A corpus study of verbs in opinion articles of The Jakarta Post and the relation with text characteristics

Ikmi Nur Oktavianti[a,1,*], Asmad Adnan[b,2]

[a,b] Universitas Ahmad Dahlan, Tamanan, Bantul, Daerah Istimewa Yogyakarta, Indonesia

[1] ikmi.oktavianti@pbi.uad.ac.id*; [2] asmad_adnan@yahoo.com

* corresponding author

ARTICLE INFO

ABSTRACT

As one of the text categories, opinion texts have distinctive characteristics compared to any other texts in newspapers, including the choice of verb usage. This study then aims at preliminarily examining the verbs used in opinion articles in *The Jakarta Post* to find out the relation between frequency and text characteristics. This study collected the opinion articles of *The Jakarta Post* comprising 47.143 words. This study was assisted by *Lancsbox* to store the corpus of opinion section texts, to identify the verb lemmas, and to count the frequency of verbs. The verbs found in this study were then classified based on Scheibman's main verb classification (which is based on Halliday's and Dixon's verb types). The results of the study show that there are three most frequent verb types used in opinion texts in *The Jakarta Post;* they are *material, verbal,* and *feeling* verb types. Meanwhile, the lesser frequent ones are *perception, possessive/relational, relational, and cognition* verbs types. Meanwhile, the least frequent verb types are *existential, corporeal and perception/relational* verbs types. As opinion text conveys the argument of the writer, it is plausible to find *feeling* verb type belongs to the third most frequent types, along with *material* type to show concrete actions and *verbal* type to report the information. These frequencies exhibit that there is a firm relationship between text characteristics and the tendency of verb choice.

1. Introduction

In English syntactic construction, verbs are salient since there will be no grammatical construction without verbs. This corresponds to the need of *tense inflection* in English clauses and to be able to have *tense inflection,* there must be verbs in the clause (Baker, 2004). In addition, verbs play a vital role to s-select the constituent allowed in the clause constructions. As argued by van Gelderen (2017), verbs are salient because they determine the internal and external arguments of the clauses. With regard to its capability, verbs are more dynamic compared to the other lexical categories, i.e. nouns, adjectives, adverbs (Clackson, 2007).

In regard to the use of language, some aspects, such as the context of usage and text characteristics, affect the use of verbs. This means that different communication modes might result in different choices of verbs. For instance, with regard to verb, according to Carter and McCarthy (2006), verb *know* is very frequent in spoken English, but it is not frequent in written English. Aside from verb use, the first-person pronoun *I* is found to be more frequent in spoken language than in written language. This proves that there are some differences in using language, depending on the mode as well as the other aspects, e.g. text genres and/or dialects (Biber & Conrad, 2009). As with the differences in spoken and written language, one of the plausible reason is the different

characteristics of the medium; spoken language tends to be more spontaneous and involves two-way interactions than written language (Biber & Quirk, 2012).

The representation of verb use in language can be seen in newspaper since the language used in newspaper is closer to everyday language (Bednarek & Caple, 2012; Biber & Conrad, 2009). Newspaper, therefore, has been very appealing for linguistic studies. Compared to literary works, the language use in them is rather manipulative, yet aesthetic for the purpose of building the imaginative world of the story (Leech & Short, 2007). As for academic text, the language used is the standard forms and use academic style (Biber & Conrad, 2009; Biber & Gray, 2016). Thus, compared to other canonical text genres, investigating verbs in news texts is obviously interesting as well as beneficial for the description of language use in general.

To be able to describe the verbs comprehensively, this study employs a corpus-based analysis, which means this study focuses on the frequency of use and uses a larger scale of data for the analysis. Several corpus studies on verbs in newspapers have been conducted to seek for pedagogical implication (Deng & Li, 2017), and corpus studies to assist critical discourse analysis of the news texts (Moon, 2016; Wang, 2015), and to depict the morphosyntactic aspects of the verbs in news texts (Bednarek & Caple, 2012). Some other verbs were examined by focusing on report verbs used by native and non-native speakers of English (Yilmaz & Ertürk, 2017), and by investigating phrasal verbs (Zarifi & Mukundan, 2013).

