

Article

Relating Lexical Access and Second Language Speaking Performance

Yu Liu

Department of Asian and Near Eastern Languages, Brigham Young University, Provo, UT 84602, USA; rachelyliu@byu.edu

Received: 11 February 2020; Accepted: 9 April 2020; Published: 13 April 2020

Abstract: Vocabulary plays a key role in speech production, affecting multiple stages of language processing. This pilot study investigates the relationships between second language (L2) learners' lexical access and their speaking fluency, speaking accuracy, and speaking complexity. Fifteen L2 learners of Chinese participated in the experiment. A task-specific, native-referenced vocabulary test was used to measure learners' vocabulary size and lexical retrieval speed. Learners' speaking performance was measured by thirteen variables. The results showed that lexical access was significantly correlated with learners' speech rate, lexical accuracy, syntactic accuracy, and lexical complexity. Vocabulary size and lexical retrieval speed were significant predictors of speech rate. However, vocabulary size and lexical retrieval speed each affected learners' speaking performance differently. Learners' speaking fluency, accuracy, and complexity were all affected by vocabulary size. No significant correlation was found between lexical retrieval speed and syntactic complexity. Findings in this study support the Model of Bilingual Speech Production, revealing the significant role lexical access plays in L2 speech production.

Keywords: lexical access; second language speech; fluency; accuracy; complexity

1. Introduction

Vocabulary is fundamental for second language learning. Vocabulary knowledge has been identified as a good indicator of language proficiency (Milton 2013; Nation 2001). Research shows that second language (L2) learners with large or better developed vocabularies have better performance in reading (e.g., Albrechtsen et al. 2008; Jeon and Yamashita 2014; Laufer 1992; Qian 1999; Stæhr 2008), listening (e.g., Stæhr 2008, 2009), writing (e.g., Engber 1995; Grant and Ginther 2000; Schoonen et al. 2003; Stæhr 2008), and speaking (e.g., De Jong et al. 2013; Uchihara and Clenton 2018; Uchihara and Saito 2019).

According to Levelt's (1989, 1999) model of speech production, lexical access plays a key role in speech production, affecting multiple stages of language processing. Different examples of evidence supporting this claim have been found in empirical studies of L2 speaking:

- (1) Vocabulary size and depth are associated with the overall scores of L2 learners' speaking proficiency (De Jong et al. 2012; Milton 2010);
- (2) Both receptive and productive vocabulary knowledge can predict L2 speaking performance, especially fluency (Uchihara and Clenton 2018; Uchihara and Saito 2019);
- (3) The speed and efficiency of lexical access affect L2 speaking fluency, allowing it to be used as a measure of L2 cognitive fluency (De Jong et al. 2013; Segalowitz and Freed 2004).

Productive vocabulary is closely related to task features such as topics and text types. L2 learners' vocabulary inventory is built up through many communicative tasks, depending on their target language contact experiences. It is important to consider task factor when we investigate the

relationship between lexical access and L2 language performance. Most previous studies related L2 learners' general vocabulary knowledge to their speaking performance in completing a small number of speaking tasks. The results might be affected by task selection. Increasing task number will improve the reliability of the experiments. However, the more tasks are included, the more overwhelmed participants will feel, which may affect the research results. Moreover, the type of word selection that best represents L2 learners' general vocabulary size is still debatable. Therefore, differently from previous studies examining general vocabulary size, in this study, we seek to narrow the investigation scope within specific tasks in order to look more closely into how lexical access interacts with L2 speaking performance within these tasks.

In this study, a task-specific vocabulary test was created based on native speakers' productive vocabulary in completing the same speaking tasks as those completed by L2 learners later. This test was used to explore L2 learners' lexical access. Two variables were used to measure lexical access: Vocabulary size (the number of words a learner knows) and lexical retrieval speed (how fast a learner recognizes and processes the words he/she knows). The purpose of this research is to investigate how lexical access relates to second language speaking performance.

2. L2 Speech Production

Levelt's model of speech production (1989, 1999) presents the process of first language (L1) speech production. The processing starts with conceptualizing the message. A pre-verbal plan consisting of concepts and language cues is generated at the conceptualizer stage. Such messages are then processed at the formulator stage through lexico-grammatical encoding, morpho-phonological encoding, and phonetic encoding. The mental lexicon, including individual words and formulaic sequences, is activated to match with conceptual specifications and the language cues. At the articulator stage, words and sentences are executed as meaningful speech. During the whole process, speakers continuously monitor both their internal speech and overt speech output, modifying their speech as needed.

Based on his theory, Kormos (2006) proposed the Model of Bilingual Speech Production (see Figure 1). Two characteristics of L2 speech production are highlighted in this model:

- (1) *L1 influence*. As proposed by Kormos (2006), L1 and L2 concepts, lexemes (word forms), lemmas (syntactic and morphological features), syllable programs, and grammatical rules are stored together. In speech production, L1 and L2 lexical items and rules are activated together and compete for selection. When L1 knowledge and encoding procedures are transferred during L2 processing, the efficiency of L2 speech production may be negatively affected.
- (2) *Serial processing*. Compared to L1 automatized speech production, language processing in L2 requires more attention and is less automatized. L2 learners hold a great amount of L2 declarative rules in their long-term memory. Before declarative rules turn into procedural knowledge, L2 learners at a lower level of proficiency have to use limited attention to process at lexical, syntactic, and phonological levels as well as in monitoring. This causes L2 speech to be less fluent and more open to L1 influence (Kormos 2006; Vieira 2017).

In addition, Kormos (2006) emphasized that the L2 speech production system is lexically driven. The mental lexicon, including concepts, lemmas, and lexemes, affects conceptualization, lexico-grammatical encoding, and morpho-phonological encoding of phrases. Lexical access in the second language is an attention-demanding cognitive process. On the one hand, L2 learners face the challenge of retrieving the correct L2 words from the competition of L1 and related L2 items. On the other hand, time constraints in real communication require learners to process language efficiently, which increases the pressure of the cognitive process. As a result, the efficiency of lexical access influences the quality of L2 speech.

