

# The Effect of Visual and Textual Accompaniments to Verbal Stimuli on the Listening Comprehension Test Performance of Iranian High and Low Proficient EFL Learners

Ataollah Maleki

English Language Department, Faculty of Medicine, Zanjan Medical Sciences University, Zanjan, Iran  
Email: atamaleki@hotmail.com

Masoud Safaee Rad

Azad University of Hamedan, Iran  
Email: masoud.safaei@gmail.com

**Abstract**—The present study aimed at comparing the effect of a series of static images with that of condensed paraphrased scripts accompanying verbal stimuli in listening comprehension tests supplemented with visual or textual input. It utilized three batteries of IELTS listening tests: one sampling 58 participants to carry out the study with, another with supplemented visual or textual aids, and the other without such aids. The participants were then divided into higher proficient listeners group (HPL) and the lower proficient listeners group (LPL) on the basis of their performance on a general listening module of the IELTS examination. The HPL was subsequently divided into two groups: one receiving static images with verbal stimuli, another receiving script presentations with the same verbal stimuli. The LPL was also divided into two groups with the same procedure. The analysis of results showed that listening tests with static images yielded better performance than script presentations, visual aids were more advantageous to lower proficient listeners while textual aids were more advantageous to higher proficient listeners, and tests with more static images yielded much better performance than those with fewer static images.

**Index Terms**—listening comprehension, verbal stimuli, visual aids, textual aids, static images, higher proficient, lower proficient

## I. INTRODUCTION

As in other fields of study, new technology has had a profound influence on language education industry. The application of such pedagogically prolific technology has been an object of research by language educationalists (see, for example, Bubel, 2008; Zielinski, 2008; Ockey, 2007; wagner, 2007; Gruba, 2006). One important, but less considered area, is listening comprehension and testing, which is now being taught and assessed using modern technology, i.e., multimedia and computerized equipment. Bachman (1990) notes that listening comprehension assessment needs to correspond with language used in the real life to be authentic. As such, listening comprehension tests offer the most challenging area of language assessment and research. The advent of reliable and high quality video recording equipment with listening texts accompanied by video texts and auditory channel juxtaposed with visual channel has eased the problem. This can be seen in the emergence of manifold listening tests as a result of different research projects conducted on the effects multimedia application on listening comprehension tests (see Robin, 2007). (Hoven, 1999; Balizet et al., 1999).

Most real life listening involves both verbal and nonverbal information; however, listening sections in current high stakes tests of English such as IELTS and TOEFL lack the non verbal (visual) component which is highly abundant in target language use. Wagner (2002) and Ginther (2001) assert that presenting a more realistic and valid test of communicative competence, increasing the face validity of the test, ensuring better representations of actual communicative situations, and enhancing the measurement of the test taker's listening comprehension may help make audiovisual presentation of the listening comprehension stimuli more advantageous to audio only presentation. On account of this, the present study is intended to address the problem by first exploring the effect of added visual component to the verbal stimuli on an IELTS listening comprehension test and, subsequently, comparing this visual aid with another help option, textual aid. Secondly, it sets out to investigate the effects of these supplemented components across different levels of test takers, viz, higher and lower proficient listeners.

The study is significant as it aims at finding out the impact and difference in the performance of two modes of modified listening comprehension test inputs. Since multidimensional representation of the input is thought to enhance

the learning of listening and its testing, the outcome of the research will contribute to improvement of teaching and testing of listening comprehension.

#### A. *Research Questions*

- a. Is there a significant difference between performance on listening comprehension tests accompanied with static images and those accompanied with condensed and paraphrased scripts?
- b. Do low proficient listeners benefit from visual and textual aids more than high proficient listeners on listening comprehension tests?
- c. Is there a meaningful difference between performance on listening tests accompanied with more static images and those accompanied with fewer static images?

#### B. *Hypotheses*

The following null hypotheses are proposed to investigate the research questions:

- a. There is no significant difference between performance on listening comprehension tests accompanied with static images and those accompanied with condensed and paraphrased scripts.
- b. There is no significant difference between low proficient listeners' benefiting from visual and textual aids and that of high proficient listeners.
- c. There is no meaningful difference between performance on listening tests accompanied with more static images and those accompanied with fewer static images.

