Profiling Word Frequency and Readability of Online Learner Dictionary Definitions

ROBERT DILENSCHNEIDER
General Education Department, School of Medicine, Jichi Medical University, Japan
PAUL HORNESS*
World Language Center, Soka University, Japan
Corresponding author email: paul@soka.ac.jp

Article information

Article history:
Received: 9 Jan 2023
Accepted: 4 Jul 2023
Available online: 25 Jul 2023

Keywords:
Online learner dictionaries
Definitions
Vocabulary
Word frequencies

Abstract

This study examined 283 online learner dictionary definitions in terms of scores based on word frequency level and readability. Results revealed three findings. First, in terms of word frequency levels, definitions from the Cambridge learner dictionary incorporated fewer non-high frequency words (mid and low frequency words) compared to Oxford, Dictionary.com and Collins COBUILD learner dictionaries. Second, in terms of readability, definitions from the Random House learner dictionary were written at a significantly higher grade level compared to the other six online learner dictionaries. Third, in terms of both level of frequency words incorporated into definitions and readability grade level definitions, the Cambridge, Merriam-Webster and Longman online learner dictionaries were easier to understand compared to online learner dictionaries from Dictionary.com, Collins COBUILD, Oxford and Random House. Overall, the findings suggest both word frequency level and readability might contribute to the difficulty of online learner dictionary definitions.

INTRODUCTION

For over the past decade, there have been a number of studies examining students' perceptions and opinions of using online dictionaries as well as their effectiveness for language learning. First, regarding students' preferences, several studies (Dashtestani, 2013; Tulgar, 2017; Wolster, 2015) have found subjects preferred online dictionaries over paper dictionaries due to their speed and convenience. In addition, Rai (2020) revealed students had favorable opinions and perceptions towards the use of mobile dictionaries due to their accessibility, availability, portability, and connectivity.

There also have been several studies examining students' opinions toward the usefulness of online dictionaries. Le and Dao (2019), for example, found that students preferred using online dictionaries because they were thought to be more useful in understanding new words compared to paper dictionaries. Likewise, Trinder (2017) reported online dictionaries were accessed more frequently than other technologies such as social networking and informational websites, and were perceived as being useful for vocabulary learning and writing. Jin and Deifell (2013) found that most foreign language college-level learners preferred using bilingual online dictionaries frequently inside and outside of classroom for reading, writing, listening, and speaking tasks. Le and Dao (2019) reported similar findings and expressed that college-level foreign language learners thought online dictionaries were beneficial for language learning.

Lastly, there have also been several studies that have examined the application of online dictionaries. For instance, there have been several studies that have looked at the use of these devices to learn new words (Chen, 2010, 2012, 2016; Dziemainko, 2010, 2011, 2012, 2017; Hamilton, 2012; Liu & Lin, 2011; Li & Xu, 2015; Rahimi & Miri, 2014). In addition, there have been studies that have examined the impact of online dictionaries on reading comprehension (Alharbi, 2016; Chun, 2001; Dilenschneider, 2017; Liu & Lin, 2011).

Although the popularity, usefulness and application of online dictionaries have been studied, there are issues that might impact the effectiveness of online dictionaries for some English language learners. Akins and Rundell (2008), for example, stated dictionaries designed for learners should employ a controlled defining vocabulary consisting of "a finite list of high-frequency words (typically the most frequent 2,000-3,000 words in the language) which the learner is expected to 'know' sufficiently well to be able to understand any definition in the dictionary" (p. 449). As such, the Longman Dictionary of Contemporary English (LDOCE), a second language learner dictionary of English, and Oxford Advanced Learner's Dictionary both use a controlled defining vocabulary of approximately 2,000-word and 3,000 words, respectively (Atkins & Rundell, 2008).

There are at least two challenges of online learner dictionaries that use a controlled defining vocabulary. First, due to a controlled vocabulary, learner dictionaries might not always precisely describe and capture the meaning of a word because "there are cases where a word cries out for the use of another word in its definition" (Atkins & Rundell, 2008, p. 449), or there is, "difficulty of attaining a natural form of expression in definitions" (Svenson, 2009, p. 248). As a result, there are cases when learner dictionaries need to incorporate words beyond the 3,000-word frequency level to adequately describe the meaning of a word. Consequently, learners whose vocabulary may be limited to the first 3,000 high frequency words, might be unfamiliar with lower frequency words incorporated into definitions and this in turn, might compromise their understanding of an unknown word.

A second challenge that might confront learners when using an online dictionary with a controlled vocabulary is the number of words used to sufficiently define a term. Too many words in a definition may hinder intelligibility because, "even if there are no difficult words in a definition, the sheer number of words may reduce clarity" (Svensen, 2009, p. 240). This is because, "defining in simple terms is usually more space-consuming than defining in more complex terms" (Svensen, 2009, p. 247). Therefore, definitions from online learner dictionaries should avoid long sentences with rare or difficult vocabulary so that learners can get to the point of understanding the meaning of unknown words (Rundell, 2022).

Word levels and definition length

One of the main challenges for learner dictionaries is to ensure the defining vocabulary within a definition is understandable and not more complex than the word that is being described (Scholfield, 2008). Therefore, when it comes to using an online learner dictionary, for its definitions to be understood easily and briefly, both the frequency level of words and the number of words used in a definition to describe unknown terms deserve an evaluation.

Word frequency levels

There are three categories of words that learners confront when they read and these consist of high-frequency words, mid-frequency words, and low-frequency words. High frequency words are typically identified as being the first 2000 (West, 1953) to 3000 (Schmitt & Schmitt, 2012) of the most commonly used words. Along with proper nouns, transparent compounds, and marginal words, high-frequency words account for roughly 90% to 95% of the running words in a typical text. Next, mid-frequency words consist of about 7,000 words beyond the first 3,000 words and range from the third 1,000 to ninth 1,000-word frequency level. These words appear less often than high-frequency words, but provide an additional 9% coverage of a typical text. Last, low-frequency words are the third category of words beyond the 9,000-word frequency level that stretch from the ninth 1,000 to the twenty-fifth 1,000-word frequency level. These words are more specialized than high or mid-frequency words and therefore, usually make up only 1-2% of the running words in a text. Words not included as being a part of the three-word categories are off-list words of and beyond the twenty-fifth 1,000-word frequency level, as well as proper nouns, transparent compounds and marginal words which together account for roughly 3-4% of the words in a typical text (Nation, 2014).

