

THE FREQUENCY DISTRIBUTION OF CENTRAL MODAL VERBS IN KURIKULUM MERDEKA'S TEXTBOOKS IN INDONESIA: A CORPUS-BASED ANALYSIS

Akhofullah¹, Ikmi Nur Oktavianti²

Universitas Ahmad Dahlan, Indonesia

¹ ahof1997@gmail.com, ² ikmi.oktavianti@pbi.uad.ac.id

Abstract

This study focuses on investigating the central modal verbs used in Kurikulum Merdeka's Textbooks in Indonesia. This study uses corpus-based analysis and categorizes it as descriptive qualitative research. The data sources are three series of Kurikulum Merdeka's textbooks for senior high school level in Indonesia. In supporting the investigation of linguistic features presented in the textbooks, this study uses the Corpus of Contemporary American English (COCA) and British National Corpus (BNC) as the control corpora to represent real language data. The result of this study shows that "can," "will," and "would" are the most frequently central modal verbs in the COCA, BNC, and kurikulum merdeka's textbooks, while the least frequent central modal verbs used among those three sources is "shall. However, "would" is not significant present in the textbooks compared to that in COCA and BNC. Moreover, the rest of the central modal verbs vary among the three sources, like "must" and "may."

Keywords: Corpus; Central Modal Verbs; Central Modal Verbs Frequency

INTRODUCTION

English textbooks have an essential role in the teaching and learning process. In providing knowledge, skills, and information, textbooks consist of learning guidelines in activities that guide students to achieve some competencies. A textbook allows the teachers and students to map out what will be learned in the classroom (Tomlinson, 2014). Hutchinson & Torres (1994) stated that textbooks are the providers of the texts, activities, and explanations that can be used for the learning process in the classroom. In line with Tayson and Woodward's (1989) in Harwood (2015), up to 90 percent of what goes on in the classroom is based on the textbooks, teachers are expected to follow the textbook less or more faithfully, and the exams are based on the textbook content. Thus, textbooks are essential for teachers and students in English Language Teaching (ELT). Despite the vital role of the textbook in ELT, many textbooks sometimes contain non-authentic materials; some words or sentences in the textbook are barely used in daily life. At the same time, the students are hoping to use English in real life.

Moreover, many materials have to be considered by the textbook developer to be applied in the textbook. One of the contents in the textbook that is commonly used in daily life but might be hard to learn by some students is modal verbs. Palmer (1979) in Rappen, Fitmaurice, & Biber (2002) argued that modals are more critical and complex than any English grammar. Holmes (1988) also stated that modal verbs have complexity because they express a wide range of meanings, and their linguistic device does not relate to a particular meaning.

Accordingly, the role of authentic materials in teaching English is expected to increase students' skills in using English in real life. Thomas (2014) stated that authentic materials are produced based on the purpose of communicating and giving information rather than teaching language. Despite its characteristics as learnable materials, it is clear that the textbook should contain

authentic materials. Hence, it is necessary to find a solution to that problem. Corpus is one solution that can be carried out to develop English textbooks materials. Kim (2002) stated that a corpus is a data collection of language converted from natural text. Also, Bowker & Pearson (2002), explained that a corpus is an extensive data collection of authentic text collected electronically based on specific criteria. At the same time, McEnery & Hardie (2012), stated that corpus is a field of studying languages that focuses on its procedures or the method. The authenticity of the corpus data can be a standard in developing a textbook and can also be used to examine the textbook's content. Romer (2010) argued that since the COBUILD project run by John Sinclair, corpus and language teaching have worked together to develop real English dictionaries and teaching materials that the learners can use in actual communicative situations. McCarthy, McCarten, & Sandiford (2014) stated that Corpora has been used to design dictionaries like the *Cambridge Dictionary of American English* using a 100-million-word sample of the Cambridge International Corpus.

