Explaining L2 learners' listening comprehension: The role of aural vocabulary knowledge[1]

Yun Wang - Northwest A&F University
yunwang@nwafu.edu.cn

Rebut / Received: 27-01-22
Acceptat / Accepted: 29-05-22

Resumen. Explicando la comprensión auditiva de los estudiantes de una segunda lengua: el rol del conocimiento de vocabulario auditivo. El conocimiento del vocabulario juega un rol importante al momento de predecir la comprensión auditiva de los estudiantes de una segunda lengua. Por mucho tiempo se han utilizado pruebas escritas de vocabulario para medir la comprensión auditiva, mientras que un número limitado de estudios han investigado el rol del vocabulario auditivo para explicar el logro de la competencia auditiva en el segundo idioma. El presente estudio adopta una prueba auditiva de vocabulario para medir el conocimiento de vocabulario de los estudiantes de segunda lengua, y un ejercicio auditivo para medir su dominio de comprensión auditiva en el segundo idioma. Los resultados muestran que el conocimiento de vocabulario auditivo de los estudiantes de segunda lengua se correlaciona con su competencia auditiva en el segundo idioma. Si bien el conocimiento de vocabulario auditivo de los estudiantes en cada rango de frecuencia (de mil a cinco mil), y su conocimiento de vocabulario académico presentan una relación significativa con su nivel auditivo en el segundo idioma, la correlación más significativa se encuentra entre el conocimiento de vocabulario académico y las habilidades auditivas, seguida de la relación entre el conocimiento de vocabulario en el rango de frecuencia de cinco mil palabras y la capacidad auditiva. Con respecto a la contribución del conocimiento

1. I would like to thank the two anonymous reviewers as well as the Editor of *Sintagma* for their valuable comments and suggestions. This work was funded by the Ministry of Education of China (Grant No. 20YJA740043), Philosophy and Social Science Foundation of Shaanxi Province, China (Grant No. 2020K010), Social Science Association of Shaanxi Province, China (Grant No. 2021ND0388), Shaanxi Province Education and Science Foundation, China (Grant No. SGH20Y1036) and Northwest A&F University, China (Grant No. 2452021185; Grant No. JY2103200; Grant No. 2452022037; Grant No. 2017RWYB30).

de vocabulario auditivo al dominio de las habilidades auditivas en el segundo idioma, los resultados obtenidos en la prueba de conocimiento de vocabulario académico muestran que estos por sí solos explican el 22 % de los cambios en la competencia auditiva del segundo idioma, mientras que los puntajes obtenidos en la prueba de conocimiento de vocabulario en el rango de frecuencia de cinco mil palabras explican el 16,7 % de los cambios. Además, se encontró que el conocimiento de vocabulario académico de los estudiantes explica la singular diferencia por encima de su conocimiento de vocabulario en el rango de frecuencia de cinco mil palabras en la competencia auditiva del segundo idioma.

Palabras claves: *comprensión auditiva en el segundo idioma; conocimiento de vocabulario auditivo; nivel de frecuencia de palabras.*

Abstract. Explaining L2 learners' listening comprehension: the role of aural vocabulary knowledge. Vocabulary knowledge plays an important role in predicting L2 learners' listening comprehension. Written vocabulary tests have long been used to measure L2 learners' vocabulary knowledge in studies on listening comprehension, while a limited number of studies have investigated the role of aural vocabulary knowledge to explain success in L2 listening. The current study adopts an aural receptive vocabulary test to measure L2 learners' vocabulary knowledge and a listening task to measure their L2 listening comprehension proficiency. The results show that L2 learners' aural vocabulary knowledge significantly correlates with their listening proficiency. Although learners' aural word knowledge at each frequency level (from 1K to 5K) and their academic word knowledge were significantly correlated with L2 listening, the strongest correlation was between academic vocabulary knowledge and listening, followed by that between word knowledge at the 5K frequency level and listening. In terms of the contribution of aural vocabulary knowledge to L2 listening, the scores achieved for the academic word knowledge measure alone explain 22% of the variance in L2 listening, and the scores achieved for the 5K word frequency level measure alone explain 16.7% of such variance. Learners' academic vocabulary knowledge is found to explain unique variance over and above their vocabulary knowledge at the 5K word frequency level in L2 listening.

Keywords: *L2 listening comprehension; aural vocabulary knowledge; word frequency level.*

1. Introduction

Listening comprehension is an important skill in language acquisition (Vandergrift & Baker, 2015); indeed, L2 listening proficiency is claimed to be one of the most difficult skills for a learner to master, and one with which it is very difficult to make progress (Graham, 2011). This is a result of the fleeting nature of listening; the incoming speech signal arrives at a speed that the listener cannot control and the sounds may be very different from those with which the listener has L1 familiarity. In addition,

unlike readers of a text, listeners cannot simply return to previous language presented in the discourse. For this reason, without adequate levels of automaticity in language processing, listeners may miss words, or entire passages, which can compromise their understanding of the text. Thus, there is growing interest in the academic literature to disentangle the roles of different variables which impact listening with a view to making recommendations for improving teaching practice.

