How extensively are the distributions and the categories of words that colligate with modal auxiliary verbs presented in Forms 1-5 Malaysian English language textbooks in line with their usage in real language use?"

2. Methodology

2.1 Population and Sampling

For the purpose of this study, the population for the English language corpus was sourced from Malaysian English language textbooks used for secondary Malaysian students of Form 1 through Form 5. The corpus used in this study was compiled by Mukundan and Analeka (2007). This pedagogic corpus consisted of 311,214 running words. This corpus can be classified as a "pedagogic corpus" as it is coined by Willis (1993) and defined by Hunston (2002) as a collection of data that "can consist of all the course books, readers etc. a learner has used" in an ESL/EFL language learning program (p.16). This textbook corpus has been used in many significant corpus-based studies (Menon, 2009; Mukundan&Roslim, 2009; Mukundan&Anealka, 2007; Mukundan, 2004; Mukundan&Khojasteh, 2011), and the results have shed light on the lexical and grammatical structures that the secondary Malaysian students are exposed to in their textbooks.

2.2 Instrumentation

The WordSmith Tools Version 4.0 was used almost entirely for the purpose of this research, because it has been recognized as a capable and suitable tool to support quantitative and qualitative data analysis by many researchers (Baker, 2006; Bondi, 2001; De Klerk, 2004; Flowerdew, 2003; Henry & Roseberry, 2001; Menon, 2009; Mukundan, 2004; Mukundan&Roslim, 2009; Nelson, 2001; Scott, 2001).

2.3 Detailed Analysis of Data

While collocation is instantly identifiable on the vertical axis of an alphabetical concordance, colligation represents a step in abstraction and is therefore less immediately recognizable unless the text is coded with precisely the required grammatical information. Since this corpus was not the annotated corpus, and the aim of this question is to analyze the modal's grammatical forms or constructions, it was essential for us to go through the coding process in order to identify the modal verb phrase structures existing in the textbook corpus. The framework used for this study to identify modal verb phrase structures in this textbook corpus is a combination of structures found by Mindt (1995)

in a corpus of fiction texts and Kennedy (2002) in the BNC. In the former, Mindt (1995) reported five modal verb phrase structures including 1) modal + bare infinitive, 2) modal + passive infinitive, 3) modal + progressive infinitive, 4) modal + perfective infinitive, 5) modal + perfect passive infinitive. Kennedy (2002) on the other hand, added four more verb phrase structures to Mindt's (1995) list including modal alone, modal + be + being + past participle (or adjective), modal + have + been + present participle, and modal + have + been + being + past participle. For this study, it was decided to keep the common structures between the two lists but only add one more structure from Kennedy's (2002) list (modal alone). It was decided not to do a query for the other structures because according to Kennedy (2002, p. 89) these structures are "extremely rare". Another reason that we decided to eliminate these three structures from our list is that Mukundan and Khojasteh (2011) found that there is not even a single instance of these structures in their corpus-based analysis on Form 1, 2 and 3 Malaysian English language textbooks.

Based on the above mentioned structures we coded the occurrence of each modal auxiliary. The reason behind the fact that we did not simply look for the grammatical choices that co-occur with modal auxiliary verbs and conversely doing this based on pre-existing structures is that Hoey, Mahlner, Stubbs and Teubert (2007) and Hunston and Francis (2000) suggest studying colligation with making use of existing grammatical terminologies. These terminologies according to Hoey *et al.*, (2007, p. 35) is the product of "pre-corpus investigations".

These structures plus a code set which was used for this study with a sample example can be seen in below:

- 1) Modal alone (0) who will go? *I will.*
- 2) Modal + infinitive (infin) Sam *can* swim./ She *must* be hungry.
- 3) Modal + be + past participle (be past ptc) It *should* be replaced.
- 4) Modal + be + present participle (be presptc) They *will* be arriving soon.
- 5) Modal + have + past participle (have ptc) He *might* have done it.
- 6) Modal + have + been + past participle (have been past ptc) It *should* have been fixed.

