Korean Journal of English Language and Linguistics, Vol 23, April 2023, pp. 285-302
DOI: 10.15738/kjell.23..202304.285

KOREAN JOURNAL OF ENGLISH LANGUAGE AND LINGUISTICS

ISSN: 1598-1398 / e-ISSN 2586-7474

http://journal.kasell.or.kr

Language Aptitude in the Development of Vocabulary Breadth and Depth of EFL Learners with Different Proficiency Levels*

Lingjie Tang (University of Seoul) **Jayeon Lim** (University of Seoul)

Received: February 10, 2023
Revised: March 27, 2023
Accepted: April 19, 2023

Lingjie Tang
(first author)
MA Graduate, Dept. of English Language and Literature, University of Seoul
Email: tanglingjie9@gmail.com

Jayeon Lim
(corresponding author)
Professor, Department of English Language and Literature, University of Seoul
Tel: 02) 6490-2521
Email: limjy@uos.ac.kr

* This research has been revised and extended from the master's thesis of the first author.

ABSTRACT

Tang, Lingjie and Jayeon Lim. 2023. Language aptitude in the development of vocabulary breadth and depth of EFL learners with different proficiency levels. *Korean Journal of English Language and Linguistics* 23, 285-302.

The purpose of this study is to examine the role of language learning aptitude in the development of vocabulary breadth and depth of adult EFL learners with different proficiency levels. To this end, sixty Chinese EFL college students were divided into two proficiency groups of high and low and participated in three online tests of measuring vocabulary breadth, depth and aptitude. The results showed that high proficiency (HP) learners outperformed low proficiency (LP) learners in three subdivided aptitudes of rote memory, grammatical sensitivity and phonetic coding ability. HP also outperformed LP in vocabulary breadth at all frequency levels, indicating a positive relationship between language aptitude and vocabulary breadth. With vocabulary depth, HP produced more valid paradigmatic associations than LP. However, no difference was found based on the two proficiency group's responses to syntagmatic and phonological associations of the vocabulary depth test. The results implicate that L2 learners' vocabulary depth developed with a more paradigmatic tendency as their proficiency increased. No difference was found in syntagmatic associations between the two proficiency groups, indicating a possibility of it continuously posing a challenge for L2 learners. As with phonological associations, it seems that L2 learners rarely rely on phonological information when learning vocabulary. Among different aptitudes, rote memory strongly contributed to vocabulary breadth at all frequency levels and to paradigmatic and syntagmatic associations in LP learners. The contribution of aptitudes in HP was unclear. The findings show that rote memorization of vocabulary may be helpful for low proficiency learners, whereas additional factors may be at play in advanced level learners. Thus, further studies are needed in investigating additional factors that may contribute to vocabulary development as learners' proficiency increases.

KEYWORDS

language aptitude, L2 vocabulary breadth and depth, English proficiency level, Chinese EFL learners

1. Introduction

The inquiry into both ESL and EFL learners' vocabulary learning has been emerging in the past two decades since a rich L2 vocabulary knowledge is considered a crucial component in L2 development (Clenton and Booth 2021; Nation 2013). Developing vocabulary knowledge can be categorized into at least two dimensions of breadth and depth (Anderson and Freebody 1981). The breadth and depth of L2 learners' vocabulary knowledge growth are both significant since the former is viewed as how large a size of vocabulary that learners achieve, and the latter is interpreted as how well learners know word associations (i.e., syntagmatic, paradigmatic, and phonological associations) (Meara 1983; Read 1993). Previous studies have provided evidence that the two show strong relationships with L2 leaner's proficiency including listening comprehension ability (e.g., Wen 2014), reading ability (Choi 2013; Rashidi and Khosravi 2010; Şen and Kuleli 2015), speaking (Enayat and Derakhshan 2021) and writing (Johnson et al. 2016).

Since the breadth and depth of vocabulary knowledge are significantly related to learners' L2 proficiency, one of the key issues is in the investigation of contributing factors (Laufer and Hulstijn 2001). Lee (2020) argued that thus there are different factors involved in L2 vocabulary knowledge and called for more research to discover contributions of different factors to dimensions of L2 vocabulary knowledge. In fact, research on L2 vocabulary during the past two decades has examined the relationships between a range of factors (e.g., input quality, motivation, anxiety, and language aptitude) and vocabulary knowledge. For instance, Sun et al. (2015) concluded that learners' L2 input quality and age of onset significantly affected their vocabulary knowledge. Also, the amount of L2 input was proved to be one of the individual difference factors to predict learners' vocabulary breadth and depth (Unsworth et al. 2014). Besides, motivation, anxiety (Ortega 2009), and foreign language aptitude (Cha and Kim 2019, Yang and Cao 2020) have been found to be potential factors that can influence L2 vocabulary knowledge.

A significance of language aptitude on L2 learning has been demonstrated in numerous recent studies. In L2 vocabulary development studies, it was shown that language aptitude was significantly correlated with learners' vocabulary breadth and depth. For instance, it was demonstrated that language aptitude significantly influenced depth of vocabulary knowledge through collocations of words (Granena and Long 2012). Lee (2020) made a claim that different subdivided aptitudes (i.e., rote memory, phonetic coding ability, and grammatical sensitivity) positively had an impact on advanced L2 learners' vocabulary breadth and depth. Additionally, Yang and Cao (2020) questioned whether language aptitude was associated with L2 learners' initial vocabulary acquisition and concluded that learner's L2 aptitude and vocabulary knowledge had only a moderate relationship.

As mentioned above, the relationship between language aptitude and L2 learning, especially vocabulary development, is inconsistent. Further research examining a relationship between L2 aptitude and vocabulary learning is called for since the previous empirical studies have focused primarily on learners at the same L2 proficiency levels. Very little research has been done on the comparison of different proficiency level learners of English (se Suárez and Gesa 2019) and most studies have only explored the role of language aptitude in vocabulary breadth and depth via receptive measurement (e.g., Lee 2020). Additionally, very little is known about the effects of different types of language aptitude (i.e., rote memory, phonemic coding, and grammatical sensitivity) on vocabulary knowledge. Hence, the purpose of the current study is to seek a more comprehensive interpretation of how the aptitude with its subdivided abilities function in affecting different proficiency learners' L2 vocabulary knowledge. This study aims to answer the following research questions:

1. Are subdivided aptitudes (rote memory, phonetic coding ability, and grammatical sensitivity) effective predictors for vocabulary breadth and depth of Chinese EFL learners?
2. If so, are there significant differences according to learners' English proficiency?

2. Literature review

2.1 Language aptitude

Learners differ widely in terms of how quickly and effectively they learn a foreign or second language (Ortega 2009). L2 researchers have generally described such differences in learning feature as L2 learners' individual differences. In addition to motivation, anxiety, learning experience and strategy, language aptitude in the context of L2 learning refers to an individual's ability for both cognitive and perceptual learning (Carroll 1981, Ortega 2009). It is often described as one's talent for mastering new languages (Robinson 2005). Carroll (1962) divided language aptitude into four discrete and quantifiable abilities: rote memory, phonetic coding ability, grammatical sensitivity, and inductive learning ability.

Language aptitude significantly influences learners' L2 development and accomplishment (Dörnyei 2005). It is viewed as a continuously tenacious individual feature that does not change significantly during the entire L2 learning process, in contrast to other individual difference factors such as motivation, engagement and anxiety that are likely to alter across L2 learning stages (Saito et al. 2019). According to Robinson (2013), students with higher aptitude perform better in the process of L2 learning, and ultimately reach higher levels of L2 proficiency.