There is extensive evidence that vocabulary knowledge is essential for reading comprehension among L2 learners (Laufer & Ravenhorst-Kalovski, 2010; Masrai, 2019a; Mecartty, 2000). A better understanding of the role of vocabulary knowledge in explaining L2 listening comprehension (henceforth L2LC) is urgently needed. While the use of written vocabulary tests to measure L2 learners' vocabulary knowledge is common among studies on L2LC, Cheng and Matthews (2018) suggest that measures of aural vocabulary are more strongly linked to L2 listening than are written measures of vocabulary knowledge. A limited number of studies has explored the impact of aural vocabulary knowledge on L2LC, as well as its impact at different word frequency levels. Certainly, according to Field (2008), the fact that a listener can understand the meaning of a word in its written form does not mean they can recognize the same word in speech or comprehend its meaning. This research focuses on the gap in our understanding of the role played by aural receptive vocabulary knowledge in explaining L2LC by providing empirical evidence about the extent to which such knowledge can explain the variance in L2LC. It also addresses the contribution of L2 learners' knowledge at different word frequency levels and on academic words to explain L2LC. Thus, 187 L2 learners from two universities were investigated. Data were collected by measuring learners' L2LC and L2 vocabulary knowledge. The findings may cast new insights for researchers on the importance of L2 learners' aural vocabulary knowledge, and of their vocabulary knowledge at various word frequency levels, in determining successful L2LC. Therefore, the findings may have important implications for L2 listening pedagogy.

2. Aural measures of L2 vocabulary knowledge and L2 listening comprehension

Field (2013) put forward a model of lower-level processes in L2LC in which lexical knowledge is positioned as an important component in the listening process. According to Field (2013), listeners' lexical knowledge helps them make the best word match from what they hear. Empirical studies on L2LC have provided evidence for the importance of lexical knowledge. Previous studies (see Tables 1 & 2) show that the contribution of vocabulary knowledge in explaining the success of L2 listening varies from 13 % (Wang & Treffers-Daller, 2017) to 59 % (Cai, 2020). Such broad differences in the variance among studies may be related to the use of different measurements of learners' L2 vocabulary knowledge and L2LC, e.g., the use of written as opposed to aural tests,

or a focus on receptive as opposed to productive knowledge (Wang & Treffers-Daller, 2017).

Table 1 provides an overview of a number of studies which provide information about the correlations between L2 written vocabulary and listening and the proportion of the variance explained by L2 written vocabulary knowledge. Although written receptive vocabulary tests were found to work well when exploring the impact of learners' vocabulary knowledge in explaining L2LC in those studies, written tests do not directly tap into L2 listeners' ability to apply vocabulary knowledge to their recognition of words in aural input in real time (Matthews & Cheng, 2015). Therefore, research which explores L2 learners' ability to recognize vocabulary in the aural form is needed.

Table 2 provides an overview of studies which examine the contribution of L2 learners' aural vocabulary knowledge to L2LC proficiency. Among the studies, Milton, Wade, and Hopkins (2010) compared the role of learners' aural and written vocabulary size in L2 listening proficiency. They found a correlation of.67 between the scores on the measure of aural vocabulary knowledge and IELTS listening scores, this was stronger than the significant correlation (.48) between the scores on the measure of written vocabulary knowledge and IELTS listening scores. They also found that the measure of aural scores explained 44 % of variance in L2LC, while those measured on written vocabulary knowledge explained 51 % of variance in L2LC. The results indicate that higher correlation between vocabulary knowledge and L2LC does not necessarily mean that it has stronger explanatory power to L2LC. In another study, Masrai (2019b) also explored the contribution of different modalities of formats of vocabulary knowledge to predict L2LC. Although the correlation patterns were comparable to Milton et al. (2010), Masrai (2019b) showed that aural vocabulary knowledge was the strongest predictor of L2LC. Vandergrift and Baker (2015) investigated the cognitive learner variables which contribute to L2LC and found a substantial correlation of.51 between L2 vocabulary knowledge and LC. They found L2 vocabulary to be a stronger predictor of listening proficiency than any other variable, although they also found that L2 learners' working memory, auditory discrimination, metacognition, and L1 vocabulary contributed indirectly to L2LC via L2 vocabulary knowledge. The latter therefore remains the key explanatory variable in their provisional model. A comparable correlation between aural vocabulary knowledge and L2LC was found by McLean, Kramer, and Beglar (2015) who developed the Listening Vocabulary Levels Test (LVLT), based on Coxhead (2000), and used it to measure L2 learners' aural vocabulary knowledge. They found that L2 learners' scores on the test were significantly correlated (r =.54, p <.001) with their performance on the TOEIC listening section. In a later study, Cheng

