Content list available at http://ijltr.urmia.ac.ir

Iranian Journal of Language Teaching Research

ORIGINAL ARTICLE

Urmia University

Examining the Effects of Task Conditions on L2 Listening Performance: A Language Alignment Model Perspective

Matthew P. Wallace [a, *], Yihan Zhang [b], Meilin Li [c], Xioalin Zhu [c], Don Yao [d]

[a] *University of Macau, Macau, China*

[b] *University College London, UK*

[c] *University of Hong Kong, China*

[d] *China Jiliang University, Hangzhou, China*

ABSTRACT

This study examined the effects of task conditions on L2 listening performance. According to the Performance System of the Language Alignment Model, performance quality is influenced by the alignment between environmental demands (e.g., text and task characteristics) and individual abilities (e.g., proficiency level). From this perspective, listening support facilitates comprehension because it lowers the environmental task demand to better align with listener abilities. However, research findings suggest that listening support may reduce, increase, or have no effect on demand. To address this inconsistency, 77 Chinese senior high school EFL learners were randomly assigned to two groups—average-speed group (N = 33) and slow-speed group (N = 44). Both groups completed listening tasks across four conditions—repeated input, vocabulary preteaching, combination repeated input and vocabulary preteaching, and listen-once. Results from within-groups analysis of variance showed that performance was significantly better under the repeated input condition than any other condition at both speeds. The combined condition performance was equal to the baseline listen-once condition performance and better than the vocabulary preteaching condition for both speeds. The vocabulary preteaching condition performance was weaker than the other methods and equal to the listen-once condition only for the slow-speed text. Results further revealed no differences in performance between the speed groups for any condition. These results indicate that repeated input reduced environmental demand, preteaching vocabulary increased demand, and the combined condition offset the demand to match the baseline.

Keywords: L2 listening support; language alignment model; repetition of input; vocabulary preview; speech rate

ARTICLE HISTORY

Received: 23 Feb. 2026 **Revised version received:** 2 July 2026

Accepted: 21 July 2026 **Available online:** 31 July 2026

* Corresponding author: Department of English, University of Macau, Macau, China
Email address: mpwallace@um.edu.mo

doi 10.30466/ijltr.2026.57044.3308

Introduction

Listening is an indispensable skill for acquiring a second language (L2; Nation & Newton, 2008; Wallace, 2022), though it is a particularly challenging one to develop (Ahmadi Safa & Beheshti, 2018). The Performance System in Wallace's (2026) Language Alignment Model (LAM) explains that language performance is influenced by the alignment of personal resources (e.g., linguistic knowledge and language processing; abilities) and environmental characteristics (e.g., text and task conditions; demands). When personal resources closely align with the environmental conditions (called demands-abilities alignment), then performance is stronger. However, when environmental demands exceed learner abilities (called misalignment), then performance is weaker or less accurate. To improve listening performance, students may increase their personal resources needed for success on listening tasks (e.g., develop vocabulary knowledge) and/or manipulate the environmental conditions (e.g., replay the audio) to narrow the demands-abilities alignment.

Notable among these environmental adjustments for L2 listening are preteaching unknown vocabulary before a listening task, replaying the audio input, and slowing the speech rate. These activities have been shown to support listening comprehension (Chang & Read, 2006; Holzknecht & Harding, 2023; Jafari & Hashim, 2012; Medina et al., 2020; Rouhi et al., 2014). Preteaching vocabulary equips listeners with the words needed to understand the aural message, repeating the input gives them an additional opportunity to segment words in case segmentation failed in the first listen, and slowing down the speech rate allows them to more easily segment words they hear during a single listen. From a LAM perspective, these interventions support comprehension because they reduce the environmental demands of the listening tasks to more closely align with learner ability levels than a baseline listen-once condition.

The current study took place in mainland China, which has arguably the largest population of English as a Foreign Language (EFL) learners. English is a compulsory subject for students from the third grade of primary school (He & Chen, 2017), and like in most EFL learning contexts, students seldom encounter English in authentic communication other than the classroom (Renandya & Hu, 2018). This limited exposure to the target language likely contributes to Chinese students' challenges when listening in English, such as the inability to deal with a fast speech rate, poor word recognition, misapplication of L1 phonological rules to the L2, and failure to make semantic inferences from the speech (Nushi & Orouji, 2020; Renandya & Farrell, 2011; Zeng, 2007). The current study examined the degree to which multiple environmental conditions facilitated more accurate listening performance for Chinese EFL learners with low-intermediate-level ability to improve our understanding of the effects of these interventions.

Literature Review

Repetition of Input (RI)

Repeating listening materials is a common approach to improving listening performance. J. Field (2015) explains that playing the audio more than once benefits listeners by reducing the cognitive demand of the task and giving them additional time to process the material. This extra listen helps them to clarify the meaning of sentences that they may not have caught in previous listens because they were focusing on decoding speech given earlier in the text. He also explains that RI can compensate for the lack of supporting visual or contextual information in listen-only tasks and help listeners get used to the voices they hear, all of which would facilitate comprehension.

These claims are supported by findings from Holzknecht and Harding (2023), who showed that Austrian secondary school students scored significantly better on listening tasks when audio input was played twice, compared to a single play. They also reported that, interestingly, comprehension on the single play was better than comprehension of the first listen of the double-play condition.

Based on their eye-tracking results, the authors explained that the participants focused on linguistic features of the input during the first listen and the overall meaning during the second. Participants in the single-listen condition focused on constructing meaning of what they heard while overcoming issues with processing the linguistic features.

