

Translanguaged practice in listening assessment: L1 vs. L2 responses in recall tasks

Shelly Xueting Ye, Matthew P. Wallace, Christine Jiaqi Dong, Lily Xiaoran Shi
& Bron Yuntian Yang

To cite this article: Shelly Xueting Ye, Matthew P. Wallace, Christine Jiaqi Dong, Lily Xiaoran Shi & Bron Yuntian Yang (06 Aug 2024): Translanguaged practice in listening assessment: L1 vs. L2 responses in recall tasks, Journal of Multilingual and Multicultural Development, DOI: [10.1080/01434632.2024.2385719](https://doi.org/10.1080/01434632.2024.2385719)

To link to this article: <https://doi.org/10.1080/01434632.2024.2385719>



Published online: 06 Aug 2024.



Submit your article to this journal [↗](#)




View related articles [↗](#)



View Crossmark data [↗](#)



Translanguaged practice in listening assessment: L1 vs. L2 responses in recall tasks

Shelly Xueting Ye ^a, Matthew P. Wallace^b, Christine Jiaqi Dong^c, Lily Xiaoran Shi^d and Bron Yuntian Yang^e

^aFaculty of Humanities and Social Sciences, City University of Macau, Macao SAR, China; ^bFaculty of Arts and Humanities, University of Macau, Macao SAR, China; ^cDepartment of Languages, Cultures and Linguistics, University of Southampton, Southampton, UK; ^dSchool of Modern Languages, University of Bristol, Bristol, UK; ^eSchool of Education, University of Adelaide, South Australia, Australia

ABSTRACT

Listening recall tasks, where test takers write L2 responses based on audio content, often inadvertently measure L2 writing skills, potentially compromising the accuracy of listening assessment. To address this concern, the present study draws inspiration from translanguaging, and employs a mixed-methods approach to investigate the potential of using L1 responses into recall tasks to mitigate the impact of L2 writing proficiency on task outcomes, thus enhancing task validity. Specifically, a listening recall task requiring L1 responses was developed and administered to 102 L2 learners in low and intermediate English proficiency. Their performance on this task was compared to a control task that required L2 responses. Follow-up interviews were also conducted. The results indicated that allowing L1 responses in the recall task led to better performance for intermediate-level L2 learners, enhanced representation of the listening construct, and garnered participants' preference as a listening assessment approach. This suggests that utilising L1 responses has the potential to improve the construct and face validity of recall tasks, addressing the initial validity concern. Consequently, this investigation contributes to the refinement of listening task development and advocates for the application of translanguaged practices in language assessment, advancing the broader concept of translanguaging in language education.

ARTICLE HISTORY

Received 20 September 2023
Accepted 24 July 2024

KEYWORDS

Translanguaging;
translanguaged practices; L2
listening assessment;
listening recall tasks; L1
responses

Introduction

The listening recall task, an assessment approach that requires L2 learners to write down what they heard (Baccouche et al. 1996; VanPatten 1990), has been observed to engage the listeners' cognitive processes in both bottom-up and top-down ways (Buck 2001), substantiating its validity in evaluating listening comprehension (Sakai 2012). In contrast to other listening assessment formats, such as multiple-choice questions, recall tasks offer several advantages. Notably, test takers are less inclined to resort to random guessing or adopt test-wiseness strategies. Additionally, the writing outputs have the potential to yield deeper insights into the examinees' listening process, helping educators to better understand learners' listening challenges and consequently benefit L2 teaching (Alderson 2005). Conversely, listening recall tasks have drawn criticism for providing scores that might inaccurately represent test takers' listening abilities. This concern is particularly relevant to L2 learners residing in non-English-speaking environments, who generally manifest delayed

development of productive skills (i.e. speaking and writing) compared to their receptive skills (i.e. listening and reading) (Rahmah and Halimi 2020). Their weaker writing proficiency is thought to hinder them from accurately writing down audio content they otherwise comprehend. This might introduce variance in scores that pertains more to writing competence than listening ability. Consequently, it could potentially result in inaccurate interpretations of test takers' listening competence, thus compromising the validity of listening assessment (Messick 1995).

A potential solution to this issue may lie in translanguaged practices, in this context, referring to the use of test takers' L1 within the L2 listening recall task, which was inspired by Cook's (2001) claim that the L1 can facilitate L2 learning and can be applied to L2 assessment. Considering that well-educated young adults generally possess advanced proficiency in their L1 writing, allowing them to write down L2 spoken content using their L1 may enhance their ability to demonstrate their listening skills, detached from any limitations posed by their less adept L2 writing abilities. Scholarly literature also provides instances where researchers (e.g. Gordon and Hanauer 1995; Shohamy 1984) noted enhanced performance and attitudinal shifts among test takers when employing L1 questions in L2 reading tasks of multiple-choice and open-ended questions (Cox, Bown, and Bell 2019). Meanwhile, the pedagogical efficacy of translanguaged practices in L2 learning has been substantiated (Qureshi and Aljanadbah 2022), although its application remains less pronounced when it comes to language assessment (Karges, Barras, and Lenz 2022). Among the limited investigations on the use of L1 in L2 assessment, most have centred on reading assessment (e.g. Cox, Bown, and Bell 2019; Karges, Barras, and Lenz 2022; Lee 1986). The realm of listening assessment, especially concerning listening recall tasks, which constitute the focus of the present study, has scarcely been discussed with the topic of translanguaged practices.

Hence, this study seeks to explore the potential effectiveness of employing translanguaged practices, specifically, using L1 responses in L2 listening recall tasks, as a means to minimise construct-irrelevant variance and consequently obtain scores that more accurately reflect test takers' listening competence, in contrast to tasks permitting only L2 usage.

Literature review

Understanding translanguaging

The notion of translanguaging was initially introduced by Williams (1994) as the fundamental concept that an individual's knowledge of one language can be leveraged to enhance the development of their proficiency in other languages. For instance, bilingual students receive input in one language and subsequently produce their output in another (Lewis, Jones, and Baker 2012). Over time, this concept has evolved, leading to a deeper understanding as evidenced by the proliferation of more and precise definitions (Qureshi and Aljanadbah 2022). Generally, translanguaging is defined as the process of employing multiple languages to convey meaning, shape experiences, and acquire knowledge (Baker 2011). A more specific definition entails learners comprehending a text and improving their linguistic proficiency by drawing upon their full linguistic repertoire (García, Johnson, and Seltzer 2017). Contrary to the traditional notion that languages are distinct and segregated entities with minimal influence from one language on the development of others (Jaspers and Madsen 2016), translanguaging is commonly practiced in educational settings with bilingual or multilingual contexts (Du, Lee, and Sok 2020). Translanguaged practices serve as tools for scaffolding learners' understanding of subject content, resources for facilitating independent language acquisition, and even have the potential to create a safe and equitable learning environment (Qureshi and Aljanadbah 2022). However, while these benefits have been extensively explored and documented for pedagogical purposes in classroom settings (Li 2018; Tai and Li 2021), the application of translanguaging in another critical educational context, assessment, remains relatively unexplored (Adhikari and Poudel 2024).