and Matthews (2018) explored the relationship between three dimensions (receptive/orthographic, productive/orthographic, and productive/phonological) of vocabulary knowledge and both listening and reading comprehension. They found that among the three forms of vocabulary knowledge, productive phonological (aural) vocabulary knowledge is strongly correlated with listening (r =.71, p <.01), and productive written vocabulary knowledge and receptive written vocabulary knowledge were moderately correlated with L2 listening (r =.55,.39, p <.01). The study also showed that aural vocabulary knowledge explained 50.6% of the variance in listening, while written vocabulary knowledge contributed 1.2 % to the explanatory power of listening. Cai (2020) provided further evidence that both lexical and syntactic measures correlated moderately with L2LC and both measures made unique contributions to the variance in the listening score. Vafaee and Suzuki (2019) used structural equation modelling to explore the relative importance of vocabulary knowledge and syntactic knowledge in L2LC. Vafaee and Suzuki (2019) found that vocabulary knowledge was a stronger predictor with an effect size being almost twice as much as that of syntactic knowledge. Also using structural equation modeling, Wallace (2022) found that among other individual differences, EFL learners' domain-specific vocabulary knowledge is the most important predictor for their LC.

the 1K and 3K frequency levels together accounted for 54% of the variance in L2LC. In a follow-up study, Matthews (2018) not only provided further evidence that word knowledge at high frequency levels, 1-2000 and 2001-3000 words, strongly correlate with L2LC (r =.64,.66, p <.001,) and may predict L2LC, but also found that mid-frequency level of 3001-5000 words explained the greatest proportion of variance in L2LC across the three word frequency levels. Therefore, Matthews (2018) concludes that aural vocabulary knowledge for words throughout the mid- and high-frequency levels are significantly associated with L2LC. In a recent study, Lange and Matthews (2020) found that L2 learners' aural knowledge of vocabulary at the first 1K word level and lexical segmentation ability could predict 34 % and 38 % of the variance in scores of two listening tests.

This summary of the literature clearly shows that L2 learners' aural vocabulary knowledge is significantly correlated with L2 listening, and that it plays an important role in explaining variance in L2 listening. However, the importance of high-frequency vocabulary (0-3000 range) and mid-frequency vocabulary (3001-5000 range) is yet to be extensively explored. In addition, the literature is limited in relation to the relative importance of high- and mid-frequency aural vocabulary knowledge for L2LC. Thus, studies which focus on the impact on L2LC of aural vocabulary knowledge at different frequency levels are likely to be valuable in the field of L2 listening.

Table 1. Studies with a focus on the role of written vocabulary knowledge in listening comprehension in L2

Author	Students' L1	Target language	N	Vocabulary components tested (task used, written/aural)	Pearson correlation with listening	Percentage (%) of variance in L2 listening comprehension explained
Andringa et al. (2012)	35 different L1s	Dutch	113	Written receptive vocabulary size	.68	--
Mecartty (2000)	Unknown but not Spanish	Spanish	77	Vocabulary knowledge, Grammar knowledge	.38	14 from vocabulary knowledge
Staehr (2008)	Danish	English	88	Receptive vocabulary size	.69	39
Staehr (2009)	Danish	English	115	Written vocabulary size and depth	.70 (size) and.65 (depth)	51
Wang & Treffers-Daller (2017)	Chinese	English	151	Written receptive vocabulary size	.44	13

Table 2. Studies with a focus on the role of aural vocabulary knowledge in listening comprehension in L2

Author	Students' L1	Target language	N	Vocabulary components tested (task used, written/aural)	Pearson correlation with listening	Percentage (%) of variance in L2 listening comprehension explained
Bonk (2000)	Japanese	English	59	Lexical recognition	.45 (Kendall's tau)	23
Cai (2020)	Chinese	English	258	Aural receptive vocabulary knowledge	.77	59
Lange & Matthews (2020)	Japanese	English	130	Aural receptive vocabulary size, lexical segmentation	.15,.12 .39,.51	34 (TOEIC listening), 38 (Eiken Pre-2 listening)
Masrai (2019b)	Arabic, Brazilian, Chinese, Iranian, Japanese	English	130	Written and aural receptive vocabulary size	.59 (XK_Lex) .67 (A_Lex)	35 (XK_Lex) 45 (A_Lex)
Matthews (2018)	Chinese	English	247	Aural word recognition	.64 -.68 at different frequency levels	52.8
Matthews & Cheng (2015)	Chinese	English	167	Aural word recognition	.73	54
Milton et al. (2010)	A range, including Arabic, Chinese and Japanese	English	30	Written and aural receptive vocabulary size	.48 (X_Lex) .67 (A_Lex)	51 (X_Lex) 44 (A_Lex)
Vandergrift & Baker (2015)	English	French	157	Aural vocabulary size	.51	49 (path coefficient)
Vafaee & Suzuki (2019)	Iranian	English	263	Aural receptive vocabulary breath/depth test	.78 (breadth) and.80 (depth)	6
Wallace (2022)	Japanese	English	226	Aural receptive vocabulary size	.78	--