Despite the benefits for RI, J. Field (2015) cautions that it is not a condition present in real-world listening encounters, so adopting it to facilitate comprehension may not benefit learners. The limited empirical research examining RI for comprehension has reported mixed findings. For example, Rouhi et al. (2014) found that listening comprehension was better when Iranian senior high school EFL learners received RI and prepared for the listening topic than when they were given a question preview or listened only once without support. The authors explained that RI likely reduced the cognitive load for the listeners, leading to better comprehension. In the Austrian context, Ruhm et al. (2016) found that for Common European Framework of Reference for Languages (CEFR) A2/B1 level eighth-grade EFL learners, double listening improved comprehension of short listening texts (under 25 seconds). However, for longer text lengths (over 60 seconds), it only facilitated understanding when the text language was not advanced and non-redundant. The authors explained that RI was beneficial when the listeners had sufficient linguistic knowledge to understand the text and complete the comprehension activities and without it, RI was unhelpful.

This notion is supported by findings from Chang and Read (2006), who compared RI with three pre-listening activities (providing background information on the listening topic, previewing the questions, and instructing vocabulary) and found that RI was effective at facilitating comprehension for higher-level listeners, but less so for lower-level listeners. Of the four conditions, RI and topic preparation supported comprehension better than previewing questions and instructing vocabulary. Chang and Read interpreted their findings to suggest that RI may be beneficial after learners have passed a threshold of linguistic proficiency, similar to interpretations made by Ruhm et al. (2016). In LAM terms, these explanations mean that RI may align better with learners who have sufficient language ability to understand the text and complete the comprehension exercises.

However, this claim is not supported by findings from Sakai (2009) who found that proficiency level did not affect how well RI helped listening comprehension for higher and lower-level Japanese university listeners. The participant levels in Sakai and Chang and Read were relatively similar, but the response format for the listening comprehension task differed between the studies. Sakai used a written recall format requiring participants to recall propositions from a spoken text and Chang and Read adopted multiple-choice questions. Sakai's recall task with RI support aligned more closely than Chang and Read's multiple-choice questions with RI support, resulting in better comprehension for Sakai's participants. For the current study, participant proficiency levels were around the CEFR B1 – B2 level, so it is unclear how well these ability levels would align with an environment consisting of RI and multiple-choice questions.

Preteaching Vocabulary

Previous research has consistently revealed a significant association between vocabulary knowledge and listening performance (e.g., Khajavy & Aghaee, 2023; Matthews & Cheng, 2015; Wallace, 2022), illustrating that unfamiliar vocabulary is an important barrier to accurate listening performance. Listeners unfamiliar with words from L2 speech tend to inaccurately decode what they hear, and in the process of working out sounds from the speech stream, they narrowly focus their attention on the parsing of sounds, thus causing them to miss subsequent speech (Wallace, 2022). Attempting to address these limitations, previewing or preteaching of key vocabulary has become a staple activity in the pre-listening phase of the listening pedagogical sequence (pre-, while-, post-listening phases).

Despite its popularity, the effect of vocabulary instruction as a pre-listening activity has been mixed. Some studies have reported that introducing vocabulary before listening tasks can lead to improved comprehension. For example, Jafari and Hashim (2012) reported that preteaching vocabulary (students were given cards of difficult words from the text, their L1 and target-language definitions, and two example sentences in the L1 and target language and studied them on their own) and introducing the main ideas prior to a listening event for Iranian university EFL learners both contributed to better listening comprehension. No statistical differences were reported between the two conditions, however. Similar findings were reported by Madani and Kheirzadeh (2022), who showed that pre-listening vocabulary instruction (new and difficult words from the listening track were introduced and productive understanding of them verified) had an effect on both lower and advanced-level listeners. Both of these studies suggest that preteaching vocabulary using activities that require receptive recognition (in the case of Jafari and Hashim) and productive use (in the case of Madani and Kheirzadeh) can support listening comprehension.

However, studies have also reported that vocabulary as a form of pre-listening support may be less effective than other methods at facilitating better comprehension. For example, Chang and Read (2006) showed that preteaching vocabulary was the least useful form of support compared with providing background information and RI, respectively. They operationalized preteaching vocabulary similarly to Jafari and Hashim, which involved (1) giving students key words from the listening text and the L1 equivalent so they could study by themselves, (2) explaining the pronunciation and lexical categories of the terms, and (3) playing short dialogues with the words in context. Participants explained in interviews that receiving instruction on key terms lowered their anxiety to listen, but that it also narrowed their attentional focus to those specific pre-taught words, preventing them from listening to the broader text. They also noted that there was insufficient time to fully learn the words, in part because they were presented in written form and then expected to understand the spoken form when listening.

In an effort to determine the best conditions for preteaching vocabulary for L2 listening, Li et al. (2012) compared three conditions for Chinese university EFL learners: key words from the listening text given one week, one day, and 30 minutes before the listening task. Results showed that none of the conditions had significant effects on listening comprehension performance. These findings suggest that, in contrast to what Chang and Read's participants claimed, the length of time required for pre-taught vocabulary to be useful for comprehension may not be as important as requiring productive use of the new words and/or giving listeners multiple opportunities to engage with the words, their meanings, and use in context. The current study, therefore, opted for this latter approach, similar to what was utilized by Jafari and Hashim.

Speech Rate

Speech rate has been frequently discussed as a vital environmental characteristic that affects EFL listening difficulty and comprehension success (Renandya & Farrell, 2011). The underlying reason for this is that the information density (the amount of information conveyed by a linguistic unit, for example, a phoneme, word, and a sentence) is expected to increase along with the speech rate, which requires greater speech processing load. Slowing down the speech rate keeps the information density manageable, thus supporting comprehension. It is also expected to address another common challenge among L2 listeners: inefficient lexical segmentation. Slowed speech supports L2 speech decoding (i.e., perceiving and parsing sounds), which allows for more complete and accurate representations of utterances to be generated.