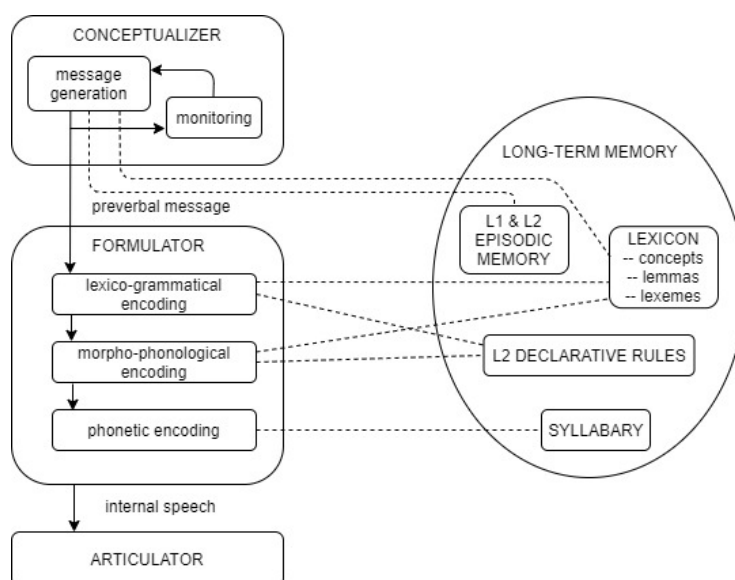


Figure 1. The Model of Bilingual Speech Production (adapted from Kormos 2006, p. 168).

3. Vocabulary Knowledge and L2 Speech

Lexical access is multifaceted. In literature, three categories have been distinguished in measuring vocabulary knowledge in literature (Anderson and Freebody 1981; Daller et al. 2007; Meara 1996; Milton and Fitzpatrick 2013):

1. *Vocabulary breadth*: The number of words a learner knows regardless of the form they are known in or how well they are known. Vocabulary breadth is also referred to as vocabulary size. Three forms of measurement have been found in literature: Yes/No format task (Uchihara and Clenton 2018), writing L2 forms corresponding to L1 meaning (Koizumi and In'nami 2013), and filling in the missing words in a sentence where the first letters are given (De Jong et al. 2013). Word selection in most studies is based on word frequency bank lists.
2. *Vocabulary depth*: How well or how completely words are known. Vocabulary depth is a rich concept that consists of various aspects. According to Nation's (2001, p. 27) description of "what is involved in knowing a word", vocabulary knowledge includes form (spoken, written, word parts), meaning (form and meaning, concepts and referents, associations), and use (grammatical functions, collocations, constraints on use). Read (2004) proposed that vocabulary involves word form and meaning, as well as associational knowledge, collocation knowledge, inflectional knowledge, and derivational knowledge. Meara and Wolter (2004) extended the vocabulary depth by including knowing the network words. Measuring vocabulary depth is less manageable because it is difficult to find a concept that holds together the variety of elements (Milton 2010).
3. *Vocabulary fluency*: The automaticity with which the words a person knows can be recognized and processed. It is also referred to as processing speed or lexical retrieval speed. Reaction time (RT) is recorded to measure vocabulary fluency in a vocabulary test (De Jong et al. 2013; Koizumi and In'nami 2013).

Previous studies have found a close relationship between L2 learners' vocabulary knowledge and their speech production. Some studies suggest that learners with larger receptive vocabulary sizes are more proficient in speaking (De Jong et al. 2013; Hilton 2008; Koizumi and In'nami 2013; Uchihara and Clenton 2018). Koizumi and In'nami (2013) found that vocabulary size explained up to 60% of the variance in speaking proficiency. However, Uchihara and Clenton (2018) found less predictive power in vocabulary size for L2 speaking, which was only 29%. The former study used an automated scoring system, whereas the latter used human ratings. Though Uchihara and Clenton (2018) found a significant correlation between receptive vocabulary size and spoken lexical use based

on human ratings, they also noticed that receptive vocabulary size and lexical sophistication measures were not significantly correlated. Their findings indicate that learners with larger vocabulary sizes do not necessarily produce more advanced words. De Jong et al. (2013) and Hilton (2008) focused on speaking fluency only. Both studies integrated lexical knowledge into a battery of tests that tested L2 learners' linguistic skills and discussed their relationship with L2 speaking fluency. The results in De Jong et al.'s (2013) study suggested that all measures of utterance fluency (e.g., speech rate, number of silent and filled pauses, repetitions, and repairs) were affected by linguistic knowledge. Hilton (2008) had similar findings in terms of temporal measures of L2 speaking fluency. He argued that the lack of lexical knowledge appeared to be the primary cause of the most serious disfluencies, and that it was the greatest impediment to L2 speaking fluency.

Among all studies, De Jong et al. (2013) and Koizumi and In'nami (2013) investigated more than one category of vocabulary knowledge. As De Jong et al. (2013) suggested, lexical retrieval speed was strongly correlated with all measures of L2 speaking fluency (e.g., number of silent pauses, filled pauses, repetitions, duration of silent pauses, speech rate). Koizumi and In'nami (2013) focused on three categories of vocabulary knowledge (e.g., size, depth, fluency). They found that vocabulary size and vocabulary depth substantially predicted L2 proficiency, while the correlation between vocabulary fluency and L2 speaking was much weaker. Koizumi and In'nami's (2013) findings were based on automatic ratings of L2 learners' speaking performance instead of human ratings or objective measures. The reliability of the automated scoring system may have had an effect on the findings.

Another area of vocabulary knowledge research measures productive vocabulary knowledge under the distinction of receptive vocabulary knowledge (passive knowledge, recognition) and productive knowledge (active knowledge, use). Only one study discussed how productive vocabulary knowledge affected L2 speaking performance (Uchihara and Saito 2019). In their study, the Lex30 test was used to investigate the productive mental lexicon. This test used a word association format, presenting learners with a list of 30 stimulus words and instructing them to write down the first four words in the target language they thought of when they read each word in the list (Meara and Fitzpatrick 2000). The data demonstrated that productive vocabulary knowledge could predict how fluently L2 learners can speak, but it was not significantly correlated with comprehensibility or accentedness. The results suggested that developed L2 lexicons led to less difficulty in retrieving L2 words, which therefore helped learners produce fluent speech.