In relation to frequency of use, there has been proven that verbs in news texts are more likely in the form of present tense and the use of active voice is more frequent than passive voice for some reasons (Bednarek & Caple, 2012). In text analysis (e.g. news text analysis), Biber (2012) points out that each text has specific lexical choice that is included as the characteristic of the text. A particular lexical unit or linguistic unit is more likely to be used in a particular text. Using *Longman Spoken and Written Corpus,* (Biber, 2012) shows that passive verbs are more frequent in academic texts than in conversation, and some verbs (e.g. *make, find*) are more frequent in academic texts than in conversation.

The study of frequency has been recognized in linguistics, such as in studying the relation of frequency and language typology (Bentz, Alikaniotis, Samardžić, & Buttery, 2017), the importance of lexical frequency for designing teaching materials (Criado & Sánchez, 2012), and the relation between frequency and language change (Feltgen, Fagard, & Nadal, 2017). In text analysis, frequency holds a salient role because each text might have its own lexical choice that results in different frequency of use (Biber, 2012). In this topic, Hardjanto (2016) has investigated the relation between text type and modal use. Although being in the same text type, Bonyadi (2011) figures out that there are different tendencies of modal use in news texts between two news media. Similarly, Hardjanto (2016) who investigates the use of modals in academic texts of five different fields (e.g. economics, linguistics, medicine, engineering, natural sciences) also describes that there are some slight differences in the choice of modals. Within the same mode of communication (i.e. written mode), there might also some notable differences in terms of linguistic unit usage such as done by Staples, Biber, and Reppen, (2018) who compare the linguistic aspects of TOEFL IBT task and disciplinary writing task.

Those studies, however, mainly focus on news text as a single register. However, there are also sub-type of news texts to be analyzed, such as hard news, soft news or feature, and opinion or editorial (Bell, 1991). There have been some studies focusing on the frequency of verbs in the more specific news texts; they are the corpus study of verbs in hard news (Oktavianti & Ardianti, 2019) and the corpus study in feature or lifestyle articles (Oktavianti & Pramesti, 2019). Biber and Conrad, (2009) also distinguish verbs used in news (both hard and soft news) and editorials. The discussion, however, is still limited to verb type based on transitivity (e.g. intransitive, transitive, ditransitive verbs) and voice (active, passive verbs). Little is known about the choice of verb types and the relation with the characteristics of opinion texts. This study then aims at examining the frequency of verbs in opinion articles. To delimit the study, *The Jakarta Post* is selected because it is a well-established and the oldest English newspaper in Indonesia. Besides, it enables free and easy access for the compilation of data so that the results of the study can be comprehensive and reliable.

2. Literature Review

This research uses a corpus method (i.e. corpus linguistics) to analyze the data. Corpus linguistics is the study of digitally stored language data (i.e. spoken and written texts) (McEnery & Hardie, 2012). Corpus is a collection of digitally stored texts, i.e. spoken and written texts (McEnery & Hardie, 2012; Stefanowitsch, 2020). Given the large amount of data compiled in a corpus, the nature of corpus is related to the capability of providing rich quantitative data as language use is not parole, but it is collective use (Burkette & Kretzschmar Jr., 2018). In accordance to the benefits it offers, corpus has been extensively used in various linguistic studies. Thus, corpus linguistics is preferably seen as a methodology to approach language rather than a branch of linguistics (McEnery & Hardie, 2012). With a large amount of data, studying language use with corpus assistance, thus, might enable comprehensive qualitative and quantitative analyses and obtain more reliable results. In linguistics, corpus studies have widely recognized for many branches (Baker, 2003; Biber, 2006; Ismail, Idrus, & Syed Sahuri, 2020; Kranich, 2010; Makamani & Mutasa, 2017; Moon, 2016; Motschenbacher, 2018; Staples et al., 2018; Wright & Brookes, 2019). Aside from linguistics, corpus can also be applied for language teaching field (Akıncı & Yıldız, 2017; Arellano, 2018; Astika, 2018; Kim, 2019; Kızıl & Savran, 2018; Lin, 2016; Phoocharoensil, 2017; Yanto & Nugraha, 2017).

