Exploring the Relationship Between L2 Listening and Metacognition After Controlling for Vocabulary Knowledge

Matthew P. Wallace

University of Macau

Correspondence concerning this article should be addressed to Matthew P. Wallace, Department of English (E21-4068), Faculty of Arts and Humanities, University of Macau, Avenida da Universidade, Taipa, Macau, China. Email: mpwallace@um.edu.mo

Metacognition is known to be important for L2 listening comprehension. However, it is unclear how much variance in listening performance it can explain after controlling for vocabulary knowledge. To examine this, data from the listening section of the TOEFL Junior test, the Metacognitive Awareness Listening Questionnaire (MALQ), and the Listening Vocabulary Levels Test were collected from 76 high school EFL learners in Japan. The MALQ measured five subscales of metacognition representing metacognitive skills and metacognitive knowledge. Representing skills, the MALQ measured perceptions of the ability to (1) plan and evaluate performance, (2) direct attention, and (3) overcome listening problems. Representing knowledge, it measured strategic knowledge of (4) avoiding mentally translating speech and person knowledge of (5) maintaining positive attitudes about listening. The descriptive results showed that participants used their metacognition moderately. Of the subscales, they directed attention the most, planned and evaluated performance least, and perceived their ability to avoid mental translation, solve problems, and maintain optimism equivalently. The results from the hierarchical regression analysis further showed that vocabulary knowledge and metacognition overall predicted listening performance. Of the MALQ subscales, only person knowledge predicted comprehension. These findings indicate that, contrary to earlier findings, metacognition was important for listening comprehension after accounting for vocabulary knowledge.

Keywords: second language listening comprehension, metacognition, vocabulary knowledge, individual differences, hierarchical regression

Introduction

Metacognition, or the ability to mentally step away from what we are thinking about to observe and evaluate our thoughts (Vandergrift & Goh, 2012), is known to be important for listening comprehension. Researchers investigating metacognition and listening performance have taken an individual differences approach to examining their relationship and the findings have consistently shown that skilled listeners use more metacognitive resources when listening than their less-skilled counterparts (Goh & Kaur, 2013; Li, 2013; Vandergrift & Goh, 2012). This is because skilled listeners plan what they will do before listening, monitor their comprehension as they listen, and evaluate their performance afterwards; unskilled listeners, by contrast, may listen in an unplanned or incidental manner (Goh, 2008). Despite the importance of metacognition for listening, there remains a dearth of empirical studies examining the relationship when vocabulary is also taken into account. This is an important oversight because current language proficiency theory (e.g., Hulstijn, 2015, 2019) suggests that metacognition may only be important for comprehension after learners have exceeded the threshold of language knowledge needed to understand a spoken text. The current study aims to address this gap in the literature and contribute additional empirical evidence on the importance of metacognition for listening.

Literature Review

Metacognition and processing

Metacognition happens at the meta-level of processing. Nelson (1996) explains that when completing cognitive activities, like comprehending speech, information is processed on two levels—an object level and a meta level. Most cognitive activity is happening at the object level, and for listening comprehension, this involves linguistic processing through three stages—perception, parsing, and utilization (Anderson, 2005). Vandergrift and Goh (2012) explain that at perception, sounds in the speech stream are distinguished among one another. These sounds are then grouped into meaningful units in the parsing stage when lexical candidates in linguistic knowledge are matched with what was perceived. Existing knowledge (e.g., topical, discourse, pragmatic knowledge) is then applied to the newly parsed unit, which is combined with previously parsed units in the utilization stage. As utterances are processed through the three stages, a mental representation of the speech is generated for use in subsequent communicative acts (e.g., responding to speech).

Meta-level processing governs the object-level processing by (1) monitoring the status of object-level processing and (2) if problems occur, controlling object-level processing by directing resources to overcome the problems in order to achieve the task goal. Veenman, Van Hout-Wolters, and Afflerbach (2006) explain that processing at the meta-level draws upon metacognitive knowledge and skills. Metacognitive knowledge refers to the interaction among person, task, and strategy knowledge during a cognitive activity. Person knowledge is what a person knows of his/her own abilities and limitations. Task knowledge is what they know about the nature and demands of tasks. Strategy knowledge is what they know about the strategies that are available and how to use them to accomplish tasks (Flavell, 1979). Metacognitive skills draw upon these three knowledge bases to regulate learning activities and solve problems as they arise by planning, monitoring, and evaluating performance. Metacognitive knowledge and skills are developed in specific contexts or domains and progress as domain-specific knowledge is increased (Pressley & Gaskins, 2006). Only after knowledge of that context or domain becomes proceduralized do metacognitive knowledge and skills transfer to other domains (Borkowski, Chan, & Muthukrishna, 2000). For language learning, this would mean that metacognitive skills for second language listening tasks develop as learners increase their knowledge of the second language and that these skills would not be useful for other language learning domains (e.g., reading) until the learners had acquired some degree of procedural knowledge of the language.

Metacognition and second language listening

Listening comprehension is defined as the ability to accurately identify information explicitly provided within speech and to make inferences based on that explicitly provided information (Wagner, 2004) in order to accomplish implicit (not consciously aware of or deliberately striving to achieve) and/or explicit listening goals. Hulstijn (2015, 2019) explains that language proficiency and, by extension, language performance are influenced by the interactions among core and peripheral variables. Core variables consist of linguistic knowledge (e.g., vocabulary, syntactic, and phonological knowledge) and access to that knowledge, while peripheral variables consist of non-linguistic knowledge (e.g., topical knowledge), metacognition, and general cognitive abilities (e.g., working memory). The core variables are stronger predictors of listening comprehension outcomes than the peripheral variables, but the peripheral variables can become important for listening as linguistic knowledge increases. This is consistent with Borkowski et al.'s (2000) notion that the amount of domain-specific knowledge may affect how much influence metacognition has on task performance. They explain that metacognition does not become helpful for cognitive activities (i.e., listening comprehension) until some domain-specific knowledge is acquired. This means that for listeners with limited linguistic knowledge, metacognition would not be helpful for overcoming problems and completing comprehension tasks because listeners are busy making sense of what they hear at the early stages of language processing (i.e., perception and parsing). This would ultimately leave little cognitive space remaining for metacognitive resources to be employed.

