Spoken Vocabulary Used by Elementary ESL/EFL Learners

Ai-Ling Kan[1]

University of Wisconsin-Madison, United States - National Tsing Huang University, Taiwan[1]

Abstract

According to language acquisition theory, the critical step in the language acquisition process is to distinguish between language purpose and language content. The development of elementary English learners' listening and speaking skills is the foundation of English acquisition. However, English as a second language/English as a foreign language (ESL/EFL) learners often have limited access to English outside of the classroom. Therefore, it is crucial to provide a word list that can help them receive useful spoken vocabulary guidance and promote the development of listening and speaking during limited classroom time.

However, no previous studies have analyzed the spoken vocabulary of beginner American ESL/Taiwanese EFL English learners or linked their main characteristics. Therefore, this study explores the ESL/EFL spoken vocabulary from a corpus-based perspective and determines whether this spoken vocabulary is related to the sight words. This study creates a spoken corpus based on classroom observations conducted between 2018 and 2020 in Taiwan and the United States to observe the spoken language used by teachers and ESL/EFL students during the English classroom interaction time.

Therefore, this study explores spoken word lists from the perspective of spoken language corpora to understand the different contexts, spoken words, and patterns used by teachers and students in the classroom. The results show that beginner ESL and EFL learners have similar oral vocabulary choices, even if their native language backgrounds are different. The characteristics of oral vocabulary can help teachers guide students to distinguish between written and oral English. Oral speech can also promote the development of learners' reading ability while improving their listening and speaking skills.

Keywords: *spoken corpus, spoken vocabulary, sight words (SW), high frequency words (HFW), English as a second language (ESL), English as a foreign language (EFL).*

1. Introduction

In the context of English as a second language (ESL) and English as a foreign language (EFL), the sight words (SWs) that teachers often teach to young learners are related to the oral expressions used in the classroom, which positively affect the development of students' listening and speaking abilities. A 2019 American education study pointed out that the reading scores of middle-low, middle- and higher-grade students in the fourth grade of elementary school are lower than in 2017. There is a gap in young learners' academic performance from non-mainstream cultural and linguistic minority backgrounds in the long run. (National Center for Educational Statistics, 2019)

Young L2 learners in Taiwan's elementary schools only learn the target language in an EFL environment [1] because they only have one to two hours of instruction each week in the target language in the classroom and have little opportunity to practice the target language outside of school. August (2005) claimed that to understand the target language within a limited time frame in the classroom, vocabulary choice plays a key role in developing a student's language ability. Hilton (2008) also pointed out that a lack of spoken vocabulary affected speech proficiency. Thus, to learn the target language efficiently according to second language acquisition (SLA) theory in an EFL or ESL environment, English Language learners should have sufficient vocabulary input.

2. Literature review

Word lists assist teachers in determining the vocabulary knowledge that is useful in the learning process of L2 students in the limited classroom time (Kan, 2020). However, an appropriate spoken vocabulary list helps not only language learners but also curriculum and textbook designers because word lists can help them select the right words to use when designing appropriate learning activities (Dang et al., 2020).

Teachers of children in kindergarten to the third grade usually use high-frequency word (HFW) lists or SWs to improve beginner English learners' elementary reading skills. While recent research has focused more on reading ability (Fry, 1980; Nation, 2013), Dolch's SW list (Dolch, 1936) uses written materials and includes spoken words commonly used by English-speaking children.

However, little research has focused on the spoken language, developed an academic word list (Dang et al., 2017), or examined the differences in English language learners' "spoken vocabulary" [2]. Additionally, no previous studies have examined the differences between the oral language and main characteristics of beginner Taiwanese EFL and American ESL learners.

Therefore, this study identifies HFWs from the perspective of ESL/EFL spoken word corpora and discusses the spoken language characteristics. Classroom observations were conducted in Taiwan and the United States between 2018 and 2020 to observe the spoken vocabulary of native English teachers and ESL/EFL students during their English classroom interactions to answer the following research questions:

[1]   What HFWs are used in the young ESL/EFL classroom?
[2]   What is the overlap ratio of the EFL/ESL HFWs and Dolch's SWs?
[3]   What is the text coverage rate of Dolch's SWs in young EFL/ESL classrooms?