Apart from the above assigned codes, another code was assigned for "none of the above items" (N) since in the textbook corpus there were incomplete utterances in exercise units that did not provide us with enough span around the node word in order to make them decide what verb phrase structure the modal carries.

After the coding, which in itself was analytical, we undertook several additional steps. These steps, too, were done within the framework of the proposed research question. First, we summarized the findings identified during the coding and then identified and articulated the patterns and relationships among our findings so that we could answer our research question. Then, we related these more involved findings to the results of major corpus-based findings on modal verb phrase structures such as Kennedy's (2002), Mindt (2000) and Mindt (1995). The last step allowed us to put our findings into perspective.

3. Results

In order to determine these shares, absolute occurrences for each possible form were retrieved from the database, applying the query strategy explained earlier. Table 1 displays the retrieved absolute numbers of all modals.

In this textbook corpus, constituting the greatest share is the modal plus infinitive (e.g. *can go*) followed by the modal + be + past participle (e.g. *can be cleaned*) structure. Conversely, all other forms (modal alone and all perfect and fragmentary forms) are comparatively rare. Out of 1288 frequency occurrences of *can*, 1,067 (82.8%) of them occurred in structure 2 and 167 (12.96%) in structure 3, which together account for almost 96% of all *can* cases throughout Forms 1 to 5. A similar trend is found for all the other modals without any exceptions, for example, (*would*, 96%; *could*, 94.5%; *will*, 92%; *should*, 90%). (1), (2) and (3) are sample sentences for structure 2 from textbook corpus.

Examples: (1) You *can* interview your family members or friends, or collect articles from the newspapers or the Internet.

(2) Sufian and his friends *would like* to invite their schoolmates to join them in their project.

(3) You *won't regret* it!

However, considering passive constructions alone, this structure (e.g. *can be used*) is only popular for some modals such as *can* (167 instances), *will* (81), and *should* (74) while it is moderately frequent in *would* with 20 instances, *could* (19), *may* (18), *must* (35). Finally, not frequent at all in this structure are the cases of *shall* and *might* with 1 and 3 instances each respectively (see (4) and (5) for sample sentences from the textbook corpus).

Examples: (4) Some of the words *can be used* more than once.

(5) Puan Lim *will be transferred* to a new school next week.

As for modals with progressive aspect (e.g. *will be seeing*), we can clearly see that this aspect is rather rare for almost all modals throughout the textbooks except for *will* with 16 instances throughout the Forms 1 - 5 textbooks. The paucity of this aspect is more obvious when we look specifically at the distribution of progressive aspect within Forms 1 - 5 textbooks. For instance, there is only one instance of this aspect in Form 1 and Form 5 and only 2 instances in the case of *must* in Form 4 (see (6) and (7) for sample sentences from textbook corpus).

Examples: (6) At first, the lessons *can be confusing* but after the whole course, you'll be fine.

(7) Define the scope of your speech - what you *will be talking* about.

Table 4.5 also shows that modals *can*, *would*, *could* and *shall* do not have any proportion in the textbook corpus in structure 5 (e.g. may have gone). However, this structure has a minor occurrence in the case of *might*, *may*, *will* and *should* throughout the textbooks. The most frequent modal marked for this structure is *must* with 14 instances (see (8) and (9) for sample sentences from textbook corpus).

Examples: (8) I know it *must have been hard* for my parents but they knew what was best for him.

(9) You may want to include the events that *might have happened* the day before.

The least common modal verb phrase structure among all is modal alone. Except for modals *can* and *will* that was marked for this aspect 6 and 5 times respectively, the rest of the modals did not show enough tendencies for structure 1 (see (10) and (11) for sample sentences from textbook corpus).

Examples: (10) Won't you dance with me? I *shouldn't*.