Readability

Another important factor to consider when learners read dictionary definitions is their readability. Readability is the ease in which learners can understand a particular text. There are several different types of readability software programs available online that profile texts and use different statistical formulas to calculate the level of education necessary for readers to understand passages. Typically, after profiling a sentence or passage, these software programs reveal an approximate grade level that learners should hold to easily read a passage. As an example, using the Flesch-Reading-east test, a high readability score of 90.0 -80.0 corresponds to an easy-to-read text appropriate for a student in the 6th grade while a low readability score of 30.0-10.0 corresponds to a highly sophisticated passage appropriate for a college graduate. Higher readability scores reduce the reading effort and speed for the reader and are important for students who do not have strong reading comprehension skills and experience difficulty in understanding complex content. Fortunately, the use of readability tests makes it possible to modify reading passages for learners.

Statement of the problem

Online learner dictionaries that incorporate a high percentage of known high-frequency words and have an easy readability grade level might be more conducive for learners in understanding a new word, as they can easily read and grasp the meaning of a new word and quickly apply it to the context of a reading passage. However, online learner dictionaries that incorporate a higher percentage of unknown mid and low-frequency words and a difficult readability grade level might not always be as favorable for learning. Such definitions may be challenging for low-proficiency learners whose vocabulary may be limited to the first 3,000 high frequency words as they need to look up the definitions of mid and low-frequency words incorporated within a dictionary definition just to understand the meaning of an unknown word. In addition, to sufficiently describe a word, the use of a controlled defining vocabulary in learner dictionaries may create lengthy definitions that "reduce clarity" and are time-consuming for learners to read before returning to the reading passage. As a result, both the level of words and the number of words used in learner dictionary definitions may impact the understanding of a word for low-level learners of English. Therefore, to examine the level of words and number of words incorporated into learner dictionary definitions, the following research questions are proposed:

1. How do the number of mid and low-frequency words used in definitions compare among online learner dictionaries?
2. How do readability grade level scores compare among online learner dictionary definitions?
3. How do both word frequency level and readability grade level affect the difficulty of online learner dictionary definitions?

METHODOLOGY

Online dictionaries

The dictionaries examined in this study were seven monolingual online learner dictionaries attained from Merriam-Webster (The Britannica Dictionary), Oxford Advanced American Dictionary, Cambridge Learner's Dictionary, Dictionary.com, Collins COBUILD Advanced Learner's Dictionary, Word Reference Random House Learner's Dictionary of American English, and Longman Dictionary of Contemporary English (LDOCE). The seven online dictionaries were chosen because they did not require a subscription fee and were listed as either recommended or popular learner dictionaries on language webpages such as Langster Languages (Langster, 2017) and Better@English (Linstruth, 2022). Furthermore, the seven online learner dictionaries can be accessed either through their websites from a computer or their respective applications. Most of these online dictionaries can also be downloaded onto a smartphone or tablet computer for students to use and access virtually anywhere and at any time. Though they are not the mainstream versions of their dictionaries, for the purposes of this study and simplicity's sake, the learner dictionaries mentioned in this study will be referred to by their household names (i.e. Merriam-Webster, Cambridge, etc.). Table 1 below shows the name of each online dictionary used in this study along with its type.

Table 1
Learner dictionary name, website and type

Learner Dictionary Name and Website	Type
*Merriam-Webster (The Britannica Dictionary) https://www.britannica.com/dictionary	Learners
Oxford Advanced American Dictionary https://www.oxfordlearnersdictionaries.com/definition/american_english/	Advanced Learners
Cambridge Learner's Dictionary https://dictionary.cambridge.org/dictionary/learner-english/	Learners
Word Reference Random House https://www.wordreference.com/	Learners
Dictionary.com https://www.dictionary.com/	Learners
Collins COBUILD Advanced Learner's Dictionary https://collinsdictionary.com/	Advanced Learners
Longman Dictionary of Contemporary English (LDOCE) https://www.ldoceonline.com/dictionary/	Learners

Note: *Merriam-Webster Learner Dictionary application defaults to the desktop version of The Britannica Dictionary.

Vocabulary profiler

The Compleat Lexical Tutor (Cobb, 2022) software program was used to examine the defining vocabulary for the target words in this study. This online software program provides several different vocabulary profilers which analyze word frequencies of reading passages. For this study, the BNC-COCA-25 vocabulary profiler was used to facilitate a thorough investigation of defining vocabulary. Specifically, this program uses a combination of one hundred million words worth written material from the British National Corpus (Nation, 2004) and a 450-million-word Corpus of Contemporary English (Davis, 2012) to analyze vocabulary. It categorizes high frequency words from the first 3,000 most frequent words, mid-frequency words from the next seven frequency word levels that stretch from the third 1,000 to the ninth 1,000-word level, low-frequency words beyond the 9,000-word level that extend fifteen word levels from the tenth 1,000 to the twenty-fifth 1,000-word frequency level, and off-list words that are identified as being above the twenty-fifth 1,000-word frequency level.

Target words

The controlled defining vocabulary used to define terms in learner dictionary definitions are high-frequency words that belong to the 1,000 to 3,000-word frequency level (Atkins & Rundell, 2008). Therefore, to find out if some online learner dictionaries incorporate more mid and low-frequency words into their definitions than others, the target words used in this study were all concrete nouns ranging from the 4,000 mid-frequency word level (e.g. emperor, nursery, tumor) to the 12,000 low frequency word level (e.g. miser, knoll, amulet) were retrieved from the Compleat Lexical Tutor list of BNC-COCA 1-25k headwords. Overall, at least 30 different target words were randomly selected for each of the nine different word frequency levels (4,000-12,000) for a total of 283 target words. No high frequency words (1,000–3,000) were selected or used in this study. Appendix A shows the target words in each of the nine different word frequency levels that were retrieved from the Compleat Lexical Tutor software program (Cobb, 2022).

