

# The Effects of Presenting Multiple-Choice Test Items in Oral and Written Modes and Item Types on Advanced EFL Learners' Listening Comprehension and Perception

**Mohammad Taghi Farvardin**

Assistant Professor,  
Ahvaz Branch, Islamic Azad University, Ahvaz  
farvardin@iauahvaz.ac.ir

**Afrooz Rezaee**

PhD Candidate,  
Shiraz Branch, Islamic Azad University, Shiraz  
Afrooz.Rezaee@gmail.com

**Amir Mashhadi**

Assistant Professor, Shahid Chamran University of Ahvaz  
Mashhadis.scu@gmail.com

## Abstract

This quasi-experimental study aimed to compare the effect of different modes and item types of multiple-choice (MC) test items on advanced EFL learners' listening comprehension and perception. To this end, 80 advanced EFL learners, aging 18 to 30, were selected. The participants took a listening test including dialogue-completion and question and answer multiple-choice items presented in written and oral modes. In addition, the participants were given a questionnaire on their perceptions of the oral and written modes. The results of two-way repeated measures ANOVA showed that there was no significant difference between the participants' scores in the oral and written modes. Moreover, they received similar scores on the two item types. However, in their questionnaire responses, most of the participants preferred the written mode to the oral mode. The results also imply that although presenting the multiple-choice test items in the oral mode may form a pure test of listening comprehension, there are acceptable reasons for presenting the multiple-choice listening items in the written mode.

**Keywords:** Advanced EFL Learners, Item Type, Listening Performance, Listening Perception, Multiple-choice Test, Oral Mode, Written Mode

## **1. Introduction**

Listening has a key role in the second/foreign language learning, and its development is of prime concern to language teachers (Rubin, 1994). For many English as a foreign language (EFL) learners, listening is stressful and difficult to undertake (Chang & Read, 2013). Mendelsohn (2002) believes that listening is not a single process but it is a set of related processes which convey information. Moreover, testing second/foreign language listening is complex and multifaceted including many factors that may affect the test-takers' comprehension and performance (Cook, 2013). For instance, processing a second language (L2) requires better memory of listeners than processing their first language (L1) (Ohata, 2006). In this case, an L2 listening test may measure test-takers' memory capacity in addition to their listening competence. Therefore, the effect of testing methods on test-takers' performance has become a crucial issue in developing listening comprehension tests (Yanagawa & Green, 2008).

Among test types measuring L2 listening comprehension, multiple-choice (MC) test items have been emphasized more (Hemmati & Ghaderi, 2014). However, the results of previous studies have revealed that there is no unanimity whether all formats and modes of MC test items are equally suitable to measure EFL learners' listening performance (Chang & Read, 2013). There are three elements of MC items for listening comprehension: the questions or item stems, the answer options and the recording. However, the order, extent and mode of presentation of these elements can completely change the nature of the task (Yanagawa & Green, 2008). In this line, Bachman (1990) proposed that "in examining the effects of test method facets on language test scores, we are also testing hypotheses that are relevant to construct validity" (p. 258). The presentation mode is one aspect of the test method, and depending on the

### *The Effects of Presenting Multiple-Choice Test...*

mode, other factors not directly related to the target constructs can affect the results (Ohata, 2006). Thus, the presentation mode is required to be viewed in test development process, and construct-irrelevant factors should be minimized (Chang & Read, 2013). In the written mode of MC test, construct-irrelevant factors include reading ability, lexical attractiveness and uninformed guessing (Freedle & Kostin, 1999). Meanwhile, some scholars (e.g., Chang & Read, 2013; Yanagawa & Green, 2008) have argued that written mode of MC test can negatively influence the validity of the listening test since the reading ability might have been measured in the written mode. This could make the written mode less valid for a listening test. Oral mode, instead, involves short-term memory capacity, which is known to be limited for L2 learners (Cook, 2013). Moreover, it has been found that the oral mode increases test-takers' anxiety which could impact their test performance (Chang & Read, 2006).