3. Research methods
This study's main participants were third-grade teachers and students in EFL elementary school classrooms in Taiwan and kindergarten to second-grade teachers and students in ESL elementary school classrooms in the United States. The 90 young learners (47 male and 43 female) in Taiwan had an average age of 10 years. The two female English teachers in the EFL class were from Michigan, USA, and Cape Town, South Africa. The participants in the ESL classes were teachers and students in 19 classes, and the average age of the students was 5 to 7 years old. Since the COVID crisis started to spread in early 2020, K–12 classrooms in the United States were transformed into online classrooms, which were accessible through the online "TCH" teaching channel. Over four semesters, 110 classroom observations were conducted.

The goal of a corpus linguistics quantitative analysis is to convert the vocabulary into a manageable list (Garnier et al., 2015). Following Garnier et al.'s (2015) first step, this research identified the HFWs commonly used in classrooms by L2 learners and teachers in different contexts and conducted a quantitative analysis of corpus linguistics to answer the first research question. The researcher then checked the overlap rate of the HFWs in each group and verified the coverage rate of SWs in different contexts to answer the second and third research questions.

4. Research results
The above process produced two spoken language corpora from the EFL and ESL classrooms. Of the 456,437 words generated from the EFL classroom samples, 6,772 types (different and distinct words) were identified. Similarly, observations of the ESL classrooms obtained 505,098 words comprising 4,883 distinct words.

Table 1. Spoken data from the EFL/ESL classrooms

Text file	File size	Tokens (running words) in text	Tokens used for word list	Types (distinct words)
EFL CLASSROOMS	456,470	81,170	81,170	6,772
ESL CLASSROOMS	505,098	93,569	93,569	4,883

After deducting the non-words, there were 6,636 words from the EFL classroom and 4,837 words from the ESL classrooms. In this study, HFWs are considered to be words that appear in the context more than 100 times (Coxhead, 2000, p. 221). After retrieving these, the researcher obtained 121 HFWs from the EFL context and 153 from the ESL context (see Appendix 1). An examination of the HFWs in the EFL/ESL groups with Dolch's SWs showed an overlap rate of 75.37%. The ESL HFWs and SWs

showed an overlap rate of 71.37%.

Table 2. Overlap ratio of the EFL/ESL HWFs and Dolch's SWs

N	Text file	File size	Tokens (running words) in text	Types (distinct words)	Type/token ratio (ttr)
1	Overall	1,964	341	257	75.37
	Dolch's SWs	1,275	220	220	100.00
	EFL HFWs	689	121	121	100.00
2	Overall	2,247	380	271	71.32
	Dolch's SWs	1,275	220	220	100.00
	ESL HFWs	972	160	160	100.00

The results showed the coverage rates of the SWs as 41.04% in the EFL spoken corpus and 37.13% in the ESL spoken corpus.

Table 3. Coverage rate of Dolch's SW and the EFL/ ESL spoken corpus

N	Text file	File size	Tokens (running words) in text	Tokens used for word list	Types (distinct words)	Type/token ratio (ttr)
1	Overall	307,947	17,496	17,496	7,180	**41.04**
	EFL spoken corpus	306,672	17,276	17,276	7,167	41.49
	Dolch's SWs	1,275	220	220	220	100.00
2	Overall	232,034	14,358	14,358	5,331	**37.13**
	ESL spoken corpus	230,759	14,138	14,138	5,326	37.67
	Dolch's SWs	1,275	220	220	220	100.00

5. Discussion

The study found the overlap rates between Dolch's SWs and the ESL and EFL high-frequency spoken vocabulary lists to be 71.37% and 75.37%, respectively. Thus, despite the different language environments, around 75% of the spoken words were the same as the visual words, which reflected the influence of visual vocabulary on oral expression (Helman, 2008). These results indicated that the SWs that teachers guide the students to read will affect their spoken performance in class. Compared with Burns' (2009) experiment in ESL classrooms, he found a significant relationship between English proficiency and English visual words' acquisition rate. The results of this study show that sight words occupy 71% to 75% of spoken English in English language classrooms, providing direct evidence for the significant relationship between English proficiency and visual vocabulary.