In regard to the main focus of a corpus study, this study examines the frequency of verbs in opinion articles. It is of importance to investigate frequency of linguistic units since frequency often tells about something. According to Baker (2010), frequency indicates something important. In other words, something that is frequent is salient to the users. In regard to language use, different texts, for example, might affect different choice of linguistic units. As studied by Biber and Quirk (2012) and Biber and Conrad (2009) on the use of modals, it is evident that core modals are less frequent in news texts than in fiction and academic texts. Emphasizing the results of those previous studies, Szudarski (2017) points out that studying frequency in a certain text (corpus) can assist the description of the text. In other words, it is noticeable that certain linguistic features or units tend to occur more frequently in a particular text.

With regard to news texts, Bell (1991) classifies types of news texts into hard news (or news), soft news (features, lifestyle), and editorial (and sort of argumentative texts). Each of the sections represents different types of text genres as well different purpose and characteristics. News report is expected to describe events with less subjectivity (Biber & Conrad, 2009). On the contrary, editorials and opinion texts are to some extent slightly different. Editorials are written in the newsroom, so they express the opinion of the newspaper. Meanwhile, opinions texts are written by outsiders or readers of the newspapers. Bell (1991) further mentions that, apart from the different perspective of the opinions, editorials and opinion texts are relatively identical. Both are basically meant to serve the same purpose to express an opinion openly and convince the readers to believe the opinion (Biber & Conrad, 2009). In other words, editorials and opinion are within the same genre (Bednarek & Caple, 2012). Since this paper focuses more on the characteristics of the linguistic features, it is possible to overlook the insignificant difference of both types of argumentative texts.

Biber and Conrad (2009) mention that what distinguishes news from opinion (or editorial) texts is factuality since opinion texts do not talk about factual events, but the opinion toward a particular event. News texts serve to report an event, while opinion texts aim at delivering an opinion (which tends to be more subjective) on an event or issue. Having distinct text characteristics and functions, therefore, opinion texts might have different linguistic features as well as different verb choices. This has been proven by Biber and Conrad (2009) showing that there are modals that are more frequently used in opinion texts compared to news report. For instance, modal *should* is more frequent in opinion texts since one of the purpose of opinion texts is to recommend what should happen in relation to the phenomenon being discussed.

3. Research Method

This study employs a corpus-based approach since it compiles a larger scale of data, uses a corpus tool, and conducts a corpus analysis (i.e. frequency analysis) (McEnery & Hardie, 2012). Corpus linguistics as one of the branches of linguistics is a methodology because it deals with how to collect and analyze linguistic data (McEnery & Hardie, 2012; Stefanowitsch, 2020). This study

compiled linguistic data (the self-compiled corpus) from opinion articles in *The Jakarta Post* published from October—November 2018 with the size of 47.143 words. The corpus was then analyzed by identifying the verbs as well as calculating the frequencies of the verbs by using a corpus tool, *Lancsbox* (Brezina, Timperley, & McEnery, 2018). In *Lancsbox*, frequency of verbs was calculated in *Whelk* feature by sorting the type into *lemma* and filtering the result box by inserting *_v to sort out the results for only verbs. The corpus tool also calculates the frequency and displays the results based on the order from the highest to the smallest ones (or vice versa) that could assist the analysis. After having the frequency results, the verbs were classified by following (Scheibman, 2001) which consists of verb classifications based on Halliday (Halliday & Matthiessen, 2004) and Dixon (Dixon, 2005). This study used the classification compiled by Scheibman since it best fits the need of the analysis of text characteristics. Table 1 displays the classification.

Table 1. Classification of verbs by Scheibman (2001)

Verb Type	Description	Examples
Cognition	cognitive activity	*know,think, remember, figure out*
Corporeal	bodily gestures, bodily interaction	*eat, drink, sleep, live, smoke*
Existential	exist, happen	*be, havve, sit, stay, happen*
Feeling	emotion, wanting	*like, want, feel, need, bother, enjoy*
Material	concrete and abstract doings and happenings	*do, go, take, teach, work, use, play, come*
Perception	perception, attention	*look, see, hear, find, notice*
Perception/ Relational	perception (subject not senser)	*look, smell, sound*
Possessive/ Relational	possession (x has a)	*have, get*
Relational	processes of being (x is a, x is AT a)	*be, get, be like* (descriptive), *become*
Verbal	saying, symbolic exchange of meaning	*say, talk, mean, tell, ask, go* (quotative), *be like* (quotative)

There are, however, some ambiguities for some words, e.g. *get* that can be classified into either *relational* or *possessive/relational.* To overcome this problem, this study uses *Lancsbox* concordance feature, *KWIC* (Keyword in Context), to distinguish which g*et* belongs to which type of verbs. This study analyzes the frequency of verbs used in opinion articles by referring to the raw frequency (including token frequency) and the percentage. To comprehensively discuss the relation between frequency and characteristics of text, this study implements *thick description* (Stake, 2010) to interpret the plausible connection.