Readability profiler

To examine the readability of definitions provided by online dictionaries, the WebFX Readability Test software program was also used in this study. This online software program provides several different readability tests to gauge the difficulty of sentences and reading passages. For this study, the Coleman-Liau Index (CLI) was used to measure readability by using a statistical formula to measure the ease in which learners can understand a dictionary definition. Like the Flesch-Kinkaid Grade Level, its output reveals the approximate public school grade level in the United States necessary to read a text. However, unlike the Flesch-Kinkaid Grade Level, which uses calculations based on the number of syllables per word in sentences, the CLI uses calculations based on the number of characters or letters per word in sentences. The CLI was created because computers can count, assess, and understand characters more easily and accurately than counting syllables and sentence length. According to Coleman and Liau (1975, p.284), "There isno need to estimate syllables since word length in letters is a better predictor of readability than word length in syllables".

Procedures

Definition word level profiling

The procedure to retrieve and organize the data for the defining vocabulary for each target word involved six steps using the online learner dictionaries, vocabulary profiler and spreadsheet software program. First, using these online learner dictionaries, the spelling of a target word was typed into the search windows of the seven online dictionaries. Second, after typing the spelling of a target word, a submission button near the search window in each online learner dictionary was used to access the definition of a target word. Third, the first presented definition of a target word displayed from each online learner dictionary was copied and pasted into the submission window of the BNC-COCA-25 vocabulary word profiler. Fourth, after the definition of a target word was pasted, a *submit* button within the BNC-COCA-25 vocabulary word profiler was used to view the output of the word frequencies for each word in the definition of the target word. Fifth, the output represented with figures for each token or word in a definition was then transferred to a Microsoft Excel spreadsheet. This procedure was performed by copying the output figures of tokens from the BNC-COCA-25 vocabulary word profiler and pasting them into their corresponding high, mid, low or off-list word frequencies cell categories within the Microsoft Excel spreadsheet. This six-step procedure was conducted for each target word for each of the seven dictionaries. Finally, this procedure was conducted a second time to verify results.

Definition readability profiling

Three steps were conducted to profile the Coleman-Liau Index (CLI) readability grade level for each definition. First, in a one-at-a-time fashion, a definition from one online learner dictionary was typed into the window of the WebFX Readability Test software program. Second, after typing the entire definition of a target word, the Calculate Readability button under the window was clicked to access the readability score for that definition. Third, the numeric output displayed from the CLI was recorded onto an Excel spreadsheet. This three-step procedure of type, submit, and record was conducted for each definition from each of the seven dictionaries. Finally, this procedure was conducted a second time to verify results.

Scoring

Frequency has been assumed to be the largest factor in determining the difficulty of learning a word (Milton, 2009; Schmitt, 2010). Hashimoto and Egbert (2019) likewise concluded that frequency was strongly related to word difficulty but also claimed there were other factors of word difficulty that need to be researched. Therefore, to explore another factor, for this study, a difficulty score based on two parts was created for each vocabulary item in a definition: word frequency level and readability. The first part was based on the dictionary's definition in relation to the word frequency level. Table 2 reveals the point value that was afforded to each word level.

Table 2
Calculating difficulty score based on word frequency level

	Word Frequency Level				
	1,000-3,000	4,000-6,000	7,000-9,000	10,000-12,000	Over 12,000
Point	1	3	5	7	9

Each word in the definition was classified by its word frequency level based on how common or rare it is based on the BNC-COCA-25 vocabulary word profiler and then ascribed a point value. However, words that have the same number of letters and syllables do not necessarily belong to the same word frequency level. For instance, although both *angry* and *irate* have the same number of letters, syllables and hold similar meanings, *angry* is a high-frequency word belonging to the 1,000-word family and *irate* is a low-frequency word belonging to the 10,000-word family. Likewise, *delighted* and *overjoyed* each have nine letters, three syllables and similar meanings, but belong to the 1,000 high-frequency word level and 10,000 low-frequency families, respectively. Therefore, higher-frequency words were afforded lower point values than low-frequency words because they are more common. The point values for all the words incorporated into a definition were then added up which gave it a difficulty score. For example, the term *garlic* displayed a different frequency score from each dictionary because both the number and word frequency level of words used to describe this term varied.

As shown from Table 3 below, the Frequency Score was determined by both the number of words in a definition and its word frequency level. For example, the Oxford Dictionary used 21 high-frequency words belonging to the 1,000–3,000-word level to define the term *garlic*. Since all words belonging to the 1,000–3,000-word level are each afforded a point value of one, the overall Frequency Score from the Oxford Dictionary in defining *garlic* was 21 (21 [1,000–3,000] x 1 Point = 21). In contrast, though the Random House Dictionary used fewer words to define the same term, the dictionary yielded a higher Frequency Score of 27 because its definition contained mid and low-frequency words that were not often used or encountered, and thus, were each afforded a higher point value (12 [1,000–3,000] x 1 Point = 12; 1 [4,000–6,000] x 3 Points = 3; 1 [7,000–9,000] x 5 Points = 5; 1 [10,000–12,000] x 7 Points = 7 or 12 + 3 + 5 + 7 = 27).

Table 3

Total frequency score from each dictionary for the word garlic

	Word Frequency Level					
	1,000-3,000	4,000-6,000	7,000-9,000	10,000-12,000	Over 12,000	
Points	1	3	5	7	9	
Learner Dictionary						Freq. Score
Merriam-Webster	27	1	0	0	0	30
Oxford	21	0	0	0	0	21
Cambridge	13	0	0	0	0	13
Dictionary.com	26	2	0	0	0	32
Collins COBUILD	26	2	0	0	0	32
Random House	12	1	1	1	0	27
Longman	14	0	0	0	0	14

The second part of the difficulty score was readability. The Coleman-Liau Index (CLI) was used to measure the readability of dictionary definitions. As mentioned previously, the CLI uses calculations based on the number of characters or letters per word in sentences to determine the approximate school grade level appropriate to read a text. For instance, upon entering the Merriam Webster's definition for the term *garlic* into the text window of the WebFX Readability Test software program, a figure of 9.4 appears. This value indicates that Merrian Webster's definition of the term *garlic* is somewhere around the 9th grade level. So, to determine the overall difficulty, both the Frequency Score and CLI Score were combined to calculate a Total Difficulty Score. Table 4 below reveals how both the Frequency Score and CLI Score were used to calculate the Total Difficulty Score for the term *garlic*.