Borkowski et al. (2000) also explain that metacognition may not be needed if students have enough knowledge to render task completion very easy. This means that for language learners with sufficient linguistic knowledge to comprehend all of the information in speech, metacognition may not be needed because comprehension tasks can be completed by drawing solely on this knowledge. This all indicates that in order for metacognition

to be helpful for listening comprehension, listeners need enough linguistic knowledge to understand some or most of the speech, but not so much that they understand all of it. Metacognition can serve in a supportive role to domain-specific knowledge. When listeners have insufficient linguistic knowledge to understand everything they hear, they utilize their metacognitive resources to help overcome these deficiencies and complete comprehension tasks (Vandergrift & Baker, 2015). The empirical literature seems to support this notion (Vandergrift & Baker, 2015, 2018; Wallace, 2020), a review of which follows below.

Metacognitive Awareness Listening Questionnaire

Research studies that have examined the relationship between metacognition and listening comprehension have used the Metacognitive Awareness Listening Questionnaire (MALQ; Vandergrift, Goh, Mareschal, & Tafaghodtari, 2006) to measure metacognition. The MALQ is a 21-item survey designed to elicit listener perceptions of five dimensions of metacognition that represent metacognitive knowledge and skills: Directed Attention, Problem Solving, Planning and Evaluation, Mental Translation, and Person Knowledge. Directed Attention items represent metacognitive monitoring and elicit perceptions of how well listeners regain their attentional focus if they lose it during a listening event. A common problem that weaker listeners experience is focusing too narrowly on a word or phrase that they mishear and ignoring the remainder of the speech. Skilled listeners, on the other hand, constantly monitor their comprehension as they listen and redirect their attention to comprehend the input as needed to accomplish their listening goals.

Problem Solving items also represent metacognitive monitoring and elicit how well listeners overcome problems as they arise during a listening event. One problem that weaker listeners experience is to assume that what they hear is consistent with their expectations of the speech's content, thus leading to miscomprehension or confusion regarding the speaker's message. Skilled listeners ensure their understanding is consistent throughout a listening event by systematically pausing to check whether their understanding is plausible based on their existing knowledge of the topic and what was said earlier in the speech. If their interpretations do not make sense, then skilled listeners will adjust their interpretations. Another problem that weaker listeners face is misunderstanding a word or stretch of speech and being stuck on working out the meaning. Skilled listeners overcome this potential problem by using their existing knowledge and/or the context provided in the speech to help them guess at the meaning of the unknown words.

Planning and Evaluation items represent planning and evaluation metacognitive skills and their connection with metacognitive knowledge. These items elicit perceptions of how well listeners plan what they will listen to before they start listening and evaluate their performance afterwards. Skilled listeners make predictions about an upcoming speaking event by thinking about (1) what they know of the topic, (2) the language they may encounter on that topic, (3) how the text may be structured, (4) how to accomplish the upcoming task, and (5) what the goal of listening will be (Vandergrift et al., 2006). Bringing this information into conscious thought makes it easier to parse what is said and then utilize it to demonstrate comprehension because it is activated before listening and does not need to be retrieved from long-term memory during a listening event. Skilled listeners will also reflect on and evaluate their performance after listening with the aim of identifying listening problems they encountered and planning how to overcome them moving forward. Reflecting on their performance increases listeners' metacognitive knowledge. They become more knowledgeable about (1) their limitations as listeners (person knowledge), (2) how to accomplish those kinds of tasks (task knowledge), and (3) how effective the strategies were that were employed (strategic knowledge).

The final two dimensions represent metacognitive knowledge. Mental Translation items elicit perceptions of how frequently listeners translate what they hear. Avoiding mentally translating speech is a strategy used by skilled listeners because translating while listening consumes valuable cognitive resources that can be devoted to other cognitive activities during a listening event (e.g., making connections between information in the input and existing topical knowledge; answering comprehension questions). Finally, the Person Knowledge dimension elicits attitudes toward listening in the target language. Skilled listeners tend to hold positive attitudes toward listening, which helps give them confidence to perform well on listening tasks and reducing their anxiety when completing them.

Empirical studies

The empirical literature has shown that despite being important for comprehension, language learners may not use very much of their metacognition when they listen. Studies examining metacognition and listening comprehension using the MALQ have shown that learners reported using a moderate amount of metacognition. For example, university language learners in Singapore (Goh & Hu, 2014), Iran (Bozorgian, 2014), and China (Li, 2013), and senior high school students in Jordan (Al-Alwan et al., 2013) reported a moderate amount of metacognition when listening (below 4.0 on a 6.0 Likert scale). Of the subscales measured on the MALQ, those representing the metacognitive skills, and monitoring in particular, are more frequently reported by learners than metacognitive knowledge. Goh and Hu's (2014) high-intermediate Chinese EFL learners in Singapore reported that directed attention and problem solving were the two most frequently used skills while listening, and planning-evaluation was more frequently used than either the person knowledge or mental translation subscales representing metacognitive knowledge.

Similar findings were reported for Li's (2013) low-level Chinese university EFL students and Al-Alwan et al.'s (2013) intermediate-level Jordanian high school EFL students. The Problem Solving, Directed Attention, and Planning and Evaluation subscales were more frequently reported than Mental Translation and Person Knowledge in these studies. One explanation for these findings may be that the learning contexts of these studies focused on improving metacognitive skills and did little to increase metacognitive knowledge. Granted, only two aspects of metacognitive knowledge were accounted for in these studies, but they do demonstrate that learners tend to mentally translate as they listen and struggle with maintaining a positive attitude toward listening in English, despite their level of proficiency. Another explanation may be that the learners may not have had sufficient listening experiences needed to increase their metacognitive knowledge. Metacognition is known to increase as learners gain more learning experiences and that their metacognition can transfer across domains. For example, if learners develop the ability to re-direct their attention if they lose it while completing a math problem, then they may transfer this to listening in a second language. If this holds for language learners, the results of existing studies suggest that, for intermediate-level listeners and below, metacognitive skills may be considered domain-general and more easily transferrable to and from other domains. However, metacognitive knowledge may be domain-specific; and in order to increase metacognitive knowledge for second language listening, learners would need more second language listening experience.

