University of Reading

Explaining listening comprehension among L2 learners of English: the contribution of general language proficiency, vocabulary knowledge and metacognitive awareness

Article

Accepted Version

Creative Commons: Attribution-Noncommercial-No Derivative Works 4.0

Wang, Y. and Treffers-Daller, J. ORCID: https://orcid.org/0000-0002-6575-6736 (2017) Explaining listening comprehension among L2 learners of English: the contribution of general language proficiency, vocabulary knowledge and metacognitive awareness. System, 65. pp. 139-150. ISSN 0346-251X doi: 10.1016/j.system.2016.12.013 Available at https://centaur.reading.ac.uk/68862/

It is advisable to refer to the publisher's version if you intend to cite from the work. See Guidance on citing.

Published version at: http://www.sciencedirect.com/science/article/pii/S0346251X16304560

To link to this article DOI: http://dx.doi.org/10.1016/j.system.2016.12.013

Publisher: Elsevier

All outputs in CentAUR are protected by Intellectual Property Rights law, including copyright law. Copyright and IPR is retained by the creators or other

University of Reading

copyright holders. Terms and conditions for use of this material are defined in the End User Agreement.

www.reading.ac.uk/centaur

CentAUR

Central Archive at the University of Reading

Reading's research outputs online

Explaining listening comprehension among L2 learners of English: the contribution of general language proficiency, vocabulary knowledge and metacognitive awareness

Wang, Yun[a/b] and Jeanine Treffers-Daller[b]

a Foreign Language Department, Northwest A&F University 53#, 22nd Xinong Road, Yangling, Shaanxi, 712100, China

b Institute of Education, London Road Campus, 4 Redlands Road, University of Reading, Reading, Berkshire, RG1 5EX, UK

Abstract

Listening comprehension constitutes a major problem for second language learners but little is known about the relative contribution of different factors to listening comprehension. Since there are still only very few studies in this area by comparison with studies focusing on the relationship between reading and vocabulary, there is a need for studies which can fill the gap in our knowledge about the specific contribution of generic and discrete-point measures of language ability to explaining listening. In the present study among 151 non-English major students at a university in Northwest China we explore what proportion of the variance in listening comprehension is explained by general language proficiency, vocabulary size and metacognitive awareness. Our results show that vocabulary size is the strongest predictor, followed by general language proficiency, while metacognitive awareness is less important. We discuss implications for the componential structure of the notion language ability, theories of listening and pedagogical practice in L2 classrooms.

Keywords: Listening comprehension, Receptive vocabulary knowledge, General language proficiency, Metacognitive awareness

1. Introduction

Listening comprehension is a complicated process because it is based on linguistic knowledge, such as vocabulary or grammar, as well as non-linguistic information, such as knowledge of the world (Buck, 2001; Vandergrift & Baker, 2015). Although there is now a considerable body of literature on listening comprehension among non-native speakers, explaining individual differences in listening has not been high

on the agenda by comparison with studies which look into individual differences in reading comprehension (Andringa, Olsthoorn, van Beuningen, Schoonen, & Hulstijn, 2012). A better understanding of the listener characteristics which determine L2 learners' success in understanding speech is urgently needed because, as pointed out by Graham (2011), "listening is a source of frustration to learners and an area in which it seems difficult to make progress" (p. 113).

This study addresses the gap in our understanding of the learner variables which impact on L2 learners' listening comprehension, by providing empirical evidence about the extent to which L2 listeners' general language proficiency, vocabulary knowledge and metacognitive awareness can explain the variance in L2 learners' listening comprehension. Studying listening comprehension is particularly relevant in the Chinese context because Chinese L2 learners often find it hard to understand native English speakers (Goh, 2000). This is not only due to typological differences between the languages, but also to important differences between the cultures of the source and the target language, which makes it more difficult for these learners to make strategic use of contextual information in the process of listening. The few available studies among Chinese listeners all point out that there is important variability among Chinese learners in their ability to understand English and that Chinese learners of English find it very challenging to improve their listening comprehension (He, 2005; He & Bao, 2006; Long & Zhao, 2009). This makes it very important to help identify the causes of the problems they encounter so that interventions can be based on solid empirical evidence.

2. Explaining listening comprehension

2.1 Factors contributing to listening comprehension

In a seminal paper, Rubin (1994) mentions five groups of factors which affect L2 learners' listening comprehension, namely text characteristics, interlocutor characteristics, task characteristics, listener characteristics and process characteristics. In the current study we will focus on individual differences in listener characteristics, as these have received little attention by comparison with studies on individual differences in reading comprehension (Andringa et al. 2012). The factors that have been studied in relation to listener characteristics include general language proficiency (J. Zuo, 2013), vocabulary knowledge (Bonk, 2000; Staehr, 2009; Vandergrift & Baker, 2015), listening strategy use (Graham, Santos, & Vanderplank, 2008), metacognitive awareness (Vandergrift, Goh, Mareschal, & Tafaghodtari, 2006; Vandergrift & Tafaghodtari, 2010; X. Y. Zuo, 2013), working memory and processing speed (Andringa et al., 2012). While we cannot do justice to all these factors in the current study, it is important to try and disentangle the contribution of a number of important factors which impact on L2 listeners' listening comprehension performance. In our choice of variables for the study we have been guided by Bachman and Palmer's (2010) model of communicative language ability. In their conceptual framework language ability is defined as "the capacity that enables language users to create and interpret discourse" (p. 33). The fact that the interpretation of discourse is explicitly mentioned in the definition of language ability

makes this an appropriate definition for a study on listening comprehension. We have chosen Bachman and Palmer's model for our study because it does not just narrowly focus on language but also on the strategic knowledge that is needed to use language in everyday life. Bachman and Palmer specify that language ability comprises language knowledge and strategic competence. Among other components, language knowledge covers grammatical knowledge (which includes vocabulary, syntax and phonology) whilst strategic knowledge covers higher-order metacognitive strategies that provide a management function in language use, as well as in other cognitive activities. In the current study we will focus on the contribution language knowledge and strategic competence make to listening comprehension among Chinese L2 learners of English. In the following sections we will briefly discuss the constructs of language knowledge and of strategic competence. Their operationalisation will be presented in the methods section.

2.2 Language knowledge

According to Bachman and Palmer's (2010) widely used model, language ability is a multidimensional construct. Most researchers nowadays share this view and there is little support for the unitary competence hypothesis associated with the work of Oller (1973). However, there is considerable evidence from the field of corpus linguistics that grammar and vocabulary cannot so easily be separated (Hunston & Francis (2000). This is also the view which emerges from Halliday's (1994) Systemic Functional Grammar. Halliday introduced the concept of lexicogrammar because

"grammar and vocabulary are merely different ends of the same continuum – they are the same phenomenon as seen from opposite perspectives" (p. 15). Under such a view holistic assessments are more likely to give a valid indication of the competencies of language users. A renewed interest in what unites the different abilities under the overarching construct is also emerging in the field of language testing. On the basis of a review of the factor structure of language test scores, Harsch (2014) argues, for example, that "language proficiency can be conceptualised as unitary *and* divisible, depending on the level of abstraction and the purpose of the assessment and score reporting" (p. 153). We support this view and assume that it is of interest to include both holistic and discrete-point measures of language proficiency in a study of listening comprehension, so that we can investigate to what extent discrete-point measures of specific components of language ability such as vocabulary tests are able to explain variance in listening comprehension over and above generic measures of language ability. To the best of our knowledge this has not been attempted so far.