Furthermore, the use of corpus as the methodology to conduct research has increased through the years, and some studies have been conducted on corpus study. Millar (2009) researched modal verbs in times; frequency changes 1923-2006. Other researchers also researched modality; Sadia & Ghani (2018) researched the use of modality in English newspapers in Pakistan. Moreover, some researchers also conducted the study related to the use of corpus as the methodology in analyzing textbooks, such as lexical (see Hsu, 2009; Norberg & Nordlund, 2018), Collocations (see Wang & Good, 2007), gender representation (see Koster, 2020; Lee, 2018), and grammar (see Murakami & Alexopoulou, 2016; Seog & Choi, 2018). Although corpus linguistics is still a new field, it has transformed language studies by offering a new style of analyzing and representing language use (Hunston, 2002; Szudarski, 2018). Nakayama (2021) investigates modal auxiliary verbs' use in Japanese EFL learners' conversations. Like that study, Oktavianti & Fajria (2021) investigated the frequency of modal verbs in the EFL textbook conversation section to find the similarities and mismatches of modal verb usage. Also, some researchers are investigating the use of modal verbs in learners' academic writing (Oh, 2007; Qian, 2017). Sudzarky points out that corpora allow the user to use specific software called concordance to arrange automatic searches and gain insight into the natural structures and patterns of the language. Corpora can also find the frequency list of words that shows how often they appear in a corpus (McEnery & Hardie, 2012). Lastly, McCarten (2007) argued that corpus could inform the frequency of some vocabulary (the most use of words and expressions, which is most frequent and rare). Based on those elaborations, the research will analyze the central modal verbs. LING (2016) stated that there are nine central modal verbs in English; will, would, can, could, may, might, shall, should, and must. However, this study attempts to answer the following questions: 1) what are the most frequent central modal verbs used in *Kurikulum Merdeka's* Textbooks, COCA and BNC? and 2) are there any similarities or differences regarding the use of central modal verbs in textbooks compared to that in control corpora, COCA and BNC?

METHOD

This study collected data from three series of *Kurikulum Merdeka* textbooks at the senior high school level in Indonesia that is; 10th grade for SMA/SMK/MA, 11th grade for SMA, and 11th grade for SMA/MA, and the size of the textbooks corpus is 37,238 words. This study used COCA and BNC as control corpora. COCA has about 560 million words of text and 20 million words collected from 1990-2017. The corpus includes spoken, popular magazines, newspapers, fiction, and academic texts (Sujatna, Sujatna, & Pamungkas, 2019). In comparison, BNC has approximately 100 million words, consisting of 90 percent written and 10 percent spoken words. The corpus includes a regional and national newspaper, specialist periodicals and

journals, academic books, published and unpublished books, many unscripted informal conversations, regional and social classes, business or government meetings, and radio shows (Baker, Hardie, & McEnery, 2006). The data (frequency of central modal verbs) from the textbooks corpus was collected using LanksBox as a corpus tool. In LanksBox, the researcher used "words" features to search the frequency of central modal verbs by entry-specific central modal verbs. For COCA and BNC, as the control, the researcher input the specific keyword in the "list" feature and sorted the part of speech (Pos) column into modal verbs. In analyzing the data, the frequency of central modal verbs in control corpora is normalized (nf) per one million words due to a large number of data. In contrast, the frequency of central modal verbs in textbooks was the raw frequency (rf) due to the small occurrences.

RESULTS AND DISCUSSION

Results

This section investigates the frequency of central modal verbs in textbook compared to that in COCA and BNC.

Table 1. Frequency of Central Modal Verbs in Textbooks

Central Modal Verbs	Frequency	Percentage
Can	123	36%
Will	95	28%
Would	32	9%
Could	31	9%
Should	20	6%
Must	16	5%
May	13	4%
Might	9	3%
Shall	4	1%

Table 1 shows the raw frequency of central modal verbs in *Kurikulum Merdeka*'s textbooks. The recorded data shows that using central modal verbs occurs 343 times in the three series of *Kurikulum Merdeka*'s textbooks. As can be seen in Table 2, *can* is dominantly used among the others with 123 occurrences, accounting for 36% of all central modal verbs. "Will" is the second most common modal verb, occurring 95 times, making up 28% of the total. The counterparts of "can" and "will," that is, "could" and "would," share the same amount percentage of 9%, though the occurrences are not the same. The word "could" appears 31 times while "would" occurs 32 times, making them the third most frequent central modal verb. Furthermore, the table displays the central modal verb "should" take the fourth place, which occurs 20 times or around 6%, followed by "must," which occurs 16 times with a percentage of 5%. In sixth place is "may," which appears 13 times, accounting for 4%. The least frequent central modal verbs are "might" and "shall," with only 9 times occurrences and 4 occurrences, respectively, as shown in Table 2.