Table 4

Combining scores for a total difficulty score in each dictionary for the word garlic

Learner Dictionary	Frequency Score	CLI Score	Total Difficulty Score
Merriam-Webster	30	9.4	39.4
Oxford	21	8.0	29.0
Cambridge	13	8.2	21.2
Dictionary.com	32	7.3	39.3
Collins COBUILD	32	7.3	39.3
Random House	27	12.5	39.5
Longman	14	6.1	20.1

RESULTS

When setting up the SPSS data, there was no hypothesis to which order the dictionaries would be in, so the input order was a matter of convenience. There were three levels of analysis conducted. First, a two-way ANOVA was conducted with the combined word difficulty score as the dependent variable and the different dictionaries and word frequency levels as the independent variables. Second, separate analyses were conducted on each element of word difficulty definition so that a clearer relationship between them could be ascertained. Finally, the Rasch Model software of Facets was used to examine the relationship of all factors on one scale of analysis.

Research question 1:

How do the number of mid and low frequency words used in definitions compare among online learner dictionaries?

A two-way ANOVA was conducted to determine whether there was a difference between dictionary type and frequency level. The dependent variable was the word frequency score and the independent variables were dictionary type and word frequency level. After eliminating z-scores beyond ± 2.96 (21 terms), 1960 terms were checked. The descriptive statistics are given in Table 5. Cambridge had the lowest mean score and the Dictionary.com had the highest mean score. Significant differences were found between the dictionary type and word frequency levels, but there was no significant interaction between the two. A Welch ANOVA was used for each independent variable as Levene's Test was significant and there were unequal group sizes.

Table 5

Descriptive statistics for word frequency difficulty scores

Dictionary	N	*M*	*SD*
Merriam-Webster	282	15.68	7.35
Oxford	279	16.78	6.87
Cambridge	283	13.86	6.85
Dictionary.com	280	17.41	7.78
Collins COBUILD	277	17.23	7.68
Random House	277	15.64	7.65
Longman	282	14.81	6.29

The Welch ANOVA Test indicated a statistically significant difference between the dictionaries, $F(6, 867.10) = 10.58$, $p < .05$, $\eta^2 = .016$. The Robust Tests of Equality of Means was significant, so follow-up pairwise comparisons were made using Games-Howell post hoc tests with Bonferroni adjustment setting alpha at .007 (.05/7). Figure 1 helps visualize the significant comparisons for Merriam-Webster (1), Oxford (2), Cambridge (3), Dictionary.com (4), Collins COBUILD (5), Random House (6) and Longman (7) dictionaries. Overall, Cambridge was significantly different than the other dictionaries except Merriam-Webster, Longman, and Random House. There were no other significant differences. In essence, there were three levels of difficulty. Cambridge is the easiest, while Oxford, Dictionary.com and Collins COBUILD dictionaries were the most difficult. In between, Merriam-Webster, Random House and Longman were not significantly different between the two groups.

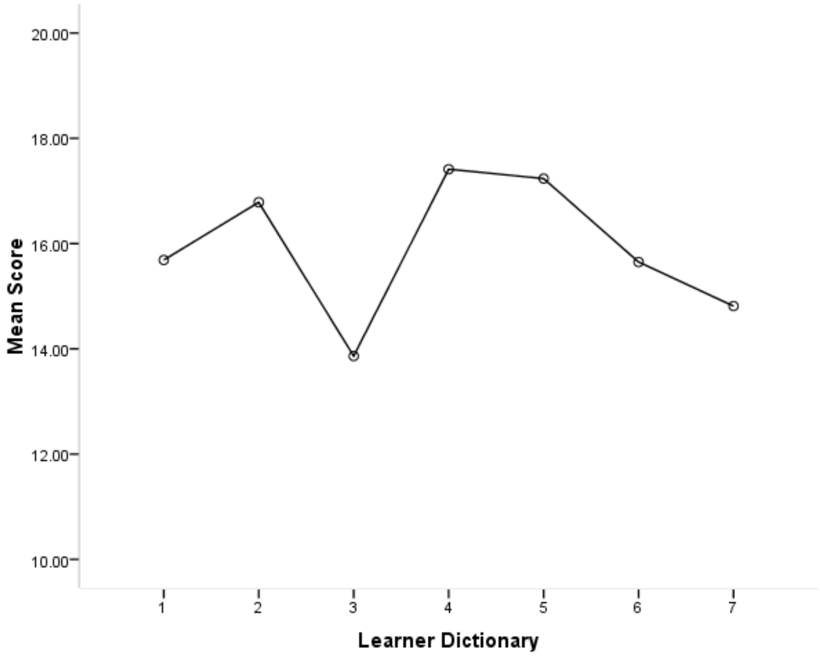

Figure 1 Frequency scores for learner dictionaries

Note: 1-Merriam-Webster, 2-Oxford, 3-Cambridge, 4-Dictionary.com, 5-Collins COBUILD, 6-Random House, 7-Longman.

Research question 2:

How do readability grade level scores compare among online learner dictionary definitions?

A two-way ANOVA was conducted again to determine whether the CLI score was different among the various dictionaries and whether the word frequency levels were different. The dependent variable was the CLI score and the independent variables were the dictionary types and word frequency levels. After eliminating z-scores beyond ± 2.96 (13 terms), 1968 terms were checked. As Levene's Test value was less than .05 and there were unequal group sizes, a Welch ANOVA was used separately for the dictionary type and frequency levels. The interaction between dictionary type and frequency level was not examined as it was not statistically significant in the initial two-way ANOVA. The Robust Tests of Equality of Means was significant, therefore the Games-Howell post-hoc tests were used for interpretation. The descriptive statistics are given in Table 6. The lowest mean score was Merriam-Webster and the highest was Random House.

Table 6
CLI descriptive statistics for dictionary types

Learner Dictionary	N	*M*	*SD*
Merriam-Webster	283	8.08	2.73
Oxford	280	8.61	2.93
Cambridge	282	8.60	3.25
Dictionary.com	281	8.73	3.07
Collins COBUILD	282	8.72	3.08
Random House	278	10.10	3.78
Longman	282	8.44	2.95

The Welch ANOVA Test indicated a statistically significant difference between the dictionaries, $F(6, 870.55) = 8.91$, $p < .05$, $\eta^2 = .017$. Follow-up pairwise comparisons were made using Games-Howell post hoc tests with Bonferroni adjustment setting alpha at .007 (.05/7). Figure 2 helps visualize the significant comparisons for the Merriam-Webster (1), Oxford (2), Cambridge (3), Dictionary.com (4), Collins COBUILD (5), Random House (6) and Longman (7) dictionaries. Only Random House was significantly different than all the other dictionaries. This reflects how Random House used more words in their definitions which increased the number of characters used to calculate CLI.