The empirical literature generally supports these contentions, showing that listening performance decreases when speech rate increases (e.g., McBride, 2011), though much of the research is dated. For example, McBride (2011) and Griffiths (1992) reported that comprehension of speech at slower speeds was higher than comprehension of faster speech for EFL learners in Chile and Oman,

respectively. Slower speech rate has not always contributed to better comprehension, though. Medina et al. (2020) examined the influence of speech rate, text length, and language proficiency on comprehension. They found that faster speech negatively affected comprehension only for higher-proficiency listeners. Medina et al. speculated that the lower-proficiency listeners did not benefit from the slower speech because they struggled with understanding any of the speech, regardless of the pace. What these findings suggest is that if listeners do not have the sufficient ability to process the linguistic features from speech, then the pace at which it is delivered may not matter. For the current study, we expected the CEFR B1 – B2 participants to be able to process the L2 features, so slowing the speech rate should improve comprehension.

Current Study

Through the LAM lens, the empirical findings allow us to make a few important observations of the effects of repeated input, preteaching vocabulary, and speech rate on comprehension. One is that all three environmental conditions can facilitate comprehension for some listeners, but they may not for all. This means that when comprehension was supported, the environmental demands aligned with the listener abilities, and vice versa when comprehension was weaker. The inconsistent findings for the effects for RI and reduced speech rate suggest that low-ability listeners remained misaligned with the comprehension task demands even when these methods were introduced. This was attributed to the limited linguistic knowledge of the listeners preventing them from understanding language in the text, regardless of the number of times they heard it or how slowly it was played. However, it was not the listener abilities alone that contributed to weaker comprehension performance. It was the misalignment between the listener abilities and the environmental conditions that these methods created. The methods were intended to reduce the environmental demands of the listening tasks, but weaker comprehension suggests that they were insufficient at doing so to adequately align with listener abilities and support performance.

The method aimed at addressing the linguistic deficiency issue, preteaching vocabulary, facilitated comprehension, but only when the instruction required participants to use the new words or receive multiple exposures to the words independently and in context. When instruction only involved receptive recognition of new terms, then comprehension was weaker. This suggests that only having listeners recognize word meanings before hearing a text may not adequately reduce the task demands to align with their listening abilities. What is noticeably absent in the literature is how these methods may affect comprehension if they were combined. Multiple methods of support should reduce the listening task environmental demands more than a single method, better aligning with listener abilities, and resulting in better comprehension. Preteaching vocabulary addresses the issue of limited linguistic knowledge, while repeated input and slowed speech address the issue of lexical segmentation. To that end, the current study examined the effects of RI and preteaching vocabulary at average and slowed speech rates. We aimed to answer the following research questions:

- RQ1: To what extent do repeated input, preteaching vocabulary, and their combination affect L2 listening performance at average and slowed rates of speech for Chinese EFL senior high school learners?

- RQ2: To what extent does listening performance within these conditions differ between average speed listening materials and a slower version?

Methodology

Participants

In total, 77 Chinese EFL students (59 female; 18 male) from two parallel classes of 11th grade students in a senior high school located in Mainland China participated in the study. The classes were considered equivalent in terms of proficiency level (CEFR B1 to B2) based on their school grades in the previous academic year (given two months prior). English was a compulsory subject for them during high school. All participants declared that they did not have listening disabilities. One class of students (N = 33) was randomly assigned to the average speed group and the other class (N = 44) to the slower speed group. The listening ability level of the participants in both groups was approximately equal, as indicated by non-significant differences in comprehension scores on the listening-once condition (the baseline measure) for both average and slower speeded texts.

Materials

The listening items were selected from the National College Entrance Examination (NCEE), generally known as *Gaokao*, which is the assessment used in Mainland China to evaluate students' academic abilities and determine their university admissions. The NCEE-English was chosen because it has been extensively validated (Davey et al., 2007; Hu et al., 2014). In the current study, four texts were selected with four items associated with each text. The items measured the ability to identify information explicitly provided in the text (e.g., details) and to make inferences from information provided within the text. Altogether, there were 16 items and each item weighed one point. All of the test items were in multiple-choice format with three choices. Characteristics of the four listening texts are reported in Table 1. The audio length range of listening materials was between 68 and 80 seconds (M = 71.5, SD = 6.24), the speed was around 3.11 syllables per second, and the lexical level was similar as most of the vocabulary was within the 2,000 most frequent words of English. We consider the length, speed, and lexical difficulty of the four listening materials to have been controlled, suggesting that the four listening materials were approximately equivalent. At the beginning of the test paper, students' demographic information (e.g., gender, age, and years of learning English) was asked. At the end of the test paper, there were several self-reported questions (e.g., whether they had seen these questions before and their English scores on the last formal English examination) to obtain a fuller understanding of students' L2 listening learning situation.

Two versions of the listening materials were provided to participants, i.e., the average-speed version and the slower-speed version. Speech rate was calculated by syllables per second. The slower-speed version was slowed by 14% of the average-speed version. We chose this rate because the average speed of the listening materials used on senior high school entrance examinations over the previous five years was 0.86 times the speed of the four listening texts in the current study.