4. Findings and Discussion

This section discusses two key analyses of this study, namely the frequency of verb types and the relation between frequency and characteristics of the text. To begin with, the following is the discussion on the frequency of verb type investigated in the self-compiled corpus.

4.1. Frequency of Verb Types in Opinion Articles

Based on the Scheibman's compilation of verb types, Table 2 presents the overall frequency of verbs in opinion articles in *The Jakarta Post.* Based on Table 2, it is found that *material verb* is the most frequent verb type in the corpus of opinion of *The Jakarta Post*. This verb type dominates the whole corpus by occupying 28,4 % of all verbs found in the corpus. The other most frequent verb types are *verbal* (18,2 %) and *feeling* verb types (14,4 %). As for the less frequent verb types, there are *possessive/relational* (9,9%)*, relational* (8,2%)*, perception* (7,3%)*, cognition verbs* (7,1%). Meanwhile, the other three types of verbs, *existential* (2,1%)*, corporeal* (2,8%) and *perception/relational* (0,8%) belong to the least frequent verb types in which the frequencies of those types are below 5% in the corpus. For the detailed description of verb type frequencies, it is listed in Table 3.

Table 2. Frequency of verbs types

Verb type	Absolute frequency	Percentage (%)
Material	225	28,4
Verbal	144	18,2
Feeling	114	14,4
Possessive/Relational	78	9,9
Relational	65	8,2
Perception	58	7,3
Cognition verbs	56	7,1
Existential	23	2,9
Corporeal	22	2,8
Perception/Relational	6	0,8
Total	**791**	**100%**

Table 3. Detailed description of the verb frequencies

Rank	Verb type	Example	Token frequency	Total frequency	Percentage (%)
1	Material	Do	14	225	28,4
		Go	31		
		Take	41		
		Teach	4		
		Work	30		
		Use	56		
		Play	17		
		Come	32		
2	Verbal	Say	65	144	18,2
		Talk	3		
		Mean	23		
		Tell	13		
		Ask	9		
3	Feeling	Like	6	114	14,4
		Want	14		
		Feel	11		
		Need	73		
		Bother	1		
		Enjoy	9		
4	Perception	Look	20	78	9,9
		See	34		
		Hear	6		
		Find	17		
		Notice	1		
5	Possessive/relational	Have	49	65	8,2
		Get	12		
6	Relation	Get	4	58	7,3
		Become	54		
7	Cognition	Know	30	56	7,1
		Think	23		
		Remember	3		
8	Existential	Stay	2	23	2,9
		Happen	21		
9	Corporeal	Eat	3	22	2,8
		Sleep	1		
		Live	18		
10	Perception/relational	Look	6	6	0,8

Table 3 shows some examples of *material* verbs found in the corpus. There are verbs, such as *take* (41 occurrences), *come* (32 occurrences), *go* (31 occurrences), *work* (30 occurrences), *play* (17 occurrences), and *teach* (4 occurrences). The second most frequent verb is *verbal* and the most frequent verb in this type is *say* (65 occurrences), followed by *mean* (23 occurrences). As with the third most frequent is *feeling* type with the verb *need* (73 occurrences) and *want* (14) as the most frequent members of the type.

The less frequent verb types also comprise several verbs as displayed in Table 3. For instance, *perception* type has the verb *see* with 34 occurrences as its most frequent member. As for *possessive/relational* type, there is *have* with 49 occurrences and *relational* type has the verb *become* with 54 occurrences as their most frequent verbs. Following these results, there are the least frequent verbs in the corpus, such as *existential, corporeal,* and *perception/relational* verbs. In *existential* type, the verb *happen* occurs 21 times, the verb *live* in *corporeal* type occurs 18 times and the verb *look* in *perception/relational* type occurs 6 times.