4. Research Questions

To our knowledge, only a limited number of studies linked L2 learners' vocabulary knowledge with their speaking fluency, accuracy, and complexity, respectively (De Jong et al. 2013; Koizumi and In'nami 2013). No such study has been done in the research of Chinese as a second language (L2 Chinese), although some studies have discussed the relationship between vocabulary size and L2 Chinese reading (Wu 2016, 2017). In view of the important role that lexical access plays in second language speech production (Kormos 2006), this study aims at exploring the dynamics between L2 learners' lexical access, measured by vocabulary size and lexical retrieval speed, and three dimensions of their speaking performance: Fluency, accuracy, and complexity.

The following research questions are addressed in the present study:

1. How does lexical access affect the fluency in L2 learners' speech in four speaking tasks?
2. How does lexical access affect the accuracy in L2 learners' speech in four speaking tasks?
3. How does lexical access affect the complexity in L2 learners' speech in four speaking tasks?

5. Methodology

5.1. Participants

Fifteen English native speakers participated in the experiment. It was a small homogeneous group. According to the background questionnaire survey, these participants had very similar

language-learning backgrounds. They have attended the same Chinese classes at the same university. All of the participants were enrolled in a third-year Chinese course at a U.S. university when the experiment was conducted. According to the instructor of the course, these participants' Chinese proficiency levels ranked at ACTFL intermediate-high to advanced-low level (ACTFL 2012). Their ages ranged between 20 and 25. Five of them were female. All of the participants replied to our call for participation on a voluntary basis. They gave their informed consent for inclusion before they participated in the study. The study was conducted in accordance with the Declaration of Helsinki, and the protocol was approved by the Ethics Committee of #15405.

5.2. Materials

A task-specific, native-referenced vocabulary test was created to investigate how lexical access interacts with L2 speaking performance within these tasks. Before the experiment, we invited six Chinese native speakers from the same university (ages 19–22, four females and two males) to complete four speaking tasks, which were also used to test fifteen L2 learners. By doing this, we were able to set up the reference on the basis of native speakers' productive vocabulary.

Four speaking tasks represented four text types, carrying four different communicative functions: Instructive, descriptive, explanatory, and argumentative. In the first task, both Chinese native speakers and L2 learners were asked to introduce the city where the university was located. In the second task, they described their first day at the university. In the third task, they were presented a data chart and were invited to explain the income gap between males and females of different age groups. In the last task, they talked about their opinions on a given topic, specifically "What kind of professors are good college professors?" These tasks were not culturally specific. The native speakers and the participants shared similar experiences at the same university. We assumed that most of the vocabulary output by these two groups should be within a limited range when they completed the same tasks.

Based on the vocabulary that the Chinese native speakers used in the four tasks, we compiled a list of vocabulary (198 items) that was most commonly used (being used at least three times by different speakers). All words were listed in a random order, controlling the effect derived from word frequency and task order. The list was then translated into English by the researcher for the vocabulary test.

5.3. Procedure

There were two parts in the experiment. The first part was a vocabulary test. In this part, participants were instructed to translate the words on the vocabulary list orally from English into Chinese as fast as possible. They were instructed to respond "I don't know" if they did not know the answer. They were not given pre-task planning time. The whole process was timed in order to record lexical retrieval speed. If participants were able to say the target words or their synonyms, the answers were rated as correct. Otherwise, the answers were rated as incorrect. Two L2 Chinese teachers rated the vocabulary test independently. There was no disagreement between the two ratings.

Participants took a five-minute break after completing the first part and then continued to finish the second part. The second part was a speaking test consisting of four monologue tasks. Participants completed four speaking tasks, which were the same as the ones completed by the six Chinese native speakers. For each task, participants had one minute to prepare and ten minutes to speak. Participants' speech was recorded through a recording software "Audacity" with the setting of stereo 44,100 Hz. The experiment was conducted in the researcher's office individually, administered by the researcher. The total time commitment for each participant in this experiment was about 1.5 to 2 h.

5.4. Measures and Statistical Procedures

L2 learners' lexical access is represented by both vocabulary size and lexical retrieval speed. Vocabulary size was measured based on the accuracy rate in the vocabulary test. Lexical retrieval

speed was measured by calculating the average response time for each word in the vocabulary test. As for speaking performance, all participants' speech samples were first transcribed by a Chinese native speaker. Afterwards, they were encoded manually by the researcher as described in detail below. Then, thirteen variables of the following three categories were measured for statistical analysis:

Fluency: We measured three facets of speaking fluency (Tavakoli and Skehan 2005): Speed fluency (speech rate, mean length of runs); breakdown fluency (mean length of silent pauses, number of silent pauses, number of filled pauses), and repair fluency (number of disfluencies). A script programmed in PRAAT (Boersma and Weenink 2009) was used to detect silent pauses. Minimum silence duration was set to 350 milliseconds. We were therefore able to measure speech rate, mean length of runs, mean length of silent pauses, and the number of silent pauses. Filled pauses, such as *en* (嗯 “um”), *ranhou* (然后 “and then”), *jiushi* (就是 “that is”), and *nage* (那个 “that”), as well as disfluencies, such as repetitions, restarts, or repairs, were extracted manually from the transcripts of the speech samples. The number of filled pauses and the number of disfluencies were then calculated.

Accuracy: All speech samples were manually divided into different AS-units (main clauses and any attached subordinate clauses or sub-clausal units). They were also manually encoded by tagging lexical errors and syntactic errors. Lexical accuracy and syntactic accuracy were then calculated.

Complexity: Syntactic complexity was measured with two methods. The first method of measurement was to calculate the number of clauses, with the second method measuring the average sentence length. Lexical complexity was measured from three aspects: Lexical diversity, word frequency, and word difficulty. Guiraud's Index (Guiraud 1960) was used to measure lexical diversity. Word frequency was measured based on the SUBTLEX-CH corpus (Cai and Brysbaert 2010), whereas word difficulty was measured based on a reference word list designed for the standardized Chinese proficiency test Hanyu Shuiping Kaoshi (HSK): The HSK word difficulty ranking (Hanban 2012).

Table 1 lists the calculation methods used to measure L2 speaking performance in this study.

Table 1. Measures of speaking performance and their calculation methods.