Table 1
Characteristics of the four listening texts

Item	Text 1	Text 2	Text 3	Text 4
Total syllables	241	203	228	220
Vocabulary level				
K1 words	89.66%	88.71%	84.38%	90.48%
K2 words	5.17%	9.14%	8.75%	3.57%
AWL words	1.15%	1.08%	0.62%	1.79%
Off-list words	4.02%	1.08%	6.25%	4.17%
Audio Length	68s	74s	65s	80s

Notes. K1 words = first 1,000-word band on the General Service List (most frequently used English words); K2 words = second 1,000-word band on the General Service List; AWL words = words on the Academic Word List

Procedures

Prior to the start of the study, we received approval from the University of Macau Panel on Research Ethics and informed consent from the participants, their parents/guardians, and the school administration. Due to the rigorous regulations of the data-collection site, pre-made videos with listening materials and instructions were played in two classes under the guidance of two English teachers. Participants were asked to complete four listening tasks under four conditions: (1) listening to the first audio once; (2) listening to the second audio twice; (3) receiving a vocabulary instruction for the third audio and then listening to it once, and (4) receiving a vocabulary instruction for the fourth audio then listening to it twice. In each of the four task conditions, students were provided 20 seconds to pre-read the comprehension questions and they were given additional time after the audio was completed to answer the questions.

For vocabulary instruction, words above the General Service List K2 band in the text were presented on lecture slides and projected on a screen. Each word was presented one by one along with its pronunciation (delivered via sound link with model pronunciation), Chinese translation, lexical category, and a sample sentence in English and its Chinese translation. When words were proper nouns of places, only the pronunciation and Chinese translation were provided. Participants were asked not to write down the words. After each word was presented, a summary slide consisting of the full list of words and their Chinese translations were presented on the slide and the researcher read aloud each word again. The session lasted four minutes for the vocabulary preteaching condition, including an introduction to the sequence and six terms, and three and a half minutes for the repeated-input vocabulary preteaching condition, which included nine terms.

Data Analysis

Students' answer sheets were manually scored and their scores were transferred to an *Excel* file for analysis. First, the data was examined to ensure it was of acceptable reliability. This involved identifying careless or guessed responses and removing them from the dataset. To do this, the item-level data was entered into Winsteps (version 4.0.1; Linacre, 2016), a Rash modelling software package that generates an ideal model of item responses and compares it with a model of actual responses from participants. Responses that did not adhere to the ideal model responses were flagged by the program for subsequent review because they indicated careless (when a strong listener missed in easier item) or guessed responses (when a weaker listener correctly answered a difficult item). The flagged responses were examined by the researchers and removed. The process resulted in around 5% of the dataset being omitted from subsequent analysis. To verify the consistency of the responses across the items, Cronbach's alpha was calculated.

After this, the data was entered into JASP (JASP Team, 2026) for descriptive and inferential analyses to answer the research questions. First, four composite variables representing each of the four conditions were created for both groups by calculating a total score: Listen Once (LO), Repeated Input (RI), Vocabulary Preteaching (VP), and the Repeated Input Vocabulary Preteaching (RIVP). Descriptive statistics, including the mean, standard deviation, skewness, and kurtosis were computed for the two groups.

To answer research question one, two within-group analyses of variance (ANOVA) were conducted to determine if there were significant differences across the four conditions within the average-speed group and the slower-speed group, respectively. Pairwise comparisons were inspected to identify significant differences in mean scores within each group and effect sizes were calculated to determine the size of the difference among the scores. Prior to conducting the analysis, the assumptions underlying ANOVA were verified. This included examining the data for the absence of outliers, a normal distribution in listening scores for each condition, and sphericity (A. Field,

2009). To determine if there were outliers in the listening scores per condition, *z*-scores were calculated. *Z*-scores outside of the range of the absolute value of 3.29, indicating that the value was outside 3.29 standard deviations from the mean, would indicate an outlier. To determine if the listening scores were normally distributed per condition, the skewness and kurtosis values were checked and the Shapiro Wilk test was conducted. Skewness and kurtosis values outside of the absolute value of 2.0 and a Shapiro-Wilk test p-value above 0.05 would indicate a non-normal distribution. For sphericity, Mauchly's Test of Sphericity was conducted. Values above 0.05 would indicate that the sphericity was assumed. The violation of normal distribution would result in Mann-Whitney being used as a non-parametric alternative to compare means.

For research question two, four independent samples *t*-tests were conducted on each of the four conditions to determine if there were differences in comprehension at the average and slower speeds. Effect sizes were also calculated to determine the size of the difference between the scores. Before conducting the analyses, the assumptions underlying *t*-tests were verified. This included the absence of outliers, a normal distribution among the listening scores per condition, and homogeneity of variances (A. Field, 2009). Homogeneity of variances was examined by conducting Levene's Test for Equality of Variances. Values above 0.05 would indicate the assumption was met. The violation of normal distribution would result in a Mann-Whitney test being used as a non-parametric alternative to compare means for this analysis.

Results

The Cronbach's alpha reliability estimate for the instrument was 0.75, suggesting an acceptable level of internal consistency among the items. The overall descriptive statistics of the four conditions for both sets of materials are presented in Table 2. After inspecting the *z*-scores, no univariate outliers were identified. The values of skewness and kurtosis were within the absolute value of 2.0, indicating a normal distribution of scores across the conditions. Results of the Mauchly's Test of Sphericity (> 0.05) indicated that sphericity was assumed. These results suggest that the dataset met the assumptions underlying the within-groups ANOVA procedure. The results from Levene's Test for Equality of Variances indicated that the equal variances were assumed for the LO, RI, and VP groups, but that it was not assumed for the RIVP group. The results from a Shapiro-Wilk test revealed the distribution of the data was non-normal, so the Mann-Whitney test was performed.