Concerning the importance of L2 listening comprehension assessment, most of extant studies on the effects of different formats of MC questions on the listening comprehension of L2 listeners, have just focused on the written mode. To the best of the researchers' knowledge, there appears to have been few published research to indicate that presenting the MC test orally leads to a challenge for EFL listeners. Furthermore, there is a gap in the literature for conducting research on the effect of oral and written modes of MC items on advanced EFL learners' listening comprehension. Moreover, there is a need for further study to consider both mode and item type, since most research on item types have used items presented only in the written mode. Therefore, investigating how the mode difference works in association with different item types is required to gain a more thorough picture of these two factors' effect. In addition, EFL learners' perceptions toward the two item modes need to be examined.

## **2. Literature Review**

### **2.1. Modes of Presenting MC Items**

Previous research has yielded mixed results on the effect of presenting L2 listening comprehension test items in written mode on test-takers' performance (e.g., Chang, 2008; Iimura, 2010; Yanagawa & Green, 2008). Some studies (e.g., Iimura, 2010; Yanagawa & Green, 2008) showing positive effects of the written mode argued that previewing items in the written mode was helpful for test-takers since it provided contextual clues of the listening input and allowed test-takers to use metacognitive strategies such as goal setting and planning. Weir (1993), on the other hand, claimed that written items distract test-takers' attention on listening input, since providing item stems and options in written form requires test-takers' reading ability.

In addition to the lexical matching strategy, it has been argued that the written answer options "provides contradictory cues and complicating planning strategies" (Yanagawa & Green, 2008, p. 110). As test-takers' cognitive processes in listening tests is affected by the features of test methods, providing the written answer options is likely to question the validity of a listening test because it may not reflect an authentic listening process (Weir, 2005). In other words, although the written mode provides test-takers with contextual clues about the listening stimuli before listening, it lets them merely match some words from the listening to select the answer and the process does not reflect the real-life listening.

### **2.2. Item Types**

Item type, as a feature of the test method, is believed to affect the test-takers' performance (Wolf, 1993). Previous studies on listening comprehension test

### *The Effects of Presenting Multiple-Choice Test...*

item types have mostly focused on multiple-choice or open-ended questions (Cheng, 2004). The inconsistency in the test results among the formats have been reported to be mainly due to the different skills that they are requiring (Cheng, 2004). Although all the formats intended to measure the same construct, the listening ability, different formats could cause other factors to affect the performance. This implies the need for further research on different item types within one response form, because even the same response format has several different item types that measure the same construct but require different skills. This study specifically focused on two different listening comprehension multiple-choice item types, dialogue-completion and question-and-answer, which are the two most frequently used item types for the multiple-choice listening comprehension questions.

In addition, no studies on the effect of oral and written mode of question/option presentation made any distinction between different item types (Chang, 2005; Chang & Read, 2013; Wu, 1998; Yanagawa & Green, 2008). Chang and Read (2013), for example, did include different item types for their listening test, but did not mention which item types they used and did not report their effect on the results. Considering item types when examining the effect of presentation mode in multiple-choice items is worth investigating, because combined characteristics of test method affect test-takers' cognitive processing and test scores in a dissimilar way (Bachman & Palmer, 1996).

In other words, since different characteristics of test methods could interact with other characteristics, the effect of any item characteristic on test-takers' performance needs a detailed analysis in terms of its interaction with other factors (Brindley & Slatyer, 2002). In this line, Cheng (2004) called for a further study that considers both mode and item type, since most research on item types, including hers, used items presented only in the written mode.

Later, Chang and Read (2013) explored the effects of MC test items in different modes on L2 listeners' performance and perceptions. The results indicated that the examinees scored almost the same in both modes with the oral (66%) and the written (68%). In spite of all these studies, there still needs to address this issue more in designing L2 listening tests.

In light of the pedagogical needs and the research gaps, the present study investigated the effect of modes of item presentation together with that of different item types of MC questions on advanced learners' listening comprehension test performance and perception. To investigate the influence of item type, the two most widely-used MC questions item types were chosen to be examined for the present study: dialogue-completion and question-and answer item type. Moreover, the interaction between types of multiple-choice items (i.e., dialogue-completion and question and answer) in listening performance of advanced EFL learners was examined. In the present study, questions are presented before the listening stimuli in the oral mode to balance the two modes regarding the effect of reviewing questions.