Among the available discrete-point measures of language ability we focus on vocabulary knowledge, as many studies show that L2 vocabulary knowledge is one of the most important predictors of listening comprehension in adult second language learners (Andringa et al., 2012; Bonk, 2000; Kelly, 1991; Mecartty, 2000; Staehr, 2009; Vandergrift & Baker, 2015). In a study of the relative contribution of grammar and vocabulary in explaining variance in listening comprehension Mecartty (2000) found that vocabulary but not grammar was a significant predictor of listening comprehension in non-native speakers. This is possibly the case because in processing

non-native speakers are mainly guided by lexical and semantic cues but not by syntactic cues (Clahsen & Felser, 2006).

Since Nation (2006) holds that a smaller vocabulary is needed to understand spoken language than written language, it is possible that vocabulary is less important in listening than in reading. In the comparisons of the contribution of vocabulary knowledge to listening and reading comprehension, Mecartty (2000) and Mehrpour and Rahimi (2010) do, in fact, reach the same conclusion that vocabulary knowledge influences learners' performance on reading comprehension more than on listening comprehension.

Further evidence for the differences between the lexical coverage needed for listening and reading can be found in van Zeeland and Schmitt (2013) who argue that lexical coverage for listening comprehension depends on the degree of comprehension desired and a lexical coverage target of 95% is suggested for adequate listening comprehension (see also Nation, 2006). In order to have a relatively good L2 listening comprehension at 95% coverage, knowledge of 2000 to 3000 of the most frequent word families is required (Adolphs & Schmitt, 2003; van Zeeland & Schmitt, 2013), which is much less than for reading. According to Laufer and Ravenhorst-Kalovski (2010), 4000-5000 word families are needed to reach a 95% coverage in reading while 8000-9000 are needed for 98% coverage (Nation, 2006).

While there appears to be a consensus that vocabulary knowledge plays a very important role in listening comprehension, how much of the variance in listening comprehension is explained by vocabulary knowledge differs per study. An added

difficulty is that the range of predictor variables included in different studies and tests used to measure vocabulary differ widely, as well as the ways in which listening is measured, which makes it difficult to compare models. Table 1 gives an overview of a number of studies which provide information about the correlations between L2 vocabulary and listening and the proportion of the variance explained by (aspects of) L2 vocabulary knowledge.

==== Table 1 approximately here==========

The correlations found range from .38 (Mecartty, 2000) to .73 (Matthews & Cheng, 2015). Interestingly these correlations are slightly lower than those reported for L2 reading and vocabulary knowledge. According to Staehr (2008), correlations between L2 reading and vocabulary knowledge range from .50 to .85 in different studies. Vocabulary knowledge can explain as much as 72% (Staehr, 2008) in L2 reading comprehension, although it is not clear to what extent this figure refers to variance *uniquely* explained by vocabulary knowledge. Staehr (2008) suggests that one of the reasons why the correlations between listening and vocabulary are generally lower is that understanding spoken input places heavy demands on the online processing system. Therefore, L2 users need to rely on strategic competence more frequently during online aural processing than during reading. Staehr (2008) also suggests that if he had used a measure of phonological vocabulary size or aural vocabulary, such as AuralLex (A_Lex) (Milton & Hopkins, 2005), stronger

correlations would have been found. A_Lex is an aural test of receptive vocabulary in which learners are presented with 120 words, randomly selected from the first five 1000 word frequency layers. It differs from X_Lex (Meara & Milton, 2003) in that the words are presented in aural form rather than in written form. The task for learners is to indicate whether they know these words or not. A correction for guessing is built into the test through the inclusion of 20 pseudowords which sound like English words but do not exist in the language. In a comparison of results obtained with X_Lex and A_Lex among Arabic and Greek learners of English, Milton and Hopkins (2006) found that aural vocabulary sizes are generally smaller than written vocabulary sizes although correlations of .68 between the two types of vocabulary knowledge were found. Milton and Hopkins conclude that learners' orthographic vocabularies are larger than their phonological vocabularies, although it is in theory possible that A_Lex was more difficult than X_Lex or that respondents knew the words in the A_Lex test but had difficulties in decoding the aural input. Milton, Wade and Hopkins (2010) provide important evidence for the validity of A_Lex in their study of the contribution of aural and written receptive vocabulary to IELTS scores. They found a correlation of .67 between A_Lex and IELTS listening scores, whilst X_Lex scores correlated significantly but less strongly (.48) with IELTS listening scores. While A_Lex measures recognition of isolated words, Matthews and Cheng (2015) studied word recognition from connected speech (WRS) and listening comprehension. In their study WRS correlated strongly and positively with IELTS listening scores (r = .73, p < .01). Knowledge of words from the first and third thousand frequency ranges

predicted 54% of the variance in listening comprehension, with measures of words from the third thousand frequency range alone predicting 52% of this variance.

The summary of the literature given above clearly shows that the mode in which vocabulary knowledge is tested matters but it is not the only variable that needs to be taken into account. Vocabulary knowledge encompasses a wide range of components which can be broadly grouped under form, meaning and use (Nation, 2001). There are no tests that can assess all the different components of word knowledge, so choices will need to be made as to the components that are most relevant for a particular study. Some researchers make a distinction between vocabulary size, that is the number of words for which language users know some aspects of their meanings (Anderson & Freebody, 1981), and vocabulary depth, that is "the quality of the knowledge" (p. 92-93) s/he has, or put differently, the number of different aspects of form, meaning and use a learner knows. For listening, Staehr (2009) found a correlation of .7 with vocabulary size and of .65 with vocabulary depth, although vocabulary size was much more important, accounting for 49% of the 51% of the variance. Similar results were obtained by M. L. Liu (2011) who reports a correlation of .54 with vocabulary size and .60 with vocabulary depth, but does not explain how much variance was uniquely explained by either of these. Further research into the contribution of specific components of language ability (and in particular the role of vocabulary in listening comprehension) is therefore urgently needed.