Regarding the theoretical underpinnings of translanguaging, two main perspectives emerge: the unitary and multilingual viewpoints. The unitary perspective views translanguaging as a unified

mechanism, termed the 'repertoire', which transcends individual languages, advocating for no boundaries among them. This perspective emphasises the fluid use of all linguistic resources within one's repertoire, disregarding labels like 'first' or 'second language' (García 2011; Otheguy, García, and Reid 2019). Learners utilise this repertoire to access various linguistic resources, enhancing language acquisition (García and Otheguy 2020), facilitating comprehension of complex texts, and employing advanced strategies (Chu 2017). However, this perspective encounters opposition from the multilingual standpoint, which considers the repertoire as multilingual or multimodal, rather than monolithic. It acknowledges boundaries among languages in the repertoire, suggesting that the use of language resources is influenced by social context and is not entirely fluid. This perspective accepts labels indicating language boundaries, such as L1 and L2 (Cenoz and Gorter 2019; MacSwan 2017).

Several reasons support the multilingual nature of the repertoire. Firstly, learners' repertoires are multidirectional, encompassing distinct attributes of different languages, which highlight internal language differences in terms of linguistic and grammatical features. This indicates the existence of boundaries among languages within the repertoire (MacSwan 2017). Understanding these boundaries helps learners acquire metalinguistic knowledge, which facilitates language acquisition and better utilisation of their repertoire for translanguaging (Cenoz and Gorter 2019). This benefit might not be achieved if the unitary viewpoint, which overlooks these boundaries, is adopted. Additional reasons supporting the existence of boundaries in the translanguaging repertoire involve social contexts. It is emphasised that translanguaging has a social dimension and the concept is associated with the use of language in specific social contexts (Cenoz and Gorter 2019). Two typical examples illustrate how translanguaged practices in social contexts suggest language boundaries in the repertoire. The first example involves translanguaging in a context with majority and minority languages (Musk 2010). It was observed that translanguaged practices in such contexts tend to utilise more resources from the majority language than the minority one. Contextual features, such as the greater number of communicative needs and interaction opportunities provided by the majority language, make people subconsciously recognise language boundaries and favour the majority language. Similarly, the second example involves translanguaging in a context involving the learning of a heritage language and a foreign language (Cenoz and Gorter 2019). People generally have a greater emotional attachment to their heritage language compared to a foreign language (Mori and Sanuth 2018). This attachment leads to a preference for learning and utilising more resources from the heritage language in their translanguaged practices. Both examples illustrate people's perceived language boundaries within the repertoire, suggesting that the use of language resources for translanguaging is not entirely free but influenced by social context, thus supporting the multilingual viewpoint. This aspect is, however, overlooked by the unitary viewpoint. Therefore, adopting the unitary viewpoint, which fails to recognise language boundaries and people's different tendencies toward using various languages as influenced by social contexts, might undermine the preservation and teaching of certain languages (e.g. minority and foreign languages). This oversight could potentially lead to the loss of these languages. Consequently, concerns arise regarding translanguaging pedagogy that advocates a unitary viewpoint (Lyster 2019).

Given the abovementioned reasons, it appears more compelling for the present study to adopt the multilingual viewpoint that acknowledge language boundaries and accepts labels such as L1 and L2, for conducting translanguaged practices. This entails employing one or multiple languages to foster the development or assessment of another language (Lewis, Jones, and Baker 2012; Qureshi and Aljanadbah 2022). Illustrative examples of translanguaging in multilingual contexts include leveraging more proficient L1 and L2 to enhance learners' morphological awareness in L3 (Leonet, Cenoz, and Gorter 2019), encouraging multilingual learners to tap into their entire linguistic repertoire, encompassing both L1 and L2 resources, as well as their increased metalinguistic awareness, to facilitate the acquisition of an additional language (Cenoz and Gorter 2022). In contrast to the relatively scarce research on translanguaging in multilingual contexts, such as in L3 acquisition (Jessner and Török 2017), more examples of translanguaged practices are observable in bilingual

contexts. These include providing L1 glosses for reading passages (Martin-Beltrán 2014), explaining grammar concepts in L1 within L2 instruction (Palmer et al. 2014), employing L1 questions in reading comprehension tasks (Vaish 2019), as well as incorporating L1 responses in listening assessments, which is the focus of the current study.

Empirical studies on translanguageing in L2 learning and assessment

Numerous pieces of evidence substantiate the pedagogical efficacy of incorporating translanguageed practices within language education. Specifically, within the context of L2 English learning, the utilisation of translanguageing pedagogy has demonstrated its capacity to enhance various aspects of L2 learners' language skills. This enhancement includes, among others, the refinement of L2 learners' reading proficiency (Chu 2017), the cultivation of L2 learners' writing competence for academic purposes (Sun and Lan 2021; Velasco and García 2014), and the facilitation of vocabulary acquisition through the implementation of translanguageing pedagogy (Galante 2020; Mwindi and Van 2015).

In contrast to the wealth of empirical support for the benefits of translanguageing in L2 learning and teaching, relatively inadequate attention has been directed towards its application in the domain of language assessment. Within the limited body of research that explores the utilisation of L1 within L2 assessment contexts, the predominant focus has centred on reading assessment. Noteworthy among these studies are the inquiries into the impact of L1 questions on L2 learners' performance in reading tasks including multiple choice questions (e.g. Gordon and Hanauer 1995; Nevo 1989), open-ended questions (e.g. Godev, Martínez-Gibson, and Toris 2002; Shohamy 1984), and reading recalls (e.g. Brantmeier 2006; Lee 1986). These investigations have consistently demonstrated that L1 questions tend to provide a more accurate reflection of test takers' reading competence, compared to L2 questions. Similarly, scant attention has been directed towards the exploration of translanguageed practices in the context of L2 listening assessment. An exception is Filipi's (2012) study which indicated that learners exhibited greater ease in responding to L1 items that gauged their global understanding of audio content. Conversely, for items demanding the comprehension of intricate and explicitly stated details, L2 items were perceived as less challenging. In Koo and Lee's (2020) study which exclusively focused on multiple-choice questions, participants consistently achieved scores that more accurately reflected their reading skills when presented with questions in Korean (L1) as opposed to those posed in English (L2). It is noteworthy, however, that a study conducted by Mihara (2015), which also delved into the utilisation of L1 questions, yielded different outcomes, indicating no significant difference in participants' performance across various task formats. The mixed results obtained from these studies regarding the impact of incorporating L1 questions into L2 listening assessments underscore the need for further research in this topic. This resonates with Mihara's (2015) call for additional investigations specifically aimed at comparing the utilisation of L1 and L2 in the context of listening comprehension.