(11) Yes, I would.

(4) Is there anything I can do to help? Yes, you can.

Structure 6 is only used once in Form 3 for the modal *could*. The only sentence found in Form 3 that represents this structure is (14).

Example: (14) My son *could* have been killed.

Finally, there are certain modals that have been tagged as 'N' because they did not follow any of the aforementioned aspects. In some cases such as *will* with 62 instances, *can* (47) and *must* (54) modals did not follow any patterns at all. Examples (12) and (13) are sample sentences from some of these modals.

Examples: (12) How often can?

(13) Give suitable responses to these sentences using 'must' or 'must not'.

4. Summary and Discussion

The results presented in Table 1 shows that the most used modal verb phrase structures for all modals in textbook corpus is structure 2 (modal + bare infinitive). On the one hand, based on the findings of Kennedy (2002), this study shows a rather one-sided picture being portrayed by Malaysian English language textbooks as regards structure 2. Kennedy (2002) reported that in BNC, the modal + bare infinitive structure accounts for 76% of all modal tokens in the corpus, whereas this study shows that the frequency count for this structure has jumped to 81%. On the other hand, this finding matches Mindt's (2000) reported figure for structure 2 that is exactly 81%. Since the discrepancy found between the results for our textbook corpus and BNC is rather marginal, it is slightly unfair to criticize the textbook authors simply on the basis of their treatment of structure 2. However, it is important to discuss further these results based on the extent which each individual modal contributes to each structure. In that case, for example, we can see that there is a mismatch between the shares each modal has for structure 2 in the results compared with real language use. Mindt (2000) and Kennedy (2002) reported that the most used modal in structure 2 is *will* followed by *can*, *would*, *could*, *may*, *should* and *shall*. However, the results of this study show that *can* with 1,067 instances outweighs that of *will* (773) and *should* (373) exceed *would* (370) and *may* (285). It is also worth mentioning that *must* with 5.72% frequency occurrences have been definitely overused in this structure throughout the textbooks because *must* with this pattern (less than 2.9%) did not even appear among other ranked modals in Mindt's (2000, p. 589) table. A tentative explanation for this bias is that proportionally *can* is the most frequently modal used in textbook corpus while this modal stands in the third place in BNC. Indeed, since Mukundan and Khojasteh (2011) came to the conclusion that the frequency and rank order of modal auxiliary verbs found in the English language textbooks used in Forms 1-5 are not empirically based, the consequences in terms of the weight given to each modal verb phrase structure are hardly surprising. *Would* for example, the second most frequent modal

auxiliary in all major corpora, stands also second (78.8%) in terms of structure 2 according to Kennedy (2002). Similarly, *would* standing in fourth place in terms of overall frequency occurrences in Malaysian English textbooks, occupies the same place (fourth) in terms of structure 2 in textbook corpus.

Regarding structure 3, Kennedy (2002) reported that in BNC *can* is the most frequently-occurring modal (27.2%) followed by *may* (19.1%), *will* (18.1%), *should* (13.7%) and *could* (10.7%). The finding from this pedagogic corpus is in line with the one in Kennedy's (2002) except for the modal *may* which was not given enough emphasis by textbook writers with its rare occurrences throughout Forms 1 to 5.

As it can be seen in Table 1, structure 4 is only used for *will* with 16 instances throughout Form 1 to 5 textbooks. This structure is notably rare in the case of the rest of the modals with for example *should* getting 6 hits and *can* with only 2 hits. When this result is compared with the frequency occurrences reported by Kennedy (2002), it is found that this structure has been under-presented in Malaysian textbooks. This structure is highly used by *will* with 41.3% and *would* (20%) in BNC, whereas in textbook corpus it is only used 1.7% and 0.25% respectively. Likewise, structure 5 does not have high use in any of the Forms 1- 5 textbooks. For some modals such as *can*, *would*, *could* and *shall* there are zero instances of this structure which leads us to conclude that structure 5 is underused in Malaysian English language textbooks. Following structures 2 and 3, structure 5 is the most frequently used structure by native speakers (Kennedy, 2002). Specifically, *would* (37%) has the highest proportion in this structure followed by the modal *will* in BNC (Kennedy, 2002). However, as it can be seen in Table 1, the results show that this structure did not even occur once for *would* in textbook corpus. This shows how textbook writers have ignored such insights from corpus-based analysis. This structure, Kennedy (2002, p. 89) posits, should be given "high priority" in language pedagogy.