Figure 2 CLI scores for the seven online leaner dictionaries

Note: 1-Merriam-Webster, 2-Oxford, 3-Cambridge, 4-Dictionary.com, 5-Collins COBUILD, 6-Random House, 7-Longman.

Research question 3:

How do both word frequency level and readability grade level affect the difficulty of online learner dictionary definitions?

A two-way ANOVA was conducted yet again to determine if there were significant differences among the dictionary types and word frequency levels. The dependent variable was the combined word difficulty score, which combines the CLI and word frequency scores. The independent variables were dictionary type and word frequency level. After eliminating z-scores beyond ± 2.96 (18 terms), 1963 terms were checked. The descriptive statistics are given in Table 7 below. Significant differences were found among the dictionaries and frequency levels, but there was no significant interaction between the dictionaries and frequency levels. A Welch ANOVA was used for each independent variable as Levene's Test was significant and there were unequal group sizes.

Table 7
Descriptive statistics for the learner dictionaries difficulty scores

Learner Dictionary	N	*M*	*SD*
Merriam-Webster	282	23.79	7.79
Oxford	279	25.52	7.37
Cambridge	283	22.51	7.66
Dictionary.com	280	26.24	8.22
Collins COBUILD	278	26.11	8.34
Random House	278	25.99	8.96
Longman	283	23.39	6.96

The Welch ANOVA Test indicated a statistically significant difference between the dictionaries, $F(6, 868.46) = 10.61$, $p < .05$, $\eta^2 = .031$. Follow-up pairwise comparisons were made using Games-Howell post hoc tests with a Bonferroni adjustment setting alpha at .007 (.05/7). Figure 3 helps visualize the significant comparisons for Merriam-Webster (1), Oxford (2), Cambridge (3), Dictionary.com (4), Collins COBUILD (5), Random House (6) and Longman (7) dictionaries. Overall, Cambridge had the lowest mean score and Dictionary.com had the highest. Cambridge and Longman Dictionaries were significantly different than all dictionaries except the Merriam-Webster. The Merriam-Webster was not significantly different than the other dictionaries except Dictionary.com.

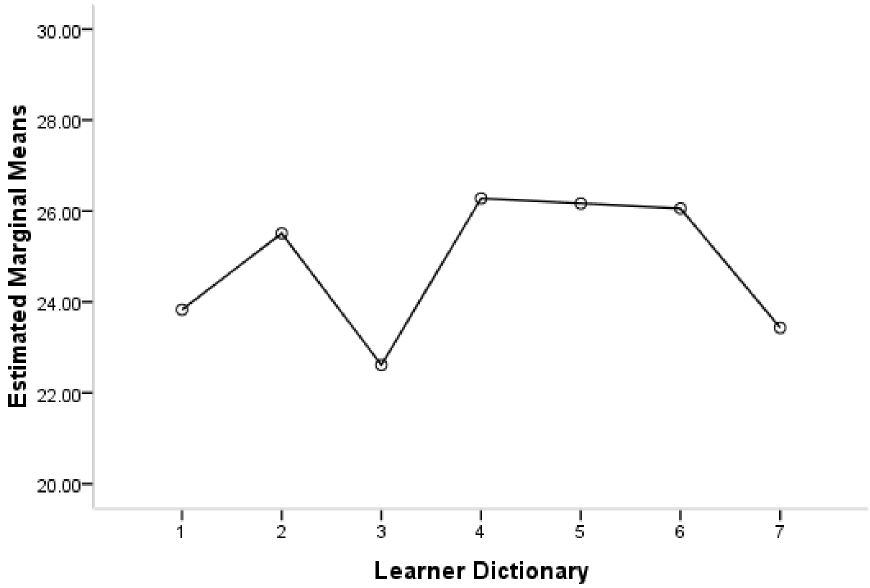

Figure 3 Estimated marginal means of difficulty scores among learner dictionaries

Note: 1-Merriam-Webster, 2-Oxford, 3-Cambridge, 4-Dictionary.com, 5-Collins COBUILD, 6-Random House, 7-Longman.

Concurrent analysis

To further confirm the findings from SPSS, a final analysis used the Rasch Model (see Bond & Fox, 2015 for a detailed explanation). Facets (Linacre, 2022) is a software that implements the Rasch Model so all the data can be placed on the same scale for observation. The collected data explained 70.32% of the variance. The reliability of the model is .96. The Fit statistics' criteria for the dictionaries were set at .5-1.5. All the dictionaries met the criteria except Random House: Outfit was 1.71 and infit was 1.64. These numbers indicate the measurement for Random House does not fit the Rasch model well. Another Rasch Model output to help us understand is the Wright Map as seen in Figure 4 below. There are several points to notice, moving from left to right. First, the vocabulary terms illustrate a bell curve, so the number of terms seems to be sufficient to compare the dictionaries. Second, the order of dictionaries reflected the results of SPSS. Cambridge was considered the easiest dictionary while Dictionary.com was considered the most difficult. Third, the word frequency level was not ranked in descending order of difficulty. The assumption that higher word frequency equals more difficulty was not exactly accurate.

```
+----------------------------------------------------------------------+
                             |Measurement
   |+terms     |+dictionary type           |+word level |Scale|
|-----+-----------+-----------------------------+------------+-----|
|   1 +           +                             +            +(60) |
|     |           |                             |            |     |
|     |           |                             |            | 56  |
|     |           |                             |            | 55  |
|     |           |                             |            | 53  |
|     |           |                             |            | 49  |
|     | .         |                             |            | 44  |
|     | **.       |                             |            | 38  |
|     | ****.     |                             | 11         | 33  |
|     | ******    | CollinsCOBUILD  Dict.com    | 12   5   7 | 29  |
*   0 * ********* * MW        Oxford   Random   * 10   8     * 26  *
|     | *****.    | Cambridge Longman           | 6    9     | 23  |
|     | ****.     |                             | 4          | 21  |
|     | ***.      |                             |            | 18  |
|     | *.        |                             |            | 17  |
|     | *         |                             |            | 15  |
|     | .         |                             |            | 13  |
|     | .         |                             |            | 12  |
|     | .         |                             |            | 11  |
|     | .         |                             |            | 10  |
|  -1 +           +                             +            + (5) |
|-----+-----------+-----------------------------+------------+-----|
|Measurement  | * = 7     |+dictionary type           |+word level
                             |Scale|
+----------------------------------------------------------------------+
```

Figure 4 Facet summary wright map

Delving deeper into the Rasch Model, it produced both an observed mean score and fair mean score for each dictionary as shown in Table 8 below. Cambridge has the lowest mean scores and was the easiest while Collins COBUILD, Dictionary.com, and Random House were considered equally as most difficult. The Rasch Model indicated that the seven dictionaries could be broken down into five distinct measurements which reflected a similar result from SPSS.

