

Linh Ai Tran
Phu Quoc High School, Vietnam

This paper aims to describe the distribution of the vocabulary presented in reading texts and listening transcripts from the new English textbook for high school students in 11th grade in Vietnam. A mini-corpus of all the reading and listening texts from the textbook was analyzed and compared with Google Books and data from the Corpus of Contemporary American English (COCA). The analysis focused on the distribution of the most frequent function and content words as well as the collocation of the three most typical content words have, parents and people appearing in the textbook. An analysis of the concordance from the COCA data is carried out to see whether the use of the vocabulary in the new English textbook for 11th grade students is in line with the English used in the United States.

Vocabulary acquisition has a great effect on higher-level language processing (Nation & Collins, 2002). Nation and Nation (2001) considered the learning of vocabulary as the heart of language acquisition as it is the social and linguistic structure of language. In other words, learners of language need a foundation of basic vocabulary in order to acquire a language (Spada, 2004).

It is clear that vocabulary is essential in language learning, but does a language learner have to acquire every word? The answer comes from the studies which focus on the significance of word frequency in language acquisition (Nation, 2004). Nation discussed the role of teaching vocabulary in intensive reading. He suggested that one of the principles for vocabulary teaching is the priority of highly frequent words. Therefore, language learners and teachers should pay considerable attention to vocabulary with high frequency (Nation, 2004).

Moreover, successful vocabulary acquisition requires a high level of exposure to the items. Spada (2004) believed that among the factors that make new vocabulary more easily learnable by second language learners is the frequency with which the word is seen, heard and understood (Nation, 2004). She suggested an estimate of at least sixteen times of exposure to a word for learners to acquire it. Nation's study also suggested that frequent engagement in new words is necessary for learners to acquire language. As a result, high frequency words and the frequent exposure to them should be emphasized for language learners.

! " ##

The new English textbook is written for students with the B3 English level and to be taught by teachers with C1 level of English. The textbook was published by Vietnamese Publish House of Education with the professional and technical cooperation of Pearson Education. The textbook is available at www.pearson.com.

some high schools throughout Vietnam" In the near future, this new English textbook is going to be applied widely around Vietnam after some revision" This paper aims to describe and analyze the vocabulary in this textbook "

\$ %

1. What are the most frequent content words in the new English 11th grade textbook in Vietnam:
2. What are the collocations of the most frequent content words in this textbook :

" & "

corpus is an electronically stored collection of language which occurs naturally in mainly spoken and written form (Leppin, 2004) For the analysis of language distribution in the new textbook of English for Vietnamese students in grade 11, three corpora were employed" The first one is a mini corpus consisting of reading and listening passages taken from 11th grade English (author, year, totaling 3,000 words and analyzed by using the software AntConc 5.3.3 (Anthony, 2011) Second, the Google Books collection retrieved from <https://books.google.com/ngrams>

Meanwhile, the top seven words in the (* (# word frequency list) retrieved from http://www.wordfrequency.info/free.asp:sl_y+ include the, &, an, /, of, a, in, and to) Figure 3+

Frequency			
rank	word/lemma	PoS	frequency
1	the	DT	1144
2	&	CC	1134
3	an	DT	1134
4	/	CC	1134
5	of	IN	1134
6	a	DT	1134
7	in	IN	1134
8	to	TO	1134
9	and	CC	1134
10	is	VB	1134

Figure 3: The top ten words in (* (#

Thus, the most frequent function words in the textbook appear in the top seven of the (* (# word frequency list" In this regard, the textbook vocabulary seems to pattern with the American English corpus"

) (" #analysis of the textbook reading and listening passages shows that the top content words are have, people, an, / parents" Figure 5 shows the frequency of these words in Google's book collection" Generally, have is used far more than people and parents, especially in the 134s" Its frequency dropped dramatically from 1144 to 114, and it has appeared consistently since 114" Meanwhile people and parents are steadily used in books through time, with people being more frequently used than parents"