2.3 Strategic competence

As pointed out in the introduction, it is not sufficient for listeners to rely exclusively on language knowledge to comprehend linguistic input. Skilled listeners and readers also integrate contextual knowledge and general knowledge about the world into their understanding of speech or writing. Indeed, as Staehr (2009) puts it, learners who are too focused on decoding the incoming speech signal "deprive themselves of the opportunity to compensate for lack of linguistic knowledge through the use of cognitive and metacognitive strategies" (p. 581). For the purposes of the current study, we define metacognition as "the awareness learners have about their general academic strengths and weaknesses, cognitive resources they can apply to meet the demands of particular tasks, and their knowledge about how to regulate engagement in tasks to optimize learning processes and outcomes" (Winne & Perry, 2000, p. 533). Despite the lack of clarity surrounding the meaning and use of the term strategy (see Dörnyei & Ryan, 2015 for a detailed critique), several studies have shown that metacognitive strategies are important in reading and listening comprehension. For example, in their study of reading comprehension among native speakers of Dutch who were learning English as a foreign language, Schoonen, Hulstijn and Bossers (1998) used a questionnaire to tap into students' metacognitive knowledge (defined as self knowledge, task knowledge, strategic knowledge and knowledge about plans and goals) and found that metacognition was more important for L2 listening comprehension among higher level students than for lower level students. Many authors make use of Vandergrift et al.'s (2006) Metacognitive Awareness Listening Questionnaire (MALQ) to measure the impact of metacognition on listening

comprehension. In their study among 226 learners of French in Canada and 115 learners of English in Iran, Vandergrift et al. demonstrated that there are five factors underlying the MALQ (problem-solving, planning and evaluation, translation, person knowledge, and directed attention) and that the scores on the MALQ could account for thirteen percent of the variance in listening scores. In their most recent study Vandergrift and Baker (2015) found person knowledge was the most important one among the metacognitive factors, although problem solving skills were also important for a subgroup in their study.

Some support for the importance of metacognition in listening can be obtained from studies which looked into the correlation between these two variables. X. Y. Zuo (2013), for example, using a questionnaire adapted from Vandergrift (1997), Vandergrift et al. (2006) and other researchers' questionnaires, found fairly strong and significant correlations between listening and metacognitive awareness (r = .52, p < .05), whereas the correlations between cognitive strategies and listening and social/affective strategies and listening were not significant. Vogely's (1995) study shows that L2 learners' perceived strategy use and the effectiveness of their strategy use differ among participants between semesters: the first-semester students perceive themselves to be the most strategic listeners and the correlation between their self-reported strategy use and the performance on listening tasks is moderately strong (r = .52, p < .05), followed by a noticeably stronger correlation (r = .42, p < .05) among the combined third- and fourth-semester students. The weakest correlation (r = .28, p < .05) appeared in the group of second semester students who perceived

themselves as listeners using the least strategies among the three groups. In Kassem (2015), learners' metacognitive awareness was measured through the Listening Strategy Questionnaire which was constructed on the basis of O'Malley and Chamot's (1990) category of learning strategies: metacognitive, cognitive and social/affective strategies. Kassem (2015) indicates that learners' self-reported overall listening strategy use is significantly correlated with their listening comprehension (r = .62, p < .01) and the frequency of use of the three categories is also significantly correlated with listening comprehension (r = .62, p < .01, for metacognitive; r = .60, p < .01, for cognitive; and r = .25, p < .05, for social/affective strategies).

Andringa et al. (2012) took a different approach to the study of the role of metacognitive abilities in listening comprehension. They were interested in the role of metacognitive abilities as measured by the complex matrices component of the Wechsler Adult Intelligence Scale (Wechsler, 1997). Interestingly, the authors found that IQ plays a more important role in listening comprehension among non-native speakers than among native speakers, probably because the task is more demanding for non-native speakers: the latter therefore need to rely more on meta-cognitive reasoning abilities in online comprehension of discourse (see also Grzegorzewska, 2015, for a detailed study of the relationship between intelligence and strategy use).

The findings reported by Vandergrift et al. (2006), Vandergrift and Baker (2015) and Staehr (2009) and also to some extent those of Andringa et al. (2012) suggest that metacognition is indeed an important variable, although the percentage of explained variance in these is generally not very high.

2.4 Summary

The available literature shows that both linguistic factors and strategic knowledge play an important role in listening comprehension, although the former appears to contribute more than the latter. As Vandergrift (2006) points out, we need more studies which disentangle the contribution of language-specific knowledge (e.g. vocabulary and general language proficiency) and general knowledge about listening (e.g. metacognitive knowledge) so that the relative contribution of each learner variable can be established.

The present study specifically explores the impact of individual differences in language knowledge and metacognitive awareness of strategies used in the listening process while keeping text characteristics, interlocutor characteristics and task characteristics constant. The research question we aim to answer is:

> What proportion of the variance in L2 learners' listening comprehension is explained by general L2 language proficiency, L2 vocabulary knowledge and metacognition?

3. Methodology

3.1 Participants

The sample consisted of 172 students from three classes of second year non-English major students at a university in Northwest China. All non-English major students at that university were taught English listening comprehension based on the National

College English Teaching Syllabus (The Ministry of Chinese Education, 2007) and all students were invited to take part. The data of only 151 participants could be used in the study because not all students completed all tests. The participants were between 18-21 years old. It is important to know that the participants seldom have contact with native speakers in authentic listening situations. Outside English class, these participants mainly practice listening comprehension through listening to materials of English tests.

3.2 Instruments

General language proficiency test

Students took the Oxford Quick Placement Test (OQPT), which is widely used in research to measure general language proficiency because it gives information about students' language ability in relation to the Common European Framework of Reference for Languages (CEFR). The OQPT is a written test which consists of 60 items. For Items 1-5, learners are required to understand notices and decide where they can see these notices. For Items 6-20 and Items 41-50, five short passages with blanks in the texts are given and learners are required to choose the word or phrase which best fits each space from the three or four given answer options. For Items 21-40 and Items 51-60, one sentence with a blank is given for each item and learners must choose the word or phrase which best complete each sentence. The OQPT measures L2 learners' English vocabulary knowledge such as word meanings, collocations, synonyms and antonyms, and phrases, and learners' grammar knowledge

such as tense, passive voice and counterfactual knowledge. The maximum score students can obtain is 60. The test results (see Table 4 for details) showed students' English language ability varied widely (M = 34.70, SD = 6.31). One student was at A1 level, 28 students were at A2 level, 86 students were at B1 level, 31 students were at B2 level, 5 students were at C1 level. No one had attained the C2 level in English.

Vocabulary size test

The Vocabulary Size Test (Nation & Beglar, 2007) was used to measure participants' vocabulary size. This receptive vocabulary test is widely used in studies of L1 and L2 acquisition and contains vocabulary drawn from different frequency layers in English, as established on the basis of the British National Corpus. The 14000 word version contains 140 multiple-choice items, with 10 items from each 1000 word family level. The number of items that a test taker chooses correctly is multiplied by 100 to compute students' total receptive vocabulary size. The Vocabulary Size Test (VST) measures learners' knowledge of words in written form, the connection of form and meaning and to a smaller degree concept knowledge (Nation, 2012). The test items appear in a single non-defining context in English, which enables students to gauge the word class to which the word belongs but does not offer further cues to the meaning of the word. Test takers are required to find the corresponding meaning from four choices in English or test takers' mother language. While it would have been preferable to have used an aural vocabulary test, individual administration of a test such as the Peabody Picture Vocabulary Task (Dunn & Dunn, 2007) or A_Lex (Milton

& Hopkins, 2006), was not feasible in the circumstances. The Listening Vocabulary Levels Test developed by McLean, Kramer, and Beglar (2015) was not yet available at the time of data collection. In addition, in the studies of Andringa et al. (2012) and Staehr (2009) a written vocabulary test worked very well in explaining variance in listening comprehension. In the present study, a shortened version of Nation and Beglar's test was used, consisting of the first 5000 words of the test. Although Adolphs and Schmitt (2003) and van Zeeland and Schmitt (2013) indicate that knowledge of 2000-3000 word families is needed for L2 learners to achieve good understanding in L2 listening, in order to avoid a ceiling effect, we tested students' knowledge of 5000 word families. This cut-off point was chosen because second-year non-English major university students in China are required to know 4795 word families according to the National College English Teaching Syllabus (The Ministry of Chinese Education, 2007). We opted for the bilingual Mandarin version of the VST because in this version respondents' ability to understand the target items is not confounded with their ability to read answer options items in the L2 as all answer options are given in Mandarin Chinese.