		Variables	Calculation Methods
Fluency	Speed fluency	Speech rate	The total number of syllables divided by total time.
		Mean length of runs	The average number of syllables produced in utterances between pauses of 0.25 s and above.
	Breakdown fluency	Mean length of silent pauses	The total length of pauses above 0.2 s divided by the total number of pauses above 0.2 s.
		Number of silent pauses	The total number of pauses over 0.2 s divided by the total amount of time spent expressed in seconds and is multiplied by 60.
		Number of filled pauses	The total number of filled pauses divided by the total amount of time spent expressed in seconds and is multiplied by 60.
	Repair fluency	Number of disfluencies	The total number of disfluencies, such as repetitions, restarts, and repairs, divided by the total amount of time expressed in seconds and multiplied by 60.
Accuracy		Syntactic accuracy	The total number of correct AS units divided by the total number of AS units.
		Lexical accuracy	The total number of correct words divided by the total number of all words.
Complexity		Number of clauses	The total number of clauses divided by the total number of AS units.

Average sentence length	The total number of words divided by the total number of AS units.
Lexical diversity	The number of types/the square root of the number of tokens.
Word frequency	The average number of word frequency ranking of all words based on the SUBLEX-CN word frequency ranking.
Word difficulty	The average number of word difficulty rankings of all words based on the HSK word difficulty ranking.

6. Results

6.1. Vocabulary Test

In the experiment, it took participants an average of 516.5 s to complete the vocabulary test. The average reaction time for each word was 2.6 s. All participants correctly translated at least half of the words in the vocabulary list, with the accuracy rate ranging from 58% to 96%. The average accuracy rate was 81%, which indicates that these learners are familiar with most of the words on the list. However, none of them could translate all of the words correctly. Table 2 shows each participant's accuracy and average reaction time for each word in the vocabulary test. This accuracy represents the learners' task-specific receptive vocabulary size. The reaction time shows how fast lexical retrieval was. The stronger the link between the conceptual messages and the L2 lexical items (e.g., the concept of causal relation "because" and the Chinese words "因为 *yinwei*"), the more words were translated correctly, and the faster the reaction time was.

Table 2. Accuracy and average reaction time in the vocabulary test.

Students (N = 15)	Accuracy	Reaction Time per Word (s)
1	0.68	3.3
2	0.58	4.3
3	0.88	2.84
4	0.85	1.68
5	0.74	2.72
6	0.84	2.96
7	0.89	1.95
8	0.91	2.13
9	0.73	2.34
10	0.78	2.43
11	0.67	3.36
12	0.95	2.03
13	0.75	2.24
14	0.95	2.08
15	0.96	2.78
Mean (SD)	0.81 (0.12)	2.61 (0.68)

6.2. Vocabulary Size and L2 Speaking Performance

Table 3 shows L2 learners' speaking performance in terms of fluency, accuracy, and complexity in four speaking tasks. To determine the degree of the relationship between vocabulary size and all measures of learners' speaking performance, Pearson's correlations were calculated. The Appendix A presents the correlations among vocabulary size, lexical retrieval speed, and all of the measures of speaking performance. A significant correlation was found between task-specific vocabulary size and

all fluency measures except for the number of filled pauses ($r = 0.012$, $p = 0.926$). In particular, L2 learners' vocabulary size was strongly correlated with speed fluency (speech rate, $r = 0.375$, $p = 0.003$; mean length of runs, $r = 0.354$, $p = 0.005$), with breakdown fluency (mean length of silent pauses, $r = -0.256$, $p = 0.048$; number of silent pauses, $r = -0.35$, $p = 0.006$), and with repair fluency (number of disfluencies, $r = -0.285$, $p = 0.027$).

Learners' vocabulary size was also significantly correlated with learner's speaking accuracy on both the lexical level ($r = 0.377$, $p = 0.027$) and the syntactic level ($r = 0.313$, $p = 0.015$).

In regard to speaking complexity, no significant correlation was found between learners' vocabulary size and syntactic complexity with two different measures (number of clauses, $r = -0.23$, $p = 0.077$; average sentence length, $r = -0.098$; $p = 0.458$). Regarding lexical complexity, mixed results were found based on different measures. The results showed that learners' vocabulary size was closely related to the lexical diversity measure ($r = 0.413$, $p = 0.001$) and to word difficulty measure ($r = 0.397$, $p = 0.002$), but not to the word frequency measure ($r = 0.145$, $p = 0.271$).

Table 3. Participants' speaking performance in four speaking tasks (N = 15).

		Task 1		Task 2		Task 3		Task 4		Average of All Tasks	
		Mean	SD	Mean	SD	Mean	SD	Mean	SD	Mean	SD
1.	Speech rate	2.49	0.47	2.69	0.64	2.14	0.49	2.62	0.49	2.48	0.52
2.	Mean length of runs	0.79	0.62	0.82	0.72	0.67	0.20	0.64	0.19	0.73	0.43
3.	Mean length of silent pauses	106.44	54.35	98.44	37.45	98.90	39.16	102.70	40.11	101.62	42.77
4.	Number of silent pauses	0.54	0.21	0.47	0.16	0.61	0.18	0.56	0.17	0.55	0.18
5.	Number of filled pauses	9.97	5.52	8.74	4.45	10.29	4.20	8.85	5.47	9.46	4.91
6.	Number of disfluencies	5.75	3.07	5.25	2.97	4.11	2.58	5.51	2.18	5.15	2.70
7.	Syntactic accuracy	0.81	0.15	0.87	0.09	0.79	0.11	0.87	0.13	0.84	0.12
8.	Lexical accuracy	0.988	0.006	0.992	0.008	0.990	0.009	0.990	0.009	0.990	0.008
9.	Number of clauses	1.43	0.27	1.40	0.22	1.49	0.25	1.64	0.44	1.49	0.30
10.	Sentence length	23.25	5.22	22.99	4.34	30.40	5.07	30.26	6.96	26.73	5.40
11.	Lexical diversity	6.78	0.85	6.95	0.70	6.27	1.18	6.55	1.05	6.64	0.95
12.	Word frequency	1924.2	908.3	1985.1	739.6	3028.3	1313.5	1930.0	632.7	2216.9	898.5
13.	Word difficulty	2.45	0.24	2.38	0.21	2.48	0.21	2.39	0.17	2.42	0.20

By using HSK word difficulty rankings to compare words produced by native speakers (NSs) and L2 learners (NNSs) in four speaking tasks, it was found that the word distributions of different groups had different patterns. Figure 2 showed the word distributions of two groups. Most of the words that L2 learners used were level 1 to level 3 words (35.06%, 22.1%, 19.36%), which were the simpler words. They also used some level 4 words (13.92%) and a few words from level 5, level 6, or above (7.03%, 2.53%), which were more advanced words. Differently from L2 learners, native speakers used mostly advanced words, especially words ranked level 3 and above (22.22%, 26.26%, 24.75%, 18.68%). Level 1 and Level 2 words were not frequently used by native speakers (1.52%, 6.57%), as opposed to L2 learners. When compared to native speakers, L2 learners generally used less advanced words. However, the more words they acquired, the more advanced and more diverse words they were likely to use. L2 learners' limited vocabulary size affected their word choice, which, as a result, might affect the quality of their speech.