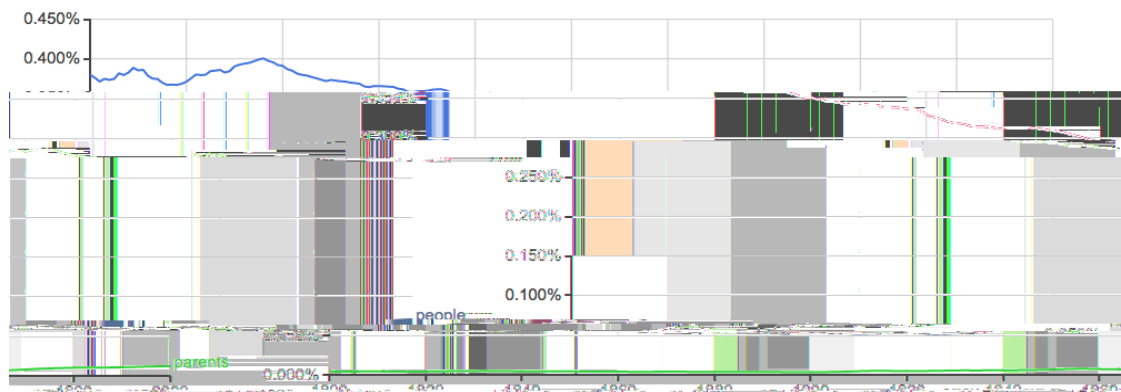


Figure 5: The frequency of have, people, and parents in Google's books; Gram Viewer, accessed 3/4/14

In the textbook, have is the first ranked content word with the frequency of 1 over 100,000. Have is also the first content word to appear in the 1000 word frequency list and ranked 1st following the top seven frequent function words with the frequency of 10,545.00

In the textbook corpus, have is used in three ways. First, it is used as a content verb with the meaning of owning something, the most frequently (15 over 100 to 100,000). Second, it is used as an auxiliary verb in the present perfect tense (37 out of 100 to 100,000). And finally, have occasionally refers to the necessity of doing something (10 out of 100 to 100,000) (see Figure D).

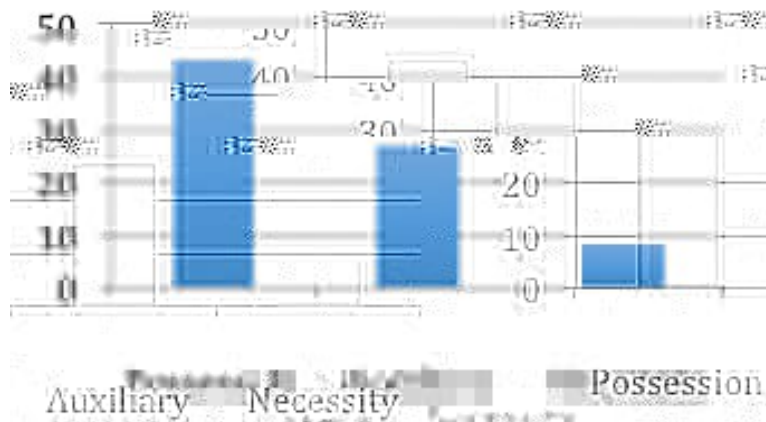


Figure D: The frequency of have by meaning in the textbook

However, data from the 1000 word frequency list show a different frequency. Have as an auxiliary verb in the present perfect tense is the most frequently used, with 13,000 instances; next frequent is have as a modal of necessity (35,000 times). The least frequent usage of have is its possessive meaning, with only 1,000 instances (Figure E). Thus, it seems that the textbook's presentation of have's usage differs in frequency compared to the 1000 word