4.2. The Relation between Frequency and Text Characteristics

This section explores the plausible relation between the frequency of verbs and the characteristics of opinion texts. As every text works distinctively, it is compelling to examine the linguistic features, i.e. the use of verbs in a particular text, i.e. opinion texts. As pointed out by Biber and Conrad (2009), the difference between news texts and opinion texts is related to factuality. In news texts, factuality plays a salient role, as it is the core of the text, while opinion texts express argument or perspective of the author as an individual or member of a certain institution or society. This is supported with the result of the studies done by Bednarek and Caple (2012) and Biber and Conrad (2009) identifying that there are some discrepancies of the use of modals among texts in newspaper (i.e. news, feature, opinion).

With regard to verb types compiled by Scheibman (2001), Oktavianti and Ardianti (2019) show that the most frequent verb type in news texts is *verbal* (e.g. say, tell) and Oktavianti and Pramesti, (2019) describe that in feature texts, *material* type is the most frequent one. In this study, *material* type dominates the opinion corpus. It can be assumed that *material* verb type takes an important role in building opinion because the author writes their opinion, argues or reacts about something that is related to concrete action. In the delivery argument or ideology, the author needs to be as clear as possible, so concrete verbs are necessary and significantly used to convey the message to the audience. In the description of the opinion, one must explain the issue or event concretely. It is in line with the nature of opinion column in newspaper that reacts to a factual event or issue containing concrete activities from people involved in the issue or event. It is not surprising then to find out that *material verbs* occupy the highest percentage of use. For instance, the verb *use* is used to describe concrete activity related to a thing or object. Likewise, other *material verbs* (e.g., *take, work, come and go*) can be found in the corpus.

The second most frequent verb type in opinion column is *verbal* type. This type has the verb *say* as its dominant verb in which this verb is prominent in delivering information as mentioned in (Lewis, 2014). Therefore, *verbal* verbs are used to describe or to assist the delivery of an idea or argument about an issue or event. Following *verbal* type in the second position, the third most frequent verb type is *feeling. Feeling* verbs serve as the verbs to state emotional aspects and wanting. The function of *feeling* verb is related to how the author's point of view about the event or issue being discussed. These verbs then suit the need of opinion articles. Biber and Conrad (2009) state that opinions express arguments of the author and try to persuade the readers to think the same way. It is not surprising that this verb type is pretty frequent in the opinion corpus.

Unlike other prior types, the verbs belong to *perception, possessive/relational, relational,* and *cognition* types are less frequently used in the corpus (but not the least ones). Based on the nature of opinion or editorial, *perception* and *possessive/relational* tend to be more subjective and personal in expressing opinion. *Perception* verbs are specifically used in relation to senses, hence irrelevant with building opinions. Meanwhile, *cognition verbs* are types of verbs expressing mental action or concerned with the act of process of knowing, perceiving, etc. This verb type is basically quite relevant with the purpose of delivering argument, but it is somehow less frequently used in the corpus. This is not because of the verb, e.g. *know,* is lesser than any other verbs since it occurs 30

times, but the number of variants in the types (found in the corpus) is limited. There are only *know, remember, think* found in the data showing that although we need them to express opinions, but we do not need various cognitive verbs. Compared to *material* type, there are more verbs belonging to the category (e.g. *do, go, work, teach, play,* etc.) and they are used as well in the texts.

The last discussion is about the least frequent verb types in opinion texts. Based on the data, there are *existential, corporeal* and *perception/relational verb* types having the smallest number of frequencies among all other verb types in the corpus. This is presumably because of the irrelevance of the function of the verb types and the characteristics of opinion text genre. For instance, *existential* verb as the verb describing the details of the event is one of the least frequent types since it is most likely to be used in hard news texts. Oktavianti and Ardianti (2019) find that *existential verb* like *happen* occupies the third most frequent type since this feature is important to report an event in news text. This type, in the contrary, is not necessary in the writing of opinion, thus it is less likely to be used.