## II. REVIEW OF THE LITERATURE

Although listening skill plays a vital role in second language learning (Field, 2002; Lynch, 2002; Peterson, 2001; Rost, 2001), it was not taken seriously until late 19th century (Rost, 2001). Rost thinks that as a result of developments in anthropology, education, linguistics, sociology and even global politics in the late 19th century, teaching listening has experienced many changes. Rost (2001) also records that 1960s witnessed the rise of Communicative Language Teaching Movement and Conversation Analysis, which regarded listening as an integral part of communicative competence (see, also, Nunan, 2002; Field, 2002; Morely, 2001). During 1970s and 1980s, spontaneous spoken language and scripted texts with a degree of authenticity and authentic recordings were regarded pertinent input (Rost, 2001; Willmington and Steinhilber, 1993). Since 1990s up to now, spontaneous spoken language and scripted texts with a degree of authenticity, face to face learner talk, expert speaker learner interaction, and authentic recordings have been the main input of this period. Here, instructional focus is the use of listening strategies for enhancing comprehension and coping with problems (Nunan, 2002). Robin (2007) asserts that over the last 25 years authentic audio and video in which real media are utilized have received considerable attention. It is now possible to deliver listening texts using video texts which extend beyond just auditory channel to encompass visual channels (Wagner, 2007). Visual information, which is richly present in real life situation, is a primary non verbal input that can aid listeners to take part in real communication (Rost, 2007; Buck, 2001).

Textual support for listening comprehension is another type of listening supplement offered to L2 listeners (Chang and Read, 2007). It has recently been investigated from many different angles including listeners' interaction with subtitles and transcripts (Grgurović and Hegelheimer, 2007), the effects of captions on listening comprehension (Robin, 2007; Winke et al., 2010), synchronous bimodal chat that includes sound and text (Blake et al., 2008), the effect of textual support on listening comprehension tests (Chang and Read, 2007) and the interaction of text and context in a multimedia environment (Kramsch and Anderson, 1999). According to Kramsch and Anderson (1999), the crucial role of textual aids in listening comprehension is incontrovertible, but there are still disputes over how they are different from visual aids (Chun and Plass, 1997).

Testing listening has posed a real dilemma for language testers (Balizet et al., 1999). According to Lynch (2002), "It is not easy to design listening comprehension tests that reflect the purpose of real life listening, partly because most listening tasks require either no response or a minimal response from the listener" (P. 42). The transient nature of the spoken language (Hughes, 2003), the internal and unobservable nature of the listening comprehension process (Chastain, 1988; Lynch, 2002), the reciprocal nature of real life situations (Lam, 2002; Nunan, 2002), the interference from knowledge or skills other than listening (Buck, 2001; Field, 2002; Lynch, 2002), the role of non verbal information (Wagner, 2002; Wagner, 2007), and the complexity of real world listening context (Buck, 2001) will certainly compound the problem.

Increasingly, listening assessment has attempted to investigate the effects of task characteristics and task conditions on learners' performance in competency based listening assessment tasks (Brindly and Slatyer, 2002), the cognitive processes involved in the listening comprehension (Bubel, 2008), the inclusion of still image or video in computer based listening tests (Ockey, 2007), test takers' viewing behavior during an L2 video listening test (Wagner, 2007), the effects of different visual types on listening comprehension tests (Ginther, 2001; Ginther, 2002), the effects of speech rate modification on listening comprehension (Griffiths, 1990), the role of technology in test construction, administration, and assessment (Chalhoub-Deville, 2001), the effects of item stems and answer options on listening comprehension tests (Yanagawa and Green, 2008), the effects of emotion on comprehension assessment (Chen et al.,

2007), the effects of multiple choice listening question on listening comprehension test performance (Yi'an, 1998), the effects of different text and question types on L2 listening test scores (Shohamy and Inbar, 1991), and the effects of test method and item preview on the measurement of listening comprehension (Buck, 1991).