Table 2. Frequency of Central Modal Verbs in Textbooks

Central Modal Verbs	Frequency (Nf)	Percentage
Can	2491922	22%
Would	2349406	21%
Will	2148852	19%
Could	1529805	14%
Should	920919	8%
May	757743	7%
Might	521997	5%
Must	427801	4%
Shall	87480	1%
Total	11235925	100%

Table 2 presents the frequency of central modal verbs in the Corpus of Contemporary American English. The four highest central modal verbs used in COCA are "can," "would," "will," and "could." "Can" is the most frequently used central modal verb, appearing 2,491,922 and constituting 22% of all central modal verbs. "Would" follows closely, which occurs 2,349,406 times (21%) in the second place. Then, "will" is in the third position with occurrences of 2,148,852 (19%), which is lower than its counterparts "would." In the fourth position, there is "could" as the "can" counterparts with the occurrences of 1,529,850 (14%). Furthermore, the data show that the rest of the central modal verbs "should," "may," "might," "must," and "shall" are the least frequent in COCA. "Should" ranks fifth, which occurs at 920,919 (8%) and followed by "may," which is slightly less frequent, with occurrences of 577,743 (7%). Then, "might" takes the seventh place, 427,801 (5%). Lastly, "must" is in the second lowest position, which only occurs at 427,801 (4%), and the lowest position is "shall," appearing only at 87,480 (1%).

Table 3. The Frequency of Central Modal Verbs in BNC

Central Modal Verbs	Frequencies (Nf)	Percentage
Would	242195	20%
Will	239466	19%
Can	226993	18%
Could	158325	13%
May	111748	9%

Should	107821	9%
Must	69293	6%
Might	58344	5%
Shall	19407	2%
Total	1233592	100%

Table 3 presents the frequency of central modal verbs in the British National Corpus (BNC) with the occurrences 1,233,592 times. The frequency is normalized per million words due to the large corpus data. The data shows that “would” is the first position as the most used central modal verb in the BNC sampler, with occurrences 242,195 (20%). “Will” and “can” closely follow with the occurrences 239,466 times (19%) and 226,993 times (18%). Then, “could” ranks fourth with the occurrences 158,325 times, accounting for 13%. The fifth and sixth positions, “may” and “should,” share the same percentage of 9% with the occurrences of “may” 111,748 times and “should” 107,821 times. Furthermore, the three lowest central modal verbs in BNC are “must,” “might,” and “shall”; “must” ranks seventh with the occurrences 69,293 times (5%), followed by “might” 58,344 times (5%) in the eighth place. Lastly, the least frequent central modal verb in BNC is “shall,” which only occurs 19,407 times (2%), as seen in Table 4.

Discussion

The findings of central modal verbs distribution show an appealing result. The usage of central modal verbs varies across these different sources. The use of “can,” “will,” and “would” are among the top three in the three corpora, textbooks, COCA, and BNC. Unsurprisingly, those modal verbs are among the top three since it is the prototype of the modal verbs category (Oktavianti & Fajria, 2021). In the textbooks, “can” is the most frequent modal verb, suggesting that it is emphasized in the instructional material, followed by “will,” which is also prominent. Surprisingly, “would” is not significant in the textbooks, it indicates there is an underused of “would” in the textbooks compare to that in COCA and BNC. However, “would” still ranks third in the textbooks. When examining COCA, “can” remains the most frequently used modal verb. However, “would” and “will” take on more significant roles, appearing prominently in this corpus. Furthermore, the BNC data, representing British English, demonstrates a slightly different distribution than COCA. “Would” takes the lead as the most common modal verb, followed closely by “will” and “can.” These results follow the result of previous studies (Bouhlal, Horst, & Martini, 2018; Khojasteh & Reinders, 2013; Oktavianti & Fajria, 2021; Qian, 2017; Yang, 2018), which present “can,” “will,” and “would” as the most frequent modal verbs. Furthermore, the three corpora mention that “shall” is in the lowest position of the central modal verbs. It possibly indicates a declining usage of the central modal verb “shall” in American and British English. These findings align with previous studies (Al-ghalban & Al-ghalban, 2015; Bouhlal et al., 2018; Khojasteh & Reinders, 2013; Qian, 2017; Yang, 2018). Lastly, the rest of the central modal verbs are not significantly used in the three corpora, however there is still a discrepancy among them. The textbooks show that “may,” “might,” and “shall” are the least frequent central modal verbs, while the COCA data reveals “might,” “must,” and “shall” as such, and the BNC data indicates “must,” “might,” and “shall” to be the least frequent. The use of “may” and “must” are suffering a decline in present-day english