The empirical literature has also shown that metacognition shares a positive relationship with listening comprehension. In several studies, scores from the MALQ correlated with listening comprehension scores (e.g., Al-Alwan, Asassfeh, & Al-Shboul, 2013; Li, 2013; Vandergrift & Baker, 2015, 2018; Wang & Treffers-Daller, 2017). This means that listeners who have more metacognition tend to comprehend speech better than those with less metacognition. However, the relationship between metacognition and listening comprehension may not be as direct as these findings indicate. When overall metacognition was examined in isolation of any other variable, it was shown to be predictive of listening comprehension. For example, results from Vandergrift et al.'s (2006) regression analysis showed that scores on the MALQ were predictive of listening test scores for language learners with varied L1 backgrounds and L2 proficiency levels. However, vocabulary knowledge was not controlled as a variable in the study and a growing trend in the empirical literature is that metacognition does not predict comprehension when vocabulary knowledge is taken into account. For example, Vandergrift and Baker's (2015, 2018) path analysis showed that MALQ scores did not have a direct effect on listening comprehension scores, but that scores from vocabulary knowledge test did for lower-level teenage French language learners. Similarly, Wang and Treffers-Daller (2017) reported that only language proficiency and vocabulary knowledge test scores were predictive of listening comprehension test scores, but metacognition scores were not for Chinese university EFL students of varied proficiency levels. Vocabulary knowledge in each of these studies was defined as vocabulary size, or the number of unique words the participants knew. Overall, these results indicate that having greater knowledge of target language vocabulary was more important for listening performance than having greater metacognition resources.

When the subscales of the MALQ were examined for their direct relationship with listening, mixed findings were reported. Al-Alwan et al. (2013) reported that metacognitive skills were most important for comprehension. Scores from all three skills subscales of the MALQ (problem solving, directed attention, planning and evaluation) were predictive of listening test scores. For higher-level listeners, Goh and Hu (2014) reported that only Person Knowledge scores and Problem Solving scores were predictive of comprehension scores. In both of

these studies, Problem Solving was the strongest predictor of the subscales for Listening Performance, suggesting that listeners who were able to overcome their problems as they listened performed better on the listening tests. However, when vocabulary knowledge was controlled for in the analysis, only person knowledge scores were reported to be predictive of listening comprehension scores (Wang & Treffers-Daller, 2017). This is an interesting finding, considering that person knowledge is consistently reported least often by participants in varied contexts across proficiency levels. Wang et al.'s findings suggest that listening success depends mostly upon the vocabulary size of the listeners, but maintaining a positive attitude toward listening (person knowledge on the MALQ) can also explain some of the variance in listening performance. These studies demonstrate that in order to examine metacognition in the context of listening comprehension, vocabulary knowledge, as one indicator of linguistic knowledge, should be controlled.

The current study

The current study aims to examine the relationship among second language listening comprehension, metacognition, and vocabulary knowledge. Data from the study comes from a larger study that examined the relationship among listening comprehension, domain-specific knowledge, and domain-general cognitive abilities (Wallace, 2020). In the larger study, metacognition was found to have an indirect relationship with comprehension through domain-specific topical knowledge. However, it was unclear how the dimensions of the MALQ may predict comprehension after controlling for vocabulary knowledge. This study aims to address this limitation and provide a more in-depth understanding of how metacognitive knowledge and skills may have influenced listening comprehension for lower-level senior high school students. Specifically, the study intends to (a) explore the self-reported levels of metacognition among Japanese high-beginner L2 listeners as reported on the MALQ, (b) examine intrapersonal differences among the subscales of the MALQ for these listeners, and (c) examine the degree to which overall metacognition and the subscales of the MALQ predict L2 listening performance after controlling for vocabulary knowledge.

Methods

Participants

A convenience sampling method was used to recruit participants from two second-year senior high school classes in Japan. In total, 76 students elected to participate in the study. The participants studied in the top language program in the school, where they received up to 10 hours of English instruction per week. Curricular aims of the program were to develop linguistic proficiency for success on high-stakes university entrance exams, which consist mainly of reading, vocabulary, and grammar activities. The predominant teaching methodology for six of the 10 hours focused on increasing vocabulary and grammatical knowledge through the translation of written English texts to Japanese. The remaining four classes aimed to improve writing and speaking skills. Listening was not given much attention in this context. The proficiency level of these students was high-beginner, around the Common European Framework of Reference (CEFR) A2 level. This was later confirmed by the descriptive statistics showing that the participants scored an average of 60% on the listening test measuring CEFR levels A2-B1.

Instruments

L2 listening comprehension test
L2 listening was measured using a practice version of the TOEFL Junior Standard test. The test was designed to measure L2 listening ability for learners ranging in the CEFR A2-B1 levels. The test consisted of monologues and dialogues among students and teachers in an academic setting. The first half of the test consisted of single items associated with a single audio track (17 items), while the second half consisted of multiple items (3-5 items) per listening track. The participants indicated their comprehension by answering multiple-choice questions with four answer options. The questions and answer choices were available in written form while the tracks were played. The items examined the ability to identify the main ideas and details of academic and non-academic texts, make inferences based on speaker intonation, comprehend idiomatic language, and understand the discourse functions of a text (Educational Testing Service, 2018). One point was assigned for each correctly answered item. Correct items were totaled to represent listening performance.

Metacognition questionnaire

Metacognition was measured with the Metacognitive Awareness Listening Questionnaire (MALQ), an instrument widely used to examine L2 listeners' metacognition (Vandergrift et al., 2006). The 21-item questionnaire measured five dimensions of metacognition representing metacognitive skills and metacognitive knowledge (see Table 1). The questionnaire was translated into Japanese to ensure the participants could clearly understand the items. It was back translated into English to ensure the translation was accurate before being administered. The questionnaire was given immediately after the listening test to best capture listening perceptions. Six items on the MALQ were reverse-coded. Three items representing the mental translation subscale were reversed because avoiding translating while listening represents better listening behavior. Two Person Knowledge items and a Directed Attention item were also reversed because they were negatively worded. Overall and scaled item groups were formed to represent the MALQ scores and subscales, respectively. All items were averaged together to represent a metacognition overall value for analysis. Items representing each subscale were averaged together to form five scales for analysis.