Based on SWs, the researchers obtained coverage rates of 41.49% and 37.67% for the EFL and ESL spoken corpora, respectively. This finding showed that only 220 Dolch's SWs can cover 37%~41% of nearly one million words of the EFL/ ESL corpora, thus emphasizing the current significance of Dolch's SWs in EFL and ESL classrooms despite being published 90 years ago. Generally speaking, young L2 learners use more of Dolch's SWs for their oral expressions, especially in the EFL context. However, since the SWs originated from written text, the results of this study also highlighted the similarities and differences between written and spoken English. Teachers should guide students to distinguish between the two types of expressions.

From a scientific perspective, the words in the EFL/ESL spoken corpora were ranked according to their frequency. Of the HFWs, 121 were used in the EFL classrooms and 160 in the ESL classrooms. These spoken vocabulary word lists are an effective way to have beginner English learners or

immigrant families participate in the young English classroom as soon as possible.

According to Nation (2016), word lists based only on corpus standards might miss words that appear less frequently in the corpus but are useful for L2 learners. Therefore, to create a suitable word list for L2 learners, researchers have recommended using corpora, such as BNC or COCA, to verify the collected data and add low- and medium-frequency words that may be beneficial for L2 learning and teaching purposes. Overall, this study suggests that vocabulary for a beginner EFL/ESL L2 environment provides the most accessible starting point for incorporating HFWs.

The main limitation of this study is that the sample was limited to the spoken vocabulary of third-grade Taiwanese learners in the EFL classroom and American K–2 learners in the ESL classroom. Thus, future research should examine L2 learners in other EFL/ESL environments to confirm the spoken vocabulary used by L2 learners in different regions and eliminate possible bias.

6. Conclusion

This study obtained an overlap rate of 71%–75% between the visual vocabulary used by the EFL/ESL teachers and young L2 students in classrooms. These results suggested that even if learners have different backgrounds, young ESL/EFL learners use similar spoken English words in the classroom. In both Taiwanese EFL and American ESL classrooms, the spoken vocabulary tended to be based on visual vocabulary; thus, teachers should guide students to distinguish the similarities and differences between written and spoken English.

References

[1] V. LoCastro, "Individual differences in second language acquisition: Attitudes, learner subjectivity, and L2 pragmatic norms.," *System,* vol. 29, no. 1, pp. 69-89, 2001.

[2] R. Hincks, "Pronouncing the Academic Word List: Features of L2 student oral presentations. In Proceedings of the 15th International Congress of Phonetics Sciences," 2003.

[3] August, D., Carlo, M., Dressler, C., & Snow, C., "The critical role of vocabulary development for English language learners.," *Learning Disabilities Research & Practice,* vol. 20, no. 1, pp. 50-57, 2005.

[4] H. Hilton, "The link between vocabulary knowledge and spoken L2 fluency.," *Language Learning Journal,* vol. 36, no. 2, pp. 153-166, 2008.

[5] Dang, T. N. Y., Webb, S., & Coxhead, A., "Evaluating lists of high-frequency words: Teachers' and learners' perspectives.," *Language Teaching Research,* 2020.

[6] I. S. Nation, Learning vocabulary in another language, Google eBook: Cambridge University Press, 2013.

[7] Dang, T. N. Y., Coxhead, A., & Webb, S., "The academic spoken word list.," *Language Learning,* vol. 67, no. 4, pp. 959-997, 2017.

[8] Garnier, M., & Schmitt, N. , "The PHaVE List: A pedagogical list of phrasal verbs and their most frequent meaning senses," *Language Teaching Research,* vol. 19, no. 6, pp. 645-666, 2015.

[9] A. Coxhead, "A new academic word list," *TESOL quarterly,* vol. 34, no. 2, pp. 213-238, 2000.

[10] I. S. P. Nation, Making and using word lists for language learning and testing, John Benjamins Publishing Company, 2016.

[11] Burns, M. K. & Helman, L. A., "Relationship between language skills and acquisition rate of sight words among English language learners.," *Literacy Research and Instruction,* vol. 48, no. 3, pp. 221-232, 2009.