(Leech, 2013). This result also indicates an overuse and underuse of some central modal verbs in the textbooks compared to those in COCA and BNC.

CONCLUSION

The usage of “can,” “will,” and “would” are the most frequently central modal verbs in the COCA, BNC, and *kurikulum merdeka*’s textbooks. However, there is an under used of “would”, which not significant present in the textbooks compared to that in COCA and BNC. While the least frequent of central modal verbs used among the three sources is “shall.” This result indicates that the distribution of central modal verbs in textbooks is similar to those in native speaker or control corpora. However, the rest of the central modal verbs vary among the three sources, like “must” and “may.”

ACKNOWLEDGMENTS

The researchers would like to express sincere gratitude to the journal supervisor who has supported and directed in conducting this study so that the researchers can complete this research. The researchers also would like to thank those who gave time and assistance for this research. However, this study is still far from perfect, and the researchers need any criticism and suggestions from the reader to improve this study. Hopefully, this research can be useful for readers and researchers with the same interest.

REFERENCES

- Al-ghalban, L. A., & Al-ghalban, L. A. (2015). *A Corpus-Based Study of Modal Verbs in Native And Egyptian Learner English*. 59(July).
- Baker, P., Hardie, A., & McEnery, T. (2006). *A Glossary of Corpus Linguistics*. Edinburg: Edinburgh University Press.
- Bouhlal, F., Horst, M., & Martini, J. (2018). *Modality in ESL Textbooks : Insights from a Contrastive Corpus-Based Analysis*. 227–252. <https://doi.org/10.3138/cmlr.3075>
- Bowker, L., & Pearson, J. (2002). *Working with Specialized Language*. London: Routledge.
- Harwood, N. (2015). English Language Teaching Textbooks: Content, Consumption, Production. In *ELT Journal* (Vol. 69). <https://doi.org/10.1093/elt/ccu066>
- Holmes, J. (1988). Doubt and certainty in ESL textbooks. *Applied Linguistics*, 9(1), 21–44. <https://doi.org/10.1093/applin/9.1.21>
- Hsu, W. (2009). College english textbooks for general purposes: A Corpus-based analysis of lexical coverage. *Electronic Journal of Foreign Language Teaching*, 6(1), 42–62.
- Hutchinson, T., & Torres, E. (1994). The textbook as agent of change. *ELT Journal*, 48(4), 315–328. <https://doi.org/10.1093/elt/48.4.315>
- Khojasteh, L., & Reinders, H. (2013). *How textbooks (and learners) get it wrong : A corpus study of modal auxiliary verbs*. 2(1), 33–44.
- Kim, S. (2002). A Corpus-based Analysis of the Words in the Elementary School English Textbooks. *English Teaching*, 57(3), 253–277.
- Koster, D. (2020). Do representations of gender and profession change over time? Insights from a longitudinal corpus study on Dutch language textbooks (1974–2017). *Journal of Gender Studies*, 29(8), 883–896. <https://doi.org/10.1080/09589236.2020.1754176>
- Lee, J. F. K. (2018). Gender representation in Japanese EFL textbooks—a corpus study. *Gender and Education*, 30(3), 379–395. <https://doi.org/10.1080/09540253.2016.1214690>
- Leech, G. (2013). *Meaning and the English Verb* (Third Edit). New York: Routledge.
- LING, Y. S. C. (2016). The Use of Modal Auxiliary Verbs in Horoscope: A Corpus-Based Study. *Issues in Language Studies*, 5(2), 82–98. <https://doi.org/10.33736/ils.1639.2016>