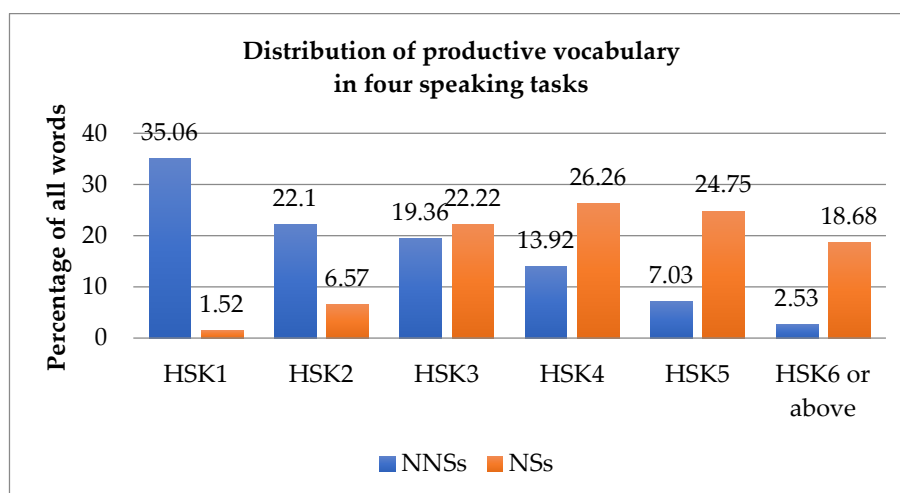


Figure 2. Distribution of words produced by second language (L2) learners (NNSs) and native speakers (NSs) based on the Hanyu Shuiping Kaoshi (HSK) difficulty ranking.

6.3. Speed of Lexical Access and L2 Speaking Performance

Pearson's correlations demonstrated a high correlation between task-specific lexical retrieval speed, which was measured by the average reaction time for each word in the vocabulary test, and speech rate ($r = -0.379$, $p = 0.003$). No significant correlation was found between lexical retrieval speed and other fluency measures, including mean length of runs ($r = 0.076$, $p = 0.565$), breakdown fluency (mean length of silent pauses, $r = 0.023$, $p = 0.862$; number of silent pauses, $r = 0.147$, $p = 0.262$; number of filled pauses, $r = -0.127$, $p = 0.332$), or with repair fluency (number of disfluencies, $r = 0.189$, $p = 0.148$).

Learners' lexical retrieval speed was significantly correlated with their speaking accuracy and complexity only on the lexical level, not on the syntactic level. Significant correlation was found between lexical retrieval speed and lexical accuracy ($r = -0.308$; $p = 0.017$), lexical diversity ($r = -0.365$; $p = 0.004$), and word difficulty ($r = -0.424$; $p = 0.001$), but not with word frequency measure ($r = 0.177$; $p = 0.176$). In terms of syntactic measures, no significant correlation was found between lexical retrieval speed and syntactic accuracy ($r = -0.192$; $p = 0.142$), syntactic complexity ($r = 0.176$; $p = 0.178$), or average sentence length ($r = 0.022$; $p = 0.867$).

6.4. Lexical Access and L2 Speaking Performance

A multiple linear regression was conducted to examine the role of L2 learners' lexical access in explaining their speaking fluency, accuracy, and complexity. Two variables of lexical access, vocabulary size and lexical retrieval speed, were entered into the regression equations to determine their contributions to the variance of thirteen measures of speaking performance. The results suggested that vocabulary size and lexical retrieval speed were significant predictors of speech fluency, explaining 16.7% of the variance of speech rate and 18.5% of the variance of the mean length of runs. Vocabulary size and lexical retrieval speed also had a significant predictive effect on breakdown fluency, accounting for 11.4% of the variance of the mean length of silent pauses and 14.2% of the variance of the number of silent pauses. Results showed that lexical access had no significant predictive effect on the number of filled pauses as well as the number of disfluencies (see Table 4).

Table 4. Prediction of L2 speaking fluency by lexical access.

Measures of L2 Speaking Fluency	Lexical Access: Vocabulary Size; Lexical Retrieval Speed	<i>b</i>	SE <i>b</i>	β	ΔR^2	F (2, 57)	<i>p</i>
1. Speech rate	Size	1.04	0.83	0.21			
	Speed	0	0	−0.23			
	Size * Speed				0.167	5.71	0.006 **
2. Mean length of runs	Size	2.56	0.72	0.59			
	Speed	0	0	0.34			
	Size * Speed				0.185	6.42	0.003 **
3. Mean length of silent pauses	Size	−176.43	65.19	−0.47			
	Speed	−0.1	0.06	−0.31			
	Size * Speed				0.114	3.68	0.031 *
4. Number of silent pauses	Size	−0.79	0.28	−0.49			
	Speed	0	0	−0.19			
	Size * Speed				0.142	4.7	0.013 *
5. Number of filled pauses	Size	−6.53	7.85	−0.15			
	Speed	−0.01	0.01	−0.23			
	Size * Speed				0.028	0.82	0.444
6. Number of disfluencies	Size	−7.21	4.28	−0.3			
	Speed	0	0	−0.02			
	Size * Speed				0.082	2.53	0.088

** Correlation is significant at the 0.01 level. * Correlation is significant at the 0.05 level.

In terms of accuracy, vocabulary size and lexical retrieval speed were found to be significant predictors of lexical accuracy, accounting for 14.6% of the variance. However, no significant predictive effect was found on syntactic accuracy (see Table 5).