Despite the myriad researches conducted, "... the assessment of listening remains far behind current views of listening" (Rost, 2001, p. 13). Attempting to bridge this gap, researchers have emphasized the importance of authentic input and authentic listening (Chang and Read, 2007; Hughes, 2003; Weir, 1990), focus on meaning (Richards, 2005), computer and technology applications in language testing (Chalhoub-Deville, 2001), and the role of context in listening assessment (Fulcher and Davidson, 2007).

Trying to investigate whether simultaneous presentation of visuals with the verbal text improved L2 listeners' performance, Ginther (2002) included content and context visuals. Ginther (2002) concluded that visuals increased the face validity of listening tests and their accompaniments to verbal stimuli ensured finer examples of actual communicative situations.

Another study which was conducted by Ockey (2007) included a series of static images or videos in computer based listening tests to see how they would affect the construct validity of listening tests. The results of the study implied that listening comprehension tests must include static images with only minimally altering the listening construct to be measured. In contrast, video implementation required a rethinking of listening construct on the grounds that test takers showed various engagements with video stimuli on computer based listening tests.

Chen et al. (2007) considered another aspect of audio visual content, efficacy of affective annotation in a multimedia listening comprehension environment. They designed Emo Player software which enabled them to make interactions upon the emotional status of the listeners through a timeline model. This specific type of modified listening input delivery, exploring both visual and auditory channels to listening, resulted in a better comprehension of a video clip in a limited period of time.

Video mediated listening tests, on balance, have gained approval in recent years (Brindley, 2001). Their efficiency has been explored from different viewpoints such as the effect of film discourse and the use of video in listening tests (Bubel, 2008), the use of videotexts for listening comprehension by adopting a media literacy perspective (Gruba, 2006), the application of listening theory in the context of visual media (Hoven, 1999), the test takers' interaction with video listening texts (Wagner, 2007), and the listeners' integration between verbal and visual information (Chun and Plass, 1997).

The important role that textual support plays in listening comprehension has been well documented in the literature (Chang and Read, 2007; Robin, 2007). Robin (2007) recommends updating the definition of authenticity and including technological possibilities supporting a text. Grgurović and Hegelheimer (2007) presented two types of textual help options, i.e., subtitles and transcripts, to test takers and wanted to explore whether subtitles or transcripts were more effective as modified listening input to L2 listeners. They believed that multimedia presentation of listening materials accompanied with textual alternatives would yield better results than paper and pencil listening comprehension tests. Results of their study showed that the participants of the study made use of subtitles more frequently and for longer periods of time than the transcripts. The subtitles were also used by higher the proficiency group more frequently and for longer period of time than the lower proficiency group. Both groups, however, interacted in a remarkably similar manner with the transcripts.

More recently, Winke et al. (2010) have sought to investigate the effects of video based listening activities with and without captioning on learners of different first language backgrounds. They concluded that captioned videos were more effective than non captioned videos partly because they improved attention and processing by reinforcing previous knowledge and mainly because "... captions facilitate language learning by helping learners visualize what they hear" (P.65). They asserted that captions were important pedagogical tools which connect auditory and visual input in authentic listening videos and could be used as "crutch" for second language listeners.

### III. METHOD

#### A. Participants

Out of a pool of 100 Iranian EFL students 62 were selected on the basis of their performance on an IELTS listening comprehension test. One participant was excluded due to his illness during the second test administration. Three others were also removed from the study because of their extreme scores. Therefore, the final number of participants was 58, which was divided into two 29 homogeneous, male female groups.

The participants who scored one standard deviation above the mean in a general listening module of the IELTS examination were regarded as the Higher Proficiency Listeners group (HPL), which included 15 females and 14 males. On the other hand, those who scored one standard deviation below the mean were considered as the Lower Proficiency Listeners group (LPL), which consisted of 14 females and 15 males.

The HPL group was then divided into two subgroups: one receiving static images along with listening texts, abbreviated as the HSI (Static Image) group and the other receiving condensed paraphrased script presentations along with listening texts, abbreviated as the HSP (Script Presentation) group. The HSI group consisted of 7 females and 8 males, but the HSP group consisted of 8 females and 6 males.

In the same way, the LPL group was divided into two subgroups: one receiving static images along with listening

texts, abbreviated as the LSI (Static Image) group and the other receiving condensed paraphrased script presentations along with listening texts, abbreviated as the LSP (Script Presentation) group. LSI group consisted of 7 females and 7 males, but LSP group consisted of 7 females and 8 males.