### **2.3. Research Questions**

To achieve the objectives of the study, the following questions were dealt with:

- Q1) Is there any significant difference between modes of presentation of multiple-choice items (i.e., oral and written) in listening performance of advanced EFL learners?
- Q2) Is there any significant difference between types of multiple-choice items (i.e., dialogue-completion and question and answer) in listening performance of advanced EFL learners?

*The Effects of Presenting Multiple-Choice Test...*

- Q3) Is there any significant interaction between modes of presentation and types of multiple-choice items (i.e., dialogue-completion and question and answer) in listening performance of advanced EFL learners?
- Q4) How do advanced EFL learners perceive oral and written modes of test item presentation?

### **3. Methods**

#### **3.1. Participants**

A total of 80 (42 male and 38 female) advanced EFL learners from four intact classes were selected. The participants were selected from the Adult Department of Iran Language Institute (ILI) in Ahvaz. The participants' age ranged from 18 to 30 ( $M=22.45$ ,  $SD=2.17$ ). The participants attended their class two days a week, for one hour and forty five minutes in the afternoon. The participants had all been learning English for an average of 4 years, beginning from Basic 1, which in turn qualifies the samples to be homogeneous. Two classes were assigned to Group A ( $n=40$ ), and the other two classes were assigned to Group B ( $n = 40$ ).

Moreover, to confirm that the two groups were homogeneous, Oxford Placement Test (OPT) which has a well-established reliability and validity (Allan, 2004), was adopted. OPT includes two sections, grammar and listening, each of which consists of 100 items. The required time to complete the test is 60 minutes. Each correct item received 1 point. Therefore, the maximum possible score was 200. The participants' scores ranged from 150 to 170, suggesting that they were at the advanced level of English proficiency ( $M= 159.33$ ,  $SD = 3.67$ ).

In addition, to confirm that the two groups were of equivalent listening ability, an independent samples *t*-test was carried out on the participants'

listening comprehension scores on the OPT. No significant difference was found between Group A and Group B.

## **3.2. Instruments**

### **3.2.1. Listening Comprehension Test**

In this study, the researchers designed a listening comprehension test in two different modes of oral and written to assess the listening comprehension performance of the participants. The test consisted of 20 four-option MC items including two item types, i.e., dialogue-completion type (10 items) and the question-and-answer type (10 items) (see Appendix). Actually, these two item types require different listening skills in that the test-takers have to complete a short conversation by choosing the most appropriate and spontaneous response for the dialogue-completion items, while they have to get the main idea or make inferences based on the conversation for the question-and-answer items. Since the stimulus material for the dialogue-completion tasks is short conversations between two people, only dialogues, not monologues, were used for the question-and-answer items in this study to keep the stimulus of the two item types the same (Table 1).

**Table 1. *Structure of the Listening Comprehension Test***

Section	Number of Items	Item Type	Presentation Mode
1	5	Dialogue-completion	Oral
2	5	Dialogue-completion	Written
3	5	Question and answer	Oral
4	5	Question and answer	Written

To counterbalance the order of the sections, two forms, Form A and Form B, were developed. For dialogue-completion items, for instance, around one half of the test-takers from each proficiency group took Form A, receiving



*The Effects of Presenting Multiple-Choice Test...*

questions 1 to 5 in the oral mode and 6 to 10 in the written mode. The other half who took Form B received questions 1 to 5 in the written mode and 6 to 10 in the oral mode. The two forms of the test and the form assignment are shown in Table 2.

**Table 2. *Forms of the Test***

	Form A	Form B
Dialogue-completion	Oral (1-5)- Written (6-10)	Written (1-5)- Oral (6-10)
Question and answer	Written (11-15)- Oral (16-20)	Oral (11-15)- Written (16-20)
Total items	20	20

Both test forms had similar test types and number of items. In the oral mode, the pre-recorded input included the stimulus texts and the multiple-choice items (both the stem and the options). In the written mode, the input merely comprised the stimulus texts and the items were presented on a test paper.

The test was piloted with 20 advanced EFL learners other than those participating in the study to determine the difficulty of the test items, the clarity of the stimulus, appropriateness of speech rates, and the length of questions. Moreover, the difficulty of the tests in the two versions was well controlled to balance the language discrepancies between spoken and written forms. They were checked by two experienced teacher at the ILI. A male native speaker of English also made the recordings.