Similar to *existential* type, *corporeal* verbs are very rare in the corpus. This verb type refers to bodily gesture and interactions (Scheibman, 2001). However, opinion articles are less likely to describe about bodily gesture or interaction individual and focus more on argumentation and the like. It is not surprising to find out that this type is rarely used in the corpus. Meanwhile*, perception/relational* type has the lowest frequency of verbs found in the opinion or editorial in the Jakarta post. This corresponds with the function *perception/relational* of type of verb to describe perception through the perspective of the object (not a human), which is to some extent not in accordance with the necessity to express someone's opinion convincingly. Thus, this type is rarely used in opinion articles (as found in the corpus). In other words, linguistic units that are not in accordance with the task of opinion to deliver argument on an issue are not pretty much needed, even though they are still used with lower frequency.

To summarize the relation between frequency and opinion text characteristics, Table 4 presents the highlighted points in relation to the most frequent verb types.

Table 4. Relation between frequency and characteristics of text

Most frequent type	The nature of the verb type	The necessity of the verb type in opinion text
Material	to express concrete actions	to portray the issue or event
Verbal	to deliver message	to assist the description of the information of factual event
Feeling	to express emotion and wanting	to express opinion toward the issue or event more personally so that the readers can get along with the feeling of the author

There are, however, some points to underline. From the data and the analysis, it is evident that verb types with lower frequencies deal with the irrelevance of the nature of the verbs with the characteristics of the texts. This result corresponds to the result of some previous studies (Biber & Conrad, 2009; Biber & Quirk, 2012). The more frequent use of certain verb types in opinion texts indicates that a particular text might need a particular linguistic unit more frequently. The result of this study is in line with those done by Hardjanto (2016) and Staples et al. (2018). This study also shows that, even within the same register, there might some slight differences among its sub-registers, as found in Biber (2012), Biber and Conrad (2009), and Hardjanto (2016). However, the lower frequency might also be related to the variants of the verb type are minimal in the corpus so that they accumulate low frequency. As in *cognition* types that are negligible in the corpus due to the limited verb variants of the type in the corpus compiled in this study. These points, therefore, should be investigated further in the following studies comprehensively using a larger amount of data and from more various newspapers.

5. Conclusion

The findings of the study show that there are three most frequent types of verbs in the opinion articles of The Jakarta Post; they are *material, verbal* and *feeling* in the corpus of opinion or editorial

column of The Jakarta Post. Along with the most frequent types, there are some less frequent ones, such as *perception, possessive/relational*, *relational,* and *cognition* types, and the least ones are *existential, corporeal* and *perception/relational* verb types. These frequencies reveal that that there is a relationship between text category and the tendency of verb choices. As opinion texts convey argument, it is plausible to find *feeling* verb type belong to the third most frequent types, along with *material* type to show concrete actions and *verbal* type to deliver the message. Meanwhile, the least frequent verb types include *existential, corporeal* and *perception/relational* types and most of them are less frequently used in opinion texts because of the irrelevance with the text purpose and/or the number of the type variants in the corpus is limited. This study shows that lower frequency can be either connected to the characteristics of texts or the limited verb variants of the type. The latter, however, needs further investigation.

References

Akıncı, M., & Yıldız, S. (2017). Effectiveness of corpus consultation in teaching verb+noun collocations to advanced ELT students. *Eurasian Journal of Applied Linguistics*, *3*(1), 91–108. https://doi.org/10.32601/ejal.461036

Arellano, R. (2018). A corpus linguistics application in the analysis of textbook as national teaching instruments of English as second language in Chile. *Actualidades Investigativas E Educacion*, *18*(1), 1–19.

Astika, G. (2018). Lemmatizing textbook corpus for learner dictionary of basic vocabulary. *Indonesian Journal of Applied Linguistics*, *7*(3), 630–637. https://doi.org/10.17509/ijal.v7i3.9813

Baker, M. (2004). *Lexical categories*. Cambridge: Cambridge University Press.

Baker, P. (2003). No effeminates please: A corpus-based analysis of masculinity via personal adverts in *gay News/Times* 1973–2000. *The Sociological Review*, *51*(1_suppl), 243–260. https://doi.org/10.1111/j.1467-954X.2003.tb03614.x

Baker, P. (2010). *Sociolinguistics and corpus linguistics*. Edinburgh: Edinburgh University Press.

Bednarek, M., & Caple, H. (2012). *News discourse*. London: Continuum.

Bell, A. (1991). *Language of news media*. Oxford: Blackwell.