The translanguageed practices in the abovementioned studies typically involve the utilisation of L1 questions. However, the utilisation of L1 in test takers' responses to questions has not been fully explored. Even though Bachman (1990) stated that the language employed in responses can influence test outcomes, there have been only a handful of efforts to investigate how L2 learners' response language affects their performance in L2 tasks. Studies by Lee (1986) and Brantmeier (2006) focusing on reading recall tasks disclosed that test takers tend to achieve higher scores when responding in L1 compared to when responding in their L2. This suggests that the utilisation of L1 responses may offer a more accurate assessment of individuals' reading abilities compared to the use of L2 responses. Similarly, Yu (2008) delved into the same topic of response language, specifically focusing on L2 learners' reading summaries. The results emphasise the potential for test scores to more accurately represent test takers' reading abilities in L1 summarisation as opposed to their performance in L2. Nevertheless, it is crucial to emphasise that these results pertain solely to reading assessments. Analogous investigations within listening assessments are notably scarce.

Until recently, only one study, conducted by Rahmah and Halimi (2020), sought to compare the performance of teenage English learners in listening tests between their L1 (Indonesian) and L2 (English). The outcomes of this study revealed higher scores when participants responded in their L1, implying that using L1 may better reflect test takers' listening competence than using L2. However, it is imperative to address certain methodological concerns within their research. Specifically, the specific listening tasks employed in their study remain unclear, and it appears that they employed the same listening text for comparative purposes. Participants listened to the same text and answered questions in Indonesian one week after they had completed the same exercise in English. This may have led to inconsistencies in the number of listening exposures administered. Consequently, it is plausible that the improved scores observed when responding in Indonesian during the second week were attributable to factors such as increased exposure to the listening text or other cognitive aspects like memory, rather than solely relying on the use of L1 for answering questions. This potential methodological flaw raises significant concerns about the accuracy of the results obtained in their study. Given the methodological limitations evident in the aforementioned study and the relatively limited attention afforded to the role of L1 in L2 listening assessments, particularly in the context of listening recall tasks – the primary focus of the present investigation – it becomes evident that further research in this domain is warranted.

Furthermore, the role of L2 learners' proficiency level has been acknowledged as potentially mediating the impact of L1 usage on their task performance. Shohamy (1984) has proposed that the advantages of incorporating L1 questions diminish as students' L2 reading proficiency progresses. It is posited that the advantages of utilising L1 in L2 assessment are most pronounced among learners with lower proficiency levels (Upton and Lee-Thompson 2001). Similarly, Brantmeier (2006) discovered that the language of response contributed to a mere 3% of the variance observed in the L2 reading recall performances of advanced learners, thereby implying its insignificantly influential role. However, a recent study by Cox, Bown, and Bell (2019) yielded conflicting outcomes, revealing that advanced learners exhibited enhanced reading performance when L1 questions were used in the tasks. This incongruity in findings underscores the necessity of accounting for learners' proficiency levels in future investigations that delve into the ramifications of translanguage practices on L2 learners' task performance.

In light of the review of existing empirical studies, several research gaps pertaining to the utilisation of translanguage practices in L2 assessment, are identified. (1) Comparative studies of L1 and L2 usage in language assessment have predominantly concentrated on reading tasks, with few in listening assessment. (2) Inadequate attention has been directed towards understanding the influence of response language on L2 learners' task performance, particularly within L2 listening recall tasks. (3) No definitive conclusions have been drawn regarding the impact of L2 learners' proficiency on this topic.

Research questions

In addition to addressing the aforementioned research gaps, the present study aims to investigate the potential efficacy of employing a translanguage practice (i.e. L1 responses) within a L2 listening recall task as a means to deal with the issue of construct-irrelevant variance, potentially stemming from the writing proficiency of L2 learners. To achieve this objective, a listening recall task was developed that required learners' L1 responses. Comparison was then conducted on the L2 learners' performance on this task against their performance on a control task that elicited responses in the L2. Besides, learners' perceptions of employing this translanguage practice in listening recall tasks, were also examined. Accordingly, two research questions have been proposed as follows:

- (1) Do low and intermediate proficiency level L2 learners perform differently on listening recall tasks with L1 and L2 responses?
- (2) How do L2 learners perceive the use of L1 responses in L2 listening recall tasks?

Methodology

Participants

102 English learners with a Chinese L1 background volunteered to participate in this study. They were all undergraduate students at a university where English is the medium of instruction. The institution offers English classes to students who need language assistance, in order to minimise the possible impact of language problems on the educational quality. Students were assigned to English classes of different levels, according to their scores on the college entrance exam and the English placement test administered by the English Learning Center. Participants in this study, who were selected from these classes, consist of 50 students in low English level and 52 students in intermediate level.

Instruments

Listening recall tasks

The listening materials utilised in this study were sourced from a practice version of a large-scale standardised English test, i.e. the TOEFL Junior test (ETS 2014). The participants, university students with English proficiency levels ranging from low to intermediate, regularly engaged in academic listening activities like listening to academic lectures. The TOEFL Junior test, assessing listening skills across CEFR levels A1 to B2 (Baron and Tannenbaum 2011), was deemed suitable for the study due to its alignment with the participants' proficiency levels. Furthermore, the test's focus on the academic domain, with listening tasks that mirror activities needed to understand academic content in English, such as comprehending academic lectures (So et al. 2015), made it a relevant choice. Thus, the decision to use the TOEFL Junior test listening materials in this study was justified by their close match in difficulty level with the participants' proficiency and their applicability to the participants' academic listening context. Consequently, the audio passages of two parallel academic listening tasks with similar difficulty levels were selected for the purpose of our investigation. The first passage, denoted as 'Text 1,' introduces the kitchen of a castle in the medieval period. Meanwhile, 'Text 2' explains how camels survive in the desert.

The influence of text features on L2 listening difficulty has been noted in previous studies (Bloomfield et al. 2011; Révész and Brunfaut 2013). Therefore, an examination of the textual characteristics of two listening passages was conducted, focusing on parameters such as type-token ratio, text easability, and syntactic simplicity. This analysis was facilitated through the utilisation of two linguistic analysis tools: the Coh-Metrix Web Tool (Graesser et al. 2004) and the vocabulary profiler of Compleat Lexical Tutor (Cobb 2023). The outcomes of these analyses revealed that Texts 1 and 2 exhibited similar text features. Furthermore, a pilot test was administered to gauge potential differences in the listening difficulty level between the two passages. The results indicated no significant distinctions in the test takers' listening performance across both passages. Consequently, Texts 1 and 2, selected for parallel listening input texts due to their comparable text characteristics and difficulty levels, were deemed suitable for employment as parallel listening materials within the present study for the development of listening recall tasks.