The reliability of the test measured through KR-20 formula was 0.84. The content validity of the test confirmed by two well-experienced ILI teachers who had more than 15 years of teaching experience at the ILI. Moreover, the concurrent validity of the test was measured by administering the listening section of TOEFL to the pilot group. Afterwards, the Pearson correlation

coefficient between the scores on both tests was measured as 0.78 which is considered a high correlation.

### 3.2.2. Questionnaire

A nine-item questionnaire was adapted from Chang and Read (2013) to examine the participants' perceptions toward the test tasks. Items 1 to 6 used a five-point Likert response scale ranging from *strongly agree* to *strongly disagree*. Items 7 and 8 dealt with test-taking strategies used by efficient listeners. The response options for these two items ranged from *always* to *never*. The last item evaluated the overall difficulty of the two modes. Moreover, the questionnaire was piloted on 20 advanced level learners to assess its clarity and intelligibility. The reliability of the questionnaire was measured through Cronbach's alpha as 0.81.

### 3.3. Data Collection Procedure

Group A was given the oral mode first, followed by the written mode, and Group B had the reverse order (see Table 3). The total administration time was 20 minutes and both groups took the test on the same day. Moreover, the participants answered the questionnaire immediately after the tests.

**Table 3. Procedure of the Study**

Group A	Group B
Taking Oral multiple-choice test	Taking Written multiple-choice test
↓	↓
Taking Written multiple-choice test	Taking Oral multiple-choice test
↓	↓
Filling out the questionnaire	Filling out the questionnaire

### 3.4. Data Analysis

SPSS 22.0 for Windows was employed for the statistical analysis. To answer the first three research questions, a repeated measure two-way ANOVA was used for analysis to examine the effect of presentation mode and item type on test-takers' L2 listening performance and perception. The two-way ANOVA was used because there were two independent variables (presentation mode and item type), and since all participants took the listening test in all different formats (both the oral and the written modes, and both the discourse-completion and the question-and-answer item types), the repeated measure was employed. Moreover, the questionnaire results were analyzed item by item.

## 4. Results

### 4.1. Results of Repeated Measures Two-way ANOVA

For the first research questions, a repeated measures two-way analysis of variance (ANOVA) was employed to examine the effect of mode and item type on the participants' scores on the listening comprehension test. The participants' mean scores and standard deviations for four different sets of listening comprehension tests are presented in Table 4. One point was given to each item and the four formats, DC-Oral, DC-Written, Q and A-Oral, and Q and A-Written, had 5 items each, so the highest score and the lowest score one could get was 5 and 0, respectively.

**Table 4. Descriptive Statistics**

	DC-Oral		DC-written		Q and A-Oral		Q and A-Written	
	M	SD	M	SD	M	SD	M	SD
Group	3.65	0.482	3.74	0.425	3.53	0.516	3.92	0.491

*Note:* DC: Dialogue-completion type, Q and A: Question-and-answer type

To make sure that the distribution was normal, Shapiro-Wilk test was used. The significance values showed that distributions were normal because the  $p$ -values were more than 0.05. To explore the effects of the presentation modes and item types on advanced learners' listening comprehension, a repeated measure two-way ANOVA was run (Table 5).

**Table 5. Results of Two-way ANOVA**

	Source	<i>SS</i>	<i>df</i>	<i>MS</i>	<i>F</i>	<i>p</i>	Partial $\eta^2$
	Item Type	2.874	1	2.874	4.921	0.127	0.091
	Mode	0.691	1	0.691	1.506	0.314	0.023
Group	IT*Mode	0.084	1	0.084	0.735	0.625	0.007
	Error	7.253	79	0.162			

As shown in Table 5, no significant main effects were shown for Item Type and Mode [ $F(1, 79)=4.921, p = 0.127$ ;  $F(1, 79)=1.506, p= 0.314$ ]. Moreover, there was no significant interaction effect between Item Type and Mode [ $F(1, 79)= 0.084, p =0.625$ ].

#### **4.2. Results of the Questionnaire**

To answer the second research question, the participants' responses to the questionnaire were analyzed. The questionnaire responses are displayed in Tables 6 and 7. Table 6 reports the responses to the first eight items.