Table 8
Rasch dictionary measurement of mean scores

Dictionary	Observed Mean	Fair Mean
Collins COBUILD	26.72	25.61
Dictionary.com	26.67	25.56
Random House	26.63	25.53
Oxford	25.81	24.84
Merriam-Webster	24.02	23.14
Longman	23.47	22.63
Cambridge	22.61	21.83

The order of word level difficulty was a little unexpected. The order of difficulty was expected to begin with 4,000-word frequency level as the easiest and followed sequentially to the 12,000-word frequency level as the hardest. Although the score for the 4,000-word level was the easiest, the 5,000-word level was positioned near the top in difficulty. Table 9 below indicates the sequential order of difficulty by the observed and fair mean scores from hardest to easiest. This indicates that the definitions are incorporating words beyond their usual word frequency level. For example, in defining the term *cassette* (level 5,000), words from lower word frequency levels (>5,000) were used, making the term more difficult to understand. Hence, the importance of using CLI.

Table 9
Rasch observed and fair mean scores for word level

Word Frequency Level	Observed Mean	Fair Mean
11,000	26.83	28.94
12,000	26.43	28.20
5,000	25.55	26.11
7,000	25.67	25.63
10,000	25.47	25.42
8,000	24.42	23.26
9,000	24.31	21.32
6,000	23.85	20.46
4,000	23.95	20.05

DISCUSSION

Overall, this study examined the frequency level of words and number of words in learner dictionary definitions. The first research question examined how the number of mid and low-frequency words used in definitions compared among online learner dictionaries. Results revealed that the Cambridge online learner dictionary was found to have the lowest mean among all dictionaries and was significantly different to the Oxford, Dictionary.com and Collins COBUILD learner dictionaries. Therefore, in sum, the Cambridge dictionary could be considered as the easiest, while the Merriam-Webster, Random House and Longman dictionaries were moderately more difficult, and Oxford, Dictionary.com and Collins COBUILD dictionaries were the most difficult.

When it comes to the controlled defining vocabulary of online leaner dictionaries, learners should know the most frequent 2,000-3,000 high frequency words to be able to understand any dictionary definition (Atkins & Rundell, 2008). This study showed that, compared to Oxford, Dictionary.com, Collins COBUILD and Random House dictionaries, the Merriam-Webster, Cambridge and Longman online learner dictionaries have fewer mid and low-frequency words in their definitions. As a result, these resources might be better for learners who want to use an online learner dictionary but are not yet well-acquainted with mid-frequency, low- frequency words or off-list words. Similarly then, the more complex dictionaries would be well-suited for learners who are keen on learning new words and expanding their vocabulary, an idea that falls in line with Hashimoto and Egbert's (2019) conclusion that repeated exposure to a word can help reduce word difficulty.

The second research question examined how readability grade level scores compare among online learner dictionary definitions. Results showed that the biggest impact of CLI is that the definitions from the Random House learner dictionary were significantly higher and more difficult to read than definitions from the other six learner dictionaries. When it comes to how definitions are written, the number of words incorporated in a definition can compromise its clarity (Rundell, 2006). This study revealed that definitions from the Random House online learner dictionary were longer which significantly contributed to more difficult CLI readability grade level scores compared to the other six online learner dictionaries. Consequently, though this online dictionary is a learner dictionary, its definitions might actually be as difficult to understand as those from an advanced learner dictionary.

The third research question examined how both word frequency level and readability grade level impacted the difficulty of online learner dictionary definitions. Based solely on word frequency levels in definitions, there were three groups of dictionaries identified. Cambridge dictionary was the easiest, the Merriam-Webster, Random House and Longman dictionaries were moderately more difficult and Oxford, Dictionary.com and Collins COBUILD dictionaries were the most difficult. However, because CLI scores from the Random House dictionary were found to be significantly higher than the other six dictionaries, the Total Difficulty Score for this dictionary shifted from the easier group of learner dictionaries (Cambridge, Merriam-Webster, and Longman) to the harder group of advanced learner dictionaries (Oxford, Dictionary.com and Collins COBUILD).

The results involving the word frequency levels in definitions might be expected from the type of learner dictionary profiled. For instance, on one hand, dictionaries from the easier group, such as Cambridge and Merriam-Webster, are presented on their websites as being a Learner Dictionary. Even information from the Longman dictionary's website says it, "uses 2000 common words in definitions to make understanding easy" (Longman Dictionary of Contemporary English Online, 2023). This coincides with Atkins and Rundell (2008) regarding the use of high-frequency words that learners are expected to know in order to understand the defining vocabulary of dictionary definitions. On the other hand, dictionaries from the harder group are possibly designed for learners with higher English language proficiency. The Oxford online dictionary used in this study, for instance, is described as advanced-level and the Collins COBUILD Dictionary is described as for advanced learners. Likewise, Dictionary.com is also an advanced learner dictionary because it is s a proprietary of Collins COBUILD.

Nevertheless, when it comes to looking up unknown words while reading, previous studies have shown that the extra time and effort learners might expend to look up unknown words can compromise their attention towards learning a word and understanding a reading passage (Dilenschneider, 2017; Dziemainko, 2017). This study has shown that definitions from the Merriam-Webster and Cambridge online learner dictionaries and the Longman online dictionary incorporate fewer mid-and low frequency words and have an easier readability grade level. Therefore, compared to the other four online learner dictionaries, these resources might be more conducive for learners to use and quickly understand the meaning of target words so that they can immediately apply them to the context of a sentence without having to look up additional unfamiliar words within a dictionary definition that might cause them to lose focus, momentum, and flow in comprehending the content of a reading passage.