Metacognitive awareness questionnaire

Metacognition was measured with the Metacognitive Awareness Listening Questionnaire, the MALQ (Vandergrift et al., 2006). We are aware that this questionnaire does not tap into respondents' actual metacognitive activity, but only into self-reported awareness of such performance, as pointed out by Vandergrift and

Baker (2015), and this is of course an important drawback of this questionnaire. Vandergrift et al. (2006) claim that the MALQ has satisfactory psychometric properties in terms of validity and reliability. It consists of 21 items which respondents are required to respond to using a 6-point Likert scale ranging from "strongly disagree" to "strongly agree". The 21 items are classified into five categories: planning and evaluation, directed attention, person knowledge, problem solving and mental translation. Each category consists of 3-6 items. All items are randomized on the questionnaire. Clear guidelines were given on the questionnaire on how to fill it in.

Listening comprehension test

The listening part of the national College English Test Band4 (CET4) was used to measure participants' listening comprehension. The CET4 is a standardized English test which is widely used in colleges and universities in China, to measure college students' English levels in listening, reading, translation and writing (Zhu & Zhu, 2007). The listening part of the CET4 contributes 35% to the total score and consists of four parts: 1) eight short dialogues, followed by one question for each dialogue; 2) two long dialogues, followed by seven questions; 3) three short monologues, followed by ten questions; 4) an aural listening task where students fill in ten blanks in the transcript. The maximum score of listening part was 249. The CET4 listening part is considered to have high validity (Zhu & Zhu, 2007). An advantage of using the CET4 was that college students are familiar with the types of listening parts of CET4 and the

topics covered in the different components are related to college students' life and study. In addition, the words and phrases are covered in the National College English Syllabus (The Ministry of Chinese Education, 2007). The test is held on the same day nationwide. In the present study, participants' CET4 listening scores were taken to be a valid reflection of their listening comprehension level in the Chinese context.

3.3 Procedure

The general proficiency test and the vocabulary size test were taken under controlled conditions during normal class time in May 2014. It took a week to complete the two tests in three different classes of students.

Although we are aware that it would have been preferable to have administered the MALQ under controlled conditions, this was not possible under the time constraints. All 172 students received a copy of the MALQ, with a request to fill it in and return it within one week. Among the 164 questionnaires which were returned after a week, 151 were filled in completely and were used in the study. Three reverse coded items (Item 3, Item 8, and Item 16) in the MALQ were recoded to facilitate the computation of the total scores. Participants' responses were totalled per category and the sum of five categories was computed to obtain a total score.

Participants' CET4 listening scores were collected directly from the university's Office for Academic Affairs which keeps a record of all undergraduates' test scores. Students took the test in December 2013, that is five months prior to the collection of the other data.

3.4 Ethical considerations

Prior to data collection, participants were informed of the research purposes and the consent was obtained from students and teachers before the start of the study.

3.5 Data analysis

Prior to data analysis some of the items in the MALQ we reverse coded a number of items. For Item 3 and Item 8 a low score meant that a learner felt anxious about listening or saw listening comprehension as much of a challenge and a high score on the two items meant that a learner did not feel anxious about listening or did not see listening comprehension as much of a challenge. We reverse coded Items 3 and 8 to ensure that a high value indicated the same type of response on each item.

We then investigated the reliability and the validity of the MALQ. The results showed that Cronbach's alpha of the MALQ as a whole was .77, which is of an acceptable level according to Field (2013), and higher than the Cronbach's alpha of .62 obtained for the MALQ in Vandergrift and Baker (2015). A principal component analysis was conducted on the 21 items with oblique rotation (direct oblimin). Factor analysis revealed that there were four factors in the data which had eigenvalues over Kaiser's criterion of 1 and in combination explained 47.63% of the variance. The questionnaire was therefore not unidimensional. Following Vandergrift et al. (2006) we labeled the four factors as problem-solving, mental translation, person knowledge and directed attention (see Table 2). Cronbach's alpha was computed for each of the variables that contributed to the four dimensions. Vandergrift et al.

distinguished five broad factors in their questionnaire, one of which was planning-evaluation. However, all the questions listed under planning-evaluation loaded highly on problem solving, except one question. Therefore we decided not to use this label. While the problem-solving subscale of the MALQ had high reliability (Cronbach alpha = .81), the other three factors had less satisfactory reliability scores (see Table 3). We could not investigate the reliability of the CET4 listening section because the Chinese government does not release the score details of individual sections of the test.[1]

==========Table 2 approximately here ==========

==========Table 3 approximately here ==========

4. Results

We computed students' vocabulary sizes by multiplying their scores on the VST by 100. As can be seen in Table 4 students knew just under 3000 word families on average. Our results indicate that students had knowledge of fewer than the 4795 word families they are required to know for the National College English Teaching Syllabus (The Ministry of Chinese Education, 2007). Since the available literature

[1] We cannot provide information about the reliability of the VST and the OQPT because the original scores of the two tests were unfortunately accidentally lost and the original test papers could no longer be accessed to compute Cronbach's alpha.

provides considerable evidence for the importance of vocabulary knowledge in L2 listening comprehension (Bonk, 2000; Matthews & Cheng, 2015; Milton et al., 2010; Staehr, 2009), many participants may have been on the cusp of understanding the listening component of the CET4, as knowledge of 2000-3000 word families is required as a minimum for listening comprehension at a 95% coverage figure (van Zeeland & Schmitt, 2013).

==========Table 4 approximately here ==========

Table 5 shows that there were significant positive correlations between listening comprehension, general language proficiency, vocabulary knowledge and metacognitive awareness. The variable which correlated most strongly with listening comprehension was vocabulary size (r = .44), followed by general language proficiency (r = .36) and metacognitive awareness (r = .19). There were also significant but modest correlations between vocabulary knowledge, general language proficiency and metacognitive awareness.