Skehan (2016) stated that the segmented abilities of language aptitude can benefit L2 learning performance in several ways throughout the SLA process. For instance, L2 learners with stronger phonetic coding ability can successfully remember newly observed sounds and can absorb and analyze language details (Yilmaz and Koylu 2016). That is, advanced phonetic coding ability permits L2 learners to use new languages more accurately and comprehend input. Additionally, learners with higher rote memory are also found to be capable of successfully mapping form-meaning associations by rapidly connecting target forms to recent learning terms as measured by the test of vocabulary breadth (Schneiderman and Desmarais 1988).

Aptitude is known to have a predictive value in L2 learning outcomes. According to Li (2016), there is a strong association between learners' L2 proficiency and language aptitude as measured by the aptitude test. In this study, language aptitude was found to be responsible for approximately 25% of the variation in L2 learners' overall proficiency. Furthermore, L2 learners' language aptitude was discovered to be predictable in a variety of characteristics of L2 learning (Li 2018). Specifically, composite aptitude scores were found to be significantly correlated in the order of L2 grammar, listening, speaking, and vocabulary ($r = .50$). Yet, it is unclear as to precisely which components of aptitude is related to the process of L2 learning since multiple components form language aptitude. Among different types of language aptitude, rote memory is defined as the capacity to not only store vocabulary information in memory but also to recall it later (Carroll and Sapon 1959). It is believed that learners with strong rote memory also have strong decontextualized learning capabilities for form-meaning correlations and better able to create associations in their memory and acquire new vocabulary quickly (Ortega 2009). Phonetic coding ability is defined as the capacity to recognize sounds, relate those sounds to their graphic symbols, and remember those sound-symbol connections. In other words, activating words' sounds with corresponded sounding-out symbols is the first step to master words' meaning when learning a foreign language (Meara 2009). Grammatical sensitivity is defined as a specific aptitude for understanding how linguistic components fit together to form linguistic wholes (Ortega 2009). That is, it enables students to recognize how certain words are used in sentences. Strong grammatical sensitivity is a crucial component of language learning since it enables learners to recognize various grammatical patterns and utilize them correctly (Meara 2009).

Modern Language Aptitude test (MLAT), created by Carroll and Sapon, is the earliest language aptitude test created for L1 English speakers and assesses rote memory, phonetic coding ability and grammatical sensitivity. The LLAMA test, created by Meara (2005), is a language aptitude test extensively used in L2 acquisition studies today (e.g., Bokander 2020, Granena and Long 2012, Yang and Cao 2020) partly due to its language neutrality (Rogers et al. 2017). It is composed of four subtests: LLAMA_B (i.e., rote memory), LLAMA_D (i.e., inductive learning ability), LLAMA_E (i.e., phonetic coding ability), and LLAMA_F (i.e., grammatical sensitivity). Unlike the MLAT, the LLAMA test is freely accessible to teachers, students, and researchers and has been used in many L2 studies. Studies have found that the results are correlated (r = .40 to r = .60) with learners' proficiency levels (Bokander and Bylund 2019, Rogers et al. 2017). Specifically, LLAMA_B, LLAMA_E, and LLAMA_F were found to be significantly correlated with students' L2 proficiency and educational levels (r = .40 to r = .60, p < .01). However, LLAMA_D, measuring inductive learning ability did not significantly correlate with L2 proficiency (p > .05) (Rogers et al. 2017).

2.2 L2 vocabulary knowledge and aptitude

There have been multiple previous attempts to illustrate a significant association between L2 aptitude and L2 vocabulary knowledge. Li (2016), a meta-analysis study on empirical studies of language aptitude explored over the previous five decades, concluded that overall aptitude as measured by the MLAT was positively associated to learners' L2 proficiency and achievement and suggested that a L2 learner with high proficiency level would have a stronger language aptitude and vice versa. However, L2 vocabulary learning was found to have a weak relationship with language aptitude. Kim (2018) explored whether L2 aptitude impacted on vocabulary learning when paired with learners' working memory capacity. In her study, 18 Korean EFL learners were asked to take a combined working memory test and two language aptitude tests administered by the overall MLAT and LLAMA_B subtest, respectively. The result showed that aptitude test correlated with working memory, but not with learners' L2 vocabulary learning. The results from both studies claim that aptitude may play only a weak role. Li (2016) went on to claim a need for further studies using a test other than the MLAT since it was created for L1 English speakers.

There are other studies that have shown contrastive results with L2 aptitude and vocabulary, especially when they measure vocabulary in terms of its breadth and depth. Granena and Long (2012) investigated the development of vocabulary depth in 65 similar proficiency level Chinese L2 learners of Spanish, as well as the influence of language aptitude. Participants were designed to three groups based on their ages: three to six years old, seven to fifteen years old, and sixteen to twenty-nine years old. In their study, vocabulary depth was measured by collocations of words, and the LLAMA test was used for language aptitude. Results found that L2 aptitude, measured by the LLAMA test, was a significant predictor of L2 vocabulary depth in 16- to 29- year-old group.

Additionally, a connection between language aptitude and vocabulary breadth was measured in Dahlen and Caldwell-Harris (2013), where 88 English speakers leaning Turkish as L2 participated. They found that aptitude, measured by MLAT, had a significant impact on L1 English speakers' capacity to learn Turkish as L2 of form and meaning word recall tasks. Cha and Kim (2019) found a significant correlation between language aptitude by LLAMA test and vocabulary breadth by young Korean EFL learners with similar proficiency levels. Suárez and Gesa (2019) study examined EFL learners in different proficiency levels by investigation the association between vocabulary learning, language aptitude, and captioned TV programs in EFL students. They discovered that low level students may rely on their L2 aptitude when learning vocabulary, whereas high level students may rely not only on language aptitude but on other factors such as learning methods. The studies here have attested at least some positive correlation with overall language aptitude and vocabulary breadth.

In addition to the research on the overall language aptitude, researchers have explored the role of subdivided aptitudes in L2 vocabulary, mainly focusing on the role of rote memory. For instance, Kormos (2012) concluded that learners with strong rote memory skills may benefit from memorizing a broader repertoire of L2 vocabulary, resulting in a richer vocabulary size. Also, Bokander (2021) explained at least conscious memorization of words (i.e., rote memory) for acquiring an L2 vocabulary appears to exist. Hackl (2018) hypothesized that those with a high working memory capacity have a higher aptitude, especially stronger rote memory, for learning new words. Yet, no published studies have so far included a correlation between the other two subdivided aptitudes (i.e., phonetic coding ability and grammatical sensitivity) and L2 vocabulary.

So far, the previous studies mentioned here have indicated the following: First, overall language aptitude seems to be positively related to L2 learners' vocabulary knowledge. Yet, some studies on examining subdivided aptitudes have mainly focused on rote memory. Second, most studies have focused on vocabulary breadth by measuring L2 leaners' size of vocabulary. Third, most studies have examined L2 leaners with same proficiency levels. Thus, there is currently no direct evidence to support the existence of these subdivided aptitudes' effects on the breadth and depth of L2 vocabulary especially with learners in different proficiency levels. It remains to be seen how different subdivided aptitudes function in affecting different proficiency L2 learners' vocabulary breadth and depth in the current study.

3. Method

3.1 Participants

Sixty Chinese EFL learners were recruited from several different colleges in China. Neither had previously lived nor studied in any English-speaking countries. They were divided into two groups of high proficiency (HP) and low proficiency (LP) based on their scores of the National College English Test (CET), which is a standardized English proficiency test in Chinese tertiary education set out to measure listening, reading, writing and translation. A total of 30 who scored above 520 out of possible 710 were assigned to the HP, whereas the other 30 who scored under 452 were grouped into the LP (N = 30). Table 1 shows that a significant difference existed between the two groups based on their CET test scores (t = 5.719, p = .000).