Subsequently, these two listening texts were combined with the requirements of L1 and L2 responses. This combination resulted in the development of four distinct formats for the listening recall tasks, namely: (1) recalling Text 1 in Chinese, (2) recalling Text 1 in English, (3) recalling Text 2 in Chinese, and (4) recalling Text 2 in English.

Procedure

Task taking

After completing the informed consent process, participants listened to the test instructions presented in Chinese. They were informed that they would need to complete two English listening

recall tasks: one requiring responses in Chinese and the other in English. To familiarise participants with the format of the tasks, a practice exercise involving brief listening excerpts, which had to be recalled in English and Chinese, was administered. The actual testing began only after all participants confirmed their understanding of the test procedures and task requirements.

While it was presumed that the linguistic characteristics and difficulty levels of the listening materials, drawn from two parallel listening tasks, would have minimal impact on the listening outcomes of L2 learners, the research design featured a counterbalancing of the listening texts and response language. Despite assertions that the listening tasks of the TOEFL Junior test do not necessitate test takers' background knowledge of specific subjects or content (So et al. 2015), as an extra precaution, this counterbalancing design was employed to further minimise the potential impact of participants' familiarity with the topics of the listening materials on their performance. This measure was implemented to ensure a more precise focus on response languages within the present study. Accordingly, participants in the lower- and intermediate-level proficiency groups were randomly divided into two subgroups. Among the lower proficiency participants, one half was instructed to recall Text 1 in English and Text 2 in Chinese, while the other half was tasked with recalling Text 1 in Chinese and Text 2 in English. A similar procedure was applied to the subgroups of intermediate proficiency participants.

During the test, each audio recording was played only once without any interruptions. Participants were prohibited from taking notes during the listening process to prevent them from directly borrowing the words as it was presented in the audio. Subsequent to the end of each audio, participants were required to write down what they had heard in the language specified in the task instructions. They were encouraged to write as much as possible on the provided answer sheets. A time limit of 20 minutes was allocated to participants for the completion of both tasks.

Scoring

This study adopts the prevalent practice of scoring recall tasks by quantifying the information recalled (Brantmeier 2006). Following Carrell's (1985) work on identifying idea units, researchers of this study (including both L1 English and Chinese speakers) undertook a rigorous text analysis procedure to determine the total number of idea units in each listening passage. This yielded 17 idea units for Text 1 and 38 units for Text 2. Subsequently, the English lists of idea units were translated into Chinese. Each corresponding unit in the English and Chinese versions underwent a meticulous examination to ensure their equivalence.

Each participants' task performance was scored by two raters, who based their evaluation on the presence of idea units in participants' written responses. Language aspects related to grammar and spelling, which pertain to learners' writing proficiency but are unrelated to the construct of listening comprehension, were excluded from consideration during the scoring process. The inter-rater correlation coefficient yielded a value of .96, indicating high scoring reliability.

Given the difference in the total number of idea units across the two listening passages, participants' listening scores for each recall task were normalised by dividing them by the total number of idea units in the corresponding task's text. The scores were then expressed as percentages, facilitating a more straightforward and accurate comparison of listening performance across recall tasks employing different listening texts, in the subsequent phases of data analysis and result interpretation.

Interview

Interviews were conducted with 14 participants, comprising seven test takers at a low English proficiency level and an equivalent number at an intermediate level. The medium of communication employed during the interviews was Chinese. Adhering to the guidelines proposed by Patton (2014) on effective interview techniques, the initial focus of each interview centred around the second research question, designed deliberately in an open-ended manner. It aimed to stimulate thoughtful considerations from the participants regarding their perspectives on the utilisation of

their Chinese L1 in English L2 recall tasks. Three questions were posed to the participants, each aimed at probing their perceptions: (1) Does the use of L1 responses facilitate improve performance in listening recall tasks? (2) Are any differences discernible to you when comparing the use of L1 against L2 responses in listening recall tasks? (3) Do you have any preferences regarding the language of response when engaged in listening recall tasks? Participants' responses garnered were accompanied by an elicitation of the reasons underlying their viewpoints.

Data analysis

Task performance data were analysed using a 2×2 mixed-subjects factorial analysis of variance (ANOVA). This analysis involved two independent variables: response language, serving as the within-subject variable, and English proficiency, operating as the between-group variable. The dependent variable was participants' performance scores on listening recall tasks. It is essential to note that prior to conducting the ANOVA, the data underwent an examination to ensure that it met assumptions, such as normal distribution and homogeneity of variances. After this confirmation, the ANOVA was performed to examine both the main effect of response language and its interaction effect with participants' proficiency levels on task scores.

The data collected from interviews were transcribed and underwent subsequent analysis utilising an inductive approach. Adhering to the guidelines proposed by Braun and Clarke (2006), two coders engaged in an iterative and recursive analysis of the data. The emphasis of this analytical process was directed towards the perceptions of L2 learners pertaining to the utilisation of L1 responses within listening recall tasks. The coding procedures underwent thorough review and discussion, prior to the attainment of consensus amongst the coders.

Results

RQ1: Do L2 learners in low and intermediate proficiency perform differently on listening recall tasks between L1 and L2 responses?

To address the first research question, a factorial ANOVA was undertaken in order to compare participants' performance on listening recall tasks when responding in their L1 (Chinese) and L2 (English). The results, as presented in Table 1, demonstrate a significant main effect of response language, $F(1, 100) = 6.63$, $p < .05$, $\eta^2 = .06$. Furthermore, an interaction effect between response language and participants' L2 proficiency levels was observed, yielding statistical significance, $F(1, 100) = 5.40$, $p < .05$, $\eta^2 = .05$. These results suggest that the response language had a medium effect on participants' performance on listening recall tasks (Cohen 1988). In other words, individuals exhibited different performance when responding in English compared to Chinese. Nevertheless, it is important to note that this effect may be contingent upon participants' English proficiency levels.

To better understand this interaction effect, simple main effects tests were conducted to examine the differences in means between the two response languages (Chinese and English) for participants in the low and intermediate proficiency groups. As shown in Table 2, the results of pairwise comparisons indicate that there was no statistically significant difference in the task scores of participants at the low proficiency level when responding in Chinese versus English. However, a significant difference was seen among participants in the intermediate proficiency group (M difference = 8.86, $p < .01$). Specifically, participants achieved higher scores ($M = 42.99$, $SD = 15.37$) when they responded in Chinese, as compared to English ($M = 34.13$, $SD = 14.15$).

Table 1. ANOVA results: within-subjects and interaction effects ($N = 102$).

	F	df	p	η^2
Response language	6.63	1	.012	.06
Response language * Proficiency	5.40	1	.02	.05

Table 2. Results of pairwise comparisons of listening recall performances with Chinese and English responses across low and intermediate L2 learners.