For structure 1 (modal alone) we can also see that except for *can* (6 instances), *will* (5) and *would* (3) with their insignificant frequency occurrences throughout the textbooks, the rest of the modals show zero frequency towards this structure. In terms of popularity of this structure among all nine central modals, Kennedy (2002) reported that 83% of the modal tokens come from *can*, *will*, *would* and *could*. In terms of learnability, like structure 3, it possibly means that Malaysian learners may have difficulty learning this structure when it is not even featured in their textbooks explicitly. Lack of proper understanding of this structure for Malaysian learners, perhaps create difficulties in producing it in spoken negative contexts because this structure accounts for 14% of the modal tokens in the aforementioned context (Kennedy, 2002). As Conrad (2000) posits, having important discourse functions in particular registers, is a very good reason for textbook writers to consult corpus linguistics in terms of descriptions and frequency information.

Structure 6 is the only structure that showed the lowest frequency occurrences for all modals. Surprisingly, there is only 1 instance in terms of modal *could* in the Form 3 textbook. However, considering the fact that this structure is not really frequent in the BNC (Kennedy, 2002), it is perhaps understandable why textbook writers did not give any share to this particular structure in Malaysian textbooks.

Finally, it is important to note that except for structure 2, none of the above mentioned structures has been taught to students in any of the Forms from 1 to 5 to inform Malaysian students about their semantic and pragmatic information. This perhaps is one of the main reasons that Malaysian learners avoid the use other structures such as modal + perfect infinitive and modal + progressive infinitive in their essays according to Manaf (2007). Although Manaf (2007) failed to report on the use of structures 3 and 4 by Malaysian students in EMAS corpus, but from the 47 examples she has provided in terms of syntactically accurate and inaccurate modal verb phrase structures, we can clearly see that none of these examples portray the use of structures 3 and 4 in Malaysian students' essays. If these structures are not taught, Decapua (2008, p. 213) believes, "a great deal of information that is not necessarily immediately obvious to students" will go unnoticed by ESL/EFL learners. In addition to that, as part of the consciousness-raising for teachers, Kennedy (2002) believes the analysis of modal verbs in corpus-derived data should be integrated in language education "not just because they exist, but because they are used often enough to justify inclusion in instruction" (p. 89).

5. Conclusion

Apart from many criteria proposed for principled selection of syllabus designs, frequency and range have been highly recommended (Kennedy, 2002; Koprowski, 2005; Mindt, 2000; Romer, 2004a; Sinclair, 1991 and many more). Nation and Waring (1997, p.17) state that applying frequency information in textbooks ensures that students are exposed to the language they most probably meet again outside the classroom walls. Likewise, Romer (2004a, p.152) believes we should always make sure that the language students are exposed to in their textbooks is as close as possible to the language they will likely to be confronted with in natural communicative situations. However, evidently, there are very great differences in the relative frequency of use of the nine verb phrase structures in which modals can occur