Table 1. Participants' CET Test Scores

Groups	*N*	*M*	*SD*	*Range*	*t*	*Sig.*
HP	30	560.24	31.44	521-624	5.719	.000
LP	30	413.65	40.57	309-452		

Note. HP = high proficiency, LP = low proficiency

3.2 Materials

Three types of tests of Vocabulary Levels Test (VLT), Word Associates Test (WAT), and LLAMA test were conducted in this study. Firstly, a modified version of VLT by Webb, Sasao and Balance (2017) was administered to measure the participants' vocabulary breadth. The test had been modified to include only the first five frequency levels from 1000 to 5000. Learners' receptive vocabulary knowledge was measured by this test of 30 target words per level. Thus, VLT was composed of 150 questions (30 x 5 frequency levels). All test words were selected from Nation (2012)'s BNC/COCA headword lists, each of which contained an equal number of nouns, adjectives, and verbs. For instance, upon being provided with verb clusters such as '*get ready*', '*make a happy sound*', and '*not remember*', they were provided with a list of options ranging from '*drink*', '*educate*', '*forget*', '*laugh*', '*prepare*', '*suit*', and '*I don't know*'. The participants then had to choose the most appropriate word for each verb cluster.

Additionally, the present study employed an open-ended version of WAT which had been modified from Read (1988) to measure EFL learners' vocabulary depth (i.e., syntagmatic, paradigmatic, phonological, and non-related associations) via an open-questioned test to measure productive vocabulary knowledge. Here, the participants had to write down the words that first came to their mind upon being presented with the stimuli words. The stimuli words were taken from Lu and Lim (2021) which excluded the words that had been used in the VLT. Thus, a total of 18 stimuli words were used: two nouns, adjectives, and verbs were chosen from the first 1000, 2000, and 3000 levels' wors respectively. All descriptions of the questions were presented in Chinese except for the stimuli words.

The responses were collected and classified by the researchers into four categories of syntagmatic, paradigmatic, phonological and non-related.

Finally, the study employed the LLAMA test (Meara 2005) to measure the participants' different types of language aptitude and operated by using the program of *PsychoPy* (2021), which is an application for creating experiments in behavioral science (Peirce et al. 2019). The original LLAMA test is composed of a 40-item and 4-section downloadable questions to assess L2 learners' aptitude levels. However, one of the sections, specifically LLMA_D, was excluded from this study since it was known to be not significantly correlated with learners' L2 proficiency and educational levels (Rogers et al. 2017). Thus, the modified version used in this study was composed of three tests of LLAMA_B, E, and F of measuring rote memory, phonetic coding ability and grammatical sensitivity respectively and constituted a total of 25 test items.

In LLAMA_B, where it measured how efficiently EFL learners could remember the names of unfamiliar objects in learning a foreign language, they were shown a series of 10 unusual objects and asked to learn their names within a minute. Then, the objects were shuffled and presented to the participants asking to identify the objects. For each correctly identified object, 1 point was awarded, with a score ranging from 0 to 10. LLAMA_E tested the ability to match familiar sounds to a new unfamiliar writing system. Participants were given 1 minute and 20 seconds to explore and learn the association between 16 unfamiliar sounds and symbols. Every symbol consisted of a number and an unfamiliar alphabet of a single syllable. The participants were then asked to listen to a new association sound that involved two syllables to choose the corresponding associations of the symbols. Again, 1 point was awardee with a score ranging from 0 to 7. LLAMA_F examined how efficiently EFL learners could learn the grammar rules of a foreign language. To this end, they were shown nine pictures of certain objects and shapes, and a sentence describing each picture. They had to figure out the similarities and differences in these pictures with their corresponding sentences in two minutes. During the test, eight pictures and four different choices for each picture were presented. The total score of this test was 8.

3.3 Procedure

The experiments were conducted via one-on-one online experiment session using Zoom with cameras on to ensure the participants' full attention to the tasks. The entire process took approximately an hour including a tutorial session. The participants who completed all tests received a monetary reward. The procedure in timetable is listed in Table 2.

Table 2. Task Procedure

Timing	Task	Duration
1	Pre-experiment instructional session 1 Practice session	2 min.
2	Vocabulary Levels Test (VLT) Word Associates Test WAT)	30 min. 5 min.
3	Break	5 min.
4	Pre-experiment instructional session 2 Practice session	3 min.
5	LLAMA Test	15 min

3.4 Data Analysis

Regarding the results of vocabulary depth, the researchers classified the participants' responses from WAT, based on Fitzpatrick (2006) and Meara (2009). Paradigmatic associations included synonym (e.g., *autumn – fall*), antonym (e.g., *dirty - clean*), and context-related responses of superordinate, subordinate, and coordinate (e.g., *liquid- water*, *article - paraphrase*). Syntagmatic associations included collocation, where the responses and the stimuli words can be put together or become phrases (e.g., *native* - *speaker*, *borrow* - *money*, *combine* - *with*). As for non-related associations, no apparent relationships existed between the stimulus and the response (e.g., *native* - *young*, *delay* - *dirt*). Phonological association consisted of two different subcategories: morphophonology and similarity in form. Morphophonology represented that the stimuli words and the responses had the same root but with different affix (e.g., *electric* - *electricity*, *soft* - *softly*). Similarity in form refers to responses that only shared the similar forms with the stimulus (e.g., *cake* and *fake*, *fiction* and *friction*). Additionally, some responses presented two or more types of associations simultaneously. Responses such as *chocolate-cake* presented both paradigmatic and syntagmatic associations since both were food and *chocolate cake* represented a kind of cake flavor. *Electric-electrical* were also a combination of paradigmatic and phonological associations since they had the same root and shared the similar meaning.

Next, a series of statistical analyses were conducted. Firstly, descriptive statistical analyses were conducted with the results of L2 aptitude and vocabulary measured by breadth and depth. Independent samples *t*-tests were also utilized to see if there were any significant differences between the LP and HP. Pearson correlation analyses were then performed to determine the relationships between vocabulary breadth, depth, and L2 aptitude. Following the establishment of the correlation, multiple regression analyses were conducted to examine the role of L2 aptitude (i.e., rote memory, phonetic coding ability, and grammatical sensitivity) in the development of vocabulary breadth and depth.

4. Results

4.1 Descriptive Statistics

4.1.1 Descriptive Statistics of Vocabulary Breadth: VLT

The VLT was a five-level test with ten cluster-questions each with the maximum possible score of 150. The descriptive statistics of the vocabulary breadth from each group are shown in Table 3, where the two groups differed significantly in terms of the vocabulary breadth at all levels. The mean score of VLT of HP was 118.33 out of 150 (*SD* = 14.57). The first 1000-level test showed the highest mean accuracy of 29.46, with the 5000-level test showing the lowest mean accuracy of 25.26. Additionally, as the vocabulary level increased, *SD*s increased accordingly (i.e., 0.77 for 1000-level to 6.13 for 5000-level). The result of LP shows that the mean was 76.6 out of 150 (*SD* = 31.12), indicating a large individual difference within the group when compared to the HP. Again, LP learners showed less accuracy as the difficulty level increased.