Proficiency	<i>N</i>	Response Language	<i>M</i>	<i>SD</i>	<i>M</i> Difference	<i>p</i>
Low	50	Chinese (L1)	23.21	15.13	0.46	.86
		English (L2)	22.75	17.16		
Intermediate	52	Chinese (L1)	42.99	15.37	8.86	.001
		English (L2)	34.13	14.15		

RQ2: How do L2 learners perceive the use of L1 responses in L2 listening recall tasks?

Interview data were analysed to address the second research question. Few discernible disparities were identified between the seven low-level and seven intermediate-level learners regarding their use of L1 responses in listening recall tasks. The analysis revealed three prominent themes that encapsulate their perceptions.

Better task performances with L1 response

Participants uniformly articulated that the language used for their responses did not influence the difficulty level of the listening phase. Nevertheless, they acknowledged that writing in their L1, Chinese, proved to be markedly easier than composing responses in English. Specifically, in the task requiring Chinese responses, participants reported a process wherein they comprehended the auditory input and stored the information in Chinese language chunks. They further emphasised, given the fact that Chinese was their L1, writing down these Chinese chunks during the recall phase was a natural and seamless process. Despite occasional loss of details during the listening phase due to memory constraints, they expressed confidence in their native-level L1 writing ability, which allowed them to compensate for these omissions by infusing their responses with missing details based on the primary content chunks, ensuring the logical coherence and completeness of their written responses. In stark contrast, the task requiring English responses after listening presented some challenges, primarily stemming from their less capable English writing. Participants frequently encountered issues such as not knowing the written forms of vocabulary, coupled with the prevailing sense that they were unable to use written English to accurately articulate their listening comprehension.

Better representation of listening comprehension construct with L1 response

Participants, despite their unanimous opinions that response language did not affect listening difficulty, offered insights into potential divergences in their listening processes between the two response language conditions. When tasked with providing responses in Chinese after listening, participants reported a tendency to prioritise comprehension and holistic understanding of the narrative, storing story or idea units in Chinese. Conversely, in tasks requiring English responses, some participants exhibited a tendency for focusing on memorising the English words and the exact sentences from the audio. During the subsequent writing phase, their efforts were directed towards recalling these words or sentences they had heard and memorised. Some participants described this task as a delayed dictation exercise, wherein comprehension held lesser significance, as long as they could accurately recall the exact sentences.

The preference for L1 response

Participants admitted that response language influenced their emotional states while undertaking the recall tasks. They reported nervousness when required to compose responses in English, juxtaposed with a sense of confidence and relaxation when responding in the L1, Chinese. The differences in the affective states during the tasks exerted an influence on their overall attitudes towards the two tasks, resulting in their preference for L1 response in listening recall tasks.

Overall, interviews conducted with low and intermediate level participants yielded insights pertaining to the aforementioned three themes, indicating a convergence in their perceptions regarding the use of L1 in L2 listening recall tasks. However, it is worth noting that several low-level participants closed their interviews by saying that response language ultimately made little difference in their listening experiences. They contended that regardless of the response language, they found it challenging to extract meaningful information from the auditory input, rendering the task of comprehending the audio inherently difficult.

Discussion

The participants' descriptions in the interviews regarding the writing challenges encountered in the recall tasks with English responses, suggest that their less advanced L2 writing proficiency may affect their performance on the listening task. This introduces construct-irrelevant variance into the scores, potentially harming the validity of the task's score interpretation (Messick 1995). This description further reinforces the concern raised in the introduction, which prompted this research. It justifies the need for the present study, which aims to identify potential methods for addressing the issue of construct-irrelevant factors in listening recall tasks.

The quantitative results of this study indicated that when participants provided responses in their L1, intermediate learners exhibited better performance in the recall task. This finding is consistent with the interview data, where participants, although not considering the L2 listening comprehension phase easier, found subsequent L1 writing easier due to their native-level proficiency. This, in turn, enhanced their performance on the recall task. Therefore, both the quantitative and qualitative findings suggest the potential effectiveness of using L1 responses to help mitigate construct-irrelevant variance introduced by test takers' L2 writing proficiency, ultimately improving the task's representation of the listening construct. While these findings are specific to our intermediate-level learners, they align with prior research investigating the impact of using L1 in reading assessment (e.g. Godev, Martínez-Gibson, and Toris 2002; Gordon and Hanauer 1995; Nevo 1989) and contribute to the limited evidence on listening assessment (Filipi 2012).

In line with prior research that demonstrated that certain task characteristics, such as the task type, can influence the cognitive processes of test takers (Gordon and Hanauer 1995; Nevo 1989), the interviews conducted in this study suggest that response language may be considered one such task feature that prompts test takers to adjust their listening process. When they needed to recall audio information in their L1, they focused on comprehending the content of the L2 audio during the listening phase. Conversely, when an L2 response was required, they prioritised memorising English words and sentences without placing as much emphasis on listening comprehension. These differences in listening processes imply that when an L1 response was used, the recall task tended to measure listening comprehension, whereas an L2 response might primarily assess test takers' memory capacity rather than their listening comprehension ability.

In addition to construct validity, face validity, which concerns stakeholders' subjective perceptions of a task's ability to accurately measure a particular skill (Holden 2010), is crucial in determining whether a task needs redesigning. When test takers lack confidence in a task's ability to measure their proficiency, their motivation to perform to the best of their abilities may decrease. In this study, participants mentioned feeling more confident and at ease when providing responses in Chinese and expressed a preference for listening recall tasks with L1 responses, indicating higher face validity compared to tasks with L2 responses. Conversely, participants reported feeling anxious when providing L2 responses. Given previous research demonstrating the detrimental impact of anxiety on L2 listening comprehension (Brunfaut and Revesz 2015; Elkhafaifi 2005; Wang and Cha 2019), it is possible that participants' anxiety during English recall could have adversely affected their task performance, thus resulting in even lower scores for English responses compared to Chinese responses in this study. The findings of this study regarding participants' attitudes and emotional states towards L1 and L2 responses in listening assessments, are consistent with Filipi's

(2012) research, indicating that lower-level L2 learners tend to favour questions in their L1 during L2 listening comprehension tasks. However, this contrasts with the study by Cox, Bown, and Bell (2019), which found no language preference among L2 learners in reading assessments within the context of an L2 Russian learning course. It is worth noting that the learners in Cox et al.'s study were advanced-level, a proficiency level not examined in our study. This implies the potential impact of learners' proficiency on these findings. Overall, these mixed results highlight the complexity of L2 learners' attitudes towards L1 use in L2 assessment, likely influenced by the type of assessment, the proficiency level of L2 participants, and the specific context.

Both low and intermediate learner groups in this study held similar perceptions regarding tasks requiring different response languages, yet they exhibited differences in task performance. Specifically, intermediate learners demonstrated significantly superior performance on the listening recall task when responding in their L1 compared to their L2. Conversely, no significant difference was observed in the performances of low-proficiency participants across the two tasks. These inconsistent results between the two groups suggest that learners' L2 proficiency may influence the impact of L1 usage on task performance, which is consistent with previous research highlighting the mediating role of L1 in reading assessment (Shohamy 1984).