=========Table 5 approximately here -==========

As person knowledge was found to be the most important factor among the different factors in the MALQ in Vandergrift and Baker (2015) we also looked at the

correlations between this factor and listening comprehension. Table 6 shows that person knowledge is indeed significantly correlated with listening comprehension (r = .29, p <.01). There were no significant correlations between any of the other components of the MALQ and listening comprehension and therefore those results are not reported here. As there is some evidence (Conrad, 1985; Rubin, 1994) that learners with relatively low proficiency levels focus more on lexical cues in the input we split the participants in a low ability (n = 73) and a high ability group (n = 78). We used the median of the OQPT (34) as the cut-off point between both groups and compared the scores on the MALQ between the low ability and the high ability groups, but this analysis did not reveal significant differences between both groups. We found that the correlation between listening comprehension and vocabulary size was similar in magnitude (i.e. moderate) for each of the proficiency groups (see Table 7).

========= Table 6 approximately here ==========

========= Table 7 approximately here ==========

We then checked whether the assumptions for carrying out a regression analysis were met. This was indeed the case because an analysis of the residuals showed that the distribution of residuals was normal and the correlation between predicted variables and residuals was non-linear.

Hierarchical multiple regression analysis was used to analyze the contribution of

the three independent variables (general language proficiency, vocabulary knowledge and metacognitive awareness) to the dependent variable (learners' listening comprehension). In analyzing the role of metacognitive awareness, we first explored the overall scores of the MALQ in the regression analysis, and then we entered the score for a subcomponent of the MALQ, namely person knowledge, and explored whether this additional factor contributed significantly to the predictive capacity of the model. We entered only this subcomponent of the MALQ because there were no correlations between other subcomponents of the MALQ and the dependent variable. In order to find out whether vocabulary knowledge explained any variance in listening comprehension over and above the contribution of general language proficiency, in the first regression analysis, we first inputted the total scores for the OQPT, then the VST and finally the MALQ. In the second regression analysis, we inputted the scores for the OQPT, the VST and one of the components of the MALQ, namely person knowledge instead of the total scores on the MALQ. Finally, as we wanted to find out what the unique contribution of general language proficiency was to the dependent variable over and above vocabulary knowledge, in the third regression analysis we reversed the order of entry of the OQPT and the VST, by entering the VST first, followed by the OQPT. In this model person knowledge was again entered last.

In the first regression analysis, we entered only general language proficiency as a predictor in the first instance (see Table 8). This model was significant (F (1,149) = 22.18, p <.001) and the OQPT explained 12% of the variance in listening

comprehension (β = .36, p < .001), but when in a second step vocabulary size was added a more powerful model emerged explaining an additional 13% of the variance. In total these two variables explain 25% of the variance in listening comprehension (F (2, 148) = 25.65, p <. 001). The difference in R squared between the models with and without vocabulary size was also significant. In this second model, the standardized beta values for general language proficiency and vocabulary size were .26 and .37 respectively. However metacognitive awareness as measured with the total scores on the MALQ did not contribute significantly to the model (β = .08, p = .29).

========== Table 8 approximately here ==========

Subsequently, in order to explore the impact of person knowledge on listening comprehension, we entered this subcomponent of the MALQ separately as an independent variable into the hierarchical regression analysis after language proficiency and vocabulary knowledge (see Table 9).

==== Table 9 approximately here ==========

The standardized predicted values and the residuals indicated that the statistical assumptions for carrying out a regression analysis were all met.[1]The ANOVA

[1] An analysis of the multicollinearity statistics reveals that the tolerance values (.90, .93 and .95) of three independent variables (general language proficiency

revealed that this model was significant (F (3, 147) = 20.62, p < .05). As in the previous models, general language proficiency and vocabulary size turned out to be significant predictors. In addition, this time person knowledge was also retained as a significant predictor. General language proficiency was found to explain 13% of the variance, vocabulary size explained an additional 13%, and person knowledge explained 4% of the variance in listening comprehension. Together these three variables explained 30% of the variance in learners' listening comprehension.

Finally, in order to establish whether vocabulary and general language proficiency are indeed separate dimensions of language ability, we also carried out a hierarchical regression analysis in which the VST and the OQPT were entered in reverse order. In this regression analysis, the VST was entered in the first instance, followed by the OQPT and person knowledge. When the VST was the only predictor variable in the model, it explained no less than 19% of the variance in listening comprehension. When the OQPT was added a further 6% of the variance is explained, and person knowledge explained a further 4%.[1] The 6% change in R squared that was triggered by the addition of the OQPT is clearly less important than the 13% change in R squared that was brought about by the addition of the VST in the previous model. This clearly demonstrates that both the OQPT and the VST make an independent contribution to listening comprehension, but that the VST explains more

vocabulary size and person knowledge) were all higher than 0.1 and VIF (1.11, 1.08, and 1.05) was lower than 10, so there was no problem of multicollinearity among independent variables (tolerance should be higher than 0.1 and VIF lower than 10 according to Field, 2013, p. 325)

[1] Due to differences in rounding up from three decimal points to two, the total explained variance differs slightly between models.

of the variance in the dependent variable than the OQPT.

5. Discussion

The results of the current study support the findings of Adolphs and Schmitt (2003), S. Liu (1995), M. L. Liu (2011), van Zeeland and Schmitt (2013) and Andringa et al. (2012) regarding the importance of vocabulary size for listening comprehension among L2 learners. The study provides clear evidence for the fact that vocabulary size explains unique variance over and above general language proficiency and that vocabulary knowledge is more important than metacognitive awareness in listening comprehension. It is possible that the VST had a stronger explanatory power than the OQPT in our study because the vocabulary items included in the VST were very carefully sampled from different frequency layers in the British National Corpus, while it is not clear how this was done in the OQPT. The correlations between vocabulary size and listening comprehension in our study were lower than those in Andringa et al. (2012) and slightly higher than in Mecartty (2000), although the strength of correlations remains difficult to compare because researchers use a wide variety of tests to measure vocabulary, as well as different tests to measure listening.

The study also throws interesting new light on the componential structure of language ability. Both general language proficiency and vocabulary knowledge were found to make a unique contribution to explaining listening comprehension, which lends support to view that these are different dimensions of language ability. As pointed out by Harsch (2014) language ability is indeed unitary as well as divisible,

and both holistic measures and discrete-point measures can therefore be used to measure this construct. As holistic and discrete-point measures are likely to share at least some variance, future research should concentrate on investigating which tests are most successful at tapping into these different dimensions or components, and identifying to what extent they overlap.

In the current study self-reported metacognitive awareness explained much less variance in listening comprehension (4%) than in Vandergrift et al.'s (2006) study where it was found to explain 13% of the variance. The correlations between metacognitive awareness and listening were also slightly lower in our study (.19) than in Vandergrift and Baker (2015), who found a correlation of .23. In the latter study, metacognitive awareness was found to only contribute indirectly to listening comprehension, via L2 vocabulary knowledge, which illustrates the relative importance of each of these once more. Slightly stronger correlations could perhaps have been obtained if the MALQ had been administered under controlled conditions directly after the students had taken a specific listening task, as in Vandergrift et al. (2006). The listening test could then have functioned as an anchor point for students in filling in the questionnaire (Graham, personal communication), and the magnitude of the correlations might have been more similar to those found in Vandergrift et al.'s study. Vogely (1995) was indeed successful in demonstrating there was a moderate correlation of .44 (Pearson) between the self-reported perceptions of strategies and listening tasks in her study where participants completed a Metacognitive Awareness Strategy Questionnaire directly after they had completed three authentic listening

comprehension tasks.