Sintagma 35, 103-118. ISSN 2013-6455 / DOI 10.21001/sintagma.2023.35.07

Differing from Milton et al. (2010) and Masrai (2019b), who compared the contribution of written/aural vocabulary knowledge, Matthews and Cheng (2015) explored the contribution of learners' ability to recognize 1K, 2K, and 3K high frequency word levels to the prediction of the proficiency of L2LC. They found learners' test scores for word knowledge at the 3K frequency level to be the strongest correlation with L2LC (r =.72, p <.01) among all frequency levels; with 52 % of the variance in L2 listening predicted by this variable alone, knowledge of words from the 3K frequency level are clearly the most important predictor. Indeed, scores on the word knowledge measure at

Research questions

This research aims to provide a clearer understanding of the correlation between L2 learners' aural vocabulary knowledge and their L2 listening. Furthermore, it explores the contribution of learners' ability to recognize high- and mid-frequency levels of aural words to predict listening proficiency. The research questions are as follows:

1. To what extent does learners' aural vocabulary knowledge correlate with L2LC among learners of English?
2. Regarding learners' aural vocabulary knowledge at different word frequency levels, which level has the strongest correlation with L2 LC?
3. To what extent does learners' aural vocabulary knowledge at each 1K frequency level contribute to explaining the variance in L2LC?

3. Method

3.1. Participants

The participants were 187 university-level Chinese learners of English (Male = 90, Female = 97) with varying learning backgrounds: 147 participants were undergraduates studying in a Chinese university, 40 participants were undergraduates studying in a UK university. Ranging from 18 to 23 years of age, they had been learning English for an average of 10.5 years. All participants took the Oxford Placement Test (OPT) listening and grammar sections (the OPT, Allan, 2004). On the CEFR scale, although the average score on the OPT of the participants studying in China was on the bottom line of the B2 level and the average score on the OPT of the participants in the UK corresponded with the top line of B2 level, generally their English proficiency was defined as varying from B1 to B2 level. Their proficiency was defined as such because according to the author's experience of teaching Chinese university undergraduates who share many characteristics with the participants in China, the output from university undergraduates in China on writing and speaking tasks can be mapped to approximately B1 level, rather than B2, although their reading level can be mapped to B2. For example, a common phenomenon among first-year undergraduates is that they used a double

object construction NP-V-NP-NP in their writing tasks, as in *He gave her a pen*. Such construction is considered a critical feature which appears from level B1 onwards, whilst learners from level B2 onwards often used object control sentences (NP-V-NP-AdjP), as in *He painted the car red* (Salamoura & Saville, 2010).

3.2. Instruments

Aural vocabulary knowledge test

The Listening Vocabulary Levels Test (LVLT, McLean et al., 2015) was used to measure participants' aural vocabulary knowledge. According to the designers, the test measures learners' English word knowledge from five 1000-word frequency levels and the Academic Word List (AWL). The test consists of six parts with 150 items in total. In each of the first five parts, 24 items are included which measure 1000 word frequency on one level. In the sixth part, 30 items are included which measure academic word knowledge. The LVLT items are chosen from the British National Corpus (BNC) / Corpus of Contemporary American English (COCA) list (Nation, 2012). According to McLean et al. (2015), the first five 1000-word frequency levels provide adequate coverage for listening across a wide range of genres and provide 96-97 % coverage of conversations (Nation, 2006). The AWL is included in the LVLT because it covers 10 % of tokens in academic texts and 4.41 % of academic spoken English.

In the LVLT, participants hear a word in a simple carrier sentence which provides the grammatical context to help listeners access the meaning of the target word. After hearing the target word and the carrier sentence, participants choose the answer which corresponds most closely to the target word in meaning. In its original form the LVLT was designed for Japanese L1 speakers, thus, for the current study of Chinese L1 speakers, a Chinese version answer sheet was adopted[2]. Each item has four answer options in Chinese to avoid a confusion between the measure of listeners' aural vocabulary knowledge and L2 reading ability (McLean et al., 2015). One example is shown below (English translations added for clarity):

1. (Test-taker hears: "time: They have a lot of time.")
   a. 钱 (money)
   b. 食物 (food)
   c. 时间 (time)
   d. 朋友 (friends)

2. I thank Dr. Stuart McLean very much for generously agreeing to share the LVLT test and the recordings designed by him and his colleagues and used in their studies.