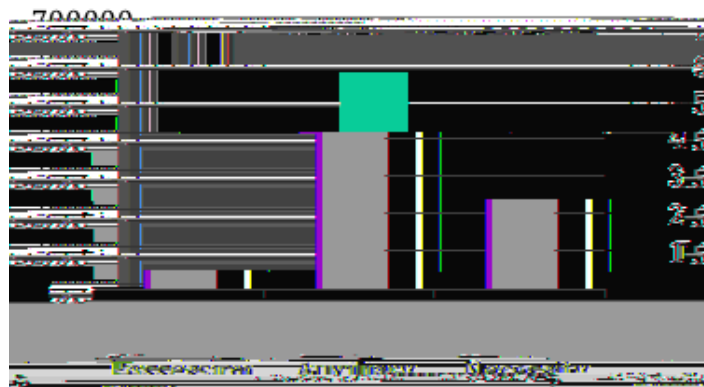
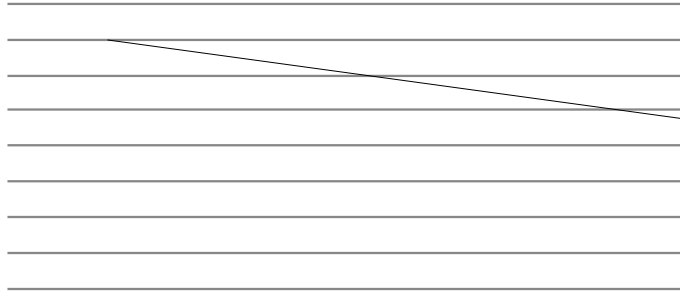


Figure E: The frequency of have by meaning in the 1000 word data

The next two content words are parents and people, with about the same frequency, 90 and 95 respectively (among the 1000 to 100,000 from the textbook or 10,000 and 10,550 per million respectively). However, there are considerable differences in the frequency of these two words in the 1000 word. Specifically, parents in its forms as parent, parents, to parent, and parenting occurs at 3.541 per one million words while

people is used at 1,9.3% per one million words" This means that in (* (#, people is far more frequent than parent and)3+ the textbook uses both people and parents more frequently than general American texts, with parent being used with much higher frequency" This suggests that the textbook and the texts in (* (# may have different content focuses? the textbook has a heavier focus on topics related to parents



Kata from (* (# show the same frequency order for yo#ng)with !5,91! to ens, after other and American# and /sa.#e /)with D. 9 to ens+ as words that go before people

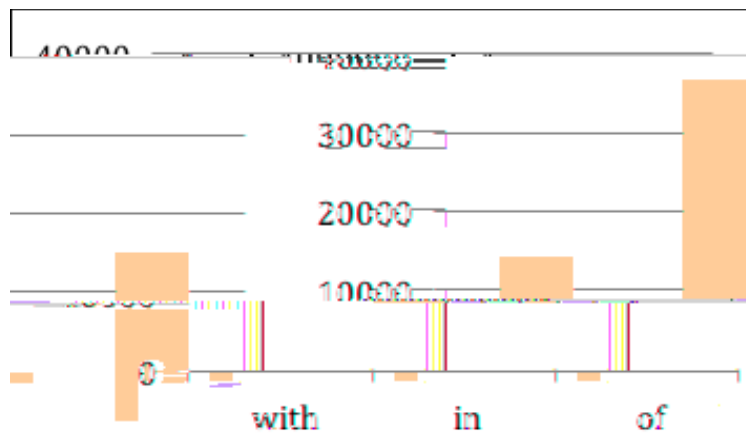


Figure 1 The distribution of with, in, and of following people from (* (#

The analysis above shows that while there are some similarities in the distribution of the most frequent lexical items in the &rade !! English textboo and general #merican English, there are also some important differences" Bbile some frequency differences have to do with the two corporas different content focus)such as the frequency of parent, people and the collocations with the word people+, some other frequency differences have to do with grammatical patterns)such as the different functions of have and the different prepositions following people+" This second type of difference may bear consequences on language learning" &iven Vietnamese learners> limited exposure to English, seeing more frequent use of the possessive meaning of have, for example, may lead learners to overuse that meaning, in contrast to the more common auxiliary meaning of the word in general #merican English" Similarly, learners may overuse the phrase people with due to the textboo s higher frequency of this phrase compared to the more common phrase people in

2c(arthy 2" M (art @")34!5+ " Vocabulary an /lang#age teaching" @outledge? ; ew Qor , - S#
; ation, 8" S" P')344!+" Learning voca#ary in another lang#age" (ambridge, - O? (ambridge - niversity Press
; ation, 8" S" P')344.+ " Teaching ESCFEAC reading and writing" ; ew Qor , ; Q? @outledge, Taylor and
Arancis"
@eppen, @")34! 4+ " Using corpora in the lang#age classroo)" ; ew Qor , ; Q? (ambridge - niversity Press"