Table 2
Overall Descriptive Statistics of Four Conditions

	Average (N = 33)			Slow (N = 44)		
	Mean (SD) 95% CI	Skewness	Kurtosis	Mean (SD) 95% CI	Skewness	Kurtosis
LO	2.09 (1.31) 1.63, 2.56	-0.18	-0.72	1.75 (1.28) 1.36, 2.14	0.35	-0.77
RI	2.94 (1.03) 2.58, 3.30	-0.97	0.79	2.57 (1.13) 2.23, 2.91	-0.48	-0.47
VP	1.49 (0.87) 1.18, 1.79	0.35	1.06	1.41 (0.97) 1.11, 1.71	-0.13	-1.01
RIVP	2.18 (1.36) 1.70, 2.66	-0.11	-1.23	1.96 (1.08) 1.63, 2.28	-0.49	-0.67

Notes. SD = Standard deviation; LO = Listen Once; RI = Repetition of input; VP = Vocabulary pre-teaching; RIVP = Combination of RI and VP

For research question one, the results from the within-groups ANOVA indicated a statistically significant difference among the conditions in the average-speed group, $F(3, 96) = 11.745$, $p < .001$, $\eta^2 = 0.259$. Pairwise comparison among the conditions indicated that comprehension scores for the RI condition was significantly higher than scores for the other three conditions. The VP condition comprehension was statistically the lowest condition. Comprehension scores for RIVP and LO were statistically equal.

The ANOVA results for the slower speeded group also indicated a statistically significant difference among the conditions, $F(3, 129) = 12.98$, $p < .001$, $\eta^2 = 0.232$. Pairwise comparisons among the conditions indicated similar results as the average-speed group. Comprehension under the RI condition was significantly better than it was under the three other conditions. The RIVP scores were significantly higher than the VP scores. Interestingly, the LO condition scores were statistically equal to the VP and RIVP condition scores. For research question two, results from the Mann-Whitney tests indicated no statistically significant difference between the average and slow-speed groups for any condition (see Table 3).

Table 3
Mann-Whitney U test comparing the average and slow-speed text comprehension for the four study variables

Variable	U	*p*
LO	846.5	.202
RI	865.5	.136
VP	741.5	.870
RIVP	799.5	.439

Notes. LO = Listen Once; RI = Repetition of input; VP = Vocabulary pre-teaching; RIVP = Combination of RI and VP

Discussion

Effects of Repetition of Input

The findings showed that the mean scores for the RI condition were significantly higher than scores for the other three conditions. This means that the participants' comprehension benefitted more from the repetition of the input than from being pre-taught challenging vocabulary alone, being pre-taught vocabulary with repeated input, and listening to the audio once. The facilitative effect of RI on comprehension corresponds to results in previous studies conducted by Sakai (2009) and Rouhi et al. (2014) for Japanese-speaking and Iranian-speaking learners of English, respectively, and particularly for higher-level listeners in Chang and Read (2006) and Ruhm et al. (2016). The participants in our study were similar to the higher-level listeners in those studies, so our findings generally align with theirs.

From a LAM perspective, we interpret these results to suggest that RI reduced the listening environmental demands to better align with listener abilities and facilitate comprehension. To explain how, we refer to Holzknecht and Harding (2023), whose eye-tracking results showed that listeners may focus their attention on the lower levels of processing (i.e., word recognition and lexical segmentation) during a single listen, but may achieve higher-order cognitive processing (i.e., using their existing knowledge to interpret the message in context) during the second playing of a listening track. The comprehension questions in our study, requiring participants to identify details and make inferences, engaged listeners' higher-order listening processes. Without the second listen, our participants may have focused on decoding the speech at lower levels. This explains why there were lower comprehension scores for participants in the LO condition, by comparison.

Effects of Vocabulary Preteaching

The results also showed that comprehension in the VP condition received the lowest score among the four conditions, though it was not significantly different from the control condition (i.e., LO) for the slowed texts. These results are consistent with findings reported by Chang and Read (2006), who operationalized VP similarly as the current study, and showed that comprehension was stronger for previewing questions and repeating input. Our findings are somewhat unexpected, given the established importance of vocabulary knowledge for listening performance (Khajavy & Aghaee, 2023; Matthews & Cheng, 2015; Wallace, 2022) and the strong emphasis that preteaching vocabulary is given to pedagogical practice.

We offer two reasons similar to Chang and Read for the apparent failure of VP to enhance students' listening performance. In line with explanations by Chang and Read's participants, the VP condition may have facilitated participants' attentional bias towards the words that were given in the preteaching session. Attentional bias occurs when attention is focused too narrowly on stimuli that is more salient or perceived to be more relevant and ignoring other, potentially relevant, stimuli (Drobes et al., 2019). During the VP listening comprehension task, our participants may have focused their attention on identifying the terms that were pre-taught, while overlooking other important information in the message. Similarly, the words that were pre-taught may have primed existing knowledge that was not relevant to the subsequent listening text. This inevitably would have obstructed their comprehension process.

Another explanation that was also proposed by Chang and Read is that the pre-taught words may not have been learned well enough to be of use for the listeners in the subsequent listening task. Although example sentences were provided to grasp how to use the pre-taught words in sentences, and it has been shown that providing vocabulary with visual and acoustic stimuli does promote short-term vocabulary memorization, requiring learners to demonstrate their understanding of the words productively may be more advantageous. This approach was taken by Madani and Kheirzadeh (2022), who reported that vocabulary preteaching condition supported comprehension, so there is some limited evidence to support this.

Of these two explanations, the one that appears to be better supported by our data is the former. For the average-speed texts, VP comprehension scores were significantly lower than the baseline LO condition. From a LAM perspective, this suggests that preteaching key vocabulary before listening increased the task environmental demand to misalign significantly more with listener abilities than the baseline condition did. Preteaching key vocabulary narrowed listeners' focus to those terms, causing them to miss important content needed to comprehend the full text.