It is interesting that person knowledge (the learners' perceptions of the difficulty and their self-efficacy in L2 listening, Vandergrift et al., 2006) was the most relevant category in the MALQ because Vandergrift et al. (2006) and Vandergrift and Baker (2015) reached the same conclusion regarding the importance of person knowledge. In the MALQ, person knowledge was measured with three items: "I find that listening in English is more difficult than reading, speaking, or writing in English", "I feel that listening comprehension in English is a challenge for me" and "I don't feel nervous when I listen to English". Since the results showed that person knowledge was significantly and positively correlated with listening comprehension, this meant that the more confident and the less anxious a listener felt, the higher scores on listening comprehension s/he had achieved. These results confirm the considerable evidence in the literature about the relationship between foreign language anxiety and a range of aspects of language learning (see Dewaele, Petrides & Furnham, 2008 for an overview) and between listening and anxiety in particular (see Mills, Pajares, & Herron, 2006). It is possible that the correlations between person knowledge and listening comprehension reached significance (by contrast with other components of the MALQ) because learners are aware of their problems with listening (Goh, 2000) and person knowledge is therefore more easily accessible to conscious inspection than other components of metacognition.

The present study shows that none of the other categories in the MALQ correlated significantly with listening comprehension. Our study therefore does not

confirm the results of X. Y. Zuo (2013) who found that metacognitive strategies were significantly correlated with listening comprehension performance. Future researchers will probably need to use other measures of the different components of metacognition which tap more directly into the mental activities under study, and are not solely based on self-report. The key problem with the MALQ, as recognised by Vandergrift and Baker (2015, p. 210), is that the MALQ does not tap into actual metacognitive activity but rather asks learners to self-report on their awareness of listening processes.

The pedagogical implications of the study can be summarised as follows: to improve learners' listening comprehension, in L2 teaching, teachers should focus on enhancing learners' vocabulary knowledge in particular. As this was not an intervention study, on the basis of the current results we cannot make recommendations regarding the types of vocabulary instruction that are most beneficial to improve listening skills, but it is likely that this can be achieved with the help of intentional vocabulary learning activities and as well as by maximizing opportunities for incidental vocabulary learning from aural and written input (see Tang & Treffers-Daller, 2016, for details). For example, since L2 learners' ability to recognise words from speech is crucial to L2 listening comprehension, Hulstijn (2003) has suggested taking advantage of multimedia software to develop L2 learners' vocabulary knowledge for L2 listening comprehension. Subsequent studies indicate that such computer-mediated approaches contribute positively to improving L2

vocabulary knowledge and L2 listening comprehension (Matthews, Cheng, & O'Toole, 2015; Matthews & O'Toole, 2013).

In the current study we did not find significant differences in metacognitive strategy use between learners with lower scores on the OQPT and learners with higher scores. This was unexpected because according to Rubin (1994) and Vandergrift and Baker (2015) lower level learners have little opportunity for metacognitive strategy use. A key reason for this is likely to be the limited processing capacity of L2 learners who spend most of their energy on decoding the incoming speech signal (Goh, 2000). It is possible that lower level learners would benefit from strategy instruction, as argued by Graham and Macaro (2008). However, according to Alderson's (1984) threshold hypothesis, knowledge of the L1 and strategic knowledge cannot be transferred to the L2 unless a certain threshold of knowledge in the L2 has been obtained. Therefore teaching such strategies to low level learners may not be that fruitful. The students in the current study had fairly small L2 vocabularies, and vocabulary size was the key explanatory variable in the study. For this reason we think that the students are more likely to benefit more from developing their vocabularies. Future research could therefore focus on establishing at which levels of L2 proficiency or from which vocabulary sizes onwards strategy instruction is most likely to benefit learners.

6. Conclusion

In this study, we explored the impact of individual differences in general language

proficiency, vocabulary knowledge and metacognitive awareness (person knowledge) on explaining learners' listening comprehension. We found that these three learner variables could explain 30% of the variance in L2 learners' listening comprehension. General language proficiency and L2 vocabulary knowledge were found to be the factors with most explanatory power (13% and 13% respectively), followed by person knowledge (4%). Limitations of the current study were that metacognition was measured through self-report rather than by measures which tap more directly into participants' cognitive functioning. Further studies will need to focus on assessing metacognition more directly, and will also need to look at the role of written versus aural vocabulary to listening comprehension. Although the current study shows that a written receptive vocabulary test can be used in a study of listening comprehension, we still know very little about the differences between the impact of aural versus written vocabulary knowledge on listening comprehension. Studies which use tests such as the Listening Vocabulary Levels Test (McLean et al., 2015) and A_Lex (Milton & Hopkins, 2006) can no doubt provide new insights into the relationship between aural vocabulary and listening.

References

Adolphs, S., & Schmitt, N. (2003). Lexical coverage in spoken discourse. *Applied Linguistics*, *24*, 425-438.

Alderson, J. C. (1984). Reading in a foreign language: A reading or a language problem? In J. C. Alderson & A. H. Urquhart (Eds.), *Reading in a foreign*

language (pp. 1–24). London: Longman.

Anderson, R. C., & Freebody, P. (1981). Vocabulary knowledge. In J. T. Guthrie (Ed.), *Comprehension and teaching: Research reviews* (pp. 77-117). Newark, DE: International Reading Association.

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

Bachman, L.F., & Palmer, A. (2010). *Language assessment in practice*. Oxford: Oxford University Press.

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

Buck, G. (2001). *Assessing listening*. Cambridge: Cambridge University Press.

Clahsen, H., & Felser, C. (2006). Grammatical processing in language learners. *Applied Psycholinguistics*, *27*, 3-42.

Conrad, L. (1985). Semantic versus syntactic cues in listening comprehension. *Studies in Second Language Acquisition*, *7*, 59-72.

Dewaele, J. M., Petrides, K.V., & Furnham, A. (2008). The effects of trait emotional intelligence and socio-biographical variables on communicative anxiety and foreign language anxiety among adult multilinguals: a review and empirical investigation. *Language Learning*, *58*(4), 911-960.

Dörnyei, Z., & Ryan, S. (2015). *The psychology of the language learner revisited.* Mahwah, NJ: Erlbaum.

Dunn, L. M., & Dunn, D. M. (2007). *Peabody picture vocabulary test (PPVT-4)* (4th ed.). Minneapolis: Pearson.

Field, A. (2013). *Discovering statistics using IBM SPSS statistics*. Los Angeles etc.: Sage.

Goh, C. C. M. (2000). A cognitive perspective on language learners' comprehension problems. *System, 28*, 55-75.

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

Graham, S., & Macaro, E. (2008). Strategy instruction in listening for lower-intermediate learners of French. *Language Learning, 58*, 747-783.

