兵庫教育大学　研究紀要　第58巻　2021年２月　pp.87-92

The Effect of Phonological Short-Term Memory on Japanese EFL Learners' Listening Skills

KONDO Akiko*

This study investigated whether phonological short-term memory (PSTM), as a language learning aptitude, is significantly related to the L2 listening skills of Japanese English as a Foreign Language (EFL) learners. To fill the present gaps in the literature, this study (1) investigated the effects of PSTM on a large sample of non-Indo-European language speakers, (2) considered the verbal and acoustic aspects of PSTM, (3) employed L2 listening variables, while excluding the effects of L2 vocabulary and grammar knowledge, and (4) identified the specific listening skills that PSTM capacity can predict. A total of 215 Japanese EFL learners participated. Their verbal short-term memory (VSTM) was measured using the Japanese-based Pseudoword Memory Span Test and their acoustic short-term memory (ASTM) using the Tonal Memory Span Test. Additionally, two separate instruments were used to measure their L2 listening skills: the English Sentence Listening Dictation Test and the Visualizing English Language Competency Test. The correlation and regression analyses revealed that both VSTM and ASTM had significant positive effects on L2 listening skills but ASTM's effects were limited to L2 sound recognition skills.

Key words : second language acquisition, individual differences, aptitude, L2 listening, Japanese EFL learners

1. BACKGROUND

In foreign or second language (L2) education, language learners' individual differences (ID) are important because of the differing abilities of learners to successfully adapt to or profit from instructions. Previous studies examining ID in second language acquisition (SLA) have focused not only on dynamic or changeable traits (e.g., motivation) but also on fixed traits (e.g., aptitudes), which are generally considered untrainable or unchangeable by instructional intervention. Therefore, language teaching practitioners have paid relatively little attention to aptitude (Skehan, 2002), despite the fact that aptitude-related ID factors significantly affect an individual's rate of SLA (Skehan, 1986; Wen & Skehan, 2011). Everyday observations of student performance in foreign language classes indicate that their efforts do not always correlate with their language skill achievement levels, suggesting that aptitudes play a significant role in language learning. In terms of students' acquisition of listening skills, many English as a Foreign Language (EFL) teachers have noticed the significant impact of ID on students learning those skills, and the influence of differing aptitudes. Thus, this study focuses on one of those aptitudes, phonological short-term memory (PSTM), because the ability to temporarily hold aurally presented sound information is an essential factor in processing the meaning of aural inputs.

In SLA research, despite the theoretical link between PSTM and listening, few studies have investigated the role of PSTM in learning a second language and the available results are inconsistent. For example, Vulchanova, Foyn, Nilsen, and Sigmundsson (2014) examined 48 Norwegian 10-year-old children and reported a significant correlation between PSTM and L2 listening skills, measured with a forward digit recall test and the English Listening Comprehension Test, respectively. Nonetheless, in their study of Hungarian school students, Kormos and Sáfár (2008) failed to find a significant relationship between PSTM and L2 listening skills, assessed with the L1-based pseudoword repetition test and the Cambridge First Certificate Exam, respectively. Furthermore, most PSTM studies have used relatively small samples of participants whose first language (L1) was Indo-European (Call, 1985; Kormos & Sáfár, 2008; Mizera, 2006; O'Brien, Segalowitz, Freed, & Collentine, 2007). As such, these study results cannot be applied with confidence to Japanese EFL learners. In addition, PSTM's non-verbal or acoustic aspects have not been addressed by previous researchers besides Call (1985)(Schulze & Koelsch, 2012; Pechmann & Mohr, 1992).