This increased demand in the VP condition appears to have been offset by two other conditions. One is the repetition of the input. For both speed texts, the VP comprehension was significantly lower than the RIVP condition. This suggests that though attentional focus was narrowed to specific words in the first listen, our participants used the second listen to construct meaning of the full text without focusing on those words. Holzknecht and Harding (2023) showed that listeners attend to linguistic features during the first listen of a repeated-input condition, but listeners typically determine what to give their attention to according to what they (mis)hear. In our study's case, the features were given to participants in advance, thus seemingly misdirecting their focus and increasing the environmental demand.

The other condition that may have offset the increased demand was the slow speech rate. There were no significant differences between the LO condition and the VP condition when the speech was slowed. We interpret this to mean that though attention was narrowed to the pre-taught terms, the slowed speech prompted listeners to abandon that strategy during their only listen. We speculate that the slowed rate may have allowed the listeners to process the speech more easily than the

average-speech rate and at some point during the listen, they realized that the words were not contributing to their interpretation of the text and started listening more globally. An alternative explanation may be that the slowed speech rate allowed the listeners to actually recognize the words that were pre-taught, allowing them to take advantage of learning the new terms. However, the statistically equal scores between LO and VP for the slow speed text suggests that even if this did happen, its effects were no better than the baseline.

For the second explanation about words not being learned sufficiently to be supported, we would have seen no difference in the LO and VP conditions or in the RI and RIVP conditions. Because VP comprehension scores were significantly lower than LO scores and RIVP scores were lower than RI scores, preteaching vocabulary appears to have had more of a detrimental effect on comprehension than a negligible effect. Altogether, our findings suggest that the VP condition increased environmental task demands to widen the misalignment with listener abilities, resulting in weaker comprehension.

This finding has important pedagogical implications because a common teaching approach in the listening classroom is to preteach the vocabulary of an upcoming text. Based on our results, only providing an L1 definition, its pronunciation, and a sample sentence appears to make comprehension more difficult than support it. If teachers want to teach key words before a listening text, they may consider reminding students that they should listen to the entire text and not focus on the specific words that were pre-taught because if they do, they will miss important information in the text.

Effects of Speech Rate

Results from the Mann-Whitney U tests showed that there was no difference in comprehension scores between the average and slower-speed texts in any condition. This suggests that reducing the speech rate by 14% of the normal rate had a negligible effect on the comprehension scores, regardless of the environmental condition (RI, RIVP, VP, LO). It contradicts our expectation that the listeners would benefit more from a slower speech rate, as we considered that slowed speech would solve potential problems arising from listening in the target language—reducing the information density and assisting lexical segmentation—to facilitate comprehension. However, we found no differences in performance.

Our findings align somewhat with previous empirical studies about the effects of slowed speech on comprehension. Similar to our results, performance was unaffected by slower speech rates for the lower-level listeners in Medina et al. (2020). The explanation offered for that result, that speaking speed may not affect comprehension if listeners are unable to process language at any speed, may also apply to our results. Taking all of this together, it appears that our participants may have struggled to process the speech at average or slowed speeds when it was presented once. When the audio was repeated, the second listen allowed the listeners to process speech at both speeds more accurately.

An alternative explanation is that the speed at which the speech was slowed was not sufficient to support processing. We slowed the speech 14%, which equates to around 2.76 syllables per second, slightly above the 2.5 syllables per second criterion set by Griffiths (1992). The average speed of 3.11 syllables per second was slightly below the 3.75 syllables per second in Griffiths. Being slightly above the slow pace and below the average pace may have narrowed the gap too much to qualitatively distinguish between the two speeds.

These results may be of use for classroom teachers who seek to improve comprehension by altering the input speed. We have shown that slowing the speech rate alone did not improve comprehension

in our study, but if it is used with another condition, like repeated input, then it may support understanding.

Conclusion

The current study findings reaffirm the previous research supporting RI as an effective form of listening support, while also showing that VP was potentially harmful for comprehension. Our findings also showed that comprehension was better under a combined VP and RI condition than a VP-only condition. Furthermore, comprehension was statistically equivalent between average and slow-speed texts across the four conditions between, suggesting that the change in speech rate had little impact on comprehension.

There are limitations in this study that need to be acknowledged. One limitation is the relatively small sample size and its makeup. The participants were EFL students from two classes in a single school. Homogeneity of the sample allowed us to ensure participants were as similar as possible, and thereby control potential confounding variables (e.g., varied L1 and proficiency level), yet it restricted the degree of generalization of results to larger EFL learner groups. Future studies may enlarge the sample size to EFL learners with different proficiency levels and backgrounds. Another limitation is that the current study was conducted in a single context in mainland China. To further improve the external validity of the findings, studies are encouraged to examine the effects of pre-listening support on comprehension in other contexts.

This study demonstrates that the Performance System of the LAM framework can be used to interpret the findings of an empirical study. The model conceptualizes performance as the product of the interaction between learner abilities and environmental demands, so differences in performance are not solely attributable to environmental characteristics or learner resources, as is commonly claimed. Though alignment was not directly measured in our study, we are able to infer from the results that environmental demands more closely aligned with our participants' abilities when comprehension scores were higher, and vice versa when they were lower. For our study examining the effects of different listening support methods, those that facilitated better comprehension above the baseline listen-once condition (i.e., repeated input) reduced the environmental demands to better align with our CEFR B1/B2 listeners' abilities. The environmental demands for the methods that were equal to the baseline condition (i.e., vocabulary pre-taught with repeated input, reduced speed), were not reduced to align with our listener abilities. Likewise, the pre-taught vocabulary condition showing weaker comprehension than the baseline likely increased demand. Moving forward, we encourage scholars to more directly examine the degree of alignment between learner abilities and environmental demands to observe its explicit effects on performances. We would also encourage researchers to consider using the LAM framework in future language studies and to support such endeavor, we provide a brief description of the framework below.