Additionally, prior suggestions that the advantage of L1 usage diminishes as L2 proficiency advances were challenged by the current study, as the benefit of L1 usage was observed solely for intermediate learners. One possible explanation for the absence of this benefit among low-proficiency learners can be found in their interview data. They reported difficulties in comprehending the listening material, regardless of the response language, and encountered challenges in recalling the audio information. Examination of their written responses revealed relatively short recalls (i.e. one to three words for recalling each idea unit), indicating their struggle to complete the task. Unlike previous studies on reading tasks that demonstrated the benefits of L1 usage for low-level learners (Shohamy 1984; Upton and Lee-Thompson 2001), the listening recall task in this study, without the option of note-taking, may have placed greater demands on test takers.

Moreover, the listening materials utilised in this study were sourced from TOEFL Junior listening tasks originally designed with selected response questions (i.e. multiple choice questions). Previous research suggests that selected response questions in listening assessments are generally less challenging than constructed response questions. The former offer listeners cues for the inferential process and impose lower demands on listeners' memory capacity for information processing (Cheng 2004). Therefore, while the input texts were chosen because their difficulty levels matched the proficiency levels of participants in this study, replacing their original questions with constructed response questions, i.e. written recalls in this study, may have increased task difficulty. As a result, it is possible that low-level learners in this study did not reach the necessary proficiency level to comprehend and retain substantial information from the audio for the writing phase. This is further supported by their task scores, which indicate that low-level participants could only recall approximately 22% of the idea units from the listening texts, suggesting a potential floor effect on their task performance. Consequently, their L2 writing ability had limited opportunity to impact recall outcomes and introduce construct-irrelevant variance into the scores. In such cases, the use of L1 responses, intended to deal with the issue of this construct-irrelevant factor, may not yield a discernible difference.

Conclusion and future research

In conclusion, this study underscores the concern that test takers' L2 writing ability may constitute a construct-irrelevant factor capable of impacting the construct validity and score interpretation of the L2 listening recall task. The investigation provides both quantitative and qualitative evidence showing that incorporating L1 responses in listening recall tasks improved L2 learners' performance, refined the measurement of listening comprehension ability, and aligned with L2 learners' preferences, thus having the potential of enhancing the construct and face validity of the L2

listening recall task. This consequently implies the considerable potential for utilising translanguaged practices in the improvement of language task design, thereby making a significant contribution to the field of language assessment. Furthermore, this study advances the preliminary notion of translanguaging, which accentuates the advantageous interplay among an individual's linguistic repertoires (Lewis, Jones, and Baker 2012).

Nevertheless, it is imperative to acknowledge that the aforementioned findings are confined to intermediate L2 learners within this study, thereby underscoring the necessity for further exploration into the influence of L2 proficiency on the effectiveness of employing translanguaged practices in language assessment. Additionally, this study acknowledges several other limitations that offer valuable insights for prospective research endeavours. Firstly, the written products of participants can furnish more intricate insights into learners' test-taking behaviours, rather than being roughly reviewed in this study to potentially illuminate the underlying reasons behind the observed quantitative outcomes among low proficiency learners. Thus, future research may consider the inclusion of additional qualitative sources, such as the in-depth examination of test takers' written products. Secondly, the interview data in this study implies that employing different response languages in listening assessments may lead to distinct listening processes among test takers. However, this aspect was not the focus of the current study and thus warrants further investigation through the adoption of a broader array of qualitative and quantitative methodologies, including questionnaires, to acquire comprehensive and large-scale data that can facilitate the understanding of potential differences in test takers' listening processes and strategy uses. Thirdly, it is essential to note that this study exclusively examined one particular type of listening assessment task (the recall task), and one specific translanguaged practice (L1 response). This limited scope restricts the generalizability of the findings. Future research may extend their inquiry to encompass various listening task types, such as open-ended questions and multiple-choice questions, which were previously examined primarily within the context of reading assessment. Additionally, diverse translanguaged practices, including permitting test takers to freely select and utilise their preferred languages, should also be investigated. In summary, the utilisation of translanguaged practices within the domain of language assessment remains a promising area warranting further exploration.

Disclosure statement

No potential conflict of interest was reported by the author(s).

ORCID

Shelly Xueting Ye  <http://orcid.org/0000-0002-3798-1944>

References

- Adhikari, B. R., and P. P. Poudel. 2024. "Countering English-Prioritised Monolingual Ideologies in Content Assessment Through Translanguaging Practices in Higher Education." *Language and Education* 38 (2): 155–172. <https://doi.org/10.1080/09500782.2023.2217804>.
- Alderson, J. C. 2005. *Assessing Reading*. Cambridge: Cambridge University Press.
- Baccouche, E., C. Castro, E. Messman, and L. Zakletskaia. 1996, March 23. "How Comprehensible is Input When Attention is Focused on Form?" *AAAL1996 Conference*, Chicago, IL. <https://eric.ed.gov/?id=ED396532>.
- Bachman, L. F. 1990. *Fundamental Considerations in Language Testing*. New York: Oxford university press.
- Baker, C. 2011. *Foundations of Bilingual Education and Bilingualism*. Bristol: Multilingual Matters.
- Baron, P., and R. Tannenbaum. 2011. "Mapping the TOEFL® Junior™ test onto the Common European Framework of Reference." ETS Research Memorandum (RM-11-07). Educational Testing Service. <https://www.ets.org/Media/Research/pdf/RM-11-07.pdf>.
- Bloomfield, A., S. C. Wayland, E. Rhoades, A. Blodgett, J. Linck, and S. Ross. 2011. *What Makes Listening Difficult? Factors Affecting Second Language Listening Comprehension*. College Park, MD: University of Maryland Center for Advanced Study of Language.