Listening comprehension task

The listening section of the Cambridge Preliminary English Test (the PET) was used to measure L2 learners' listening proficiency. The PET is a comprehensive exam developed by Cambridge English Language Assessment (2014). According to the designer, the PET is suitable for language learners who need to communicate in English in a practical, everyday way. The test measures test-takers' skills in reading, listening, writing, and speaking. The PET listening section aims to assess test-takers' abilities to comprehend dialogues and monologues in both informal and neutral settings on a range of everyday topics. There are four parts to the listening section. Part 1 comprises seven short neutral or informal monologues or dialogues, each of which is followed by a three-option multiple-choice item with pictures. Part 2 is a longer interview, followed by six three-option multiple-choice items. Part 3 is a gap-fill task with six items for each of which test-takers write one or more words in the gap. Part 4 is a longer informal dialogue followed by six True/False items. This test, rather than other higher-level tests, was used for the study mainly because the majority of the participants were defined as approximately at B1 level on the CEFR scale.

3.3. Procedure

The data collection took place in two sessions. In the first session, the participants took the PET listening test. The test was administered in its paper-and-pen form in a language lab during the participants' after-class time. In accordance with the official PET administration procedures, the recordings were played only once. The test lasted 35 minutes, including six minutes given to transfer answers to an answer sheet. In the second session, the participants took the LVLT test. The test was also administered in its paper-and-pen form in a language lab. The test lasted 35 minutes.

3.4. Ethical considerations

Prior to data collection, the participants were informed of the research purpose, the tasks they would be asked to complete, and the time the tasks would be expected to take. They were told their data would only be used for academic research purposes and were assured that no real names would be used in any subsequent publications. All participants signed consent forms before they started the tasks.

3.5. Data analysis

The scores used in the analyses were percentages of correct responses in the PET listening test, in the vocabulary scale, and in each word frequency level and the AWL in the LVLT. Each correct response was awarded one point. There were 25 items for the PET listening test in total, 24 items for each word frequency level, and 30 items for the

AWL for the vocabulary scale. The reliability of the different instruments used in the study was found to be acceptable or high (Field, 2017), i.e., the Cronbach's alpha of the PET listening section was.75 and for the LVLT it was.86. The assumptions for carrying out a regression analysis were checked and revealed that the assumptions associated with multivariate normality, multicollinearity, and homoscedasticity were all met. To answer RQ1, the score on the vocabulary scale was correlated to the score on the listening test. To answer RQ2, the scores on each word frequency level of the vocabulary scale were correlated to the score on the listening test. To answer RQ3, the listening score was regressed on the scores of each word frequency level of the vocabulary scale.

4. Results

Table 3 presents descriptive statistics of the measures obtained from the five levels and the AWL of the LVLT, the total score of the LVLT and the listening comprehension test. As can be seen from Table 3, the mean score of the LVLT at the 1K frequency level was greater than those of the LVLT at the other frequency levels and the mean score of the AWL. What was unexpected was that the mean score of the LVLT at the 5K frequency level was higher than that at the 4K frequency level. A possible reason might be that the learners happened to have learned those words at the 5K frequency level and therefore knew them.

Table 3. Descriptive statistics of the vocabulary scale and the listening scale

	Mean (%)	**Maximum (%)**	**Minimum (%)**	**SD (%)**
1K frequency level	95.2	100	79	5.7
2K frequency level	79.6	100	54	12.9
3K frequency level	61.6	92	29	16.0
4K frequency level	51.6	79	25	14.1
5K frequency level	59.1	83	33	13.3
The AWL	57.9	93	23	17.8
The LVLT	67	90	44	11.3
The listening test	53.3	88	20	16.5

To answer RQ1, correlation analysis was undertaken. The results showed that the score for the LVLT significantly correlated with the score for the PET listening test (r =.47, p <.01). The results mean that learners' aural vocabulary knowledge has a significant correlation with their L2LC.

In order to answer RQ2, the correlations between the six sections of aural vocabulary knowledge in the LVLT and the PET listening were investigated. The results displayed in Table 4 reveal that the scores of the LVLT at the six levels significantly correlated with the dependent variable. The strongest correlation was between academic vocabulary knowledge and listening (r =.47, p <.01), followed by the correlation between word knowledge at the 5K frequency level and listening (r =.41, p <.01). The lowest significant correlation was between the score at the first 1K frequency level and the dependent variable (r =.18, p <.05). It is found that the correlation between learners' academic vocabulary knowledge and L2 listening and the correlations between learners' mid-frequency level words and L2 listening were medium, while the correlations between learners' high-frequency level words and L2 listening were small.