Unified Language Alignment Model

Wallace's (2026) LAM explains how complex dynamic systems involved in language learning and assessment can be operationalized across the four macro-skills of language learning. Drawing on the concepts of person-environment fit in organizational behavior (Kristof-Brown, 2005) and complex dynamic systems in second language acquisition (Verspoor & Lowie, 2021), the core LAM claim is that the interaction between the individual and the environment explains language learning and performance.

The LAM conceptualizes performance and learning as related and independent systems, called the Performance System and Learning System, respectively. The Performance System explains why performances on language assessments may differ. Optimal performance alignment is when the

demands of the environment (e.g., language task difficulty) align closely with the abilities of the individuals (called demands-abilities alignment). When demand is too low or too high relative to learner ability, there is a floor (e.g., overly easy tasks lead to very strong performance) or ceiling effect (e.g., overly difficult tasks result in poor performance), respectively. The Performance System helps interpret why performance on language assessments may vary across participants, both in the pedagogical context and for research, as we have demonstrated in the current study.

The Learning System describes how the interaction between the learner and environment affect knowledge acquisition. Under this system, learning is most efficient when the alignment between learner needs and environmental supplies is close (called needs-supplies alignment). However, when the environment is over-supplied (e.g., excessive scaffolding) or under-supplied (e.g., insufficient scaffolding), learning becomes dependent or frustrating, respectively. For L2 listening, the comprehension approach, a common pedagogical technique involving activating schemata and giving comprehension exercises (J. Field, 2009), provides an under-supplied environment, particularly for low-level learners whose needs are to develop their perception (distinguishing among speech sounds) and parsing (grouping perceived chunks into meaningful units) processes (Goh & Vandergrift, 2021). It insufficiently addresses bottom-up processing needs in the lesson sequence because its emphasis is accurate comprehension of a text.

The next logical question becomes what to do if teachers/learners recognize that the environment is over- or under-supplied? The LAM offers a developmental process through which teachers can adjust the environment to best meet their learners' needs and ultimately develop independent learners. Teachers first set a task with the desired conditions to observe an over- or under-supply (Diagnosis). Their responsibility is to then change the environment (i.e., task conditions) to better align with learner needs based on their evaluation of the diagnostic task (Adjustment). Learners then successfully engage with the task with the support of the adjusted environment (Activation). Processing is effortful in this stage and learners are able to complete tasks with the assistance of the support. This Activation Stage aligns with Vygotsky's (1978) Zone of Proximal Development, noted as being the ideal balance between challenge and ability for knowledge and skill acquisition. Over time, learners' needs to complete such tasks reduce and the teacher likewise reduces the support given (Withdrawal). The reduction in support continues until the learner performs the task independently (Independence). Then a new, more challenging task is introduced and the cycle continues until the learner is able to reach independence.

This is a brief description of the LAM and as we have explained here, it can be used to interpret the effectiveness of performance on language assessments and the effectiveness of language learning pedagogical approaches to support learning. We recommend researchers and language practitioners consider using the model for their respective needs because it bridges the important theory-practice gap in the language learning, teaching, and assessment literature.

Acknowledgements:

This research project was funded by the Specialized Subsidy Scheme for Macao Higher Education Institutions in the Area of Research in Humanities and Social Sciences (HSS-UMAC-2021-03) and the Research Services and Knowledge Transfer Office, University of Macau (MYRG2022-00235-FAH).

References

Ahmadi Safa, M., & Beheshti, S. (2018). Interactionist and interventionist group dynamic assessment (GDA) and EFL learners' listening comprehension development. *Iranian Journal of Language Teaching Research*, *6*(3), 37-56.

Chang, A. C. S., & Read, J. (2006). The effects of listening support on the listening performance of EFL learners. *TESOL Quarterly, 40*(2), 375-397. https://doi.org/10.2307/40264527

Davey, G., Lian, C. D., & Higgins, L. (2007). The university entrance examination system in China. *Journal of Further and Higher Education, 31*(4), 385-396. https://www.doi.org/10.1080/03098770701625761

Drobes, D. J., Oliver, J. A., Correa, J. B., & Evans, D. E. (2019). Attentional bias and smoking. In V.R. Preedy (ed.) *Neuroscience of nicotine: Mechanisms and treatment*, (pp. 145-150). Academic Press. https://doi.org/10.1016/B978-0-12-813035-3.00018-6

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

Field, J. (2009). *Listening in the language classroom*. Cambridge University Press.

Field, J. (2015). *The effects of single and double play upon listening test outcomes and cognitive processing.* British Council.

Goh, C. C. M., & Vandergrift, L. (2021). *Teaching and learning second language listening: Metacognition in action* (2nd ed.). Routledge.

Griffiths, R. (1992). Speech rate and listening comprehension: Further evidence of the relationship. *TESOL Quarterly, 26*(2), 385-390. https://doi.org/10.2307/3587015

He, L., & Chen, D. (2017). Developing common listening ability scales for Chinese learners of English. *Language Testing in Asia*, *7*(4). https://doi.org/10.1186/s40468-017-0033-4

Holzknecht, F. (2019). *Double play in listening assessment.* [Doctoral dissertation], Lancaster University.