Table 3. Descriptive Statistics of Vocabulary Breadth, by Group

VLT	HP (*n*=30)		LP (*n*=30)		*t*	*p*
	M	*SD*	*M*	*SD*		
Overall	118.33	14.57	76.6	31.12	6.652	.000***
1000-level	29.46	0.77	25.26	6.06	3.764	.001**
2000-level	27.13	2.59	18.9	7.66	5.571	.000***
3000-level	25.53	3.22	15.33	8.49	6.151	.000***
4000-level	19.86	4.91	10.26	6.52	6.440	.000***
5000-level	16	6.13	5.83	5.54	6.739	.000***

Note. ***p*<.01, ****p*<.001

4.1.2 Descriptive Statistics of Vocabulary Depth: WAT

HP provided 540 actual responses whereas LP provided 465 out of possible 540 since they did not provide any responses to 65 stimuli. Additionally, 3 responses from HP and 10 from HP were excluded due to incorrect spellings that led to difficulty in understanding. Minor spelling errors were disregarded. The final valid responses were 537 from HP and 465 from LP, which then were classified into paradigmatic, syntagmatic, phonological and non-related. Some responses were classified into two or more associations; Overall, HP and LP provided 61 and 30 of such associations respectively. Specifically, HP and LP answered 38 and 21 responses, respectively, of a combination of the paradigmatic and syntagmatic association. The number of responses that had a combination of paradigmatic and phonological association from HP and LP were 17 and 7, respectively. In both groups, the third combinations of syntagmatic and phonological associations were 6 and 2. Thus the total number of valid associations were 598 for HP and 495 for LP.

During the WAT analysis, HP (N = 598) provided more valid associations than LP (N = 495), indicating that HP had a greater vocabulary depth than LP. Similarly, paradigmatic associations accounted for the greatest proportion of responses in both groups (HP, 54%; LP, 40%). In HP, the second most common response was syntagmatic associations (25%), followed by non-related (12%) and phonological associations (9%). In contrast to HP, participants in LP created comparable proportions of syntagmatic (28%) and non-related associations (25%). Furthermore, responses with phonological association (7%) were less frequent in both groups than other associations. The t-test was employed to determine the differences between two groups. As presented in Table 4, two groups performed differently. HP produced significantly more responses than LP (t = 4.037, p = .000). However, the two groups' vocabulary depth measured by the WAT was inconsistent. Among the four association types, responses from paradigmatic (t = 5.043, p = .000) and non-related associations (t = -2.484, p = .017) showed significant differences between the two groups. There were no significant differences with responses from syntagmatic and phonological associations.

Table 4. Descriptive Statistics of Vocabulary Depth, by Group

Association Types	HP (n = 30)				LP (n = 30)				t	p
	N	%	M	SD	N	%	M	SD		
Overall	598	100	19.93	1.01	495	100	16.5	4.54	4.037	.000***
Paradigmatic	323	54	10.77	2.15	199	40	6.63	3.94	5.043	.000***
Syntagmatic	149	25	4.97	2.08	138	28	4.6	2.84	0.571	.570
Phonological	52	9	1.73	1.46	35	7	1.17	1.58	1.444	.154
Non-related	74	12	2.47	1.85	123	25	4.1	3.09	-2.484	.017*

4.1.3 Results of Language Aptitude: LLAMA

In terms of the language aptitude measured by the LLAMA test, a significant difference between the two groups were also found. Table 5 shows that HP outscored LP by 6.3 points in the overall LLAMA test (t = 7.149, p = .000). The two groups differed significantly in each subdivided test. HP scored considerably higher than LP (t = 5.938, p = .000) in LLAMA_B, implying that learners with higher English proficiency have better rote memory. In the LLAMA_E, HP consistently scored higher than LP (t = 6.966, p = .000), suggesting that high proficiency L2 learners tend to have stronger phonetic coding ability. HP also outperformed LP (t = 4.412, p = .000) in the LLAMA_F test, indicating a better grammatical sensitivity than LP.

Table 5. Descriptive Statistics of LLAMA, by Group

LLAMA test	HP (*n* = 30)		LP (*n* = 30)		*t*	*p*
	M	*SD*	*M*	*SD*		
Overall	15.63	3.34	9.33	3.49	7.149	.000***
LLAMA_B	7.77	1.5	5.03	2.03	5.938	.000***
LLAMA_E	3.37	1	1.5	1.07	6.966	.000***
LLAMA_F	4.53	1.68	2.8	1.35	4.412	.000***

Note. ***p<.001

4.2 Results of Correlation Analyses

4.2.1 Correlations of Language Aptitude and Vocabulary Breadth

To see if there were any correlations between all measurements in the two proficiency groups, a series of Pearson correlation analyses were conducted. Firstly, the vocabulary breadth of the high proficiency group was compared to their subdivided aptitudes as determined by the LLAMA test. As indicated in Table 6, each level of the VLT were correlated with three aptitude tests. Rote memory measured by LLAMA_B was found to be significantly correlated with 1000 (r = .511, p < .01) and 5000 levels (r = .371, p < .05) of vocabulary breadth, but not with 2000, 3000, and 4000 levels. Furthermore, LLAMA_F was only found to be significantly correlated with 1000 level (r = .438, p < .05) of vocabulary breadth. However, no significant correlation was found between LLAMA_E and each level of vocabulary breadth.

Table 6. Correlation between L2 Aptitude and Vocabulary Breadth, HP

	HP (*n*=30)				
	1000-level	2000-level	3000-level	4000-level	5000-level
LLAMA_B	.511**	.291	-.002	.225	.371*
LLAMA_E	.261	.206	.141	.348	.248
LLAMA_F	.438*	.356	.233	.160	.292

Note. LLAMA_B: rote memory; LLAMA_E: phonetic coding ability; LLAMA_F: grammatical sensitivity; *p<.05, **p<.01

Further, the Pearson correlation analyses were conducted with LP and their language aptitude and vocabulary breadth (see Table 7). First, there was a significant correlation between the LLAMA_B and each level of the VLT in LP, with the 2000 level showing the closest correlation (r = .658, p < .001), followed by 1000 (r = .611, p < .01), 3000 (r = .477, p < .01), 4000 (r = .467, p < .01), and 5000 (r = .397, p < .05) levels. The weakest correlation was found with the 5000 level. In contrast, a weak correlation was only found between the LLAMA_E and 5000 level (r = .362, p < .05) of the VLT. Additionally, it was discovered that the 1000 (r = .567, p < .01) and 2000 (r = .548, p < .01) levels of the VLT had a positive correlation with the LLAMA_F.

Table 7. Correlation between L2 Aptitude and Vocabulary Breadth, LP

	LP (n=30)				
	1000 level	2000 level	3000 level	4000 level	5000 level
LLAMA_B	.586**	.658***	.477**	.467**	.397*
LLAMA_E	.164	.299	.17	.28	.362*
LLAMA_F	.567**	.548**	.337	.308	.221

Note. LLAMA_B: rote memory; LLAMA_E: phonetic coding ability; LLAMA_F: grammatical sensitivity; *p<.05, **p<.01, ***p<.001

4.2.2 Correlations of Language Aptitude and Vocabulary Depth

No significant correlation was found between vocabulary depth and each aptitude from HP. Thus, a Pearson correlation analysis was conducted only with LP, to analyze the relationship between L2 aptitude and vocabulary depth in that group. Table 8 provides the results. The LLAMA_B was found to have a significant correlation with paradigmatic (r = .438, p < .05), syntagmatic (r = .477, p < .01), and non-related (r = -.441, p < .05) associations. However, no significant correlation was discovered between the LLAMA_E and four association types, or between the LLAMA_F and the four associations.