Table 4. Correlations between LC and the six sections of the LVLT

	1K frequency level	**2K frequency level**	**3K frequency level**	**4K frequency level**	**5K frequency level**	**The AWL**
The PET listening	.18*	.37**	.35**	.38**	.41**	.47**

*Note: **p <.01; *p <.05.*

In order to answer RQ3, hierarchical regression analyses were conducted with the vocabulary variables at different frequency levels. Based on the strength of the correlations between the scores on the PET listening and the scores on each section of the LVLT, we entered the scores at the five levels and the AWL section of the LVLT as six predictors. The results (see Table 5) indicate that in the first model (F(1, 185) = 52.08, p <.05), the score at the AWL explained 22 % of the variance in LC (β =.469, p <.05). When, in a second step, the score at the 5K frequency level was added, a further 2.4% of the variance was explained in the listening score. The addition of this variable made a significant change to the explained variance of the model (F(2, 184) = 29.59, p <.05). Then, when the score for each of the other four sections was added, respectively, to the model, the addition of each explained variance (together less than 1%) was very little and did not make a significant change to the model. In order to discover whether the score for the AWL explained any variance in listening comprehension over and above the contribution of the score at the 5K frequency level, we reversed the order of entering the two variables. In the second regression analysis, the score at the 5K frequency level was input first, followed by the score at the AWL. The results indicated that in the first model, the score at the 5K frequency level explained 16.7 % of the variance in LC (β =.408, p <.05). In a second step, when the score at the AWL was added, a further 7.7 % of the variance was explained. The addition of this variable made a significant change to the explained variance of the model (F(2, 184) = 29.59, p <.05). The results mean that,

individually, the score at the AWL and the score at the 5K frequency level are predictors of learners' L2LC. The results also mean that learners' academic vocabulary knowledge explains unique variance over and above their vocabulary knowledge at the 5K word frequency level in L2LC.

TABLE 5. REGRESSION MODELS EXPLAINING L2LC WITH THE SCORES OF THE LVLT AT THE AWL AND THE 5K WORD FREQUENCY LEVEL AS PREDICTORS

Model	*R*	*R squared*	*Adjusted R squared*	*Std. error of the estimate*	*R squared changed*	*Sig. F change*
1	*.469*[a]	*.220*	*.215*	*.15*	*.220*	*.000*
2	*.493*[b]	*.243*	*.235*	*.14*	*.024*	*.017*

[a] Predictors: (Constant), the scores of the AWL

[b] Predictors: (Constant), the scores of the AWL, the scores of the 5K word frequency level

5. Discussion

The first point of significance in this research is the overall association observed between the aural vocabulary test score and the listening comprehension score (r =.47, p <.01). The strength of association was lower than that found by Milton et al. (2010), (A_Lex), Cheng and Matthews (2018), Masrai (2019b), Cai (2020), Vafaee and Suzuki (2019), and Wallace (2022), but was higher than that found by Bonk (2000) and Lange and Matthews (2020). However, it is difficult to compare the strength of correlations because different tests were used to measure learners' aural vocabulary knowledge and their listening proficiency. This suggests that vocabulary tests which require recognition of words from speech reflect the constructs of word knowledge which are associated with higher-level listening proficiency.

Secondly, in terms of word knowledge at different frequency levels, of particular significance was the strength of association observed between the AWL scores, the 5K word frequency level scores, and L2LC (r =.47,.41, p <.01). Henceforth, the ability to recognize academic words and mid-frequency words from speech has good utility in the prediction of L2LC. For the participants in the current study, the greater their capability to recognize academic words and mid-frequency words from speech and to understand word meaning, the higher the probability that they could comprehend L2 speech input successfully. The findings suggest that L2 learners' ability to recognize academic words and mid-frequency words from speech and to understand word meaning is indispensable to successful L2LC. The findings of the present study differ from those of Matthews and Cheng (2015) who found high correlations between aural vocabulary knowledge for words at high frequency levels (from 1K to 3K) and L2LC. If their study had explored vocabulary knowledge on words at mid-frequency levels, the results may have been