- Brantmeier, C. 2006. "The Effects of Language of Assessment and L2 Reading Performance on Advanced Readers' Recall." *The Reading Matrix* 6 (1): 1–17.
- Braun, V., and V. Clarke. 2006. "Using Thematic Analysis in Psychology." *Qualitative Research in Psychology* 3 (2): 77–101. <https://doi.org/10.1191/1478088706qp0630a>.
- Brunfaut, T., and A. Revesz. 2015. "The Role of Task and Listener Characteristics in Second Language Listening." *Tesol Quarterly* 49 (1): 141–168. <https://doi.org/10.1002/tesq.168>.
- Buck, G. 2001. *Assessing Listening*. Cambridge: Cambridge University Press.
- Carrell, P. L. 1985. "Facilitating ESL Reading by Teaching Text Structure." *TESOL Quarterly* 19 (4): 727–752. <https://doi.org/10.2307/3586673>.
- Cenoz, J., and D. Gorter. 2019. "Multilingualism, Translanguaging, and Minority Languages in SLA." *The Modern Language Journal* 103 (S1): 130–135. <https://doi.org/10.1111/modl.12529>.
- Cenoz, J., and D. Gorter. 2022. "Enhancing Bilingual Resources in Third Language Acquisition: Towards Pedagogical Translanguaging." *Educational Linguistics* 1 (2): 338–357. <https://doi.org/10.1515/eduling-2022-0009>.
- Cheng, H. F. 2004. "A Comparison of Multiple-Choice and Open-Ended Response Formats for the Assessment of Listening Proficiency in English." *Foreign Language Annals* 37 (4): 544–553. <https://doi.org/10.1111/j.1944-9720.2004.tb02421.x>.
- Chu, C. 2017. "Translanguaging in Reading Comprehension Assessment: Implications on Assessing Literal, Inferential, and Evaluative Comprehension among ESL Elementary Students in Taiwan." *NYS TESOL Journal* 4 (2): 20–34.
- Cobb, T. 2023. "Compleat Lexical Tutor [Computer software]." <https://www.lex tutor.ca/>.
- Cohen, J. 1988. *Statistical Power Analysis for the Behavioral Sciences*. Hillsdale, NJ: Lawrence Erlbaum Associates.
- Cook, V. 2001. "Using the First Language in the Classroom." *The Canadian Modern Language Review* 57 (3): 402–423. <https://doi.org/10.3138/cmlr.57.3.402>.
- Cox, T. L., J. Bown, and T. R. Bell, eds. 2019. "In Advanced L2 Reading Proficiency Assessments, Should the Question Language Be in the L1 or the L2?: Does It Make a Difference?" In *Foreign Language Proficiency in Higher Education*, Vol. 37. Springer International Publishing. <https://doi.org/10.1007/978-3-030-01006-5>.
- Du, Q., J. W. Lee, and S. Sok. 2020. "Using China English, Creating Translingual Space." *World Englishes* 39 (2): 275–285. <https://doi.org/10.1111/weng.12460>.
- Elkhafafi, H. 2005. "Listening Comprehension and Anxiety in the Arabic Language Classroom." *The Modern Language Journal* 89 (2): 206–220. <https://doi.org/10.1111/j.1540-4781.2005.00275.x>.
- ETS. 2014. *Official Guide for the TOEFL Junior Standard Test*. Princeton, NJ: Educational Testing Service.
- Filipi, A. 2012. "Do Questions Written in the Target Language Make Foreign Language Listening Comprehension Tests More Difficult?" *Language Testing* 29 (4): 511–532. <https://doi.org/10.1177/0265532212441329>.
- Galante, A. 2020. "Translanguaging for Vocabulary Development: A Mixed Methods Study with International Students in a Canadian English for Academic Purposes Program." In *Envisioning TESOL Through a Translanguaging Lens: Global Perspectives*, edited by Z. Tian, L. Aghai, P. Sayer, and J. L. Schissel, 293–328. Springer International Publishing. https://doi.org/10.1007/978-3-030-47031-9_14.
- García, O. 2011. "Theorizing Translanguaging for Educators." In *Translanguaging [A CUNY-NYSIEB Guide for Educators]*, edited by C. Celic and K. Seltzer, 1–6. New York: CUNY-NYSIEB.
- García, O., S. Johnson, and K. Seltzer. 2017. *The Translanguaging Classroom: Leveraging Student Bilingualism for Learning*. Philadelphia: Caslon.
- García, O., and R. Otheguy. 2020. "Plurilingualism and Translanguaging: Commonalities and Divergences." *International Journal of Bilingual Education and Bilingualism* 23 (1): 17–35. <https://doi.org/10.1080/13670050.2019.1598932>.
- Godev, C. B., E. A. Martínez-Gibson, and C. C. M. Toris. 2002. "Foreign Language Reading Comprehension Test: L1 Versus L2 in Open-Ended Questions." *Foreign Language Annals* 35 (2): 202–221. <https://doi.org/10.1111/j.1944-9720.2002.tb03155.x>.
- Gordon, C. M., and D. Hanauer. 1995. "The Interaction between Task and Meaning Construction in EFL Reading Comprehension Tests." *TESOL Quarterly* 29 (2): 299–324. <https://doi.org/10.2307/3587626>.
- Graesser, A. C., D. S. McNamara, M. M. Louwerse, and Z. Cai. 2004. "Coh-Metrix: Analysis of Text on Cohesion and Language." *Behavior Research Methods, Instruments, & Computers* 36 (2): 193–202. <https://doi.org/10.3758/BF03195564>.
- Holden, R. R. 2010. "Face Validity." In *The Corsini Encyclopedia of Psychology*, 1–2. John Wiley & Sons, Ltd. <https://doi.org/10.1002/9780470479216.corpsy0341>.
- Jaspers, J., and L. M. Madsen. 2016. "Sociolinguistics in a Linguagised World: Introduction." *Applied Linguistics Review* 7 (3): 235–258. <https://doi.org/10.1515/applirev-2016-0010>.
- Jessner, U., and V. Török. 2017. "Strategies in Multilingual Learning: Opening new Research Avenues." In *Future Research Directions for Applied Linguistics*, edited by J. Navrascics and S. E. Pfenninger, 192–214. Multilingual Matters. <https://doi.org/10.21832/9781783097135-011>.
- Karges, K., M. Barras, and P. Lenz. 2022. "Assessing Young Language Learners' Receptive Skills: Should we ask the Questions in the Language of Schooling?" In *Current Research Into Young Foreign Language and English as an Additional Language Learners' Literacy Skills*, edited by S. Frisch and J. Rymarczyk, 83–112. Berlin: Peter Lang.