Bentz, C., Alikaniotis, D., Samardžić, T., & Buttery, P. (2017). Variation in word frequency distributions: definitions, measures and implications for a corpus-based language typology. *Journal of Quantitative Linguistics*, *24*(2–3), 128–162. https://doi.org/10.1080/09296174.2016.1265792

Biber, D. (2012). Register as a predictor of linguistic variation. *Corpus Linguistics and Linguistic Theory*, *8*(1), 9–37. https://doi.org/10.1515/cllt-2012-0002

Biber, D. (2006). *University language: a corpus-based study of spoken and written registers*. Amsterdam ; Philadelphia: J. Benjamins.

Biber, D., & Conrad, S. (2009). *Register, genre, and style*. Cambridge: Cambridge University Press.

Biber, D., & Gray, B. (2016). *Grammatical complexity in academic English : linguistic change in writing*. Cambridge: Cambridge University Press.

Biber, D., & Quirk, R. (Eds.). (2012). *Longman grammar of spoken and written English* (10. impression). Harlow: Longman.

Bonyadi, A. (2011). Linguistic manifestations of modality in newspaper editorials. *International Journal of Linguistics*, *3*(1), 1–13. https://doi.org/10.5296/ijl.v3i1.799

Brezina, V., Timperley, M., & McEnery, T. (2018). *Lancsbox*. Lancaster: Lancaster University

Burkette, A., & Kretzschmar Jr., W. A. (2018). *Exploring linguistic science: Language use, complexity, and interaction* (1st ed.). Cambridge University Press.

Carter, R., & McCarthy, M. (2006). *Cambridge grammar of English*. Cambridge: Cambridge University Press.

Clackson, J. (2007). *Indo-European linguistics: An introduction*. Cambridge: Cambridge University Press.

Criado, R., & Sánchez, A. (2012). Lexical frequency, textbooks and learning from a cognitive perspective: A corpus-based sample analysis of ELT materials. *Volumen Monográfico*, *2012*, 77–94.

Deng, F., & Li, X. (2017). A computer corpus-based study of FOREIGN in English Newspapers and its pedagogical implications. *Eurasia Journal of Mathematics, Science and Technology Education*, *13*(10), 6799–6806. https://doi.org/10.12973/ejmste/78278

Dixon, R. M. W. (2005). *A semantic approach to English grammar*. Oxford: Oxford University Press.

Feltgen, Q., Fagard, B., & Nadal, J.-P. (2017). Frequency patterns of semantic change: corpus-based evidence of a near-critical dynamics in language change. *Royal Society Open Science*, *4*(11), 170830. https://doi.org/10.1098/rsos.170830

Halliday, M. A. K., & Matthiessen, C. M. I. M. (2004). *An introduction to functional grammar* (3rd ed). London : New York: Arnold ; Distributed in the United States of America by Oxford University Press.

Hardjanto, T. D. (2016). Hedging through the use of modal auxiliaries in English academic discourse. *Jurnal Humaniora*, *28*(1), 37. https://doi.org/10.22146/jh.v28i1.11412

Ismail, H., Idrus, M. M., & Syed Sahuri, S. N. (2020). Masculinity through "emotion:" A corpus-assisted discourse analysis on sport news and implications for social workers in sport. *Journal of Human Behavior in the Social Environment*, *30*(1), 54–70. https://doi.org/10.1080/10911359.2019.1664358

Kim, H. (2019). The perception of teachers and learners towards an exploratory corpus-based grammar instruction in a Korean EFL primary school context. *The Korea Association of Primary English Education*, *25*(1), 123–152. https://doi.org/10.25231/pee.2019.25.1.123

Kızıl, A. Ş., & Savran, Z. (2018). The integration of corpus into EFL speaking instruction:A study of learner perceptions. *International Online Journal of Education and Teaching*, *5*(2), 376–389.

Kranich, S. (2010). *The progressive in modern English: A corpus-based study of grammaticalization and related changes*. Amsterdam ; New York, NY: Rodopi.

Leech, G. N., & Short, M. (2007). *Style in fiction: a linguistic introduction to English fictional prose* (2nd ed). New York: Pearson Longman.

Lewis, A. (2014). *Oxford dictionary of English*. Oxford: Oxford University Press.