*The Effects of Presenting Multiple-Choice Test...*

**Table 6. Responses of Participants to Questions**

Item no.	Frequency of rating (percentages)				
	1 (strongly agree)	2 (agree)	3 (neutral)	4 (disagree)	5 (strongly disagree)
1	I felt oral MC was easy because I did not have to worry about not understanding the questions and answer options.				
	4 (5%)	10 (13%)	28 (35%)	33 (41%)	5 (6%)
2	I felt written MC was easy because I did not have to worry that I might not aurally comprehend the oral questions and answer options.				
	9 (11%)	31 (38%)	29 (37%)	8 (10%)	3 (4%)
3	I felt oral MC was difficult because I could not read the questions and options before hearing the input.				
	6 (8%)	35 (43%)	24 (30%)	12 (15%)	3 (4%)
4	I felt written MC was difficult because I could not finish reading the questions and options.				
	2 (2%)	6 (8%)	31 (39%)	34 (42%)	7 (9%)
5	I had no difficulty remembering the questions and options while doing the oral MC questions.				
	7 (9%)	13 (16%)	25 (32%)	31 (38%)	4 (5%)
6	Doing written MC was difficult because I had to read and listen at the same time.				
	2 (2%)	5 (6%)	28 (35%)	37 (47%)	8 (10%)
	<b>Frequency of rating (percentages)</b>				
	<b>1 (always)</b>	<b>2 (often)</b>	<b>3 (sometimes)</b>	<b>4 (rarely)</b>	<b>5 (never)</b>
7	When doing oral MC items, I did not have to wait until the speaker finished all the options. I chose the right one once I heard it.				
	7 (9%)	24 (30%)	23 (28%)	20 (25%)	6 (8%)
8	When doing written MC, I did not have to finish reading all options. I chose the right one once I heard it.				
	11 (14%)	30 (36%)	25 (32%)	10 (13%)	4 (5%)

Items 1, 3, 5, and 7 dealt with the participants' perceptions toward the oral test items, while Items 2, 4, 6, and 8 focused on the written mode. The results showed that a large number of the respondents perceived oral items as difficult (Items 1 and 3). About 43% of participants responded that they had problem in remembering the whole test item (Item 5). This indicated the role of memory in answering the oral items. The responses to Item 7 were mixed; the responses were spread almost evenly on both sides.

On the other hand, the responses to the written items inclined to the positive side of the scale, with 49% strongly agreeing or agreeing that the written mode was easy (Item 2). Only did 10% report that they had difficulty reading the whole question (Item 4). In addition, 8% strongly agreed or agreed that listening simultaneously to oral input and reading test items was difficult (Item 6). In Item 8, half of the participants agreed or strongly agreed that they often chose their response option immediately, once they heard it. In addition to the eight specific items, the participants were asked to assess whether one mode was easier than the other (see Table 7).

**Table 7. Which Mode is Easier? (Oral, Written, No Difference)**

	Frequency
Oral mode	9 (11%)
Written mode	41 (51%)
No difference	30 (38%)
Total	80

The results presented in Table 7 show that more than half of participants viewed written MC items easier than the oral MC ones. Generally, most participants considered test items in the written mode easier than the oral mode.

## **5. Discussion and Conclusion**

The results revealed that the advanced EFL learners performed equally well whether the questions and answers were recorded or written. Moreover, no significant main effects were shown for Item Type and Mode. Furthermore, there was no significant interaction effect between Item Type and Mode. This result is consistent with earlier studies which compared the scores of oral and written mode between participants with higher and lower proficiency levels (Chang & Read, 2013; Yanagawa & Green, 2008). Moreover, the participants' scores on the question-and-answer type items were similar to those of dialogue-completion items. While the input listening stimuli for the dialogue-completion items consisted of 3 turns, that for the question-and-answer items consisted of 6 turns. The answer options for the question-and-answer items were also relatively longer. The average number of words in an option for a question-and-answer item was 9.2, whereas that for a dialogue-completion item was 7.5. Generally, the two modes of presenting MC items had no significant difference in participants' performance. Moreover, advanced EFL learners perceived the written mode as easier than the oral mode since the written mode presented questions and answer options for them to read.