Table 8. Correlation between L2 Aptitude and Vocabulary Depth, LP

	LP (*n*=30)			
	Paradigmatic	Syntagmatic	Phonological	Non-related
LLAMA_B	.438*	.477**	.246	-.441*
LLAMA_E	.134	.079	.132	.026
LLAMA_F	.316	.312	-.032	-.045

Note. LLAMA_B: rote memory; LLAMA_E: phonetic coding ability; LLAMA_F: grammatical sensitivity; *p < .05, **p < .01

4.3 Results of Regression Analyses

Firstly, a series of Cronbach's tests were conducted to ascertain the reliability of further tests. The Cronbach's alphas for three subdivided aptitude tests, five language breadth tests, and four language depth tests were .864, .831, and .861 respectively. Thus, to investigate whether the three subdivided L2 aptitudes had a predictive power on the development of vocabulary breadth and depth, a series of multiple regression analyses were performed with the data where significant correlations were found. The dependent variables were the vocabulary breadth in five levels and vocabulary depth in four association types. The three independent variables were: (1) rote memory (i.e., LLAMA_B), (2) phonetic coding ability (i.e., LLAMA_E), and (3) grammatical sensitivity (i.e., LLAMA_F).

4.3.1 Language Aptitude and Vocabulary Breadth

Since no significant correlations were found between the three subdivided aptitudes and the vocabulary breadth along with 2000, 3000, 4000 levels in HP, the multiple regression analyses were conducted to further identify three subdivided abilities' predictive power of 1000 and 5000 levels of vocabulary breadth only.

Table 9. Multiple Regression Analyses for Aptitudes and Vocabulary Breadth, HP

		HP (*n*=30)						
		B	*SE*	*Beta*	*t*	*adj.* R^2	R^2 *Change*	*F*
LLAMA_B	1000	.264	.084	.511**	3.146**	.235**	.261**	9.896**
LLAMA_E		.096	.133	.124	.722	-	-	
LLAMA_F		.093	.094	.222	.994	-	-	
LLAMA_B	5000	1.515	.717	.371*	2.113*	.107*	.138*	4.466*
LLAMA_E		.940	1.131	.153	.832	-	-	
LLAMA_F		.327	.809	.123	.404	-	-	

Note. LLAMA_B: rote memory; LLAMA_E: phonetic coding ability; LLAMA_F: grammatical sensitivity; Dependent variables: 1000 level, 5000 level of vocabulary breadth; **p <. 01, *p < . 05

As shown in Table 9, HP's rote memory accounted for 26.1% of the variance in 1000 level of vocabulary breadth when it was entered (R^2 = .235, *F* (1, 28) = 9.896, *p* < .01). As a result of the regression coefficient test on rote memory, it was found that HP's rote memory had a statistically significant effect on the 1000-level of vocabulary breadth (*t* =3.146, *p* < .01). In addition, the beta value of rote memory (*β* = .511) indicated that the stronger the rote memory, the better the performance on the 1000-level of vocabulary breadth. However, phonetic coding ability (*β* = .124, *p* = .477) and grammatical sensitivity (*β* = .222, *p* = .330) did not account any variance to the 1000-level of vocabulary breadth when the two variables were considered as indicated by the *p* values. Table 9 further presents the results for LLAMA_ B, LLAMA_E, and LLAMA_F in 5000 level of vocabulary breadth. The Model showed HP's rote memory explained 13.8% of the variance at the 5000 level (R^2 = .107, *F*(1, 28) = 4.466, *p* < .05). It also revealed that HP's rote memory had a statistically significant influence on the 5000 level of vocabulary breadth (*t* = 2.113, *p* < .05). According to the beta value of rote memory (*β* = .371), the performance on the 5000-level of vocabulary breadth improved with stronger rote memory. However, phonetic coding ability (*β* = .153, *p* = .413) and grammatical sensitivity (*β* = .123, *p* = .569) did not account any variance to the 1000 level of vocabulary breadth.

Based on the correlation analyses mentioned above, a series of the multiple regression analyses were also administrated in LP to explore three L2 aptitudes' predictive power in vocabulary breadth with each level, as is evidenced in Table 10.

Table 10. Multiple Regression Analyses for Aptitudes and Vocabulary Breadth, LP

LLAMA	LP (*n* = 30)							
		B	*SE*	*Beta*	*t*	*adj.* R^2	R^2 *Change*	*F*
LLAMA_B	1000	1.755	.717	.586**	3.828**	.320**	.344**	10.685***
LLAMA_E		-.166	.929	-.029	-.178	-	-	
LLAMA_F		1.713	.771	.381*	2.221*	.385*	.105*	
LLAMA_B	2000	2.489	.539	.658***	4.618***	.412***	.432***	21.323***
LLAMA_E		.683	1.086	.096	.629	-	-	
LLAMA_F		1.585	.932	.289	1.700	-	-	
LLAMA_B	3000	1.998	.696	.477**				
LLAMA_E		.135	1.413					
LLAMA_F		.813	1.269					
LLAMA_B	4000	1.502	.538	.467**	2.791**	.190**	.218**	7.789**
LLAMA_E		.874	1.079	.144	.810	-	-	
LLAMA_F		.339	.975	.096	.348	-	-	
LLAMA_B	5000	1.086	.474	.397*	2.288*	.127*	.158*	5.236*
LLAMA_E		1.343	.927	.261	1.448	-	-	
LLAMA_F		-.100	.839	.027	.132	-	-	

Note. LLAMA_B: rote memory; LLAMA_E: phonetic coding ability; LLAMA_F: grammatical sensitivity; Dependent variables: 1000 to 5000 levels of vocabulary breadth; ****p* < .001, ***p*<.01, **p*<.05

Table 10 shows LP's regression analyses for rote memory, phonetic coding ability, and grammatical sensitivity in the 1000 to 5000 levels of vocabulary breadth. The results show that rote memory can significantly explain 34.4% of the variance in the 1000-level of vocabulary breadth (R^2 = .320, *F*(1, 28) = 14.657, *p* < .01). Grammatical sensitivity accounted for 10.5% of the variance (R^2 = .385, *F*(1, 26) = 4.935, *p* < .05). Additionally, the regression coefficient of rote memory (*t* = 3.828, *p* < .01) and grammatical sensitivity (*t* = 2.221, *p* < .05) confirmed this. The beta values indicated that rote memory and grammatical sensitivity both had significant effects (*β* = .586; *β*

= .381) with possibly a stronger effect of rote memory. However, phonetic coding ability (β = -.029, p = .137) did not provide any variance to the vocabulary breadth. With 2000 level of vocabulary breadth, it was shown that rote memory can significantly predict 43.2% of the variance in the 2000 level of vocabulary breadth (R^2 = .412, F(1, 28) = 21.323, p < .001). The beta value of rote memory (β = .658) displayed that rote memory strongly influenced LP's performance on the 2000 level's vocabulary breadth whereas phonetic coding ability (β = .096, p = .534) and grammatical sensitivity (β = .289, p = .080) did not predict any variance. For 3000 level, LP's rote memory accounted for 22.7% of the variance in their vocabulary breadth (R^2 = .200, F(1, 28) = 8.277, p < .01). The regression coefficient test of rote memory revealed that LP's rote memory had a statistically significant effect (t = 2.868, p < .01). The beta value of rote memory (β = .477) indicated that LP participants with stronger rote memory mastered more vocabulary at the 3000-level. However, phonetic coding ability (β = .017, p = .925) and grammatical sensitivity (β = .128, p = .659) did not predict any variance for the 3000 level in LP since no valid values were found. With 4000 level of vocabulary breadth, only rote memory significantly predicted 21.8% of the variance (R^2 = .190, F(1, 28) = 7.789, p < .01). The beta value (β =.467) revealed that LP participants with greater rote memory showed better learning in this level. However, phonetic coding ability (β = .144, p = .425) and grammatical sensitivity (β = .096, p = .629) did not predict any variance at this level. The rote memory significantly contributed to LP's performance on the 5000 level of vocabulary breadth, accounting for 15.8% variance (R^2 = .127, F(1, 28) = 5.236, p < .05). LP's rote memory had a statistically significant effect on their performance on the 5000 level of vocabulary breadth (t = 2.288 and p < .05). The beta value (β = .397) also showed that the stronger one's rote memory, the larger one's vocabulary size at the 5000-level. However, phonetic coding ability (β = .261, p = .159) and grammatical sensitivity (β = .027, p = .896) did not predict any variance. Overall, rote memory provided significant contribution to LP's vocabulary breadth at most levels whereas grammatical sensitivity only exerted a limited contribution to LP's vocabulary at 1000-level.