Holzknecht, F., & Harding, L. (2023). Repeating the listening text: Effects on listener performance, metacognitive strategy use, and anxiety. *TESOL Quarterly*, *58*(1), 451 – 478. https://doi.org/10.1002/tesq.3249

Hu, W., Li, F., & Gan, L. (2014). Does China's national college entrance exam effectively evaluate applicants? *Frontiers of Economics in China, 9*(2), 174-182. https://www.doi.org/10.3868/s060-003-014-0010-7

JASP Team (2026). JASP (Version 0.97.1)[Computer software].

Jafari, K., & Hashim, F. (2012). The effects of using advance organizers on improving EFL learners' listening comprehension: A mixed method study. *System, 40*(2), 270-281. https://doi.org/10.1016/j.system.2012.04.009

Khajavy, G. H., & Aghaee, E. (2023). Background knowledge as a moderator of the relationship between vocabulary knowledge and foreign language listening comprehension. *Iranian Journal of Language Teaching Research*, *11*(1), 117-139.

Kristof-Brown, A. L., Zimmerman, R. D., & Johnson, E. C. (2005). Consequences of individuals' fit at work: A meta-analysis of person–job, person–organization, person–group, and person–supervisor fit. *Personnel Psychology*, *58*(2), 281-342.

Li, L., He, K., & He, Q. (2012). Effect of vocabulary preparation on students' vocabulary and listening comprehension. In *International Conference on Information Computing and Applications* (pp. 516-523). Springer.

Linacre, J. M. (2016). Winsteps® Rasch measurement computer program (Version 4.0.1) [Computer software]. Winsteps.com.

Madani, B. S., & Kheirzadeh, S. (2022). The impact of pre-listening activities on EFL learners' listening comprehension. *International Journal of Listening*, *36*(1), 53-67. https://doi.org/10.1080/10904018.2018.1523679

Matthews, J., & Cheng, J. (2015). Recognition of high frequency words from speech as a predictor of L2 listening comprehension. *System, 52*, 1-13. https://doi.org/10.1016/j.system.2015.04.015

McBride, K. (2011). The effect of rate of speech and distributed practice on the development of listening comprehension. *Computer Assisted Language Learning, 24*(2), 131-154. https://doi.org/10.1080/09588221.2010.528777

Medina, A., Socarras, G., & Krishnamurti, S. (2020). L2 Spanish listening comprehension: The role of speech rate, utterance length, and L2 oral proficiency. *The Modern Language Journal*, *104*(2), 439-456. https://doi.org/10.1111/modl.12639

Nation, I. S. P., & Newton, J. (2008). *Teaching ESL/EFL listening and speaking*. Routledge.

Nushi, M., & Orouji, F. (2020). Investigating EFL teachers' views on listening difficulties among their learners: The case of Iranian context. *Sage Open*, *10*(2), 1 – 16. https://doi.org/10.1177/2158244020917393

Renandya, W. A., & Farrell, T. (2011). Teacher, the tape is too fast! Extensive listening in ELT. *ELT Journal, 65*(1), 52-59. https://doi.org/10.1093/elt/ccq015

Renandya, W. A., & Hu, G. (2018). L2 listening in China: An examination of current practice. In A. Burns & J. Siegel (Eds.) *International perspectives on teaching the four skills in ELT* (pp. 37-50). Palgrave Macmillan.

Rouhi, A., Nabavi, S. M., & Mohebbi, H. (2014). The effects of previewing questions, repetition of input and topic preparation on listening comprehension of Iranian EFL learners. *Iranian Journal of Language Teaching Research*, *2*(2), 73-85.

Ruhm, R., Leitner-Jones, C., Kulmhofer, A., Keifer, T., Mlakar, H., & Itzlinger-Bruneforth, U. (2016). Playing the recording once or twice: Effects on listening test performances.

International Journal of Listening, *30*(1), 67-83. https://doi.org/10.1080/10904018.2015.1104252

Sakai, H. (2009). Effect of repetition exposure and proficiency level in L2 listening tests. *TESOL Quarterly*, *43*(2), 360-372. https://www.jstor.org/stable/27785016

Verspoor, M., & Lowie, W. (2021). Complex Dynamic Systems Theory and second language development. In H. Mohebbi & C. Coombe (Eds.), *Research Questions in Language Education and Applied Linguistics: A Reference Guide* (pp. 799–803). Springer. https://doi.org/10.1007/978-3-030-79143-8_138

Vygotsky, L. (1978). *Mind in society: The development of higher psychological processes*. Harvard University Press.

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. https://doi.org/10.1111/lang.12424

Wallace, M.P. (2026). *Language learning strategy use as adaptive behavior: A demands-abilities fit perspective* [Manuscript submitted for publication]. Department of English, University of Macau.

Zeng, Y. (2007). *Metacognitive instruction in listening: A study of Chinese non-English major undergraduates*. [Doctoral dissertation], Nanyang Technological University.

Matthew P. Wallace is an Associate Professor in the Department of English at the University of Macau. His research interests include individual differences in second language listening comprehension, second language listening pedagogy, and language assessment fairness.

Yihan Zhang obtained her master's degree in Language Sciences at University College London. Her research interest covers second language acquisition, children's language development, and speech recognition.

Meilin Li obtained her master's degree in Linguistics at the University of Hong Kong. Her research interest covers sociolinguistics, multilingualism, and second language acquisition.

Xiaolin Zhu obtained her master's degree in English studies at the University of Hong Kong. Her research interest covers semiotics, multilingualism, and second language acquisition.

Don Yao is a Lecturer working in the Department of English at the China Jiliang University. His research interest mainly focuses on language assessment and L2 teaching and learning.