The advanced learners performed similarly in the written and the oral items. However, in the questionnaire, the majority of them (51%) responded that written items were easy because they did not have to worry about not comprehending the oral prompts. One reason can be lack of opportunity in the oral mode to pre-read the questions and answer options (Chang & Read, 2006; Yanagawa & Green, 2008). Few participants (18%) regarded the oral items as easy. One reason can be the fact that the oral mode is less familiar to EFL learners. In fact, this listening test mode is not as widely used as the written mode. Moreover, the EFL learners are not able to have an access to the clues

from previewing the items in the oral mode, which has been viewed as a key factor (Yanagawa & Green, 2008). More than half (57%) of participants disagreed that reading and listening simultaneously caused difficulties while responding to items in the written mode (Item 6). This is in line with studies showing that higher level students are also good readers (Chang, 2005). Advanced learners also seemed to be strategically skillful in responding to test items (Items 7 and 8) in both modes, and in particular the written items.

Taking into account the participants' listening test results, it seemed that when a learner possesses both good listening and reading skills, their performance will be less affected by the mode of testing, as Yanagawa and Green (2008) also noted. From a psychological perspective, EFL learners feel more secure when they can visualize the written form of a word, as many studies have shown that a source of difficulty for L2 listeners is that they are slow at matching written form with spoken form (Chang & Read, 2006, 2013). From a strategic perspective, the written mode allows test-takers to have more control over their strategies. For these reasons, it is understandable that the majority of the participants preferred the written to the oral mode.

It can be concluded that both oral and the written modes were found to have strengths and weaknesses with respect to construct validity. The influence of reading ability on test-takers' listening comprehension performance could be avoided with the oral presentation mode by providing all questions and options orally, but this imposed an additional memory burden on the test-takers. On the other hand, test-takers felt less anxious with written questions and options, because they did not have to remember them. However, the written mode required a certain level of reading proficiency and entailed other reading related construct-irrelevant factors, such as predicting without listening and using word matching strategies.



### *The Effects of Presenting Multiple-Choice Test...*

Based on participants' performance and perception on the different formats of listening tests, some suggestions can be made for EFL test developers and teachers. It is recommendable to give the MC items orally, because it does not require reading ability which could impede the test-takers with low reading proficiency from fully demonstrating their listening ability. This is mainly advisable for the dialogue-completion items, since the test-takers have to choose a response that is a part of the whole orally-given conversation. However, the written mode can be more proper when the options are too long for the oral mode and too difficult for the test-takers to process with their memory capacity only by listening. The written mode can also have a positive effect on the test-takers by reducing their test anxiety. However, if the questions and options are to be delivered in written form, they should be written in easy language not to require a high level of reading ability.

There were some limitations in this study. This study did not control for item difficulty, e.g., the order of the correct option in the oral mode, the location of necessary information for correct options, and the lexical overlap between input text and test items and all variables that were shown to be significant in previous research (Kostin, 2004; Yanagawa & Green, 2008). Moreover, to relieve the memory burden of the aural mode not by giving the written options but by reducing their number, the interaction effect between the number of options and the presentation mode needs to be investigated. Finally, the appropriate length or complexity of the language for the options in the written mode can be examined in relation to the test-takers' proficiency level.

## **Acknowledgements**

The authors wish to thank the Ahvaz Branch, Islamic Azad University for the financial support. Moreover, the authors would like to thank all the editors and reviewers for their comments in the development and improvement of this paper.

## **References**

- Allan, D. (2004). *The Oxford placement test*. Oxford: Oxford University Press.
- Bachman, L. F. (1990). *Fundamental considerations in language testing*. Oxford, UK: Oxford University Press.
- Bachman, L. F., & Palmer, A. S. (1996). *Language testing in practice: Designing and developing useful language tests*. Oxford, UK: Oxford University Press.
- Brindley, G., & Slatyer, H. (2002). Exploring task difficulty in ESL listening assessment. *Language Testing, 19*(4), 369-394.
- Buck, G. (1991). The testing of listening comprehension: An introspective study. *Language Testing, 8*(1), 67-91.
- Buck, G. (2001). *Assessing listening*. Cambridge: Cambridge University Press.
- Chang, A. C. (2005). The perceived effectiveness of question preview in listening comprehension tests by EFL learners. *New Zealand Studies in Applied Linguistics, 11*, 75-99.
- Chang, A. C. (2008). Listening strategies of L2 learners with varied test tasks. *TESL Canada Journal, 26*(1), 1-26.
- Chang, A. C., & Read, J. (2006). The effects of listening support on the listening performance of EFL learners. *TESOL Quarterly, 40*(2), 375-397.
- Chang, A. C., & Read, J. (2013). Investigating the effects of MC listening test items in the oral versus written mode on L2 listeners' performance and perceptions. *System, 41*, 575-586.