Table 1

Metacognitive knowledge and skills, MALQ subscales, and number of items per subscale

	Metacognitive skills			Metacognitive knowledge	
Skill/knowledge type	Monitoring	Monitoring	Planning/ evaluation	Strategy knowledge	Person knowledge
MALQ subscale	Directed Attention	Problem Solving	Planning & Evaluation	Mental Translation	Person Knowledge
Number of items	4	6	5	3	3

Vocabulary knowledge test

Vocabulary knowledge was measured using the Listening Vocabulary Levels Test (McLean, Kramer, & Begler, 2015), a paper-based auditory vocabulary size test developed specifically for Japanese EFL learners. The test consisted of six sections. The first five sections (24 words each) consisted of the first five-thousand most frequently occurring words in English based on the British National Corpus and Corpus of Contemporary American English. A final section (30 words) measured knowledge of words extracted from the Academic Word List (Coxhead, 2000). Participants heard English words followed by a sentence not indicating its meaning. They matched the English words with their Japanese equivalent from four answer options. One point was assigned for each correctly answered item. Correct items on the test were totaled to represent vocabulary knowledge.

Data collection procedures

The listening test was administered in its paper-and-pencil form after school in the students' classrooms. The participants heard the recordings once and responded to the questions in 40 minutes. Immediately after the test, the participants completed the MALQ within 15 minutes and were directed to focus their reflections on how they listened during the listening test. The vocabulary test was administered on a different day. The participants heard the recordings once and responded to the items in 50 minutes. The responses to the listening and vocabulary tests were scored by hand once by the researcher and again by a research assistant. The test scores and MALQ responses were input into Microsoft Excel and imported into SPSS for analysis.

Data analysis and screening

The scores used in the analyses were the correct responses on the listening test and vocabulary test, the average scores of all of the MALQ items, and the average scores for items representing each subscale of the MALQ. The maximum score was 40 for the listening test, 6 for the MALQ overall score and each subscale score, and 150 on the vocabulary test. Overall, there were eight variables included in the analyses: Listening Performance, Metacognition Overall, Directed Attention, Mental Translation, Person Knowledge, Planning and Evaluation, Problem Solving, and Vocabulary Knowledge. Reliability estimates were calculated first to ensure the instruments were internally consistent. Cronbach's alpha values above .60 are considered acceptably reliable in applied linguistics research (Dornyei, 2007).

To answer research question one, descriptive statistics were calculated to determine how much metacognition the participants reported having. To answer research question two, a one-way within-subjects analysis of variance was conducted to determine significant differences among the subscales of the MALQ. To answer research question three, Pearson product moment correlation coefficients were first calculated for the listening test scores and the MALQ scores to establish a linear relationship among the variables. Only those relationships that were statistically significant were included in the subsequent hierarchical multiple regression analyses. Two hierarchical regression analyses were then conducted. In the first analysis, a Listening Performance variable was regressed onto a Vocabulary Knowledge variable in the first step. In the second step, the Listening Performance variable was regressed onto the Metacognition Overall variable. The second analysis followed the same procedure as the first hierarchical regression analysis except that in the second step, the Listening Performance variable was regressed onto variables representing each of the MALQ subscales. The analyses were conducted using SPSS version 24.

Before conducting the analyses, the assumptions underlying the statistical procedures were verified. In line with Goh and Hu (2014), the sample size was evaluated for its adequacy in detecting a medium-sized effect between predictor and outcome variables using Green's (1991) formula: $N \geq (8/f2) + (m-1)$. In the formula, N is the number of participants, m is the number of predictor variables, and f 2 is the effect size. A medium effect size of 0.13 was entered into the formula. According to the formula, the minimum sample size needed to detect a medium effect would be 63 for the first regression analysis and 66 for the second analysis. This indicates that the sample size of 76 in the current study was adequate to detect a medium-sized effect.

The data subsequently underwent several analyses recommend by Field (2009) to verify that it met the assumptions underlying the procedures. To confirm the assumption of univariate normality, descriptive statistics were inspected. Skewness and kurtosis values outside the absolute value of 2.0 were considered outliers and removed from the dataset. It has been argued that outliers may not need to be removed from a dataset (e.g., Larson-Hall & Herrington, 2010). However, alternative methods proposed to address the outlier issue also call for data to be removed, albeit in a different way. Therefore, three outliers were removed, accounting for less than 4% of the data. For multivariate normality, Mahalanobis distances were inspected for the predictor variables. Those significant at 0.001 were removed from the analysis. To test the assumption of sphericity for the analysis of variance, Mauchly's Test of Sphericity was conducted. If the assumption is violated, the Greenhouse-Geisser correction would be estimated. Assumptions underlying regression were also verified for both regression analyses. To check the assumption of independent observations, a Durbin-Watson statistic was calculated. Values below 1.0 and above 3.0 would violate this assumption. To verify the absence of multicollinearity, a variance inflation factor was calculated. Values below 0.1 would indicate multicollinearity among the predictor variables. The test whether the residuals were normally distributed, the Shapiro-Wilk test was conducted. If the results of the test are not significant, the residuals would be normally distributed. Finally, homoscedasticity was confirmed by running a Breush-Pagan test. If the results of the test showed the predictor variables not affecting the residual values, the assumption would be met.

Results

Table 2 reports the mean, standard deviation, skewness, kurtosis, and reliability coefficient (Cronbach's alpha) for each of the measures. The Cronbach's alphas ranged from .63 to .85. Most alphas were above .70, indicating good internal consistency. The Planning and Evaluation subscale (.63) was below .70, though it is acceptably reliable (Cohen, Manion, & Morrison, 2011; Dornyei, 2007). The lower alpha value was likely due to the limited sample size. The absolute value of the skewness and kurtosis estimates were within 2.0, suggesting that the scores were approximately normally distributed.