To address the research gap, Kondo (2019) investigated the effects of PSTM on the L2 listening skills of 223 Japanese EFL learners, a comparatively large sample in this field. The author demonstrated the positive influence of both verbal and acoustic aspects of PSTM on L2 sound recognition skills. However, the study focused only on sound recognition skills and did not examine PSTM's influence on comprehensive L2 listening skills. Additionally, in measuring L2 listening skills, previous researchers, including Kondo (2019), did not eliminate the influence of L2 vocabulary and grammatical

＊Hyogo University of Teacher Education, Graduate School,Advanced Professional Development in School Education,Language Studies. Associate Professor

令和2年10月22日受理

knowledge. Therefore, the present study aimed to investigate the effect of both verbal and acoustic aspects of PSTM on Japanese learners' L2 sound recognition and comprehensive listening skills, while eliminating the effects of L2 vocabulary and grammatical knowledge. To examine these factors, the following research questions were formulated:

1. Is verbal short-term memory (VSTM) related to Japanese EFL learners' L2 listening skills? If so, to what degree does VSTM capacity explain the variance in Japanese EFL learners' L2 listening skills, including sound recognition and comprehension?
2. Is acoustic short-term memory (ASTM) related to Japanese EFL learners' L2 listening skills? If so, to what degree does ASTM explain the variance in Japanese EFL learners' L2 listening skills, including sound recognition and comprehension?

2. METHOD

2.1 Participants

A total of 215 Japanese EFL learners from two institutions agreed to participate in this study. As they had similar demographic features (age, nationality, native language, and English learning experience), the two groups were merged, yielding a larger sample size and greater variance in terms of gender and academic major, to increase generalizability of the results. The participants were 18–23 years of age, their native language was Japanese, and their English proficiency was low- to high-intermediate, according to the placement test taken when they entered the school.

As student participation was voluntary, the candidates were selected based on availability. To recruit students, the study information was sent to students via the school e-mail system. Those who wished to participate registered on the study website and chose a day to take the tests. However, students with more than six months of experience in a foreign country were excluded from the study to control for the influence of a significant English learning experience. Finally, the participants received a reward of 5000 yen upon completing the study.

2.2 Instruments

2.2.1 Instruments used to measure L2 listening skills

To measure the participants' L2 (English) listening skills, two types of tests were adopted, as this study focused on both sound recognition and comprehension skills.

First, the English Sentence Listening Dictation Test (ESLDT) was developed to evaluate L2 sound recognition skills by measuring the ability to recognize words that appear in connected speech, such as contractions, reductions, linking, assimilation, and deletion. The items consisted of 18 English sentences borrowed from *A New Threshold for the TOEIC Test Listening* (Kurata, Thorpe, & Fujikura, 2007). Each sentence comprised 8 to 15 words from the first three 1,000-word frequency levels of the British National Corpus, which the target participants were expected to know. After participants listened to the instructions in Japanese and completed one practice item, they listened to each sentence, which was spoken twice at a natural speed with a 10-second pause in between, and were asked to write the entire sentence on the answer sheet. This test took approximately 15 minutes to complete. Two graduate students with high school English teacher licenses subsequently graded the participant responses.

Second, the listening section (VTL) of Kinseido's (English textbook publisher in Tokyo) Visualizing English Language Competency Test (VELC) was implemented to measure comprehensive L2 listening skills. This test was administered because it is a well-known method for measuring the range of participant proficiency (Kumazawa, Shizuka, Mochizuki, & Mizumoto, 2016). The VTL consists of three parts, each with 20 items (60 items total). In the first section, the test-takers must select the English word that corresponds to the Japanese word they hear. The answer is entered on a sheet that only displays item numbers and four letter options (no Japanese or English words). The second section measures the ability to recognize aspects of English sounds. The test-takers listen to an incomplete sentence and choose a word or phrase from four options to complete the sentence. The last section measures the ability to understand short sentences or passages and predict the storylines, which corresponds to what listeners often do in normal conversation. The test-takers listen to a sentence or short passage in which a beep replaces the last word. Then, they must choose the last word from four options based on the context of each sentence or passage. The VTL scores provided by the testing company were modified such that the mean score was 500 and the standard deviation was 100. This helped to indicate whether the score was higher or lower than the mean score of Japanese university students (Shizuka & Mochizuki, 2012).

2.2.2 Instruments used to measure VSTM and ASTM

To measure the participants' VSTM, a Japanese-based Pseudoword Memory Span Test (JPMST) was designed. Instead of real words, pseudowords were included to avoid the effect of familiarity or individual experiences.