All participants were native speakers of Persian. Their age ranged from 16 to 26 with the mean of 21.22. Their English language learning experience ranged from 3 to 7 years with the mean of 4.6. Mean age of the HPL group was 22.16 and that of the LPL group was 20.31. The English language learning experience mean of the HPL group was 4.64 years and that of the LPL group was 4.54 years.

#### B. Instrumentation

In order to obtain measurable data to get the intended results of the study, we used the following instruments, which were adopted from Harrison and Whitehead (2005):

1. An IELTS listening module test
2. IELTS listening test one—a four section forty question IELTS listening module test
3. IELTS listening test two—a four section forty question IELTS listening module test with two different versions of modified input supplemented to the listening text—one with visual aid consisting of a series of still images related to the verbal stimuli of the listening text and another with textual aid consisting of a series of condensed and paraphrased script presentations related to the verbal stimuli of the same text.

The Kurder Richardson 20 formula was used to estimate the reliability of tests one and two ( $r_1 = 0.754$ ;  $r_2 = 0.743$ ).

#### C. Design

Two different types of modified listening input, namely visual and textual supplementations to the verbal stimuli, are provided to be administered via a computerized multimedia IELTS listening comprehension test (Listening test 2). These tests results are then compared with each other and with the results obtained from another IELTS listening comprehension test with no modifications, that is, an audio only listening test module (Listening test 1). Both tests, however, are similar, except for the modifications in the listening test 2.

### IV. RESULTS

Table 1 presents data on the performance of the LSI, LSP, HSI, and HSP, HPL, and LPL groups on the listening test one.

TABLE 1  
DESCRIPTIVE STATISTICS FOR THE IELTS LISTENING TEST ONE SCORES OF THE LSI, LSP, HSI AND HSP, HPL, AND LPL GROUPS

	Mean	Median	Maximum	Minimum	Std Deviation	Variance
LSI	15.57	15.50	21	10	2.74	7.49
LSP	15.13	15.00	25	10	3.70	13.70
HSI	25.53	25.00	31	19	3.44	11.84
HSP	21.64	22.00	27	17	2.68	7.17
HPL	23.58	23.5	31	17	3.06	9.5
LPL	15.35	15.25	26	10	3.22	10.59

As depicted in table 1, the HSI group scored the highest mean (25.53), but the LSP group scored the lowest mean (15.13). Moreover, the lowest average disparity among the scores on the listening test one belongs to the HSP group ( $SD = 2.68$ ) while the highest average disparity among the scores on the listening test one belongs to the LSP group ( $SD = 3.70$ ).

Table 1 also shows the performance of the HPL and LPL groups on the listening one test. Here, the highest mean score belongs to the HPL group (23.58). The HPL group's standard deviation ( $SD = 3.06$ ), on the other hand, is lower than that of the LPL group ( $SD = 3.22$ ).

Table 2 displays data on the performance of the LSI, LSP, HSI, and HSP groups on the listening test two.

TABLE 2  
DESCRIPTIVE STATISTICS FOR THE IELTS LISTENING TEST TWO SCORES OF THE LSI, LSP, HSI, AND HSP GROUPS

	Mean	Median	Maximum	Minimum	Std Deviation	Variance
LSI	19.36	19.00	29	14	3.89	15.17
LSP	15.40	16.00	23	9	4.10	16.83
HSI	26.73	27.00	33	20	4.30	18.50
HSP	23.64	23.50	28	19	2.56	6.55

According to table 2, the highest mean score in the listening two belongs to the HSI group (26.73) and the lowest mean score belongs to the LSP group (15.40). In addition, the lowest average disparity among the scores in listening test two belongs to the HSP group ( $SD = 2.56$ ) while the highest average disparity belongs to the HSI group ( $SD = 4.30$ ).

Section one of the IELTS listening test is a non academic dialogue which comprises the first ten items of the test. Table 3 shows the performance of the LSI, LSP, HSI, and HSP groups on this section.