Graham, S., Santos, D., & Vanderplank, R. (2008). Strategy clusters and sources of knowledge in French L2 listening comprehension. *Innovation in Language Learning and Teaching, 4*, 1-20.

Grzegorzewska, L. (2015). *Exploring the relationship between intelligence and the use of vocabulary learning strategies*. Presentation at the Language in Focus conference, Cappadocia, March 2015.

Halliday, M. A. K. (1994). *An introduction to functional grammar* (2nd ed.). London: Arnold.

Harsch, C. (2014). General language proficiency revisited: Current and future issues. *Language Assessment Quarterly, 11*(2), 152-169.

He, M. Y., & Bao, K. J. (2006). 重点大学与普通院校学生英语听力策略比较:个案研究(A case study on English learners' listening strategies). *外语研究*

(Foreign Languages Research), (2), 46-49.

He, Z. J. (2005). 英语听力理解教学中元认知策略培训的实验研究(An experimental study on meta-cognitive strategy training in listening class). *外语电化教学 (Computer-assisted Foreign Language Education), April*, 56-61.

Hulstijn, J. H. (2003). Connectionist models of language processing and the training of listening skills with the aid of multimedia software. *Computer Assisted Language Learning*, *16*, 413-425.

Hunston, S., & Francis, G. (2000). *Pattern grammar: A corpus-driven approach to the lexical grammar of English* (Studies in Corpus Linguistics 4). Amsterdam etc.: Benjamins.

Kassem, H. M. (2015). The relationship between listening strategies used by Egyptian EFL college sophomores and their listening comprehension and self-efficacy. *English Language Teaching*, *8*, 153-169.

Kelly, P. (1991). Lexical ignorance: The main obstacle to listening comprehension with advanced foreign language learners. *IRAL: International Review of Applied Linguistics in Language Teaching*, *2*, 135-149.

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

Liu, M. L. (2011). 英语词汇知识和听力理解的关系研究 (The relationship between English vocabulary knowledge and listening comprehension). *中国电力教育 (China Electric Power Education), (17),* 193-196.

Liu, S. (1995). 英语词汇量与听力理解关系的研究 (The relationship between English vocabulary size and listening comprehension). *海南大学学报 (Journal of Hainan University (Social Science)), (2),* 84-89.

Long, Y. F., & Zhao, P. (2009).大学英语听力教学中元认知策略与多模态交互研究 (To construct a body of organized knowledge for and lead the development of the field). *外语电化教学(Computer-assisted Foreign Language Education), July*, 58-62.

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

Matthews, J., Cheng, J., & O'Toole, J. M. (2015). Computer-mediated input, output and feedback in the development of L2 word recognition from speech. *ReCALL*, *27*, 321-339.

Matthews, J., & O'Toole, J. M. (2013). Investigating an innovative computer application to improve L2 word recognition from speech. *Computer Assisted Language Learning*, *28*, 364-382.

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

Meara, P., & Milton, J. (2003). *X_Lex. The Swansea levels test*. Newbury, UK: Express.

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

Mehrpour, S., & Rahimi, M. (2010). The impact of general and specific vocabulary knowledge on reading and listening comprehension: A case of Iranian EFL learners. *System, 38,* 292-300.

Mills, N., Pajares, F., & Herron, C. (2006). A reevaluation of the role of anxiety: SE, anxiety and their relation to reading and listening proficiency. *Foreign Language Annals, 39,* 276-293.

Milton, J., & Hopkins, N. (2005). *Aural Lex*. Swansea, UK: Swansea University.

Milton, J., & Hopkins, N. (2006). Comparing phonological and orthographic vocabulary size: do vocabulary tests underestimate the knowledge of some learners. *The Canadian Modern Language Review*, *63*(1),127-147

Milton, J., Wade, J., & Hopkins, N. (2010). Aural word recognition and oral competence in a foreign language. In R. Chacón-Beltrán, C. Abello-Contesse, & M. Torreblanca-López (Eds.). *Further insights into non-native vocabulary teaching and learning* (pp. 83-98). Bristol: Multilingual Matters.

Nation, I. S. P. (2001). *Learning vocabulary in another language*. Cambridge: Cambridge University Press

Nation, I. S. P. (2006). How large a vocabulary is needed for reading and listening? *Canadian Modern Language Review. 63*(1), 59-81.

Nation, I. S. P. (2012). Vocabulary Size Test Instructions and description. Retrieved April 30, 2014, from http://www.victoria.ac.nz/lals/about/staff/publications/paul-nation/Vocabulary-Size-Test-information-and-specifications.pdf.

Nation, I. S. P., & Beglar, D. (2007). A vocabulary size test. *The Language Teacher*, *31*, 9-13.

Oller, J. W. Jr. (1973). Cloze tests of second language proficiency and what they measure. *Language Learning, 23,* 105–118.

O'Malley, J. M., & Chamot, A. U. (1990). *Learning strategies in second language acquisition*. Cambridge: Cambridge University Press.

Rubin, J. (1994). A review of second language listening comprehension research. *The Modern Language Journal*, *78,* 199-221.

Schoonen, R., Hulstijn, J., & Bossers, B. (1998). Metacognitive and language-specific knowledge in native and foreign language reading comprehension. An empirical study among Dutch students in grades 6, 8 and 10. *Language Learning*, *48*, 71-106.

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

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

Tang, C. & Treffers-Daller, J. (2016). Assessing incidental vocabulary learning by Chinese EFL learners: testing the Involvement Load Hypothesis. In G. Yu & Y. Jin (Eds.). *Assessing Chinese Learners of English: Language Constructs, Consequences and Conundrums* (pp. 121-148). Basingstoke: Palgrave Macmillan.

The Ministry of Chinese Education. (2007). *College English curriculum syllabus*.

Shanghai: Shanghai Foreign Language Education Press.

van Zeeland, H., & Schmitt, N. (2013). Lexical coverage in L1 and L2 listening comprehension: The same or different from reading comprehension? *Applied Linguistics, 34*, 457-479.

Vandergrift, L. (1997). The comprehension strategies of second language (French) listeners: A descriptive study. *Foreign Language Annals, 30*, 387-409.

Vandergrift, L. (2006). Second language listening: Listening ability or language proficiency? *The Modern Language Journal*, *90*, 6-18.

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

Vandergrift, L., Goh, C. C. M., Mareschal, C. J., & Tafaghodtari, M. H. (2006). The metacognitive awareness listening questionnaire: Development and validation. *Language Learning*, *56*, 431-462.

Vandergrift, L., & Tafaghodtari, M. H. (2010). Teaching L2 learners how to listen does make a difference: An empirical study. *Language Learning, 60*, 470-497.

Vogely, A. (1995). Perceived strategy use during performance on three authentic listening comprehension tasks. *The Modern Language Journal*, *79*, 41-56.

Wechsler, D. (1997). *The Wechsler memory Scale-III manual*. San Antonio, Tx: The Psychological Corporation.

Winne, P., & Perry, N. (2000). Measuring self-regulated learning. In M. Boekaerts, P. Pintrich, & M. Zeidner (Eds.), *Handbook of self-regulation* (pp. 531–566).