Table 2

Descriptive statistics and reliability estimates of measures (N = 76)

Variables (total items)	M	SD	Min.	Max.	Skewness	Kurtosis	alpha
Listening Performance (40)	23.58	7.01	11.00	39	0.248	-0.968	.85
Metacognition Overall (6)	3.90	0.61	1.95	5.5	-0.039	0.956	.84
Directed Attention (6)	4.40	0.81	1.50	6.0	-0.593	1.690	.72
Mental Translation (6)	3.94	1.08	1.00	6.0	-0.273	-0.064	.73
Person Knowledge (6)	3.92	1.08	1.00	6.0	-0.273	-0.064	.70
Planning and Evaluation (6)	3.26	0.83	1.60	5.8	0.280	0.439	.63
Problem Solving (6)	4.08	0.84	2.00	5.7	0.006	-0.432	.72
Vocabulary Knowledge (150)	100.91	9.84	75.00	132	0.618	1.190	.74

Results from analysis of variance

To answer the second research question, a one-way within-subjects analysis of variance was calculated. The results from Mauchly's Test of Sphericity indicate that the assumption was violated. Therefore, the Greenhouse-Geisser correction was estimated. The results show that there were statistically significant differences among the MALQ subscales, $F (3.12, 234.09) = 21.44$, $p < .001$. Post hoc tests using Bonferroni correction revealed that Directed Attention ($M = 4.40$, $SD = 0.81$) was more frequently reported than the other four subscales on the MALQ and the differences were statistically significant. Problem Solving ($M = 4.08$, $SD = 0.84$) was reported more frequently than Mental Translation ($M = 3.94$, $SD = 1.08$) and person knowledge ($M = 3.92$, $SD = 1.08$), but the differences among the three subscales were not statistically significant. Planning and Evaluation was reported less frequently than the other four subscales ($M = 3.26$, $SD = 0.83$) and the differences were statistically significant.

Results from correlation analysis

Bivariate correlations were calculated for the variables. The assumptions underlying correlation (e.g., absence of outliers, interval-scaled data, linearity, normal distribution) were found to meet the requirements for the procedure. The correlations and their associated confidence intervals are presented in Table 3.

Table 3

Bivariate correlations among the observed variables (N = 76)

	1	2	3	4	5	6	7
Listening Performance	1						
Metacognition Overall	.350** [.13, .57]	1					
Directed Attention	.233* [.01, .46]	.789** [.65, .93]	1				
Mental Translation	.302** [.08, .52]	.428** [.22, .64]	.239* [.01, .46]	1			
Person Knowledge	.385** [.17, .60]	.673** [.50, .84]	.455** [.25, .66]	.335** [.12, .55]	1		
Planning and Evaluation	.049 [-.18, .28]	.618** [.44, .80]	.348** [.13, .57]	-.046 [-.28, .19]	.195 [-.03, .42]	1	
Problem Solving	.260* [.04, .49]	.821** [.69, .95]	.634** [.46, .81]	.113 [-.12, .34]	.403** [.19, .62]	.429** [.22, .64]	1
Vocabulary Knowledge	.484** [.28, .69]	.219 [-.01, .45]	.165 [-.06, .39]	.263* [.04, .49]	.145 [-.09, .37]	.045 [-.19, .28]	.151 [-.08, .38]

Notes. **$p<.01$,*$p<.05$; 95% confidence intervals in brackets

The results in Table 3 show that the Listening Performance variable shared a moderate correlation with the Metacognition Overall variable (r = .350, 95% CI = .13, .57) and the Vocabulary Knowledge variable (r = .484, 95% CI = .28, .69). Each variable representing the subscales of the MALQ correlated with the Listening Performance variable except for the Planning and Evaluation variable (r = .049, 95% CI = -.18, .28).

Results from regression analyses

To answer the third research question, two hierarchical regression analyses were conducted. The data met the assumptions underlying the two regression analyses (Metacognition Overall and MALQ subscales). In both analyses, the Durbin-Watson statistic was between 1.0 and 3.0, the variance inflation factor was between 0.1 and 2.0 for each variable for each analysis, the Shapiro-Wilk test was not significant, and the Breush-Pagan test results showed that no predictor variable had a significant effect on the residual values.

In the first analysis, the Vocabulary Knowledge variable was added in the first step and the Metacognition Overall variable was added in the second step. The regression model was significant, R^2 = .297, adjusted R^2 = .278, F (1, 73) = 6.505, p < .01. The Vocabulary Knowledge variable (β = .428, 95% CI: 0.16, 0.45) and Metacognition Overall variable (β = .256, 95% CI: 0.65, 5.25) were significant predictors of the Listening Performance variable. The Vocabulary Knowledge variable explained 22.4% of the variance in Listening Performance scores and the Metacognition Overall variable explained an additional 5% of variance. The results of the analysis are presented in Table 4.

Table 4

Hierarchical regression predicting listening performance from vocabulary knowledge and the MALQ subscales (N = 76)

	B	*SE B*	*β*	*sr²*	*R²*	*adjusted*	*Δ R²*
Step 1					.235	.224	
Vocabulary Knowledge	0.35	0.07	.484*** [0.20, 0.49]				
Step 2					.297	.278	.05
Vocabulary Knowledge	0.31	0.07	.428*** [0.16, 0.45]	.418			
Metacognition Overall	2.95	1.26	.256** [0.65, 5.25]	.250			

Notes. ***p <.001, **p <.01, 95% confidence intervals in brackets

In the second analysis, variables representing the subscales of the MALQ replaced the Metacognition Overall variable. Because the Planning and Evaluation variable failed to share a linear relationship with the Listening Performance variable, it was removed from the analysis for violating the linearity assumption underlying regression. The data met the other assumptions underlying regression analysis. The regression model was significant, R^2 = .351, adjusted R^2 = .304, F (5, 70) = 7.557, p < .001. The Vocabulary Knowledge variable was a significant predictor of the Listening Performance variable (β = .410, 95% CI: 0.15, 0.44) in the model. Of the subscales added in the second step, only the Person Knowledge variable (β = .273, 95% CI: 0.29, 3.27) was a significant predictor of the Listening Performance variable. The Vocabulary Knowledge variable explained 22.4% of the variance in Listening Performance scores and the Person Knowledge variable explained an additional 8% of variance. The results of the analysis are presented in Table 5.