Additionally, to avoid the influence of L2 proficiency, a total of 17 L1-based test items were designed, comprising two to five pseudowords that sound like Japanese words but have no meaning. After participants listened to the Japanese instructions and completed one example item, they listened to recordings on headphones and orally reproduced and recorded the pseudowords in the same order using a microphone. The study used a partial-credit scoring method in which each successfully reproduced phoneme was accorded a point. Thus, if the participants reproduced three phonemes, they were given three points, but if they only reproduced one of the three sounds, they received one point.

To measure the participants' ASTM, the study administered the Tonal Memory Span Test (TMST), designed to assess receptive memory spans for tonal sound information. The participants listened to a set of melodies played on a piano and marked which note was different on the answer sheet. This test consisted of 33 items and the sequences for each item were gradually lengthened from three to six notes. A correct answer received one point and an incorrect answer zero points, as there was only one correct answer for each item.

2.2.3 Instruments used to measure L2 vocabulary and grammatical knowledge

One significant feature of this study is its focus on L2 listening skills, which eliminates the variance caused by L2 vocabulary and grammatical knowledge. To measure the participants' knowledge of English vocabulary and grammar, the first and second parts of the VELC Test's Reading Section (VTR) were used. The first part (VTR1) measured written receptive English vocabulary knowledge and consisted of 20 items. From four options, the participants chose the English word that best corresponded to a Japanese word. Next, to measure participants' knowledge of English syntactic structure, the second part of the VELC Reading Section (VTR2), which also consisted of 20 items, was used. From four options, the participants selected the appropriate placement of the target word to complete an English sentence.

2.3 Data Collection and Analysis

The data were collected in a computer room that contained 50 computers. Using this location allowed the students to listen to sound files, record their voices, and create an mp3 file that included both the test input and responses. Prior to testing, the participants were provided general instructions, including the study purpose, payment information, and overall test schedule. The five tests were administered in the following order: VTL, VTR, ESLDT, TMST, and JPMST. Following data collection, a Rasch analysis was performed to determine instrument validity for the originally designed tests, such as the ESLDT, JPMST, and TMST. In addition, the Rasch person ability estimates were calculated to reveal item difficulties that raw scores simply cannot demonstrate. After the instrument validity analysis and data screening, the main analysis (correlation and regression analyses) was performed to answer the research questions.

3. RESULTS

3.1 Preliminary Analysis

Table 1 summarizes the Rasch analysis results for the three originally designed tests (ESLDT, JPMST, and TMST) used in this study. Overall, the tests appear to function efficiently.

Table 1. *Rasch Analysis Summary*

	ESLDT	JPMST	TMST
k	18	17	33
Items			
Separation	13.40	13.28	6.54
Reliability	.99	.99	.98
Person			
Separation	3.54	2.26	2.28
Reliability	.93	.84	.84
Unidimensionality			
Variance measure (%)	64.8	60.7	35.7
Variance items (%)	61.1	56.4	20.3
Eigenvalue unexplained 1st contrast	1.7	1.8	2.0

3.2 Creating L2 Listening Variables that Eliminate the Effects of L2 Vocabulary and Grammar Knowledge

To create the L2 sound recognition variable (the ESLDT Rasch person ability measure that controls for the influence of L2 vocabulary and grammar knowledge), the unstandardized residuals were computed with the regression model. These unstandardized residuals represented the dependent variable's [VTL] values after removing the value predicted by L2 vocabulary (VTR1) and grammar knowledge (VTR2). For the comprehensive L2 listening variables, as with the former L2 listening variable, the unstandardized residuals (VTL values after removing the value predicted by VTR1 and VTR2) were calculated using the regression model.

3.3 Descriptive Statistics

Table 2 illustrates the descriptive statistics for the four tests (ESLDT, VTL, JPMST, and TMST), including the mean, standard deviation, and the minimum and maximum scores. The two dependent variables (ESLDT-RES and VTL-RES) were the listening variables obtained after removing the variance explained by VTR1 and VTR2. The two independent variables (JPMST and TMST) were the Rasch person ability measure (logit).