- Koo, J., and E.-J. Lee. 2020. "The Effects of Differing Language Test Presentation Methods on EFL College Students' Listening Test Performance." *Journal of Pan-Pacific Association of Applied Linguistics* 24 (2): 115–133. <https://doi.org/10.25256/PAAL.24.2.6>.
- Lee, J. F. 1986. "On the Use of the Recall Task to Measure L2 Reading Comprehension." *Studies in Second Language Acquisition* 8 (2): 201–211. <https://doi.org/10.1017/S0272263100006082>.
- Leonet, O., J. Cenoz, and D. Gorter. 2020. "Developing Morphological Awareness Across Languages: Translanguaging Pedagogies in Third Language Acquisition." *Language Awareness* 29 (1): 41–59. <https://doi.org/10.1080/09658416.2019.1688338>.
- Lewis, G., B. Jones, and C. Baker. 2012. "Translanguaging: Origins and Development from School to Street and Beyond." *Educational Research and Evaluation* 18 (7): 641–654. <https://doi.org/10.1080/13803611.2012.718488>.
- Li, W. 2018. "Translanguaging as a Practical Theory of Language." *Applied Linguistics* 39 (1): 9–30. <https://doi.org/10.1093/applin/amx039>.
- Lyster, R. 2019. "Translanguaging in Immersion: Cognitive Support or Social Prestige?" *The Canadian Modern Language Review* 75 (4): 340–352. <https://doi.org/10.3138/cmlr.2019-0038>.
- MacSwan, J. 2017. "A Multilingual Perspective on 'Translanguaging.'" *American Educational Research Journal* 54 (1): 167–201. <https://doi.org/10.3102/0002831216683935>.
- Martin-Beltrán, M. 2014. "What Do You Want to Say?" How Adolescents Use Translanguaging to Expand Learning Opportunities." *International Multilingual Research Journal* 8 (3): 208–230. <https://doi.org/10.1080/19313152.2014.914372>.
- Messick, S. 1995. "Standards of Validity and the Validity of Standards in Performance Assessment." *Educational Measurement: Issues and Practice* 14 (4): 5–8. <https://doi.org/10.1111/j.1745-3992.1995.tb00881.x>.
- Mihara, K. 2015. "An Analysis of the Differences among L2 Listening Comprehension Test Formats." *Language Testing in Asia* 5 (1): 12. <https://doi.org/10.1186/s40468-015-0021-5>.
- Mori, J., and K. K. Sanuth. 2018. "Navigating between a Monolingual Utopia and Translingual Realities: Experiences of American Learners of Yorùbá as an Additional Language." *Applied Linguistics* 39 (1): 78–98. <http://dx.doi.org/10.1093/applin/amx042>.
- Musk, N. 2010. "Code-switching and Code-Mixing in Welsh Bilinguals' Talk: Confirming or Refuting the Maintenance of Language Boundaries?." *Language, Culture and Curriculum* 23 (3): 179–197. <https://doi.org/10.1080/07908318.2010.515993>.
- Mwinda, N., and D. W. C. Van. 2015. "From 'English-Only' to Translanguaging Strategies: Exploring Possibilities." *Per Linguam* 31 (3): 100–118. <https://doi.org/10.5785/31-3-620>.
- Nevo, N. 1989. "Test-taking Strategies on a Multiple-Choice Test of Reading Comprehension." *Language Testing* 6 (2): 199–215. <https://doi.org/10.1177/026553228900600206>.
- Otheguy, R., O. García, and W. Reid. 2019. "A Translanguaging View of the Linguistic System of Bilinguals." *Applied Linguistics Review* 10 (4): 625–651. <https://doi.org/10.1515/applirev-2018-0020>.
- Palmer, D. K., R. A. Martínez, S. G. Mateus, and K. Henderson. 2014. "Reframing the Debate on Language Separation: Toward a Vision for Translanguaging Pedagogies in the Dual Language Classroom." *The Modern Language Journal* 98 (3): 757–772. <https://doi.org/10.1111/modl.12121>.
- Patton, M. Q. 2014. *Qualitative Research & Evaluation Methods: Integrating Theory and Practice*. Thousand Oaks, CA: Sage publications.
- Qureshi, M. A., and A. Aljanadbah. 2022. "Translanguaging and Reading Comprehension in a Second Language." *International Multilingual Research Journal* 16 (4): 247–257. <https://doi.org/10.1080/19313152.2021.2009158>.
- Rahmah, I. P., and S. S. Halimi. 2020. "A Comparison of Using L1 (Indonesian) and English to Answer English Listening Comprehension Questions." *Jet (Journal of English Teaching) Adi Buana* 5 (02): Article 02. <https://doi.org/10.36456/jet.v5.n02.2020.2612>.
- Révész, A., and T. Brunfaut. 2013. "Text Characteristics of Task Input and Difficulty in Second Language Listening Comprehension." *Studies in Second Language Acquisition* 35 (1): 31–65. <https://doi.org/10.1017/S0272263112000678>.
- Sakai, T. 2012. "Evaluation with Informational and Navigational Intents." *Proceedings of the 21st International Conference on World Wide Web*, 499–508. <https://doi.org/10.1145/2187836.2187904>.
- Shohamy, E. 1984. "Does the Testing Method Make a Difference? The Case of Reading Comprehension." *Language Testing* 1 (2): 147–170. <https://doi.org/10.1177/026553228400100203>.
- So, Y., M. K. Wolf, M. C. Hauck, P. Mollaun, P. Rybinski, D. Tumposky, and L. Wang. 2015. "TOEFL Junior® Design Framework." *ETS Research Report Series* 2015 (1): 1–45. <https://doi.org/10.1002/ets2.12058>.
- Sun, Y., and G. Lan. 2021. "Enactment of a Translingual Approach to Writing." *TESOL Quarterly* 55 (2): 398–426. <https://doi.org/10.1002/tesq.609>.
- Tai, K. W., and W. Li. 2021. "Co-learning in Hong Kong English Medium Instruction Mathematics Secondary Classrooms: A Translanguaging Perspective." *Language and Education* 35 (3): 241–267. <https://doi.org/10.1080/09500782.2020.1837860>.
- Upton, T. A., and L.-C. Lee-Thompson. 2001. "The Role of the First Language in Second Language Reading." *Studies in Second Language Acquisition* 23 (4): 469–495. <https://doi.org/10.1017/S0272263101004028>.

- Vaish, V. 2019. "Translanguaging Pedagogy for Simultaneous Biliterates Struggling to Read in English." *International Journal of Multilingualism* 16 (3): 286–301. <https://doi.org/10.1080/14790718.2018.1447943>.
- VanPatten, B. 1990. "Attending to form and Content in the Input: An Experiment in Consciousness." *Studies in Second Language Acquisition* 12 (3): 287–301. <https://doi.org/10.1017/S0272263100009177>.
- Velasco, P., and O. Garcia. 2014. "Translanguaging and the Writing of Bilingual Learners." *Bilingual Research Journal* 37 (1): 6–23. <https://doi.org/10.1080/15235882.2014.893270>.
- Wang, S. Y., and K. W. Cha. 2019. "Foreign Language Listening Anxiety Factors Affecting Listening Performance of Chinese EFL Learners." *The Journal of AsiaTEFL* 16 (1): 121–134. <https://doi.org/10.18823/asiatefl.2019.16.1.8.121>.
- Williams, C. 1994. "Arfarniad o ddulliau dysgu ac addysgu yng nghyd-destun addysg uwchradd ddwyieithog—ProQuest [Bangor University]." https://www.proquest.com/openview/037459c4eede79ac6ff9de7eed57cc8b/1?pq-origsite=gscholar&cbl=51922&diss=y&casa_token=BeTW-zZ35lQAAAAA:-URMxnPprXEEFOyTdlfkn9LTqdezGNwwl-7FaR4YxxCZAMDDLEV2O16unxG_oPdEvlZraKuU.
- Yu, G. 2008. "Reading to Summarize in English and Chinese: A Tale of two Languages?" *Language Testing* 25 (4): 521–551. <https://doi.org/10.1177/0265532208094275>.