Table 5. Prediction of L2 speaking performance by lexical access.

Measures of L2 Speaking Accuracy	Lexical Access: Vocabulary Size; Lexical Retrieval Speed	<i>b</i>	SE <i>b</i>	β	ΔR^2	F (2, 57)	<i>p</i>
7. Syntactic accuracy	Size	0.39	0.2	0.35			
	Speed	0	0	0.06			
	Size * Speed				0.099	3.15	0.051
8. Lexical accuracy	Size	0.02	0.01	0.32			
	Speed	0	0	−0.08			
	Size * Speed				0.146	4.88	0.011 *

** Correlation is significant at the 0.01 level. * Correlation is significant at the 0.05 level.

As for complexity, vocabulary size and lexical retrieval speed were found to be significant predictors of lexical complexity, explaining 18.1% of the variance of lexical diversity and 19.9% of vocabulary difficulty. No significant predictive effect was found on vocabulary frequency. The results also suggested that lexical access had no significant predictive effect on syntactic complexity (see Table 6).

Table 6. Prediction of L2 speaking performance by lexical access.

Measures of L2 Speaking Accuracy	Lexical Access:	<i>b</i>	SE <i>b</i>	β	ΔR^2	F (2, 57)	<i>p</i>
	Vocabulary Size; Lexical Retrieval Speed						
9. Number of clauses	Size	−0.58	0.5	−0.21	0.053	1.61	0.209
	Speed	0	0	0.03			
	Size * Speed						
10. Sentence length	Size	−12.7	10.44	−0.22	0.026	0.76	0.475
	Speed	−0.01	0.01	−0.18			
	Size * Speed						
11. Lexical diversity	Size	2.64	1.44	0.31	0.181	6.31	0.003 **
	Speed	0	0	−0.15			
	Size * Speed						
12. Vocabulary frequency	Size	−361.51	1653.43	−0.04	0.032	0.95	0.395
	Speed	1.16	1.43	0.15			
	Size * Speed						
13. Word difficulty	Size	0.35	0.3	0.19	0.199	7.08	0.002 **
	Speed	0	0	−0.29			
	Size * Speed						

** Correlation is significant at the 0.01 level. * Correlation is significant at the 0.05 level.

7. Discussion

The results in this study revealed that task-specific lexical access was closely related to all three dimensions of L2 speech: Fluency, accuracy, and complexity, though it is mainly on the lexical level. Vocabulary size and lexical retrieval speed were found to be significant predictors of speech rate, silent pauses, lexical accuracy, and lexical complexity. The explanatory power ranged from 11.4% to 19.9%. Among the three above outlined categories of speaking performance, fluency was most easily affected by vocabulary size and lexical retrieval speed.

In response to the first research question, “How does lexical access affect the fluency in L2 learners’ speech in four speaking tasks?”, the results showed that both vocabulary size and lexical retrieval speed were found to be highly correlated with speech rate ($p < 0.01$). L2 learners’ speech rate was most easily affected by lexical access. A significant correlation was found between vocabulary size and most of the fluency measures, whereas lexical retrieval speed was not significantly correlated with other measures of speaking fluency. Those who knew more L2 words and had faster processing speed produced more fluent speech. This finding is aligned with previous studies (De Jong et al. 2013; Hilton 2008; Koizumi and In’nami 2013; Uchiyama and Clenton 2018). This can be explained by the theory of automatization in Kormos’s model (2006). L2 lexical access is an attention-demanding task. This is because in the L2 lexico-semantic system, the link between conceptual messages and L2 lexical items is weaker than that of L1. In addition, L1 and L2 words, as well as related L2 words, are activated and compete for selection. Moreover, the syntactic rules of lexical forms are stored as declarative knowledge, which requires attention control to be executed. Under the impact from the above factors, the L2 speech production process is serial, rather than automatized, which causes L2 speech to be less fluent.

Answering the second research question, “How does lexical access affect the accuracy in L2 learners’ speech in four speaking tasks?”, the results demonstrated that both vocabulary size and lexical retrieval speed were significantly correlated with lexical accuracy ($p < 0.05$). Syntactic accuracy was significantly correlated with vocabulary size ($p < 0.05$), but not with lexical retrieval speed. Those who knew more L2 words and had faster processing speed spoke more accurately. It should be noticed that, although in this study knowing more words resulted in more accurate sentences, language processing on the syntactic level is very complicated because knowing L2 words not only means matching the concepts and the L2 lexical forms, but also means being aware of these words’ grammatical rules. Since L2 rules are stored as declarative knowledge in learners’ long-term memory, learners often encounter challenges in selecting the correct words among synonyms and using them correctly in sentences. It is necessary for future studies to investigate the relationship between vocabulary depth and L2 speaking performance.

In terms of the third question, “How does lexical access affect the complexity in L2 learners’ speech in four speaking tasks?”, it was found that vocabulary size was highly correlated with lexical diversity and word difficulty ($p < 0.01$). Those who acquired more L2 words and had faster processing speed produced more diverse and more advanced words as they spoke. This finding differs from that of Uchihara and Clenton’s study (2018). In their study, the receptive vocabulary size and lexical sophistication measures were not significantly correlated. The reason may be that the word selection methods used in the vocabulary test were different in both studies. In their study, researchers used 100 real and 100 imaginary words to create the word list. Participants were asked to judge which words were real and which ones were imaginary. Judging real words reflected L2 learners’ word recognition ability, whereas translating words from L1 to L2 reflected the effectiveness of mapping concepts and L2 words.

Furthermore, two different patterns of word distribution in L1 and L2 were found in completing the same tasks. L2 learners tended to use more lower-level words: The more advanced the words, the less they were used. Word distribution was presented as a curve from high to low. On the contrary, most of the words used by native speakers were more advanced words (HSK level 3 and above). The word distribution curve went from low to high and turned into a flat line, showing a balanced distribution from level 3 to higher-level words.