Lin, M. H. (2016). Effects of corpus-aided language learning in the EFL grammar classroom: A case study of students' learning attitudes and teachers' perceptions in Taiwan. *TESOL Quarterly*, *50*(4), 871–893. https://doi.org/10.1002/tesq.250

Makamani, R., & Mutasa, D. E. (2017). A corpus-based critical discourse analysis (CDA) of the linguistic encoding of HIV and AIDS discourse by the *Kwayedza* newspaper in Zimbabwe. *South African Journal of African Languages*, *37*(1), 85–98. https://doi.org/10.1080/02572117.2017.1316933

McEnery, T., & Hardie, A. (2012). *Corpus linguistics*. Cambridge: Cambridge University Press.

Moon, R. (2016). A corpus-linguistic analysis of news coverage in Kenya's Daily Nation and The Times of London. *International Journal of Communication*, *10*, 2381–2401.

Motschenbacher, H. (2018). Corpus linguistics in language and sexuality studies: Taking stock and looking ahead. *Journal of Language and Sexuality*, *7*(2), 145–174. https://doi.org/10.1075/jls.17019.

Oktavianti, I. N., & Ardianti, N. R. (2019). A corpus-based analysis of verbs in news section of The Jakarta Post: How frequency is related to text characteristics. *JOALL (Journal of Applied Linguistics & Literature)*, *4*(2), 203–214. https://doi.org/10.33369/joall.v4i2.7623

Oktavianti, I. N., & Pramesti, Z. A. (2019). Frequency of verbs in lifestyle column in The Jakarta Post and the relation to text characteristics: A corpus-based analysis. *IDEAS: Journal on English Language Teaching and Learning, Linguistics and Literature*, *7*(2), 233–246. https://doi.org/10.24256/ideas.v7i2.1038

Phoocharoensil, S. (2017). Corpus-based exploration of linking adverbials of result: Discovering what ELT writing coursebooks lack. *3L The Southeast Asian Journal of English Language Studies*, *23*(1), 150–167. https://doi.org/10.17576/3L-2017-2301-11

Scheibman, J. (2001). Local patterns of subjectivity in person and verb type in American English conversation. In *Frequency and the emergence of linguistic structure* (pp. 61–89). Amsterdam: John Benjamins Publishing Company.

Stake, R. E. (2010). *Qualitative research: studying how things work*. New York: Guilford Press.

Staples, S., Biber, D., & Reppen, R. (2018). Using corpus-based register analysis to explore the authenticity of high-stakes language exams: A register comparison of TOEFL iBT and disciplinary writing tasks. *The Modern Language Journal*, *102*(2), 310–332. https://doi.org/10.1111/modl.12465

Stefanowitsch, A. (2020). Corpus linguistics: a guide to the methodology Berlin: Language Science Press..

Szudarski, P. (2017). *Corpus linguistics for vocabulary: A guide for research* (1st ed.). Routledge.

van Gelderen, E. (2017). *Syntax: An introduction to minimalism*. Amsterdam: John Benjamins Publishing Company.

Wang, H. (2015). A corpus-based contrastive study of online news reports on economic crisis: A critical discourse analysis perspective. *Journal of Language Teaching and Research*, *6*(3), 627–632. https://doi.org/10.17507/jltr.0603.20

Wright, D., & Brookes, G. (2019). "This is England, speak English!": a corpus-assisted critical study of language ideologies in the right-leaning British press. *Critical Discourse Studies*, *16*(1), 56–83. https://doi.org/10.1080/17405904.2018.1511439

Yanto, E. S., & Nugraha, S. I. (2017). The implementation of corpus-aided discovery learning in English grammar pedagogy. *Journal of ELT Research*, *2*(2), 66–83. https://doi.org/10.22236/JER_Vol2Issue2pp66-83

Yilmaz, M., & Ertürk, Z. Ö. (2017). A contrastive corpus-based analysis of the use of reporting verbs by native and non-native ELT researchers. *Novitas-ROYAL (Research on Youth and Language)*, *11*(2), 112–127.

Zarifi, V., & Mukundan, J. (2013). Phrasal verb combinations in corpus-based studies: A critical review. *International Journal of Applied Linguistics & English Literature*, *2*(4), 212–217. https://doi.org/10.7575/aiac.ijalel.v.2n.4p.212