different. In addition, the differences in the findings of the two studies might have been caused by the use of different measures of learners' vocabulary knowledge. That is, Matthews and Cheng (2015) adopted an aural productive vocabulary test which focused on measuring learners' word recognition and production, while the present study used an aural receptive vocabulary test which focused on measuring learners' meaning comprehension. In a further study, Matthews (2018) found that although L2 learners' aural vocabulary knowledge at the levels of 0 - 2000 words, 2001 - 3000 words, and 3001 - 5000 words contributed uniquely to the prediction of L2 listening for the entire cohort, only measures of the last two levels could uniquely predict L2 listening for relatively high proficiency participants. Findings from the current study partially support Matthews' (2018) work, however, both studies found that L2 learners benefit from possessing knowledge of words at mid-frequency level, i.e., the 5K frequency level, when comprehending L2 speech input. Although both the present study and Lange and Matthews (2020) adopted the same measure for learners' L2 vocabulary knowledge from 1K to 5K word frequency levels, the findings are different on the importance of high- and mid-frequency level words. The differences might be attributed to the measures for learners' listening proficiency and English proficiency.

6. Conclusion

This research provides further evidence for the fact that aural vocabulary knowledge has a significant correlation with L2 learners' LC and it is this knowledge which explains a proportion of the variance in L2 listening. It was also found that although learners' aural word knowledge at each frequency level from 1K to 5K and academic word knowledge were significantly correlated with L2LC, the strongest correlation was between academic vocabulary knowledge and listening, followed by the correlation between word knowledge at the 5K frequency level and listening. In terms of the contribution of aural vocabulary knowledge to L2LC, it was found that the scores on the academic word knowledge measure alone explain 22% of the variance in L2LC and the scores on the 5K word frequency level measure alone explain 16.7% of the variance in L2LC. Learners' academic vocabulary knowledge was found to explain unique variance over and above their vocabulary knowledge at the 5K word frequency level in L2LC.

The study has two limitations. The first is that only receptive aural vocabulary knowledge was assessed, but since listening and aural vocabulary are both receptive skills it makes sense to use a receptive task in such a study. Since Cheng and Matthews (2018) found that L2 learners' aural productive word knowledge was more important than their receptive aural word knowledge in predicting L2LC, further studies are needed to explore the role of receptive/productive aural word knowledge in the success of comprehending L2 speech input. The other limitation is that although the LVLT is a good test, there are some concerns around the construct validity of multiple-choice test formats, i.e., learners' vocabulary knowledge may be overestimated (Kremmel &

Schmitt, 2016; Stoeckel, McLean, & Nation, 2021). Therefore, researchers need to be careful when interpreting scores from vocabulary size tests (Kremmel & Schmitt, 2016). In addition to the meaning-recognition test formats, i.e., the LVLT, researchers may try alternative test formats to measure the construct of vocabulary tests. For example, the meaning-recall test formats are suggested by Stoeckel et al. (2021). In a recent study, Cheng, Matthews, Lange, and McLean (2022) found that a single word aural meaning-recall test could predict 38 % of the variance observed in L2 learners' listening scores, which provides evidence for the usefulness of the meaning-recall test formats in related studies. It is likely that the results of the current study could have been different if a meaning-recall test had been adopted. Comparatively, aural meaning-recall tests have the potential advantage on measuring lexical employability (Cheng et al., 2022) over meaning-recognition tests.

This research has some implications for teaching L2 listening. Since learners' academic word knowledge and knowledge for words at the mid-frequency level were found to contribute to predicting their listening proficiency, it is suggested that L2 teachers focus on improving learners' ability to recognize words from speech, especially the ability to recognize aural academic words and words at mid-frequency levels. Vocabulary knowledge plays an important role in determining the success of L2 learners' listening comprehension, but since listening comprehension involves language knowledge, world knowledge and cognitive abilities, future research would do well to focus on L2 listeners' phonological knowledge, topical knowledge, attention, and emotions while attempting listening comprehension.

References

Allan, D. (2004). *Oxford Placement Test 1*. Oxford University Press.

Andringa, S, Olsthoorn, N., Van Beuningen, C., Schoonen, R. y Hulstijn, J. (2012). Determinants of success in native and non-native listening comprehension: An individual differences approach. *Language Learning*, *62*, 49-78.

Bonk, W. J. (2000). Second language lexical knowledge and listening comprehension. *International Journal of Listening*, *14*, 14-31.

Cai, H. (2020). Related lexical and syntactic knowledge to academic English listening: The importance of construct representation. *Frontiers in Psychology*, *11*, 1-11. DOI: 10.3389/fpsyg.2020.00494.

Cambridge English Language Assessment. (2014). *The Preliminary English Test (PET)*. Cambridge University Press.

Cheng, J., & Matthews. J. (2018). The relationship between three measures of L2 vocabulary knowledge and L2 listening and reading. *Language Testing*, *35*, 3-25.

Cheng, J., Matthews, J., Lange, K., & McLean, S. (2022). Aural single-word and aural phrasal verb knowledge and their relationships to L2 listening comprehension. *TESOL Quarterly*. 1-29 DOI: 10.1002/tesq.3137.