in real language and the ones in Malaysian English language textbooks (Form 1-5). For example, despite the fact that *will* is the most frequently used modal in structure 2 (modal + infinitive) in real language use, in textbook data *can* is predominantly used in this structure. The survey result also reveals that secondary learners are not really exposed to other verb structures, particularly structures with passive, progressive and perfect aspects. Perhaps that would be one of the reasons why Rosli and Edwin (1989) in their study of errors in Form 4 student compositions found that verb forms and the verb aspects of modals are the most problematic for Malaysian learners. According to Kennedy (2002), the above mentioned structures including structure 1 (modal alone) should take priority in language pedagogy. Accordingly, explicit teaching of these structures is highly recommended in Malaysian English textbooks in order for students to differentiate between the verb forms used in each. This, in turn, Manaf (2007) suggests, will help Malaysian students to have a better understanding of the structures used particularly in writing. From 1 to Form 5 Malaysian English textbooks indeed enjoy many positive features; their coverage of modal verb phrase structures is only a small part of the books. However, the most salient facts reflected from natural language corpora should not be ignored in the textbooks.

References

- Barbieri, F., & Eckhardt, S. (2007). Applying corpus-based findings to form-focused instruction: The case of reported speech. *Language Teaching Research*, 1(3), 319–346. <http://dx.doi.org/10.1177/1362168807077563>
- Baker, P. (2006). *Using Corpora in Discourse Analysis*. London: Continuum.
- Biber, D., S. Conrad, R. Reppen, P. Byrd, and M. Helt. (2002). Speaking and writing in the university: A multi-dimensional comparison. *TESOL Quarterly*, 36, 9-48
- Bondi, M. (2001). Small corpora and language variation: Reflexivity across genres. In M. Ghadessy, A. Henry, & R. Roseberry (Eds.), *Small corpus studies and ELT* (pp. 135-174). Amsterdam/ Philadelphia: John Benjamins Co.
- Celce-Murcia, M., & Larsen-Freeman, D. (1999). *The Grammar book: an ESL/EFL teacher's course*. 2nd ed. Boston: Heinle&Heinle.
- Decapua, A. (2008). *Grammar for teachers: A Guide to American English for Native and Non-Native Speakers*. New Rochelle: Springer.
- De Klerk, V. (2004). The use of 'actually' in spoken Xhosa English: A corpus study. *World Englishes*, 24(3), 275-288. <http://dx.doi.org/10.1111/j.0883-2919.2005.00410.x>
- Flowerdew, L. (2003). A combined corpus and systematic-functional analysis of the problem solution pattern in a student and professional corpus of technical writing. *TESOL Quarterly*, 37(3), 489-511. <http://dx.doi.org/10.2307/3588401>
- Gavioli, L., & Aston, G. (2001). Enriching reality: Language corpora in language pedagogy. *ELT Journal*, 55(3), 238-246. <http://dx.doi.org/10.1093/elt/55.3.238>
- Gilmore, A. (2004). A comparison of textbook and authentic interactions. *ELT Journal* 58(4), 363–71. <http://dx.doi.org/10.1093/elt/58.4.363>
- Henry, A., & Roseberry, R. L. (2001). Using a small corpus to obtain data for teaching a genre. In M. Ghadessy, A. Henry, & R. Roseberry (Eds.), *Small corpus studies and ELT* (pp. 93-113). Amsterdam/ Philadelphia: John Benjamins Co.
- Hoey, M., Mahlner, M., Stubbs, M., & Teubert, W. (2007). *Text, discourse and corpora: Theory and analysis*. London: Continuum.
- Hunston, S., & Francis, G. (2000). *Pattern Grammar. A Corpus-driven Approach to the Lexical Grammar of English*. Amsterdam: John Benjamins.
- Hunston, S. (2002). *Corpora in applied linguistics*. Cambridge, England: Cambridge University Press.
- Koprowski, M. (2005). Investigating the usefulness of lexical phrases in contemporary coursebooks. *ELT Journal*, 59(4), 322-332. <http://dx.doi.org/10.1093/elt/cci061>
- Meunier, F., & Gouverneur, C. (2009). New types of corpora for new educational challenges. In K. Aijmer (Ed.), *Corpora and language teaching* (pp.179-201). Amsterdam: John Benjamins.
- Mindt, D. (1995). *An Empirical Grammar of the English Verb: Modal Verbs*. Berlin: Cornelsen.
- Mindt, D. (2000). *An Empirical Grammar of the English Verb System*. Berlin: Cornelsen.
- Mukundan, J. (2004). A Composite Framework for ESL Textbook Evaluation. *PhD Thesis*. Universiti Putra Malaysia.