San Diego: Academic Press.

Zhu, S., & Zhu, L. X. (2007). 改革后的大学英语四级考试的测试质量分析与评价 (Testing analysis and evaluation of new CET4). *绥化学院学报* (*Journal of Suihua University)*, *27*(4), 119-121.

Zuo, J. (2013). 关于图式知识和语言水平对英语听力理解影响的实证性研究 (A study on the effect of schemata and language proficiency on EFL listening comprehension). *海外英语 (Overseas English)*, *January*, 102-103.

Zuo, X. Y. (2013). 英语听力理解策略与听力水平的相关性研究 (The relevant research on English listening comprehension strategies and listening levels). *学科探索* (*Disciplines Exploration)*, *July*, 50-51.

Table 1

Studies which focused on the role of vocabulary in listening comprehension in L2

Author	Students' L1	Target language	N	Vocabulary component tested (task used)	Pearson correlations with listening	Percentage (%) of variance in L2 listening comprehension explained
Andringa et al. (2012)	35 different L1s	Dutch	113	Written vocabulary size (multiple choice)	.68	†
Bonk (2000)	Japanese	English	59	Lexical recognition (dictation task)	.45 (kendall's tau)	23
Liu, M. L. (2011)	Chinese	English	90	Receptive vocabulary size and vocabulary depth	.54 (size) and .60 (depth)	†
Liu, S. (1995)	Chinese	English	168	Reading vocabulary size (multiple choice) and listening vocabulary size (multiple choice)	.57 (reading) and .66 (listening) respectively	†
Matthews & Cheng (2015)	Chinese	English	167	Aural word recognition (partial dictation task)	.73	54
McLean, Kramer, & Beglar (2015)	Japanese	English	214	Listening vocabulary levels (aural)	.54	†
Mecartty (2000)	Unknown but "not Spanish"	Spanish	77	Written word association and antonyms	.38	14
Milton, Wade, & Hopkins (2010)	A range, including Arabic, Chinese and Japanese	English	30	Written and oral receptive vocabulary size test (X_Lex and A_Lex)	.48 (X_Lex) .67 (A_Lex)	51 (X_Lex) 44 (A_Lex)
Staehr (2008)	Danish	English	88	Receptive vocabulary size	.69	39 (Nagelkerke's R^2)

Staehr (2009)	Danish	English	115	Written vocabulary size and depth	.7 (size) and .65 (depth)	51 (size: 49; depth: 2)
Vandergrift & Baker (2015)	English	French	157	Oral vocabulary size (PPVT)	.51	49 (path coefficient)

†No indication is given of the proportion of the variance explained by vocabulary knowledge.

Table 2

Factor loadings and communalities based on a principal component analysis with oblimin rotation for 21 items from the MALQ (N=151)

Items	Factors				Communalities
	Problem -solving	Mental translation	Person knowledge	Directed attention	
As I listen, I compare what I understand with what I know about the topic.	.77				.55
I use my experience and knowledge to help me understand.	.74				.53
I use the words I understand to guess the meaning of the words I don't understand.	.65				.45
I use the general idea of the text to help me guess the meaning of the words that I don't understand.	.64				.53
When my mind wanders, I recover my concentration right away.	.59				.43
I have a goal in mind as I listen.	.51				.41
I try to get back on track when I lose concentration.	.49		-.323		.50
I translate key words as I listen.	.46				.31
After listening, I think back to how I listened, and about what I might do differently next time.	.45				.30
When I guess the meaning of a word, I think back to everything else that I have heard, to see if my guess makes sense.	.44	-.43			.46
Before listening, I think of similar texts that I may have listened to.	.44				.25
As I listen, I periodically ask myself if I am satisfied with my level of comprehension.		-.67		-.318	.60
I translate word by word, as I listen.		-.63			.43
I translate in my head as I listen.		-.57			.44
I find that listening in English is			.79		.64

more difficult than reading, speaking, or writing in English.					
I feel that listening comprehension in English is a challenge for me.			.69		.54
I focus harder on the text when I have trouble understanding.				.75	.41
When I have difficulty understanding what I hear, I give up and stop listening.	-.31			.57	.45
As I listen, I quickly adjust my interpretation if I realize that it is not correct.	.45			.55	.65
Before I start to listen, I have a plan in my head for how I am going to listen.	.39			.45	.46
I don't feel nervous when I listen to English.		.40		.43	.50

Note. Factor loadings < .3 are suppressed.

Table 3 Descriptive statistics for the four MALQ factors (N=151)

Factors	Number of items	Cronbach alpha
Problem-solving	11	.81
Mental translation	3	.52
Person knowledge	2	.53
Directed attention	5	.62

Table 4 Scores on the VST, the OQPT and the listening test

	M	SD	minimum	maximum
VST	2949.01	614.15	1500	4700
OQPT	34.70	6.31	16	52
Listening test scores	145.71	24.06	92	219

Table 5

Correlations between listening comprehension, general language proficiency, vocabulary knowledge and metacognitive awareness (N=151)

		Correlation Coefficients		
	Variables	General language proficiency	Vocabulary knowledge	Metacognitive awareness
DV	Listening comprehension	.36**	.44**	.19*
IVs	General language proficiency		.27**	.19*
	Vocabulary knowledge			.19*

Note. *$p < .05$. **$p < .01$.

Table 6

Correlations between listening comprehension, general language proficiency, vocabulary knowledge and person knowledge (N=151)

Variables	Correlation Coefficients		
	Vocabulary size	General language proficiency	Person Knowledge (MALQ)
Listening comprehension	.44**	.36**	.29**
Vocabulary size		.27**	.11
General Language proficiency			.21*

Note. *$p < .05$. **$p < .01$.

Table 7

Pearson correlations with listening comprehension according to levels of language ability

	Vocabulary size	Overall metacognitive awareness
Low ability	.42**	-
High ability	.45**	-

Note. *p < .05. **p < .01.

Table 8

Regression models explaining listening comprehension with general language proficiency and vocabulary size as predictors

Model	R	R squared	Adjusted R Squared	Std. Error of the Estimate	R squared changed	Sig. F change
1	.36[a]	.13	.12	22.52	.13	< .001
2	.51[b]	.26	.25	20.87	.13	< .001

Note. [a]Predictors: (Constant), general language proficiency, [b]Predictors: (Constant), general language proficiency and vocabulary size.

Table 9

Regression models explaining listening comprehension with general language proficiency, vocabulary size and person knowledge as predictors

Model	R	R squared	Adjusted R Squared	Std. Error of the Estimate	R squared change	Sig. F change
1	.36[a]	.13	.12	22.52	.13	<.001
2	.51[b]	.26	.25	20.87	.13	<.001
3	.54[c]	.30	.28	20.39	.04	.005

Note. [a]Predictors: (Constant), general language proficiency, [b]Predictors: (Constant), general language proficiency, vocabulary size, [c]Predictors: (Constant), general language proficiency, vocabulary size, person knowledge.