Field, J. (2008). *Listening in the language classroom*. Cambridge: Cambridge University Press.

Field, J. (2013). Cognitive validity. In A. Geranpayeh, & L. Taylor (Eds.), *Examining listening: Research and practice in assessing second language listening* (pp. 77-151). Cambridge University Press.

Field, A. (2017). *Discovering statistics using IBM SPSS statistics*, 5th ed. London: SAGE.

Graham, S. (2011). Self-efficacy and academic listening. *Journal of English for Academic Purposes*, *10*, 113-117.

Kremmel, B., & Schmitt, N. (2016). Interpreting vocabulary test scores: What do various item formats tell us about learners' ability to employ words?. *Language Assessment Quarterly*, *13*(4), 377-392.

Lange, K., & Matthews, J. (2020). Exploring the relationships between L2 vocabulary knowledge, lexical segmentation, and L2 listening comprehension. *Studies in Second Language Learning and Teaching*, *10*(4), 723-749.

Laufer, B., & Ravenhorst-Kalovski, G. C. (2010). Lexical threshold revisited: Lexical text coverage, learners' vocabulary size and reading comprehension. *Reading in a foreign language*, *22*, 15-30.

Masrai, A. (2019a). Vocabulary and reading comprehension revisited: Evidence for high-, mid-, and low-frequency vocabulary knowledge. *SAGE Open,* 1-13. DOI: 10.1177/2158244019845182.

Masrai, A. (2019b). Exploring the impact of individual differences in aural vocabulary knowledge, written vocabulary knowledge and working memory capacity on explaining L2 learners' listening comprehension. *Applied Linguistics Review*, *10*(1), 1-25.

Matthews, J. (2018). Vocabulary for listening: Emerging evidence for high and mid-frequency vocabulary knowledge. *System*, *72*, 23-36.

Matthews, J., & Cheng, J. (2015). Recognition of high frequency words from speech as a predictor of L2 listening comprehension. *System*, *52*, 1-13.

McLean, S., Kramer, B. & Beglar, D. (2015). The creation and validation of a listening vocabulary levels test. *Language Teaching Research*, *19*, 741-760.

Mecartty, F. H. (2000). Lexical and grammatical knowledge in reading and listening comprehension by foreign language learners of Spanish. *Applied Language Learning*, *11*, 323-348.

Milton, J., Wade, J., & Hopkins, N. (2010). Aural word recognition and oral competence in a foreign language. In, R. Chacón-Beltrán, C. Abello-Contesse, C., & M. Torreblanca-López, *Further Insights into Non-native Vocabulary Teaching and Learning,* 83-98. Multilingual Matters.

Nation, P. (2006). How large a vocabulary is needed for reading and listening? *Canadian Modern Language Review*, *63*, 59-81.

Nation, P. (2012). The BNC/COCA word family lists (17 September 2012). Unpublished paper. Available at: www.victoria.ac.nz/lals/about/staff/paul-nation.

Salamoura, A., & Saville, N. (2010). Exemplifying the CEFR criterial features of written learner English from the English Profile Programme. In Bartning, I., Martin, M., & Vedder, I. (Eds.), *Communicative proficiency and linguistic development: Intersections between SLA and language testing research*. Eurosla Monographs Series 1, 101-132.

Staehr, L. S. (2008). Vocabulary size and the skills of listening, reading and writing. *The Language Learning Journal*, *36*, 139-152.

Staehr, L. S. (2009). Vocabulary knowledge and advanced listening comprehension in English as a foreign language. *Studies in Second Language Acquisition*, *31*, 577-607.

Stoeckel, T., McLean, S., & Nation, P. (2021). Limitations of size and levels tests of written receptive vocabulary knowledge. *Studies in Second Language Acquisition*, *43*(1), 181-203.

Vafaee, P., & Suzuki, Y. (2019). The relative significance of syntactic knowledge and vocabulary knowledge in second language listening ability. *Studies in Second Language Acquisition*, *42* (2), 383-410. DOI: 10.1017/S0272263119000676.

Vandergrift, L., & Baker, S. (2015). Learner variables in second language listening comprehension: An exploratory path analysis. *Language Learning*, *65*, 390-416.

Wallace, M. P. (2022). Individual differences in second language listening: Examining the role of knowledge, metacognitive awareness, memory, and attention. *Language Learning*. *72*(1), 5-44 DOI:10.1111/lang.12424.

Wang, Y., & Treffers-Daller, J. (2017). Explaining listening comprehension among L2 learners of English: the contribution of general language proficiency, vocabulary knowledge and metacognitive awareness. *System, 65*, 139-150.