4.3.2 Language Aptitude on Vocabulary Depth

Since no correlations were found between language aptitude and vocabulary depth with HP, regression analyses were not conducted with this group. The regression analyses showing the predictive value of three subdivided L2 aptitudes in LP's vocabulary depth are shown in Table 11.

Table 11. Multiple Regression Analyses for Aptitudes and Vocabulary Depth, LP

		LP (n = 30)						
		B	*SE*	*Beta*	*t*	*adj.* R^2	R^2 *Change*	*F*
LLAMA_B	Paradigmatic	.852	.331	.438*	2.575*	.163*	.192*	6.633*
LLAMA_E		-.032	.671	-.009	-.048	-	-	
LLAMA_F		.388	.603	.127	.637	-	-	
LLAMA_B	Syntagmatic	.667	.233	.477**	2.869**	.200**	.227**	8.234**
LLAMA_E		-.223	.470	-.085	-.474	-	-	
LLAMA_F		.249	.423	.094	.482	-	-	
LLAMA_B	Non-related	-.673	.259	-.441*	-2.604*	.166*	.195*	6.779*
LLAMA_E		.544	.514	.189	1.058	-	-	
LLAMA_F		.483	.456	.241	1.238	-	-	

Note. LLAMA_B: rote memory; LLAMA_E: phonetic coding ability; LLAMA_F: grammatical sensitivity; Dependent variable: WAT (i.e., vocabulary depth); *p < .05

Multiple regression analyses were performed to see the effect of three subdivided abilities on each association type of vocabulary depth. In Table 11, LP's rote memory accounted for 19.2% of the variance in participants' final valid responses with paradigmatic association (R^2 = .163, F (1, 28) = 6.633, p < .05). That is, participants with stronger rote memory produced more valid responses with paradigmatic association (t = 2.575, p < .05). Additionally, the beta value of rote memory (β = .438) shows that LP's rote memory significantly explained their numbers of paradigmatic responses. However, when phonetic coding ability (β = -.009, p = -.048) and grammatical sensitivity (β = .127, p = .637) were entered, no significant variance was found.

The variance in participants' final valid responses with syntagmatic connection was explained by LP's rote memory in 22.7% of the cases (R^2 = .200, F (1, 28) = 8.234, p < .01). LP learners with better rote memory can generate more reliable syntagmatic association responses (t = 2.869, p < .01). Additionally, the rote memory of LP significantly and favorably explained their quantities of syntagmatic associations, as evidenced by the rote memory's beta value (β = .477). However, with phonetic coding ability (β = -.085, p = .639) and grammatical sensitivity (β = .094, p = .633), no significant variances were found to predict the numbers of responses on syntagmatic association. Regarding non-related responses, rote memory significantly predicted 19.5% of the variance in LP's responses with non-related association (R^2 = .166, F (1, 28) = 6.779, p < .05). That is, better rote memory contributed to LP's fewer production of non-related association responses (t = -2.604, p < .05). In addition, the beta value (β = -.441) showed that rote memory had a significantly negative impact on LPs' responses with non-related association in their vocabulary depth. However, no predictive powers were observed with both phonetic coding ability (β = .189, p = .299) and grammatical sensitivity (β = .241, p = .226).

To summarize the findings of regression analyses, LP's rote memory was responsible for 34.4%, 43.2%, 22.7%, 21.8%, and 15.8% of the positive differences in their vocabulary breadth at all frequency levels. Additionally, it was discovered that 10.5% of their grammatical sensitivity predicted vocabulary breadth at the 1000-level. With vocabulary depth, rote memory was found to be positively predictive in valid responses from LP participants for 19.2% and 22.7% with paradigmatic and syntagmatic associations. Furthermore, rote memory had 19.5% negative predictive power of the numbers of response from the LP that had non-related associations, showing higher rote memory predicted less non-related responses.

5. Discussion and Conclusion

This study investigated the role of language aptitude in Chinese EFL learners' vocabulary knowledge. Specifically, it examined the relationship between language aptitude and L2 vocabulary breadth at different frequency levels, and the relationship between language aptitude and L2 vocabulary depth categorized into four main association types. It also examined how subdivided aptitudes function in affecting different proficiency learners' L2 vocabulary breadth and depth.

First, the results showed that participants in the two proficiency groups differed in their performance on language aptitude and vocabulary breadth and depth, with HP outperforming LP in all three. Regarding vocabulary breadth, HP had a larger vocabulary size in each level test than LP. Regarding vocabulary depth, participants in HP produced more valid associations than LP in the vocabulary depth test thereby showing that HP's vocabulary depth outperformed that of LP. The findings suggest that L2 proficiency has a considerable impact on both vocabulary breadth and depth performance, as was claimed in Meara (2009). Specifically, HP produced more valid paradigmatic associations than LP, as was also attested in previous studies (Lu and Lim 202, Meara 2009, Read 1993). It seems that as learners develop their vocabulary depth, their paradigmatic associations develop. Regarding syntagmatic association, no difference was found between the two proficiency groups' responses. It is possible that syntagmatic association continues to pose challenges for EFL leaners even when their proficiency increases. Additionally, considering that syntagmatic association is generally created when learners encounter less frequent words (Meara 2009), the result from this study may have been found since its stimulus words had been selected from the first three most frequent levels of the BNC/COCA headword list. Thus, familiarity may be the possible reason that both HP and LP have rarely exhibited syntagmatic associations. Further research by incorporating words beyond the high frequency levels may substantiate the claim that syntagmatic association is less likely to be observed with high frequency vocabulary. Lastly, in the present study, the two proficiency groups produced the least phonological associations during the vocabulary depth test. The finding is in partial agreement to those of Read (1993) and Meara (2009) who found that L2 adult learners' response is rarely based on phonological association, unlike L2 child learners who tend to ignore meanings to produce phonological association such as *cook-coke*, *go-goat*, and *hum-him* (p.7). It is unclear whether the age alone is responsible for this difference, but it seems to be the case that phonological associations disappear as they are replaced by paradigmatic associations. To further substantiate the claim, learners in different age ranges need to be incorporated in future studies.