TABLE 3  
DESCRIPTIVE STATISTICS FOR THE IELTS LISTENING TEST TWO (SECTION ONE) SCORES OF THE LSI, LSP, HSI, AND HSP GROUPS

	Mean	Maximum	Median	Minimum	Std Deviation	Variance
LSI	6.36	10	6	5	1.60	2.55
LSP	5.60	9	6	3	2.03	4.11
HSI	6.87	10	7	1	2.13	4.55
HSP	6.43	8	6	3	1.28	1.65

As depicted here, the highest mean score in section one of the listening test two belongs to the HSI group (6.87) and the lowest belongs to the LSP group (5.60).

Section two of the IELTS listening test is a non academic monologue which comprises the second ten items of the test. Table 4 shows the performance of the LSI, LSP, HSI, and HSP groups on this section.

TABLE 4  
DESCRIPTIVE STATISTICS FOR THE IELTS LISTENING TEST TWO (SECTION TWO) SCORES OF THE LSI, LSP, HSI, AND HSP GROUPS

	Mean	Maximum	Median	Minimum	Std Deviation	Variance
LSI	6.07	8	6.50	2	1.59	2.53
LSP	4.47	7	5.00	1	1.51	2.27
HSI	7.80	10	8.00	4	1.66	2.74
HSP	6.50	8	7.00	3	1.34	1.81

According to table 4, the highest mean score in section two on listening test two belongs to the HSI group (7.80) and the lowest belongs to the LSP group (4.47).

Section three of the IELTS listening test is an academic discussion which consists of the third ten items in the test. Table 5 shows data on the performance of the LSI, LSP, HSI, and HSP groups on this section.

TABLE 5  
DESCRIPTIVE STATISTICS FOR THE IELTS LISTENING TEST TWO (SECTION THREE) SCORES OF THE LSI, LSP, HSI, AND HSP GROUPS

	Mean	Maximum	Median	Minimum	Std Deviation	Variance
LSI	5.21	7	5.50	3	1.12	1.26
LSP	4.07	6	5.00	1	1.39	1.92
HSI	6.80	9	7.00	3	1.66	2.74
HSP	5.71	8	6.00	3	1.44	2.07

As depicted in table 5, the highest mean score in section three belongs to the HSI group (6.80) and the lowest belongs to the LSP group (4.07).

Section four of the IELTS listening test is a mini lecture monologue which consists of the fourth ten items in the test. Table 6 displays data on the performance of the LSI, LSP, HSI, and HSP groups on this section.

TABLE 6  
DESCRIPTIVE STATISTICS FOR THE IELTS LISTENING TEST TWO (SECTION FOUR) SCORES OF THE LSI, LSP, HSI, AND HSP GROUPS

	Mean	Maximum	Median	Minimum	Std Deviation	Variance
LSI	1.71	7	1	.00	2.13	4.53
LSP	1.27	3	1	.00	.96	.92
HSI	5.27	10	5	1.00	2.43	5.92
HSP	5.00	7	5	2.00	1.47	2.15

Here, the highest mean score is 5.27 which belongs to the HSI group and the lowest belongs to the LSP group (1.27).

## V. DATA ANALYSIS

Based on the information in table 2 for the LSI and LSP groups' performance on listening test 2, an independent samples t test was run to measure the differences between the scores of the two groups (Table 7).

TABLE 7  
INDEPENDENT SAMPLES T TEST OF THE LSI AND LSP GROUPS' PERFORMANCE ON LISTENING TEST TWO  
**Independent Samples Test**

	Levene's Test for Equality of Variances		t-test for Equality of Means						
	F	Sig.	t	df	Sig. (2-tailed)	Mean Difference	Std. Error Difference	95% Confidence Interval of the Difference	
								Lower	Upper
Equal variances assumed	1.50	.486	2.50	27	.013	3.96	1.49	.90	7.0
Equal variances not assumed			2.50	27	.013	3.96	1.49	.90	7.0

According to table 7, there was a significant difference in the scores of the LSI group (M= 19.35, SD= 3.89) and the LSP group (M= 15.40, SD= 4.10),  $t(27) = 2.66$ ,  $P < .05$ . The magnitude of the difference in the means was very large

(eta squared= .207), with the static images accompanying the verbal stimuli explaining 20.7 percent of the variance in scores over the LSI and LSP groups.