Table 2. *Descriptive Statistics for the Four Tests*

Variables	*M*	*SD*	*Min*	*Max*
ESLDT-RES	.00	.06	-.16	.14
VTL-RES	.00	46.26	-114.43	93.56
JPMST	.00	.04	-.10	.10
TMST	.00	1.27	-3.32	3.34

Note. $N = 215$

3.4 VSTM Capacity and L2 Listening Skills

The first research question examined the relationship between Japanese EFL learners' VSTM capacities and L2 listening skills, and the degree to which VSTM capacity can explain the variance in learners' L2 listening skills. Regarding the L2 sound recognition skills, a correlation coefficient of .32 ($p < .001$) was obtained between the ESLDT residuals, after eliminating the L2 vocabulary and grammar knowledge effect (ESLDT-RES), and the Rasch person ability measure for the JPMST (M-JPMST). Next, a standard simple linear regression analysis was performed, with ESLDT-RES as the dependent variable and M-JPMST as the independent variable. The results showed that the R for regression was significantly different from zero, $R^2 = .10$, $F(1, 213) = 23.45$, $p < .001$. The R^2 value indicated that the VSTM capacity, determined by the JPMST, predicted 10% of the variability in the L2 listening skills. The regression coefficients for the JPMST differed significantly from zero ($\beta = .32$, $p < .001$).

The relationship between VSTM and comprehensive L2 listening skills, which eliminated the variance of L2 vocabulary and grammar knowledge, was also examined. The result showed a correlation coefficient of .16 ($p < .05$) between the VTL residuals, after eliminating VTR1 and VTR2 variance (VTL-RES), and the JPMST's Rasch person ability measure (M-JPMST). The regression analysis (VTL-RES as the dependent variable and M-JPMST as the independent variable) revealed that the R was significantly different from zero, $R^2 = .03$, $F(1, 213) = 5.56$, $p < .05$. The R^2 value indicated that the VSTM capacity predicted 3% of the variability in the L2 listening skills measured by ESLDT. Lastly, the regression coefficients for M-JPMST differed significantly from zero ($\beta = .16$, $p < .05$).

3.5 The ASTM Capacity and L2 Listening Skills

The second research question assessed the relationship between Japanese EFL learners' ASTM capacities and L2 listening skills, as well as the degree to which ASTM capacity can explain the variance in their L2 listening skills. Pearson's correlation analysis produced a correlation coefficient of .25 ($p < .001$) between the unstandardized residuals of the ESLDTs Rasch person ability measure, after removing the dependent variables, L2 vocabulary and grammar knowledge variance (ESLDT-RES), and the TMST's Rasch person ability measures (M-TMST). After confirming the significant correlation between the two variables, a standard simple linear regression was performed with the ESLDT-RES as the dependent variable and the M-TMST as the independent variable. The results showed that R was significantly different from zero, $R^2 = .06$, $F(1, 213) = 13.94$, $p < .001$. The R^2 value indicated that the TMST capacity predicted 6% of the variability in L2 listening skills indicated by the ESLDT-RES. The regression coefficients for M-TMST differed significantly from zero ($\beta = .25$, $p < .001$).

This study also examined the relationship between the L2 comprehensive listening skills, indicated by the VTL total scores after removing the L2 vocabulary and grammar knowledge variance, and the ASTM capacity. Pearson's correlation analysis revealed a .03 ($p = n.s.$) coefficient between the VTL's unstandardized residuals, after removing the variance for the dependent variable, and the TMST's Rasch person ability measures. As these two variables were not significantly correlated, the regression analysis was not performed.

4. DISCUSSION

4.1 L2 Listening and VSTM Capacity

The results showed that VSTM plays a significant role in explaining the individual variability in L2 sound recognition skills, supporting Kondo's (2019) findings. Even after eliminating the L2 vocabulary and grammatical knowledge variance, the effects of VSTM were significant—a result that greatly contributes to this field of research. For example, VSTM not only explained the sound recognition skills indicated by the ESLDT but also the variance in comprehensive L2 listening skills. However, the effect was smaller than for the L2 sound recognition skills, potentially due to the different test formats or the VTL's complex answering format. For instance, in the second part of the VTL,