Regarding language aptitude, HP outperformed LP in both overall language aptitude and three subdivided aptitudes tests, supporting the previous studies in that L2 proficiency is significantly and positively related to one's performance on language aptitude (Dörnyei 2005, Robinson 2013). In the present study, rote memory was found to be most strongly correlated with L2 vocabulary among the three subdivided aptitudes. Here, it can be claimed that LP learners with stronger rote memory might have strong decontextualized learning capabilities for form-meaning correlations, as was claimed by Ortega (2009). Specifically, results of LP in the study suggest that except for 1000- and 2000-levels, predictive power decreased linearly with as frequency decreased. In other words, vocabulary breadth is more significantly predicted by rote memory skills in low level learners with more frequent levels of vocabulary.

LP's language aptitude was correlated significantly and positively with vocabulary depth. Specifically, it was positively correlated with valid responses of paradigmatic and syntagmatic associations and negatively correlated with the non-related association. A possible explanation might be that LP learners with strong rote memory are not only better at increasing vocabulary size, but also better at creating more valid word associations (i.e., the largest proportion for paradigmatic followed by syntagmatic and non-related for the least) as claimed by Ortega (2009). The finding partially supports Granena and Long's (2012) study where L2 aptitude was a significant predictor of L2 vocabulary depth in the adult group. Yet, their study did not divide participants based on proficiency levels.

In terms of HP in the present study, no significant relationship was found between language aptitude and vocabulary depth. Similar to the results of HP in vocabulary breadth, it is also possible that their vocabulary depth was affected by factors other than aptitude. Such a claim is substantiated by HP results where the rote memory was shown to account for 26.1% and 13.8% of the positive variations in their vocabulary breadth only at the 1000 and 5000 levels. Thus, the role of rote memory remains rather unclear for HP. Since no positive variation was observed in 2000 to 4000 levels for HP, there is a possibility that more skills may be needed when L2 proficiency as well as vocabulary size increase (Lee 2020, Suárez and Gesa 2019). In other words, high-proficiency L2 learners' vocabulary learning may not be explained solely by language aptitude; it is possible that other factors such as learning strategy, motivation, and L2 exposure (cf. Lee 2020) are at play, rather than only language aptitude that the low-level learners generally rely on. Another possibility is that advanced-level learners' aptitude, as mediated by motivation and strategy use, can contribute to their L2 learning results, as was demonstrated by Winke (2013). The aptitude of advanced-level L2 Chinese learners was investigated in this study to determine its impact on L2 learning in reading, listening, and speaking. As a result, learners' aptitude had an indirect effect on their L2 learning, which was mediated by strategy use and motivation. In order to investigate this mediation effect, future studies need to include factors such as strategy and motivation as variables in examining learners at different proficiency levels.

Overall, learners in the present study seemed not to have activated phonetic coding ability to recognize the written words during the aptitude tests. Yet, LP had a higher proportion of phonological association than HP. According to Schmitt (2000), L1 children exhibit more phonological association and a tendency to shift to more paradigmatic association in their vocabulary development. It may be that learners in the present study show a similar pattern. Since phonological association in both LP and HP exhibited the least proportion, the absence of statistical significance may have resulted due to small amount of data, so the above claim cannot be substantiated here. Future studies thus need to examine the exact effects on lower frequency words (i.e., words frequency that over 5000 level) to find out a comprehensive understanding of the relationship between phonetic coding ability and L2 vocabulary.

Regarding grammatical sensitivity, LP's grammatical sensitivity predicted 10.5% variance of vocabulary breadth at the 1000 level unlike in HP. According to Skehan (2002), different subdivided aptitudes perform different roles based on the following four L2 learning stages: the first stage being the beginning of L2 acquisition when learners start to realize or input new language knowledge (i.e., Noticing), the second stage (i.e., Patterning) where learners are expected to summarize some knowledge such as grammar rules. Unlike the first and second stages, the third stage (i.e., Controlling) where an upper level for further developing proficiency from low to the top, and the final stage (i.e., Lexicalizing) where learners output language fluently without errors. The reason that grammatical sensitivity is not correlated with HP's but LP's vocabulary breadth may be because it is considered to only play its role at the beginning of the second stage. In other words, learners higher than the beginning level may not use the grammatical sensitivity to develop their vocabulary depth. It can be claimed that HP learners administer some other cognitive abilities such as working memory and attentional processing, as claimed by Kim (2018) and Skehan (2002). Similarly, Robinson (2005) and Artieda and Muñoz (2016) claimed that cognitive skill (i.e., aptitude) alone might not be predictive of advanced learners' L2 acquisition and the current language aptitude assessments are not sensitive to the cognitive skills that are relevant at higher levels of L2 development. In line with their claims, HP in this study may possess abilities that LP do not have to develop their vocabulary knowledge.

To summarize, this study discovered that the development of high-frequency words such as those in the first 1000-frequency level in their isolated forms was significantly predicted by rote memory. This suggests that for less-skilled EFL learners, practicing vocabulary by rote memorization can help them recognize and recall words at least in their early stages. However, because lower-frequency words were not as susceptible to rote memorizing, educators should keep in mind that rote memorization is not as helpful in developing more advanced vocabulary. As for the advanced learners, stronger language aptitude may not be enough for their vocabulary development. Therefore, educators may wish to help learners find additional ways to develop their L2 vocabulary. They may need to provide various opportunities to improve learner motivation and engagement inside and outside of the classroom. In addition to classroom input, extracurricular L2 exposure made available via offline and online settings should be useful for vocabulary development.

There are some limitations to the present study. First, in terms of high-proficiency learners, language aptitude might yield stronger effects by including additional cognitive skills such as working memory and attentional processing. Thus, future studies may need to explore additional factors rather than aptitude alone to find out the abilities that are relevant at higher levels of L2 development of vocabulary. Second, in terms of the function of phonetic coding ability, increasing frequency levels of vocabulary may provide a more comprehensive understanding of its influence. Third, in terms of the participants, the role of language aptitude at different L2 proficiency should be varied. Future studies may include more diverse levels of participants (e.g., low, intermediate, high, and near-native levels) to discover the relationships between language aptitude and its effect on L2 vocabulary more precisely.

References

Anderson, R. C. and P. Freebody. 1981. Vocabulary knowledge. In J. T. Guthrie, ed., *Comprehension and Teaching: Research Reviews*, 77-117. Newark, DE: International Reading Association.

Artieda, G. and C. Muñoz. 2016. The LLAMA tests and the underlying structure of language aptitude at two levels of foreign language proficiency. *Learning and Individual Differences* 50, 42-48,

Bokander, L. 2020. Language aptitude and crosslinguistic influence in initial L2 learning. *Journal of the European Second Language Association* 4(1), 35-44.

Bokander, L. 2021. *Validity Considerations in the Study of Language Learning Aptitude*. Växjö, Sweden: Linnaeus University Press.

Bokander, L. and E. Bylund. 2019. Probing the internal validity of the LLAMA language aptitude tests. *Language Learning* 70(1), 11-47.

Carroll, J. B. 1962. The prediction of success in intensive foreign language training. In R. Glaser, ed., *Training Research and Education*, 87-136. New York: University of Pittsburgh Press.

Carroll, J. B. 1981. Twenty-five years of research on foreign language aptitude. In K. C. Diller, ed., *Individual Differences and Universals in Language Learning Aptitude*, 83-118. Rowley, MA: Newbury House.

Carroll, J. B. and S. M. Sapon. 1959. *Modern Language Aptitude Test (MLAT)*. New York: Psychology Corporation.

Cha, S. and J. Kim. 2019. Correlation between language aptitude and vocabulary learning instruction. *Korean Association for Learner-Centered Curriculum and Instruction* 19(18), 75-94.

Choi, H. 2013. Effects of depth and breadth of vocabulary knowledge on English reading comprehension among Korean high school students. *Language Research* 49(2), 419-452.