Table 5

Hierarchical regression predicting listening performance from Vocabulary Knowledge and MALQ subscales variables (N = 76)

	B	SE B	β	sr²	R²	adjusted	Δ R²
Step 1					.235	.224	
Vocabulary Knowledge	0.35	.072	.484*** [0.20, 0.49]				
Step 2					.351	.304	.08
Vocabulary Knowledge	0.29	.072	.410*** [0.15, 0.44]	.392			
Directed Attention	-0.47	1.14	-.045 [-2.73, 1.80]				
Mental Translation	0.62	0.70	.096 [-0.77, 2.01]				
Person Knowledge	1.78	0.75	.273* [0.29, 3.27]	.230			
Problem Solving	1.10	1.12	.131 [-1.12, 3.33]				

Notes. ***p <.001, *p <.05, 95% confidence intervals in brackets

Discussion

Metacognition overall and the MALQ subscales

The descriptive statistics show that the students reported having a moderate amount of metacognition. This means that they used their metacognitive skills or metacognitive knowledge some of the time during the listening test. These findings are consistent with earlier studies showing that second language learners in varied contexts and of wide-ranging proficiency levels use a moderate amount of metacognition when listening (e.g., Al-Alwan et al., 2013; Goh & Hu, 2014; Li, 2013). A common explanation offered for this result is that pedagogical approaches for listening do not focus on improving metacognition, despite the reported importance it has on learning in general and for improving listening comprehension more specifically. In many language learning contexts, the comprehension approach, which refers to having students repeatedly listen to audio texts and answer comprehension questions (Field, 2008) is used overwhelmingly. While this approach gives learners practice at listening in the target language and perhaps training in improving the ability to answer questions eliciting a certain type of listening (e.g., listening for details), it does little to help improve the ability to plan before a listening task, monitor comprehension, or evaluate performance afterwards; nor does it help listeners use their strategic knowledge or person knowledge to benefit their understanding of the target language speech. For the present study, it is possible that the participants may not have received that much listening instruction. Most of their pedagogical focus is oriented on the written word through reading, grammar, and vocabulary, and the spoken input they encounter is limited to hearing their classroom teachers speak.

The Directed Attention variable was the most frequently reported on the MALQ. This means that during the listening test, the participants monitored their attentional focus and redirected it to best complete the listening task more consistently than other metacognitive skills or knowledge. This result is consistent with Goh and Hu (2014), who reported that Directed Attention was the most frequently reported subscale of the MALQ for their high-intermediate Chinese EFL students. In their study, Directed Attention scores were statistically equal to the Problem Solving scores, suggesting that their listeners reported using their monitoring skills more frequently than any other subscale of the MALQ. Problem Solving was the second highest reported subscale in the current study, but the difference was statistically significant from the Directed Attention subscale. This means that participants were more aware of how to direct their attentional resources than they were in using their existing linguistic knowledge and listening experience to overcome their comprehension problems.

The Problem Solving variable was statistically equal to the two metacognitive knowledge variables, Mental Translation and Person Knowledge. This means that participants used their problem-solving skills, strategic knowledge, and person knowledge equivalently during the listening test. These findings were somewhat unexpected since earlier studies have reported that metacognitive skills, and monitoring in particular, were used more frequently than metacognitive knowledge (e.g., Al-Alwan et al., 2013; Goh & Hu, 2014). These results may be due to the strong connection with knowledge among each of these subscales. When overcoming listening problems, participants used their linguistic knowledge and listening experience to help them. When the participants reported on their strategic knowledge for the Mental Translation variable, they drew upon their knowledge for how to avoid translating speech as they listened; and they used what they knew of themselves as listeners (Person Knowledge variable), particularly about maintaining a positive attitude, when listening. These three MALQ subscales drew upon different knowledge bases than either the Directed Attention subscale, which is involved with sensory and perception processes, or the Planning and Evaluation subscale, which is involved with goal setting and achievement processes. Altogether, these results indicate that the participants used their perceptive processes most frequently as they listened and drew upon their knowledge resources to a lesser extent.

The most surprising result was that the Planning and Evaluation variable was the least frequently reported. This unexpected result differs strikingly from earlier studies showing that Planning and Evaluation was among the most frequently reported subscale of the MALQ (e.g., Al-Alwan et al., 2013; Goh & Hu, 2014; Goh & Kaur, 2013; Li, 2013). It means that the participants did not frequently plan how they were going to listen before the listening tasks began or evaluate their performance afterwards, and suggests the participants were unskilled listeners. Skilled listeners plan their approach to an upcoming listening task in order to best prepare them for achieving their implicit or explicit listening goals. By not doing this, listeners run the risk of continuing to struggle with listening tasks because they come into the listening event cold and they may miss key details of the speech while familiarizing themselves with what is being discussed. Skilled listeners also appraise their performance after listening to identify their listening strengths and weaknesses and to consider how to adjust their approach in subsequent listening tasks. Not doing this may cause them to make the same mistakes repeatedly because they are not learning from their errors. It is also interesting that the participants infrequently planned for or evaluated the listening tasks because the test-centered nature of the learning context nearly ensures that they have considerable experience in learning test-taking strategies, like previewing the questions and answer options before a task begins.

The results show that the participants utilized the metacognitive skills to different degrees during the listening test—they monitored frequently, but they only sometimes planned and evaluated their performance. An explanation for this result may be that the students had not received much explicit instruction on how to plan and evaluate their listening performance (Goh & Hu, 2014). When instructional focus is given to developing metacognitive skills and knowledge, students reported using them frequently during listening tasks (Fahim & Fakhri Alamdari, 2014; Graham & Macaro, 2008; Rahimirad & Shams, 2014). However, the Japanese EFL curriculum concentrates on improving reading comprehension and grammatical accuracy, and increasing vocabulary knowledge (Nishino & Watanabe, 2008). Strategic knowledge is addressed in the form of preparing learners to complete tasks similar to those they would encounter on high-stakes language assessments, such as university entrance exams.