- Mukundan, J., & Anealka, A. H. (2007). A forensic study of vocabulary load and distribution in five Malaysian Secondary School Textbooks (Forms 1-5). *Pertanika Journal of Social Science and Humanities*, 15(2), 59-74
- Mukundan, J., & Khojasteh, L. (2011). Modal Auxiliary Verbs in Prescribed Malaysian English Textbooks. *English Language Teaching*, 4(1), 79-89
- Mukundan, J., & Roslim, N. (2009). Textbook Representation of Prepositions. *English Language Teaching*, 2(4), 123-130
- Nation, I. S. P., & Macalister, J. (2010). *Language Curriculum Design*. New York: Routledge.
- Nation, P., & R. Waring.(1997). Vocabulary size, text coverage and word lists. In N. Schmitt, and M. McCarthy (Eds), *Vocabulary: Description, Acquisition and Pedagogy* (pp. 6–19). Cambridge: Cambridge University Press.
- Nelson, M. (2001). *A corpus based study of business English and business teaching*. Oxford: Oxford University Press.
- Neuendorf, K. A. (2002). *The content analysis guidebook*. Thousand Oaks, California: Sage Publications.
- Nordberg, T. (2010). *Modality as portrayed in Finish upper secondary school EFL textbooks: A corpus-based approach*. Master's Thesis. University of Helsinki. [Online] Available: <https://helda.helsinki.fi/handle/10138/19357>
- Palmer, F. R. (1974). *The English verb*. London: Longman.
- Romer, U. (2004a). A corpus-driven approach to modal auxiliaries and their didactics. In J. Sinclair (Ed), *How to Use Corpora in Language Teaching* (pp. 185-199). Amsterdam: John Benjamins.
- Romer, U. (2004b). Comparing real and ideal language learner input: the use of an EFL textbook corpus in corpus linguistics and language teaching. In G. Aston, S, Bernardini, & D, Stewart (Eds.), *Corpora and Language Learners* (pp. 151-168). Amsterdam: John Benjamins.
- Romer, U. (2005). *Progressives, patterns, pedagogy. A Corpus-driven Approach to English Progressive Forms, Functions, Contexts and Didactics*. Amsterdam: John Benjamins.
- Rosli, T., & Edwin, M. (1989). Error analysis of Form Four English compositions. *The English Teacher*, XVIII, 110-124
- Sinclair, J. M. (1991). *Corpus, Concordance and Collocation*. Oxford: Oxford University Press.
- Scott, M. (2001). Comparing corpora and identifying key words, collocations and frequency distributions through WordSmith Tolls suite of computer programs. In M. Ghadessy, A. Henry, & R. L. Roseberry (Eds.), *Small Corpus Studies and ELT* (pp. 47-67). Amsterdam / Philadelphia: John Benjamins Publishing Co.
- Willis, D. (1993). *Syllabus, corpus and data-driven learning*. IATEFL Conference Report: Plenaries.

Table 1. Verb-phrase Structures of Modals in Form 1-5 Textbooks

Modal Structures	Can	Could	Will	Would	Should	Shall	May	Might	Must
S1	6	-	5	3	-	-	-	-	-
S2	1067	207	775	370	373	17	285	49	238
S3	167	19	79	20	72	1	18	3	35
S4	2	3	16	1	6	1	2	1	2
S5	-	-	1	-	6	-	5	5	14
S6	-	1	-	-	-	-	-	-	-
None of the above	8	10	62	6	36	3	34	28	54