CONCLUSION

This study statistically examined how the types of words (word frequency level) and number of words (readability grade level) might theoretically impact the difficulty of online learner dictionary defintions. Overall, analyses from both SPSS and Rasch Model indicated that different online learner dictionaries were not equal in difficulty based on the word level frequency score and CLI readability score. Results from both SPSS and Rasch indicated that there were two main groups. Cambridge, Longman, and Merriam-Webster were in the easier group while Oxford, Random-House, Collins COBUILD, and Dictionary.com were in the harder group. The first research question showed that word frequency level had a clear influence and role in determining the difficulty score among learner dictionaries. However, the second research question showed that CLI readability can differ among dictionaries. Together, the third research question showed that both the word frequency levels and the readability from CLI scores might both be an integral component that can contribute to the understanding of dictionary definitions, because even though online learner dictionaries might employ simple defining vocabulary, the way they are written might make them challenging for learners to understand.

There are certainly other factors such as the repetition of a word (Schmitt, 2000, 2010), the context in which a word in a passage is presented (Nation, 2001) and other lexical characteristics

(Hashimoto & Egbert, 2019) that contribute to learning vocabulary when learners read a text. Therefore, this study does not claim that one online learner dictionary might be necessarily better for all types of learners. Rather, factors such as the level of defining vocabulary and readability grade level together might impact the difficulty of learners understanding definitions. That said, future resarch might consider the following limitations to better understand the application of this study for language learning. First, the target words used in this study were mostly concrete nouns. To determine whether findings from this study can be applicable to different parts of speech, future studies could examine how verbs, adverbs and adjectives are defined among online learner dictionaries. Second, the profiling for the words in this study occurred at the beginning of the year 2022. As such, it is possible that the definitions of certain words in the online learner dictionaries used in this study might have already been modified or changed since then. Third, only the first definitions were taken from each of the words profiled in each dictionary. The findings from this study could potentially be different if secondary definitions from each dictionary were profiled. Fourth, the online Oxford dictionary used is in this study was the Advanced American Dictionary, an *advanced-level monolingual dictionary for learners of American English*. Overall, this dictionary was not found to be among the group of easier dictionaries; however, its ranking might be different if definitions were instead retrieved from the Oxford Advanced Learner's Dictionary, an *advanced-level dictionary for learners of English* even though both online dictionaries seem to share many of the same definitions. Finally, additional studies might explore how students of different English language proficiencies respond to definitions from different online learner dictionaries. For example, qualitative studies might examine to what extent the incorporation of mid to high-frequency words included in a definition or the number of words in a definition, hinders or frustrates learners in their learning of a new word when reading sentences and passages under the dynamic of a time constraint.

THE AUTHORS

Robert Dilenschneider is an associtate professor at Jichi Medical University in Tochigi, Japan. He received his Ed.D. in Educaiton from Temple University, Japan. He has published articles related to the fields of study abroad, vocabulary acquisition and computer assisted language learning.
bdilenschneider@jichi.ac.jp

Paul Horness is an associate professor at Soka University in Tokyo, Japan. He received his Ph.D. in Education from Temple University, Japan, and teaches research methods and assessment in the ILE TESOL program. He has published articles relating to study abroad programs, formulaic language, and international relations.
paul@soka.ac.jp

REFERENCES

Alharbi, M. A. (2016). Using different types of dictionaries for improving EFL reading comprehension and vocabulary learning. *JALT CALL Journal, 12*(2), 123-149.

Atkins, B. T., & Rundell, M. (2008). *The Oxford guide to practical lexicography*. Oxford University Press.

Bond, T. G., & Fox, C. M. (2015). *Applying the Rasch model: Fundamental measurement in the human sciences* (3rd ed.). Routledge/Taylor & Francis Group.

Chen, Y. (2012). Dictionary use and vocabulary learning in the context of reading. *International Journal of Lexicography, 25*, 216-247.

Chen, Y. (2016). Dictionary use for collocation production and retention: A call-based study. *International Journal of Lexicography, 30*(2), 225-251.

Cobb, T. (2023). *Compleat lexical tutor software*. https://www.lextutor.ca/vp/comp/

Coleman, M., & Liau, T. L. (1975). A computer readability formula designed for machine scoring. *Journal of Applied Psychology, 60*, 283–284.

Chun, D. (2001). L2 reading on the web: Strategies for accessing information in hypermedia. *Computer Assisted Language Learning, 14*, 367-403.

Dashtestani, R. (2013). EFL teachers' and students' perspectives on the use of electronic dictionaries for learning English. *CALL-EJ, 14*(2), 51-65.

Davis, M. (2008). *The corpus of contemporary American English: 425 Million words, 1990-present*. http://corpus.byu.edu/coca/

Dilenschneider, R. F. (2017). Examining the conditions of using an on-line dictionary to learn words and comprehend texts. *ReCALL, 30*(1), 4-23.

Dziemianko, A. (2010). Paper or electronic? The role of dictionary form in language reception, production and the retention of meaning and collocations. *International Journal of Lexicography, 23*(3), 257-273.

Dziemianko, A. (2011). Does dictionary form really matter? In K. Akasu & S. Uchida (Eds.), *ASIALEX2011 Proceedings, Lexicography: Theoretical and Practical Perspectives* (pp. 92-101). Asian Association for Lexicography.

Dziemianko, A. (2012). Why one and two do not make three: Dictionary form revisited. *Lexikos, 22*, 195-216.

Dziemianko, A. (2017). Dictionary form in decoding, encoding and retention: Further insights. *ReCALL, 29*(3), 335-356.

Hamilton, H. (2012). The efficacy of dictionary use while reading for learning new words. *American Annals of the Deaf, 157*(4), 358-372.

Hashimoto, B. J., & Egbert, J. (2019). More than frequency? Exploring predictors of word difficulty for second language learners. *Language Learning, 69*(4), 839-872.

Heuberger, R. (2016). Learner's dictionaries: History and development; current issues. In P. Durkin (Ed.), *The Oxford handbook of lexicography* (pp. 25–43). Oxford University Press.

Jin, L., & Deifell, E. (2013). Foreign language learners' use and perception of online dictionaries: A survey study. *MEROLT Journal of Online Learning and Teaching, 9*(4), 515-533.

Langster. (2017). Best English online dictionaries for beginners. *Langster Languages*. https://langster.org/en/blog/best-english-online-dictionaries-for-beginners/

Le, T. K., & Dao, T. T. (2019). An investigation into students' perception of utilizing online dictionaries in translation-interpretation. *Espacios, 40*, 1-8.