*The Effects of Presenting Multiple-Choice Test...*

- Cheng, H. F. (2004). A comparison of MC and open-ended responses formats for the assessment of listening proficiency in English. *Foreign Language Annals*, 37, 544-555.
- Cook, V. (2013). *Second language learning and language teaching*. London, UK: Routledge.
- Freedle, R., & Kostin, I. (1999). Does the text matter in a multiple-choice test of comprehension? The case for the construct validity of TOEFL's minitalks. *Language Testing*, 16, 2-32.
- Gilakjani, A., & Ahmadi, M. (2011). A study of factors affecting EFL learners' Listening comprehension and the strategies for improvement. *English Language Teaching and Research*, 2, 977-988.
- Hemmati, F., & Ghaderi, E. (2014). The effect of four formats of multiple-choice questions on the listening comprehension of EFL learners. *Procedia-Social and Behavioral Sciences*, 98, 637-644.
- Hughes, A. (2003). *Testing for language teachers* (2nd ed.). Cambridge University Press, Cambridge.
- Imura, H. (2010). Factors affecting listening performance on multiple-choice tests: The effects of stem/option preview and text characteristics. *Language Education & Technology*, 47, 17-36.
- Kostin, I. (2004). *Exploring item characteristics that are related to the difficulty of TOEFL dialogue items*. Princeton, NJ: Educational Testing Service.
- Mendelsohn, D. (2002). The lecture buddy project: An experiment in EAP listening comprehension. *TESL Canada Journal*, 20(1), 64-73.
- Ohata, K., (2006). Auditory short-term memory in L2 listening comprehension processes. *Journal of Language Learning*, 5, 21-28.
- Rost, M. (2002). *Teaching and researching listening*. London: Longman.
- Rubin, J. (1994). A review of second language listening comprehension research. *The Modern Language Journal*, 78(2), 199-221.
- Teng, H. C. (1998). The effects of text types and task types on English listening comprehension. *English Teaching Journal*, 23, 5-18.

- Vandergrift, L. (2003). From prediction through reflection: Guiding students through the process of L2 listening. *The Canadian Modern Language Review*, 59(3), 425-440.
- Weir, C. J. (1993). *Understanding and developing language tests*. Upper Saddle River, NJ: Prentice Hall.
- Weir, C. J. (2005). *Language testing and validation*. London, UK: Macmillan.
- Wolf, D. F. (1993). A comparison of assessment tasks used to measure foreign language reading comprehension. *The Modern Language Journal*, 77(4), 473-489.
- Wu, Y., (1998). What do tests of listening comprehension test? A retrospection study of EFL test-takers performing a MC task. *Language Testing*, 15, 21- 44.
- Yanagawa, K., Green, A. (2008). To show or not to show: The effects of item stems and answer options performance on a MC listening comprehension test. *System*, 36(1), 107-122.

## **Appendix**

### *Listening Comprehension Test*

- (A) The two classes meet in an hour and a half.
  - (B) The class meets three hours per week.
  - (C) Each half of the class is an hour long.
  - (D) Two times a week the class meets for an hour.
- 
- (A) He hasn't yet begun his project.
  - (B) He's supposed to do his science project next week.
  - (C) He needs to start working on changing the due date.
  - (D) He's been working steadily on his science project.
- 
- (A) The professor drowned the cells in a lab.
  - (B) The lecture was long.
  - (C) The lecturer divided the lecture into parts.
  - (D) The biologist tried to sell the results of the experiment.
- 
- (A) Housing within his budget is hard to locate.
  - (B) It's hard to find his house in New York.
  - (C) He can't afford to move his house to New York.
  - (D) Housing in New York is unavailable.
- 
- (A) The boisterous students made the teacher mad.
  - (B) The teacher angered the students with the exam results.
  - (C) The students were angry that the teacher was around.
  - (D) The angered students complained to the teacher.
- 
- (A) It's rained unusually hard this year.
  - (B) There hasn't been any rain for many years.
  - (C) It's been many years since it rained.
  - (D) He doesn't like rain.