Further, as depicted in table 2, the HSI group displayed a noticeably higher mean score for listening test two than the HSP group (26.73 V 23.64). To see whether the difference between the two means was by chance, we run an independent sample t test (Table 8).

TABLE 8  
INDEPENDENT SAMPLES T TEST OF THE HSI AND HSP GROUPS' PERFORMANCE ON LISTENING TEST TWO

	Levene's Test for Equality of Variances		t-test for Equality of Means						
	F	Sig.	t	df	Sig. (2-tailed)	Mean Difference	Std. Error Difference	95% Confidence Interval of the Difference	
								Lower	Upper
Equal variances assumed	3.80	.061	2.40	27	.028	3.09	1.32	.369	5.82
Equal variances not assumed			2.40	23.06	.027	3.09	1.30	.393	5.79

The analysis proved that the difference between the means was significant ( $t(27) = 2.32, P < .05$ ). The magnitude of the difference in the means was large (eta squared= .167), with the static images accompanying the verbal stimuli explaining 16.7 percent of the variance in scores over the HSI and HSP groups.

The statistics in tables 1 & 2 for the LSI group display a noticeably higher mean score for the listening test two than the listening test one (19.36 V 15.57). To determine whether this difference reaches statistical significance, we carried out a paired samples test. Table 9 summarizes the data.

TABLE 9  
PAIRED SAMPLES T TEST OF THE LSI GROUP'S PERFORMANCE ON LISTENING TEST ONE AND TWO

	Paired Differences					t	df	Sig. (2-tailed)
	Mean	Std. Deviation	Std. Error Mean	95% Confidence Interval of the Difference				
				Lower	Upper			
Test1-Test2	-3.78	3.30	.884	-5.69	-1.87	-4.3	13	.001

The information in table 9 indicates that there was a significant difference between the LSI group scores for test two ( $M = 19.36, SD = 3.89$ ) and test one ( $M = 15.57, SD = 2.74$ ),  $t(13) = -4.28, P < .05$ .

In a similar way, based on the statistics in tables 1 & 2, the HIS group displayed a somewhat higher mean score for the listening test two than the listening test one (26.73 V 25.53). Table 10 depicts the statistical significance of the tests means.

TABLE 10  
PAIRED SAMPLES T TEST OF THE HSI GROUP'S PERFORMANCE ON LISTENING TEST ONE AND TWO

	Paired Differences					t	df	Sig. (2-tailed)
	Mean	Std. Deviation	Std. Error Mean	95% Confidence Interval of the Difference				
				Lower	Upper			
Test1-Test2	-1.20	2.27	.587	-2.45	.059	-2.04	14	.060

The information in table 10 indicates that there was no significant difference between the HSI group scores for the test two ( $M = 26.73, SD = 4.3$ ) and the test one ( $M = 25.53, SD = 3.44$ ),  $t(14) = -2.04, P > .05$ .

According to tables 1 & 2, the LSP group performed somewhat better on the test two than the test one (15.40 V 15.13). The paired samples t test, table 11, proved that there was no significant difference between the LSP group scores for the test two ( $M = 15.40, SD = 4.10$ ) and the test one ( $M = 15.13, SD = 3.70$ ),  $t(14) = -.48, P > .05$ .

TABLE 11  
PAIRED SAMPLES T TEST OF THE LSP GROUP'S PERFORMANCE ON LISTENING TEST ONE AND TWO

	Paired Differences					t	df	Sig. (2-tailed)
	Mean	Std. Deviation	Std. Error Mean	95% Confidence Interval of the Difference				
				Lower	Upper			
Test1-Test2	- .266	2.12	.547	-1.44	.907	-.487	14	.634

As it was shown in tables 1 & 2, the mean difference between the test two scores and the test one scores for the HSP group was significant (23.64 V 21.64). This was proved by results of the paired samples tests in table 12 ( $M = 23.64$ ,  $SD = 2.56$ / $M = 21.64$ ,  $SD = 2.68$ ),  $t(13) = -3.22$ ,  $P < .05$ ).