test-takers listened to an English sentence and identified, from four options, the appropriate word to fill in the blank. This process is more complex than the ESLDT test format in which the participants listened to a recording and transcribed the sentence they heard. Even if they were able to recognize all of the aurally presented English words in the VTL, some test-takers may have made an error in determining word placement. Furthermore, the third part of the test required the ability to comprehend the content and infer what the speaker would say next, which involves higher cognitive abilities than sound recognition skills. Hence, memorizing verbal information may not have a strong enough influence on L2 listening skills, as suggested by Kormos and Sáfár (2008). These authors studied 121 Hungarian EFL learners' L2 listening skills, using the Cambridge First Certificate Exam (comprehensive test type) and revealed the significant effects of working memory, which involves higher cognitive processing than VSTM. To summarize, their study results showed that the effects of the VSTM were more salient in basic or primitive L2 listening skills (e.g., sound recognition). Although the present study demonstrated VSTM's statistically significant effects on comprehensive listening skills, its minimal effects on L2 listening skills (e.g., content comprehension) indicated that the latter require a higher level of cognitive ability than just identifying sound information.

4.2 L2 Listening and ASTM Capacity

The other innovative feature of the present study is its focus on not only the verbal aspects but also the acoustic aspects of PSTM. The ASTM explained the significant variance in L2 sound recognition skills, even after eliminating the L2 vocabulary and grammatical knowledge variance. Although the explained variance was small, the acoustic or non-verbal aspects notably influenced L2 listening skills. This implies that sensitivity to tonal sound information played a role in recognizing English sentences, or that test-takers used tonal information, such as intonation, when processing English at the sentence level, as required in the ESLDT tasks.

However, the results also showed that ASTM capacity was not significantly related to comprehensive listening skills, as indicated by the VTL. This result contradicts Call's (1985) study in which the tonal memory capacities of 41 Spanish- and Arabic-speaking students explained 4% of the variance in L2 listening skills measured with the Michigan Test of Aural Comprehension (MTAC). Although the L2 listening test used by Call was comprehensive, as was the VTL in the present study, the MTAC's test format differed from that of the VTL, which could explain the different results. Additionally, the difference in the participants' native language could also explain the contradictory results between studies. Moreover, while the ESLDT focused on sound recognition skills, the VTL, which consisted of three sections, required various types of skills. For example, in the third section, test-takers had to understand an entire passage and infer what the speaker would say next, which involved various abilities beyond simply recognizing the sound. To achieve this task, recognition of acoustic information expressed as intonation may play a role but the effect would not be large enough to be statistically significant. In addition, the test-takers focused on single words during the VTL's first and second tasks and, as a result, they may not have used as much tonal information as in the ESLDT task.

5. CONCLUSION

Although the explained variance of the regression analyses was rather small, this study nonetheless revealed that PSTM was significantly related to L2 listening, a relationship that has not been extensively researched thus far. The findings provided significant insight into these effects of PSTM in SLA through the study's large sample size and its demographic of non-Indo-European language speakers. Although the large sample increased the generalizability of the results, further research investigating different types of language speakers is necessary to conclude that PSTM's significant effects are significant regardless of the learners' native language.

Another contribution of this study was the investigation of PSTM effects on L2 listening skills, while eliminating the influence of learners' L2 vocabulary and grammar knowledge. Specifically, previous studies demonstrated that PSTM explained significant variance in L2 vocabulary and grammar knowledge (French, 2006; French & O'Brien, 2008; Hummel, 2009; Martin & Ellis, 2012), suggesting that these publications used the L2 listening variable without excluding the effects of this knowledge. Thus, as they did not show PSTM's effects on the specific skill factors included in the variable, the present study makes a meaningful contribution by focusing on listening-specific skills and excluding the influence of this L2 knowledge.

Furthermore, this study revealed the effects of PSTM's acoustic aspects on L2 listening. As both VSTM and L2 listening are similar in terms of verbal information processing, it is understandable that VSTM capacity had causal effects on L2 listening skills. However, the results remarkably demonstrated the role of acoustic or non-verbal aspects in L2 listening skills (in particular, sound recognition skills), which has neither drawn much attention nor been investigated in the

SLA research field.