- Mccarten, J. (2007). *Teaching Vocabulary: Lessons from the Corpus , Lessons for the Classroom*. New York: Cambridge University Press.
- McCarthy, M., McCarten, J., & Sandiford, H. (2014). *Student Book Touchstone*. Cambridge University Press.
- McEnery, T., & Hardie, A. (2012). *Corpus Linguistics: Method, Theory and Practic*. New York: Cambridge University Press.
- Millar, N. (2009). Modal verbs in TIME. *International Journal of Corpus Linguistics*, 14(2), 191–220. <https://doi.org/10.1075/ijcl.14.2.03mil>
- Murakami, A., & Alexopoulou, T. (2016). L1 Influence on the Acquisition Order of English Grammatical Morphemes. *Studies in Second Language Acquisition*, 38(3), 365–401. <https://doi.org/10.1017/S0272263115000352>
- Nakayama, S. (2021). Modal Auxiliary Verbs in Japanese EFL Learners ' Conversation : A Corpus-based Study. *Asia Pacific Journal of Corpus Research*, 2(1), 23–34.
- Norberg, C., & Nordlund, M. (2018). A Corpus-based Study of Lexis in L2 English Textbooks. *Journal of Language Teaching and Research*, 9(3), 463. <https://doi.org/10.17507/jltr.0903.03>
- Oh, S.-Y. (2007). A Corpus-based Study of Epistemic Modalityin Korean College Students' Writings in English. *English Teaching*, 62(2), 147–175. <https://doi.org/10.15858/engtea.62.2.200706.147>
- Oktavianti, I. N., & Fajria, A. (2021). Modal Verbs in a Curriculum-Based EFL Textbook of Senior High School in Indonesia: A Corpus-Based Study. *TESOL International Journal*, 16(5), 171–204.
- Qian, L. (2017). *A Corpus-Based Study of Modal Verbs Use in English Writing by EFL Learners*. 13(11), 31–35. <https://doi.org/10.3968/9975>
- Rappen, R., Fitmaurice, Su. M., & Biber, D. (2002). *Using Corpora to Explore Linguistics Variation*. <https://doi.org/10.1075/scl.9.17gei>
- Romer, U. (2010). Using General and Specialized Corpora in English Language Teaching: Past, Present and Future. In *Corpus-Based Approaches to English Language Teaching*. Retrieved from <https://www.ptonline.com/articles/how-to-get-better-mfi-results>
- Sadia, S., & Ghani, M. (2018). Modality in Editorials of Pakistani English Newspapers: A Corpus Based Study. *International Journal of English Linguistics*, 9(1), 144. <https://doi.org/10.5539/ijel.v9n1p144>
- Seog, D. S. Y., & Choi, I. (2018). Interlanguage development of young Korean EFL learners' modal usage: A learner corpus study. *Linguistic Research*, 35(Specialissue), 83–103. <https://doi.org/10.17250/khisli.35..201809.004>
- Sujatna, M. L., Sujatna, E. T. S., & Pamungkas, K. (2019). Exploring the Use of Modal Auxiliary Verbs in Corpus of Contemporary of Ameriscan English (Coca). *Sosiohumaniora*, 21(2), 166–172. <https://doi.org/10.24198/sosiohumaniora.v21i2.19970>
- Szudarski, P. (2018). *[Routledge corpus linguistics guides] Paweł Szudarski - Corpus Linguistics for Vocabulary_ A Guide for Research (2018, Routledge) - libgen.li.pdf*. London: Routledge.
- Thomas. (2014). *Meeting EFL Learners Halfway by Using Locally Relevant Authentic Materials*.
- Tomlinson, B. (2014). Material Evaluation. In *Developing materials for language teaching*.
- Wang, J. T., & Good, R. (2007). The Repetition of Collocations in EFL Textbooks: A Corpus Study. *Online Submission*, 1–13.
- Yang, X. (2018). *A Corpus-based Study of Modal Verbs in Chinese Learners ' Academic Writing*. 11(2), 122–130. <https://doi.org/10.5539/elt.v11n2p122>