Li, L., & Xu, H. (2015). Using an online dictionary for identifying the meanings of verb phrases by Chinese EFL learners. *Lexikos, 25*, 191-209.

Linacre, J. M. (2022). *Winsteps 5.2: Rash analysis and rasch measurement software*. https://www.winsteps.com/

Linstruth, L. (2022). The best free online dictionaries for English Learners. *Better@English*. https://www.betteratenglish.com/the-best-free-online-english-dictionaries/

Liu, T., & Lin, P. (2011). What comes with technological convenience? Exploring the behaviors and performances of learning with computer-mediated dictionaries. *Computers in Human Behavior, 27*, 373-383.

Longman Dictionary of Contemporary English Online. (2023, June 28). *About LDOCE Online: Longman English Dictionaries*. https://www.ldoceonline.com/about.html

Milton, J. (2009). *Measuring second language vocabulary acquisition* (Vol 45). Multilingual Matters.

Nation, I. S. P. (2001). *Learning vocabulary in another language*. Cambridge University Press.

Nation, I. S. P. (2004). A Study of the most frequent word families in the British National Corpus. In P. Bogaards & B. Laufer (Eds.), *Vocabulary in a second language: Selection, acquisition, and testing* (pp. 3-13). John Benjamins Publishing Company.

Nation, I. S. P. (2014). How much input do you need to learn the most frequent 9,000 words? *Reading in a Foreign Language, 26*, 1-16.

Rahimi, M., & Miri, S. (2014). The impact of mobile dictionary use on language learning. *Procedea – Social and Behavioral Sciences, 98*, 1469-1474.

Rai, H. K., (2020). *Students' perception towards the use of mobile dictionary app for vocabulary learning* [Masters Thesis, Tribhuvan University], 1-47.

Rundell, M. (2006). More than one way to skin a cat: Why full-sentence definitions have not been universally adopted. *In Proceedings of the 12th Euralex International Congress*, 323-337.

Rundell, M. (2023). *What makes a good dictionary example?* Future learn: Understanding English dictionaries. https://www.futurelearn.com/info/courses/understanding-dictionaries/0/steps/62849/

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

Schmitt, N. (2010). *Researching vocabulary: A vocabulary research manual*. Palgrave Macmillan.

Schmitt, N., & Schmitt, D. (2012). A reassessment of frequency and vocabulary size in L2 vocabulary teaching. *Language Teaching*, 1-20.

Scholfield, P. (2008). *Dictionaries*. Centre for Languages, Linguistics & Area Studies. https://web-archive.southampton.ac.uk/www.llas.ac.uk/resources/gpg/229.html/

Svensen, B. (2009). *A handbook of lexicography*. Cambridge University Press.

Trinder, R. (2017). Informal and deliberate learning with new technologies. *ELT Journal, 71*(4), 401-412.

Tulgar, A. T. (2017). Dictionary use of undergraduate students in Foreign Language Departments in Turkey at present. *Universal Journal of Educational Research, 5*(12), 51-57.

West, M. (1953). *A general service list of English words*. Longman, Green & Co.

Wolster, L. (2015). *Dictionary use and preferences of L2 English learners in an intensive English context* [Unpublished PhD thesis, St. Cloud State University].

WebFX (2023). *Readability test software*. https://www.webfx.com/tools/read-able/

Appendix A

Word Frequency Levels (Number of words)								
4000 (31)	5000 (32)	6000 (31)	7000 (32)	8000 (30)	9000 (33)	10,000 (32)	11,000 (31)	12,000 (31)
amateur	balcony	admiral	acorn	axle	alcove	bayonet	anvil	albino
ballot	cassette	banquet	barricade	bandit	banister	canary	beaker	banjo
canal	depot	cavern	canyon	cactus	canister	diaper	cyanide	contraband
dam	enzyme	deacon	detergent	daisy	deluge	executor	armada	ditty
emperor	fungus	embargo	elm	eel	edifice	frigate	epitaph	elixir
fog	globe	famine	fable	fetus	fawn	gist	flotilla	forceps
garlic	hay	garrison	gangster	goblin	goblet	hamster	glade	gazebo
habitat	ivory	hare	harp	hermit	holster	inferno	hatchet	hovel
inmate	jockey	infantry	inlet	ivy	insomnia	jaundice	kerosene	knoll
jungle	lesion	jeep	jargon	jubilee	juncture	lentil	lectern	lullaby
lemon	marsh	lizard	lagoon	lavatory	lilac	noose	mangrove	miser
mist	nun	maiden	malaria	maverick	morphine	ointment	nemesis	notary
nursery	orphan	nylon	nomad	nicotine	nape	partridge	oboe	obelisk
olive	paddle	orchard	obituary	otter	odyssey	quark	paragon	pandemic
patriot	quest	pamphlet	panther	pantry	prodigy	rampart	quorum	quiche
receipt	reef	quilt	raven	quail	quartz	scalpel	rodeo	regatta
salmon	swamp	raft	spade	revolver	ravine	talon	schooner	sleigh
tumor	tomb	saloon	tapestry	scribe	smog	vigilante	trinket	typhoon
virgin	vase	thorn	venom	thicket	turnip	wafer	viola	viceroy
widow	witch	vault	watt	veranda	virtuoso	chalet	welt	walrus
medal	ham	wig	antenna	waltz	wimp	awning	aqueduct	amulet
bruise	cushion	barge	concert	clover	fudge	parapet	gong	imposter
dome	clown	bunker	denim	podium	mermaid	cog	kilt	motorcade
needle	pony	cockpit	hive	tundra	oar	dungeon	legume	epilogue
bridge	logo	den	melon	baton	tycoon	grotto	machete	harpoon
surgeon	aquarium	lava	oats	flask	chasm	moron	vermin	orifice
tutor	stove	mole	puddle	kayak	dimple	hemp	malady	pontoon
cord	vase	ore	foyer	pentagon	acne	ladle	latrine	skiff
dock	badge	plum	rodent	rink	hoe	placard	midget	tampon
rack	ranch	wrench	hamlet	yolk	moat	penthouse	cairn	phalanx
ferry	yacht	buddy	raven		sauna	tempest		islet
	meadow		dagger		sentry			
					cyst			