L2 listening, metacognition overall, and the MALQ subscales

Two hierarchical regression analyses were conducted to examine the relative contributions that scores from the MALQ overall and scores from the subscales of the MALQ made on listening comprehension scores after controlling for vocabulary knowledge. Results from the first analysis showed that the Vocabulary Knowledge variable had a moderate effect on the Listening Performance variable and that the Metacognition Overall variable made a weak contribution beyond that explained by Vocabulary Knowledge. This means that vocabulary knowledge was the most important for listening, but that metacognition also contributed to listening performance, albeit to a lesser degree. This finding is consistent with the Core-Peripheral language proficiency model (Hulstijn, 2015, 2019) showing that core linguistic knowledge and processing variables are more important for language performance than peripheral variables like metacognition. However, this finding is inconsistent with previous studies showing that when vocabulary knowledge is controlled for that metacognition does not explain variances in listening. For example, metacognition was not predictive of

listening comprehension when Wang and Treffers-Daller (2017) included vocabulary knowledge and language proficiency variables into their regression model. Similarly, Vandergrift and Baker (2015; 2018) controlled for vocabulary knowledge for teenage French L2 learners and reported that metacognition scores were not predictive of comprehension scores. One explanation for this may be that the vocabulary level of the participants in the current study was adequate for them to process the speech at lower levels (perception and parsing), but not so high that they understood all of the language on the listening test. Because processing was not consumed at these lower levels, the participants would have had enough cognitive resources to devote to their metacognition. This explanation aligns with Borkowski et al. (2000), who claimed that metacognition becomes important for performance after some domain-specific knowledge has been acquired. In the current study, this domain-specific knowledge is target language vocabulary.

The results from the second hierarchical regression analysis showed that only the Person Knowledge variable was predictive of the Listening Performance variable. This means that after controlling for vocabulary size, the ability to maintain a positive attitude about listening in the target language had a direct effect on comprehension. This finding is consistent with Wang and Treffers-Daller (2017), who also reported that Person Knowledge scores were the only MALQ subscale to predict comprehension scores after controlling for vocabulary knowledge. Wang and Treffers-Daller argued that their results reinforce the importance of high self-efficacy and low anxiety for language performance and similar claims can be made about the current study's findings. Of the five subscales measured by the MALQ, Person Knowledge may be the only one that is considered an affective variable, or one related to feelings of self-esteem. The items measuring person knowledge ask how listeners feel about listening in the target language, which differs from how items measuring the other four subscales elicit perceptions of what listeners actually do during a listening event. The other four variables may then be considered cognitive in nature, meaning that they involve how listeners process information, and not how they regulate their feelings about listening. The results showing that the Person Knowledge variable was the only subscale to predict the Listening Performance variable suggest that the affective aspect was more important than the cognitive components of metacognition for the participants in this study. This result is unsurprising considering the importance that language anxiety has been shown to play in language learning in general, and for second language listening in particular (Elkhafaifi, 2005). To address this in the classroom, language teachers may consider including methods to reduce anxiety when listening in the target language within their lessons. Doing so will likely lead to improved listening performance.

Another explanation may be that the cognitive components of the MALQ become more important for listening when listeners have larger vocabulary levels for the given listening tasks. Goh and Hu (2014) reported that Problem Solving scores were the strongest predictor of listening comprehension test scores for high/intermediate-level Chinese EFL learners, and that Person Knowledge scores also predicted comprehension scores. Al-Alwan et al. (2013) showed that scores from the three subscales of the MALQ representing metacognitive skills (problem solving, directed attention, and planning-evaluation) were the only predictors of comprehension scores for low-intermediate-level EFL learners. The participants in both of those studies were reportedly higher than those of the current study, which may explain why the strongest predictors were the cognitive subscales of the MALQ in those studies. However, because vocabulary knowledge was not controlled for in those studies, this claim is speculative. To test the claim's veracity, future research should compare the relationships among the MALQ subscales and listening comprehension between high- and low-proficiency learners after controlling for vocabulary knowledge.

Conclusion

The results showed that the metacognitive monitoring skills were more frequently reported than the planning and evaluation skills and metacognitive knowledge. The results further showed that the Vocabulary Knowledge variable was the strongest predictor of the Listening Performance variable, but that the Metacognition Overall variable and the Person Knowledge variable, in particular, also predicted Listening Performance scores beyond what was explained by the Vocabulary Knowledge variable. These findings support the Core-Peripheral model of language proficiency, which posits that core linguistic knowledge and processing variables are most important for language performance, but that peripheral, non-linguistic variables can influence performance to a lesser degree. The findings also provide tentative support for Borkowski et al.'s (2000) claims that having some domain-specific knowledge is necessary for metacognitive resources to be engaged during a task. Among

the MALQ subscales, the one associated with anxiety (Person Knowledge) had the strongest effect on Listening Performance, highlighting the importance of affect for listening comprehension.

The study is not without its limitations, though. The narrow range of language proficiency levels among the participants limits the degree to which the findings can be generalized to the greater language learner population. Further, the use of correlational analytical methods limits the degree to which causality among the variables can be established. Although this study continued a tradition in the empirical literature with its approach, future studies may consider utilizing a longitudinal design to observe changes among the relationships over time. Adding a qualitative measure (i.e., diaries) would also be advantageous for this purpose. Future studies may also consider broadening the scope of metacognition to include task knowledge and metacognitive experience. Metacognitive experience, which refers to the previous experiences listeners have had using their metacognitive knowledge to accomplish listening tasks, has yet to be adequately measured in the empirical literature, despite its reported importance for task performance. Future studies are encouraged to consider developing a tool to measure metacognitive experience to allow for observations between it and listening comprehension to be made.

References

Al-Alwan, A., Asassfeh, S., & Al-Shboul, Y. (2013). EFL learners' listening comprehension and awareness of metacognitive strategies: How are they related? *International Education Studies*, *6*, 31-39.