According to the Model of Bilingual Speech Production (Kormos 2006), conceptualization, lexico-grammatical encoding, and morpho-phonological encoding are all directly affected by lexical access. L2 speech production is a lexical-driven process. Retrieving vocabulary from long-term memory in the process of second speech production has been claimed to be less automatic, serial processing instead of automatic, parallel processing. The evidence in this study supports this model. Learners’ performance in the vocabulary test revealed their ability to retrieve L2 vocabulary, which was shown to be significantly correlated with the quality of their speaking performance. The more efficient lexical retrieval was, the faster the lexico-grammatical encoding and morpho-phonological encoding were, and the more fluent speech was produced by L2 learners. When L2 Learners encountered difficulty in retrieving the corresponding L2 words or it took longer to successfully retrieve them, their speech was less fluent, less accurate, and simpler. The evidence in this study provides details of how L2 learners’ fluency, accuracy, and complexity are affected by their capability or incapability of retrieving vocabulary in their L2 speech. The importance of vocabulary learning in second language acquisition is echoed in this study.

8. Limitations and Directions for Future Research

In this study, we adopted a task-specific, native-referenced approach in vocabulary test design, which is rarely seen in current related literature. Rather than testing L2 learners’ general vocabulary size, we focused on learners’ vocabulary size within specific tasks. We believe that by using this more focused approach, the dynamics between L2 learners’ lexical access and their speaking performance can be presented more clearly. However, there are also limitations in doing so. One limitation is that the number of vocabularies for investigation is limited. Future studies can include more tasks with different topics and text types. It is also necessary to consider individual differences of productive vocabulary in completing the same tasks, especially the “vocabulary gap” between native speakers and non-native speakers in relation to native-referenced vocabulary test design. To improve the reliability of the experiment for future work, a larger number of participants at different proficiency levels can be included to control the effect from individual differences. With a larger scale of investigation and duplicated studies, we would be able to further the discussion of how lexical access affects L2 speech. It would also be useful to compare the word distribution of native speakers and non-native speakers by using a computer-coding approach, in order to explore more deeply the link between receptive vocabulary knowledge and productive vocabulary knowledge.

Funding: This research received no external funding.

Conflicts of Interest: The author declares no conflict of interest.

Appendix A

Table A1. Pearson's correlations among vocabulary size, lexical retrieval speed, and all of the measures of speaking performance.

	Accuracy of Vocabulary Test	Speed of Lexical Access	Speech Rate	Mean Length of Runs	Mean Length of Silent Pauses	Number of Silent Pauses	Number of Filled Pauses	Number of Disfluencies	Syntactic Accuracy	Lexical Accuracy	Number of Clauses	Sentence Length	Lexical Diversity	Lexical Richness	Lexical Complexity
Accuracy of vocabulary test															
Speed of lexical access	−0.703 **														
Speech rate	0.375 **	−0.379 **													
Mean length of runs	0.354 **	−0.076	0.255 *												
Mean length of silent pauses	0.256 *	0.023	−0.236	−0.732 **											
Number of silent pauses	−0.350 **	0.147	−0.568 **	−0.526 **	0.461 **										
Number of filled pauses	0.012	−0.127	−0.068	−0.404 **	0.589 **	0.12									
Number of disfluencies	−0.285 *	0.189	0.098	−0.097	0.131	0.008	−0.212								
Syntactic accuracy	0.313 *	−0.192	0.123	0.008	−0.062	−0.302 *	0.257 *	−0.297 *							
Lexical accuracy	0.377 **	−0.308 *	0.248	0.085	−0.081	−0.166	0.222	−0.306 *	0.245						
Number of Clauses	−0.23	0.176	0.071	−0.06	−0.024	0.105	−0.239	0.304 *	−0.482 **	−0.377 **					
Sentence length	−0.098	−0.022	−0.128	−0.165	−0.063	0.21	−0.126	0.21	−0.284 *	−0.155	0.685 **				
Lexical diversity	0.413 **	−0.365 **	0.228	0.181	−0.139	−0.071	−0.008	−0.407 **	0.358 **	0.039	−0.200	−0.290 *			
Word frequency	−0.145	0.177	−0.047	0.06	−0.292 *	−0.034	−0.192	−0.04	−0.243	−0.133	0.131	0.235	−0.222		
Word difficulty	0.397 **	−0.424 **	0.325 *	0.440 **	−0.327 *	−0.268 *	−0.222	−0.137	0.005	−0.115	0.2	0.19	0.316 *	−0.024	