Lastly, the pedagogical significance of the present study lies in its indication that PSTM significantly affects primitive language skills (e.g., sound recognition) more than complex language processing (e.g., inferring the next word based on the heard sentence). Sound recognition is the most basic skill that students need to acquire when beginning their language learning process. This implies that the ID in PSTM create larger gaps among the listening skills of beginners. Thus, L2 teachers should design listening activities for beginners by considering that the PSTM differences can partly explain the ID in L2 listening skills. However, to confirm this implication, future studies should compare how PSTM affects beginners and advanced learners.

REFERENCES

Call, M. E. (1985). Auditory short-term memory, listening comprehension, and the input hypothesis. *TESOL Quarterly, 19,* 765–781. doi:10.2307/3586675

French, L. M. (2006). *Phonological working memory and second language acquisition: A developmental study of Francophone children learning English in Quebec.* New York, NY: Edwin Mellen Press.

French, L. M., & O'Brien, I. (2008). Phonological memory and children's second language grammar learning. *Applied Psycholinguistics, 29*(3), 463-487.

Hummel, K. M. (2009). Aptitude, phonological memory, and second language proficiency in nonnovice adult learners. *Applied Psycholinguistics, 30*(2), 225-249. doi:10.1017/S0142716409090109

Kondo, A. (2019). The effects of verbal and acoustic short-term memory on Japanese EFL learners' sound recognition skills. *Language Education & Technology, 56,* 77–101.

Kormos, J., & Sáfár, A. (2008). Phonological short-term memory, working memory and foreign language performance in intensive language learning. *Bilingualism: Language and Cognition, 11,* 261–271. doi:10.1017/S1366728908003416

Kumazawa, T., Shizuka, T., Mochizuki, M., & Mizumoto, A. (2016).Validity argument for the VELC Test® score interpretations and uses. *Language Testing in Asia 6*(2), 1–18.

Kurata, M., Thorpe, T., & Fujikura, N. (2007). *A new threshold for the TOEIC test listening.* Tokyo, Japan: Seibido.

Martin, K. I., & Ellis, N. C. (2012). The roles of phonological short-term memory and working memory in L2 grammar and vocabulary learning. *Studies in Second Language Acquisition, 34*(3), 379-413. doi:10.1017/S0272263112000125

Mizera, G. J. (2006). *Working memory and L2 oral fluency* (Unpublished doctoral dissertation). University of Pittsburgh, Pittsburgh.

O'Brien, I., Segalowitz, N., Freed, B., & Collentine, J. (2007). Phonological memory predicts second language oral fluency gains in adults. *Studies in Second Language Acquisition, 29,* 557–582. doi:10.1017/S027226310707043X

Pechmann, T., & Mohr, G. (1992). Interference in memory for tonal pitch: Implications for a working-memory model. *Memory & Cognition, 20,* 314-320.

Schulze, K., & Koelsch, S. (2012). Working memory for speech and music. *Annals of the New York Academy of Sciences, 1252,* 229–236. doi:10.1111/j.1749-6632.2012.06447.x

Shizuka, T., & Mochizuki, M. (2012). VELC Test for testing competency: Verification of reliability and validity. Retrieved from http://www.velctest.org/contact/VelcTest-for-TestingCompetency.pdf.

Skehan, P. (1986). Cluster analysis and the identification of learner types. In V. Cook (Ed.), *Experimental approaches to second language learning* (pp. 81–94). Oxford, England: Pergamon.

Skehan, P. (2002). Theorising and updating aptitude. In P. Robinson (Ed.), *Individual differences and instructed language learning* (pp. 69-94), Philadelphia, PA: Benjamins.

Vulchanova, M., Foyn, C. H., Nilsen, R. A., & Sigmundsson, H. (2014). Links between phonological memory, first language competence and second language competence in 10-year‐old children. *Learning and Individual Differences, 35,* 87–95. doi:10.1016/j.lindif.2014.07.016

Wen, Z., & Skehan, P. (2011). A new perspective on foreign language aptitude research: Building and supporting a case for working memory as language aptitude. *A Journal of English Language, Literatures in English and Cultural Studies, 60,* 15–44.