Anderson, J. R. (2005). *Cognitive psychology and its implications* (6 ed.). New York: Worth Publishers.

Borkowski, J. G., Chan, L. K., & Muthukrishna, N. (2000). A process-oriented model of metacognition: Links between motivation and executive functioning. In G. Schraw & J. C. Impara (Eds.), *Issues in measurement of metacognition* (pp. 1-41). Lincoln, NE: Buros Institute of Mental Measurements.

Cohen, L., Manion, L., & Morrison, K. (2011). *Research methods in education*. New York: Routledge.

Coxhead, A. (2000). A new academic word list. *TESOL Quarterly*, *34*, 213-238. https://doi.org/10.2307/3587951

Dornyei, Z. (2007). *Research methods in applied linguistics*. New York: Oxford University Press.

Elkhafaifi, H. (2005). Listening comprehension and anxiety in the Arabic language classroom. *The Modern Language Journal*, *89*, 206-220. https://doi.org/10.1111/j.1540-4781.2005.00275.x

ETS. (2018). *Handbook for the TOEFL Junior tests*. Retrieved from http://www.ets.org/s/toefl_junior/pdf/toefl_junior_tests_handbook.pdf

Fahim, M., & Fakhri Alamdari, E. (2014). Maximizing learners' metacognitive awareness in listening through metacognitive instruction: An empirical study. *International Journal of Research Studies in Education*, *3*, 79-91. https://doi.org/10.5861/ijrse.2014.762

Field, A. (2009). *Discovering statistics using SPSS* (3rd ed.). London: Sage.

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

Flavell, J. H. (1979). Metacognition and cognitive monitoring: A new area of cognitive-developmental inquiry. *American Psychologist*, *34*, 906-911. https://doi.org/10.1037/0003-066X.34.10.906

Goh, C. C. M. (2008). Metacognitive instruction for second language listening development: Theory, practice and research implications. *RELC Journal*, *39*, 188-213. https://doi.org/10.1177/0033688208092184

Goh, C. C. M., & Hu, G. (2014). Exploring the relationship between metacognitive awareness and listening performance with questionnaire data. *Language Awareness*, *23*, 255-274. https://doi.org/10.1080/09658416.2013.769558

Goh, C. C. M., & Kaur, K. (2013). Insights into young learners' metacognitive awareness about listening. *The European Journal of Applied Linguistics and TEFL*, *2*, 5-26.

Graham, S., & Macaro, E. (2008). Strategy instruction in listening for lower-intermediate learners of French. *Language Learning*, 58(4), 747-783. https://doi.org/10.1111/j.1467-9922.2008.00478.x

Green, S. B. (1991). *How many subjects does it take to do a regression analysis? Multivariate Behavioral Research*, *26*, 499-510. https://doi.org/10.1207/s15327906mbr2603_7

Hulstijn, J. H. (2015). *Language proficiency in native and non-native speakers: Theory and research*. New York: John Benjamins Publishing Company.

Hulstijn, J. H. (2019). An individual-differences framework for comparing nonnative with native speakers: Perspectives from BLC Theory. *Language Learning*, *69*, 157-183. https://doi.org/10.1111/lang.12317

Larson-Hall, J., & Herrington, R. (2010). Improving data analysis in second language acquisition by utilizing modern developments in applied statistics. *Applied Linguistics*, *31*, 368-390. https://doi.org/10.1093/applin/

amp038

Li, W. (2013). A study of metacognitive awareness of non-English majors in L2 listening. *Journal of Language Teaching and Research*, *4*, 504-511.

McLean, S., Kramer, B., & Begler, D. (2015). The creation and validation of a listening vocabulary levels test. *Language Teaching Research*, *19*, 1-20. https://doi.org/10.1177/1362168814567889

Nelson, T. O. (1996). Consciousness and metacognition. *American Psychologist*, *51*, 102-116. https://doi.org/10.1037/0003-066X.51.2.102

Nishino, T., & Watanabe, M. (2008). Communication-oriented policies versus classroom realities in Japan. *TESOL Quarterly*, *42*, 133-138. https://doi.org/10.2307/40264432

Pressley, M., & Gaskins, I. W. (2006). Metacognitively competent reading comprehension is constructively responsive reading: How can such reading be developed in students? *Metacognition and Learning*, *1*, 99-113. https://doi.org/10.1007/s11409-006-7263-7

Rahimirad, M., & Shams, M. R. (2014). The effect of activating metacognitive strategies on the listening performance and metacognitive awareness of EFL students. *International Journal of Listening*, *28*, 162-176. https://doi.org/10.1080/10904018.2014.902315

Vandergrift, L., & Baker, S. C. (2015). Learner variables in second language listening comprehension: An exploratory path. *Language Learning*, *65*, 390-416. https://doi.org/10.1111/lang.12105

Vandergrift, L., & Baker, S. C. (2018). Learner variables important for success in L2 listening comprehension in French immersion classrooms. *Canadian Modern Language Review*, *74*, 79-100. https://doi.org/10.3138/cmlr.3906

Vandergrift, L., & Goh, C. C. M. (2012). *Teaching and learning second language listening: Metacognition in action*. New York: Routledge.

Vandergrift, L., Goh, C. C. M., Mareschal, C. J., & Tafaghodtari, M. H. (2006). The metacognitive awareness listening questionnaire: Development and validation. *Language Learning*, *56*, 431-462. https://doi.org/10.1111/j.1467-9922.2006.00373.x

Veenman, M. V., Van Hout-Wolters, B. H., & Afflerbach, P. (2006). Metacognition and learning: Conceptual and methodological considerations. *Metacognition and Learning*, *1*, 3-14.

Wagner, E. (2004). A construct validation study of the extended listening sections of the ECPE and MELAB. *Spaan Fellow Working Papers in Second or Foreign Language Assessment*, *2*, 1-26.

Wallace, M. P. (2020). Individual differences in second language listening: Examining the role of knowledge, metacognitive awareness, memory, and attention. *Language Learning*, Advance online publication. https://doi.org/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. https://doi.org/10.1016/j.system.2016.12.013