** Correlation is significant at the 0.01 level. * Correlation is significant at the 0.05 level.

References

- ACTFL. 2012. *ACTFL Proficiency Guidelines—Speaking*, 3rd ed. Alexandria: American Council on the Teaching of Foreign Languages. Available online: <https://www.actfl.org/publications/guidelines-and-manuals/actfl-proficiency-guidelines-2012/english/speaking> (accessed on 29 January 2019).
- Albrechtsen, Dorte, Kirsten Haastrup, and Birgit Henriksen. 2008. *Vocabulary and Writing in a First and Second Language: Processes and Development*. New York: Palgrave Macmillan.
- Anderson, Richard C., and Peter Freebody. 1981. Vocabulary knowledge. In *Comprehension and Teaching: Research Reviews*. Edited by Guthrie John T. Newark: International Reading Association, pp. 77–117.
- Boersma, Paul, and David Weenink. 2009. *Praat: Doing Phonetics by Computer* [computer program]. Version 5.1.17. Available online: <http://www.praat.org> (accessed on 27 January 2019).
- Cai, Qing, and Marc Brysbaert. 2010. SUBTLEX-CH: Chinese word and character frequencies based on film subtitles. *PLoS ONE* 5: e10729. doi:10.1371/journal.pone.0010729.
- Daller, Helmut, James Milton, and Jeanine Treffers-Daller, eds. 2007. *Modelling and Assessing Vocabulary Knowledge*. Cambridge: Cambridge University Press.
- De Jong, Nivja H., Margarita P. Steinel, Arjen F. Florijn, Rob Schoonen, and Jan H. Hulstijn. 2012. Facets of speaking proficiency. *Studies in Second Language Acquisition* 34: 5–34.
- De Jong, Nivja H., Margarita P. Steinel, Arjen Florijn, Rob Schoonen, and Jan H. Hulstijn. 2013. Linguistic skills and speaking fluency in a second language. *Applied Psycholinguistics* 34: 893–916.
- Engber, Cheryl A. 1995. The relationship of lexical proficiency to the quality of ESL compositions. *Journal of Second Language Writing* 4: 139–55.
- Grant, Leslie, and April Ginther. 2000. Using computer-tagged linguistic features to describe L2 writing differences. *Journal of Second Language Writing* 9: 123–45.
- Guiraud, Pierre. 1960. *Problèmes et méthodes de la statistique linguistique*. Paris: Presses Universitaires de France.
- Hanban. 2012. HSK (Chinese as a second language proficiency test: Hanyu Shuiping Kaoshi) Vocabulary. [新汉语水平考试 (HSK) 词汇 (2012 年修订版)]. Available online: <http://www.chinesetest.cn/userfiles/file/HSK/HSK-2012.xls> (accessed on 27 January 2019).
- Hilton, Heather. 2008. The link between vocabulary knowledge and spoken L2 fluency. *Language Learning Journal* 36: 153–66.
- Jeon, Eun Hee, and Junko Yamashita. 2014. L2 reading comprehension and its correlates: A meta-analysis. *Language Learning* 64: 160–212.
- Koizumi, Rie, and Yo In'nami. 2013. Vocabulary knowledge and speaking proficiency among Second Language Learners from Novice to Intermediate Levels. *Journal of Language Teaching & Research* 4: 900–13.
- Kormos, Judit. 2006. *Speech Production and Second Language Acquisition*. Abingdon: Routledge.
- Laufer, Batia 1992. How much lexis is necessary for reading comprehension? In *Vocabulary and Applied Linguistics*. Edited by Arnaud Pierre J. L. and Henri Béjoint. London: Macmillan International Higher Education, pp. 126–32.
- Levelt, Willem J. M. 1989. *Speaking: From Intention to Articulation*. Cambridge: MIT Press.
- Levelt, Willem J. M. 1999. Producing spoken language. In *The Neurocognition of Language*. Edited by Brown, Colin M., and Peter Hagoort, Oxford: Oxford University Press, pp. 83–122.
- Meara, Paul. 1996. The Vocabulary Knowledge Framework. Available online: <http://www.swan.ac.uk/cals/calsres/vlibrary/pm96d.htm> (accessed on 28 January 2019).
- Meara, Paul, and Tess Fitzpatrick. 2000. Lex30: An improved method of assessing productive vocabulary in an L2. *System* 18: 19–30.
- Meara, Paul, and Brent Wolter. 2004. V_Links: Beyond vocabulary depth. *Angles on the English Speaking World* 4: 85–96.
- Milton, James. 2010. The development of vocabulary breadth across the CEFR levels. In *Communicative Proficiency and Linguistic Development: Intersections between SLA and Language Testing Research*. Edited by Bartning Inge, Maisa Martin and Ineke Vedder. Eurosia Monograph Series, 1, pp. 211–32. Available online: <http://eurosia.org/monographs/EM01/EM01home.html> (accessed on 28 January 2019).
- Milton, James. 2013. Measuring the contribution of vocabulary knowledge to proficiency in the four skills. In *L2 Vocabulary Acquisition, Knowledge and Use*. Edited by Bardel Camilla, Christina Lindqvist and Batia Laufer. Amsterdam: EUROSIA-the European Second Language Association, pp. 57–78.
- Milton, James, and Tess Fitzpatrick, eds. 2013. *Dimensions of Vocabulary Knowledge*. London: Macmillan International Higher Education.
- Nation, Paul. 2001. *Learning Vocabulary in Another Language*. Cambridge: Cambridge University Press.
- Qian, David. 1999. Assessing the roles of depth and breadth of vocabulary knowledge in reading comprehension. *Canadian Modern Language Review* 56: 282–308.

- Read, John. 2004. Plumbing the depths: how should the construct of vocabulary knowledge be defined? In *Vocabulary in a Second Language: Selection, Acquisition and Testing*. Edited by Bogaards Paul and Batia Laufer. Amsterdam: John Benjamins Publishing, vol. 10, pp. 209–27.
- Schoonen, Rob, Amos van Gelderen, Kees de Gloppe, Jan Hulstijn, Annegien Simis, Patrick Snellings, and Marie Stevenson. 2003. First language and second language writing: The role of linguistic knowledge, speed of processing, and metacognitive knowledge. *Language Learning* 53: 165–202.
- Segalowitz, Norman, and Barbara F. Freed. 2004. Context, contact, and cognition in oral fluency acquisition: Learning Spanish in at home and study abroad contexts. *Studies in Second Language Acquisition* 26: 173–99.
- Stæhr, Lars Stenius. 2008. Vocabulary size and the skills of listening, reading and writing. *Language Learning Journal* 36: 139–52.
- Stæhr, Lars Stenius. 2009. Vocabulary knowledge and advanced listening comprehension in English as a foreign language. *Studies in Second Language Acquisition* 31: 577–607.
- Tavakoli, Parvaneh, and Peter Skehan. 2005. Strategic planning, task structure and performance testing. In *Planning and Task Performance in a Second Language*. Edited by Ellis Rod. Amsterdam: John Benjamins Publishing, pp. 239–73.
- Uchihara, Takumi, and Jon Clenton. 2018. Investigating the role of vocabulary size in second language speaking ability. *Language Teaching Research* 1–17. doi:10.1177/1362168818799371.
- Uchihara, Takumi, and Kazuya Saito. 2019. Exploring the relationship between productive vocabulary knowledge and second language oral ability. *The Language Learning Journal* 47: 64–75.
- Vieira, Gicle Vergine. 2017. Lexical access in L2 speech production: A controlled serial search task. *Ilha do Desterro* 70: 245–64.
- Wu, Sina. 2016. Influence of vocabulary, syntactic and metacognitive awareness of Japanese students on their Chinese reading comprehension [词汇、句法和元认知策略对日本学生汉语阅读理解的影响]. *Language Teaching and Linguistic Studies* [语言教学与研究] 178: 59–66.
- Wu, Sina. 2017. Contribution of vocabulary knowledge, morphological awareness and lexical inference ability to reading comprehension: Evidence from SEM [语素意识、语法知识与汉语二语阅读理解——来自结构方程模型的证据]. *Chinese Teaching in the World* [世界汉语教学] 31: 420–32.



© 2020 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<http://creativecommons.org/licenses/by/4.0/>).