Clenton, J. and P. Booth. 2021. *Vocabulary and the Four Skills : Pedagogy, Practice, and Implications for Teaching Vocabulary*. London: Routledge.

Dahlen, K. and C. Caldwell-Harris. 2013. Rehearsal and aptitude in foreign vocabulary learning. *The Modern Language Journal* 97(4), 902-916.

Dörnyei, Z. 2005. *The Psychology of the Language Learner: Individual Differences in Second Language Acquisition*. London: Routledge.

Enayat, M. J. and A. Derakhshan. 2021. Vocabulary size and depth as predictors of second language speaking ability. *System* 99(3), 102521.

Fitzpatrick, T. 2006. Habits and rabbits: Word associations and the L2 lexicon. *EUROSLA Yearbook* 6, 121-145.

Granena, G. and M. H. Long. 2012. Age of onset, length of residence, language aptitude, and ultimate L2 attainment in three linguistic domains. *Second Language Research* 29(3), 311-343.

Hackl, H. 2018. Aptitude for vocabulary acquisition. In S. M. Reiterer, ed., *Exploring Language Aptitude: Views from Psychology, the Language Sciences, and Cognitive Neuroscience*, 43-56. Berlin: Springer.

Johnson, M. D., A. Acevedo and L. Mercado. 2016. Vocabulary knowledge and vocabulary use in second language writing. *TESOL Journal* 7(3), 700-715.

Kim, H. 2018. Working memory and language aptitude with focus on L2 vocabulary learning. In S. M. Reiterer, ed., *Exploring Language Aptitude: Views from Psychology, the Language Sciences, and Cognitive Neuroscience*, 57-71. Berlin: Springer.

Kormos, J. 2012. The role of individual differences in L2 writing. *Journal of Second Language Writing* 21(4), 390-403.

Laufer, B. and J. H. Hulstijn. 2001. Incidental vocabulary acquisition in a second language: The construct of task-induced involvement. *Applied Linguistics* 22(1), 1-26.

Lee, S. 2020. Examining the roles of aptitude, motivation, strategy use, language processing experience, and gender in the development of the breadth and depth of EFL learners' vocabulary knowledge. *SAGE Open* 10(4), 1-15.

Li, S. 2016. The construct validity of language aptitude. *Studies in Second Language Acquisition* 38(4), 801-842.

Li, S. 2018. Language aptitude. In A. Burns and J. Richards, eds., *The Cambridge Guide to Learning English as a Second Language*, 63-72. Cambridge: Cambridge University Press.

Lu, X. and J. Lim. 2021. Categories of word association in L2 mental lexicon of Chinese EFL learners. *Modern Studies in English Language & Literature* 65(1), 331-351.

Meara, P. M. 1983. Word associations in a foreign language: A report on the Birkbeck vocabulary project. *Nottingham Linguistics Circular* 11, 29-38.

Meara, P. M. 2005. *The LLAMA Tests*. Available at https://www.lognostics.co.uk/tools/LLAMA_3.

Meara, P. M. 2009. *Connected Words: Word Associations and Second Language Vocabulary Acquisition.* Amsterdam: John Benjamins.

Nation, I. S. P. 2013. *Learning Vocabulary in Another Language* (2nd ed.). Cambridge: Cambridge University Press.

Ortega, L. 2009. *Understanding Second Language Acquisition.* London: Routledge.

Peirce, J. W., J. R. Gray, S. Simpson, M. R. MacAskill, R. Höchenberger, H. Sogo, E. Kastman and J. Lindeløv. 2019. PsychoPy2: experiments in behavior made easy. *Behavior Research Methods* 51, 105-203.

PsychoPy. Release 2021.2.3. Available at http://github.com/psychopy/psychopy/releases/tag/2021.2.3

Rashidi, N. and N. Khosravi. 2010. Assessing the role of depth and breadth of vocabulary knowledge in reading comprehension of Iranian EFL learners. *Journal of Pan-Pacific Association of Applied Linguistics* 14(1), 81-108.

Read, J. 1988. Measuring the vocabulary knowledge of second language learners. *RELC Journal* 19(2), 12-25.

Read, J. 1993. The development of a new measure of L2 vocabulary knowledge. *Language Testing* 10(3), 355-371.

Robinson, P. 2005. Aptitude and second language learning. *Annual Review of Applied Linguistics* 25, 46-73.

Robinson, P. 2013. Abilities and aptitudes for second language learning and performance. *Kanto JACET Journal* 5, 1-5.

Rogers, V., P. Meara, T. Barnett-Legh, C. Curry and E. Davie. 2017. Examining the LLAMA aptitude tests. *Journal of the European Second Language Association* 1(1), 49-60.

Saito, K., Y. Suzukida and H. Sun. 2019. Aptitude, experience, and second language pronunciation proficiency development in classroom settings: a longitudinal study. *Studies in Second Language Acquisition* 41(1), 201-225.

Schmitt, N. 2000. *Vocabulary in Language Teaching.* Cambridge: Cambridge University Press.

Schneiderman, E. I. and C. Desmarais. 1988. The talented language learner: some preliminary findings. *Interlanguage Studies Bulletin (Utrecht)* 4(2), 91-109.

Şen, Y. and M. Kuleli. 2015. The effect of vocabulary size and vocabulary depth on reading in EFL context. *Procedia - Social and Behavioral Sciences* 199, 555-562.

Skehan, P. 2002. Theorising and updating aptitude. In P. Robinson, ed., *Individual Differences and Instructed Language Learning*, 69-93. Amsterdam: John Benjamins Publishing.

Skehan, P. 2016. Foreign language aptitude, acquisitional sequences, and psycholinguistic processes. In G. Granena, D. Jackson and Y. Yilmaz, eds., *Cognitive Individual Differences in L2 Processing and Acquisition*, 15-38. Amsterdam: John Benjamins Publishing.

Suárez, M. del M. and F. Gesa. 2019. Learning vocabulary with the support of sustained exposure to captioned video: Do proficiency and aptitude make a difference? *Language Learning Journal* 47(4), 497-517.

Sun, H., R. Steinkrauss., J. Tendeiro. and K. De Bot. 2015. Individual differences in very young children's English acquisition in China: Internal and external factors. *Bilingualism: Language and Cognition* 19(3), 550-566.

Unsworth, S., L. Persson, T. Prins and K. De Bot. 2014. An investigation of factors affecting early foreign language learning in the Netherlands. *Applied Linguistics* 36(5), 527-548.

Webb, S., Y. Sasao and O. Balance. 2017. The updated Vocabulary Levels Test: Developing and validating two new forms of the VLT. *International Journal of Applied Linguistics* 168(1), 33-69.

Wen, W. 2014. Assessing the roles of breadth and depth of vocabulary knowledge in Chinese EFL learners' listening comprehension. *Chinese Journal of Applied Linguistics* 37(3), 358-372.

Winke, P. 2013. An investigation into second language aptitude for advanced Chinese language learning. *The Modern Language Journal* 97(1), 109-130.

Yang, Y. and X. Cao. 2020. Effects of task involvement load on L2 vocabulary acquisition and their association with language aptitude. *The Asia-Pacific Education Researcher* 30(5), 421-430.

Yilmaz, Y. and Y. Koylu. 2016. The interaction between phonetic coding ability and feedback exposure condition. In G. Granena, D. Jackson and Y. Yilmaz, eds., *Cognitive Individual Differences in L2 Processing and Acquisition*, 303-326. Amsterdam: John Benjamins Publishing.

Examples in: English
Applicable Languages: English
Applicable Level: Tertiary