*Iranian Journal of Applied Language Studies, Special Issue, Autumn 2017*

- (A) The agent was standing in line with his passport.  
(B) The line to get new passports is very long.  
(C) The woman must wait her turn to get her passport checked.  
(D) He can check her passport instead of the agent.
- (A) All the lawyer's preparation did no good.  
(B) The lawyer prepared nothing for the case.  
(C) It wasn't work for the lawyer to prepare for the case.  
(D) The lawyer didn't work to prepare for the case.
- (A) He's not really happy.  
(B) The contractor's work was satisfactory.  
(C) He would rather work with the contractor himself.  
(D) He was already contacted about the work.
- (A) She'd like some pie.  
(B) It's easy to buy it.  
(C) The task the man's working on isn't difficult.  
(D) It's easier to prepare pie than do what the man is doing.
- (A) "I cannot understand you because you speak too fast. Please speak more slowly."  
(B) "I think you should speak more slowly in class. You're speaking very fast."  
(C) "I was wondering if you could slow down a little. I have a hard time following you."  
(D) "I think you speak too fast, so I can't understand you."
- (A) "Sorry. I want to buy a new copy."  
(B) "I have spilled a cup of coffee over your magazine. Do you still want it?"  
(C) "I ruined your magazine so I'll replace it."  
(D) "Sorry, Sean, I ruined your magazine. I'll buy you a new copy."
- "No, because I want to read it myself."  
"Sorry, I'll need it myself this afternoon."  
(C) "No, I'm sorry, I can't."  
(D) "Sure, here you are."

*The Effects of Presenting Multiple-Choice Test...*

- (A) "Nikita, I need to go to the doctor this afternoon. Could you possibly take my shift?"
- (B) "Nikita, could I go to the doctor this afternoon? I haven't been feeling well."
- (C) "Nikita, would you mind taking my shift this afternoon? I'd really appreciate it." (D)
- "Nikita, I really haven't been feeling very well. I think I'll go see my doctor this afternoon."

- (A) "Sorry, I know I'm late, I had a problem at the office."
- (B) "Sorry, I got held up at the office. Have you been waiting long? "
- (C) "I got stuck at the office. Sorry to keep you waiting."
- (D) "Hey, Jack. How long have you been here? I'm late because I had a problem at the office."

- (A) "Well, actually I'm busy now, but how about tonight?"
- (B) "I am sorry. I have to go to my parent's house now."
- (C) "It is too heavy to move it by myself. Maybe we can get some help later."
- (D) "Of course. I'd be happy to give you a hand with that."

- (A) "Mark, I'm really sorry to bother you but could I possibly do an interview with you?"
- (B) "Mark, I need to do an interview with somebody for class. Could I do one with you real quick?"
- (C) "Mark, do you have the time to answer a few questions? It's for a real important class project."
- (D) "Mark, I have to get this interview for class done tonight. Could I just ask you a couple of questions?"

- (A) "I know, I'm really sorry. I just couldn't get away at work. I'll explain to Professor Johnson."
- (B) "Oh, I am sorry. I had to work late yesterday. But I'll just work extra on it now."
- (C) "Sorry about that. I guess you'd like me to go and tell Professor Johnson that it's all my fault."
- (D) "I had to stay a bit later at work and that's why I did not come yesterday. I really apologize."

*Iranian Journal of Applied Language Studies, Special Issue, Autumn 2017*

- (A) “Oh, Kevin, I’m sorry. You don’t even have \$ 200?”
  - (B) “Sorry, I wouldn’t even feel comfortable lending anybody \$ 200.”
  - (C) “Sorry, I don’t have that much right now. How about \$ 200?”
  - (D) “Sorry, Kevin, but I only have \$200. I could give you some of that.”
- 
- (A) “I’ll pay you whatever it costs.”
  - (B) “Oh, I’m terribly sorry. How very clumsy of me.”
  - (C) “Well, it’s just a vase. But I’ll give you the money for it.”
  - (D) “I’m so sorry. I’ll buy you a new one, of course.”