TABLE 12  
PAIRED SAMPLES T TEST OF THE HSP GROUP'S PERFORMANCE ON LISTENING TEST ONE AND TWO

Paired Samples Test								
	Paired Differences					t	df	Sig. (2-tailed)
	Mean	Std. Deviation	Std. Error Mean	95% Confidence Interval of the Difference				
				Lower	Upper			
Test1-Test2	-2.00	2.32	.62	-3.33	-.66	-3.22	13	.007

Table 13 summarizes the information related to the LSI and HSI groups mean performance on sections two and four of the listening test two.

TABLE 13  
THE LSI AND HSI GROUPS' PERFORMANCE ON SECTIONS TWO AND FOUR OF THE LISTENING TEST TWO

Paired Samples Statistics				
LSI+HSI	N	Mean	Std. Deviation	Std. Error Mean
Section 2	29	6.94	1.80	.33
Section 4	29	3.50	2.80	.53

As it is shown in the table, the mean score for section two is approximately twice the mean score for section four (6.94 V 3.50). Results of a paired samples t test (table 14) proves that the difference was statistically significant ( $M = 6.94$ ,  $SD = 1.80$ / $M = 3.50$ ,  $SD = 2.80$ ),  $t(28) = 6.04$ ,  $P < .05$ ).

TABLE 14  
PAIRED SAMPLES T TEST OF THE LSI AND HSI GROUPS' PERFORMANCE ON SECTIONS TWO AND FOUR OF THE LISTENING TEST TWO

Paired Samples Test								
LSI+ HSI	Paired Differences					t	df	Sig (2-tailed)
	Mean	Std. Deviation	Std. Error Mean	95% Confidence Interval of the Difference				
				Lower	Upper			
Section 2-Section 4	3.40	3.00	.56	2.25	4.57	6.04	28	.000

## VI. DISCUSSION AND CONCLUSION

Drawing on the results obtained to answer the first question of the study, there was a significant difference between performance on listening comprehension tests accompanied with static images and those accompanied with condensed and paraphrased scripts. In the lower proficient listeners group, the LSI group outperformed the LSP group and in the higher proficient listeners group, the HSI group outperformed the HSP group. Therefore, results for the first question suggest that listeners' performance on listening comprehension tests supplemented with static images is better than their performance on the same listening tests supplemented with condensed and paraphrased scripts. The finding is in line with the research results reached by researchers such as Ginther (2002) and Ockey (2007).

Regarding the second research question, it was found that lower proficient listeners (the LSI group) benefited from visual aids accompanied with the verbal stimuli in the listening test two and the listening test one. However, higher proficient listeners (the HSI group) did not benefit from visual aids accompanied with the verbal stimuli in the listening test two and the listening test one. It was also found that lower proficient listeners (the LSP group) did not benefit from textual aids accompanied with the verbal stimuli in listening test two and listening test one. The higher proficient listeners (the HSP group), however, did benefit from textual aids accompanied with the verbal stimuli in the listening test two and the listening test one. Visual aids, on the other hand, helped lower proficient listeners (the LSP group) more than the higher proficient listeners (the HSP). Other researchers have also supported these findings in one way or another (see, for example, Brindley, 2001; Hoven, 1999; Chun and Plass, 1997).

Finally, answer to the third research question revealed that listening tests accompanied with more static images resulted in much better performance by the LSI and HSI groups than listening tests accompanied with fewer static images. Therefore, it can be concluded that higher frequency of static images accompanied with verbal stimuli enhances test takers' performance on listening comprehension tests. These findings, also, get support from other studies (e.g. Robin, 2007; Grgurović and Hegelheimer, 2007; Winke et al., 2010).

Pedagogically, the present study has practical implications for listening comprehension materials development. Multidimensional presentation of listening input at most, and bimodal presentation of listening input at least, will surely duplicate the real life language atmosphere in the classroom in which both listening comprehension and viewing comprehension are accentuated.

Taking communicative listening assessment into consideration, the study implies that authentic performance based assessment needs to be more representative of target language use domain by supplementing non verbal components to verbal stimuli in listening comprehension tests. Their inclusion may also add to the validity and interactiveness of listening tests.

## REFERENCES

- [1] Bachman, L. F. (1990). *Fundamental considerations in language testing*. Oxford: Oxford University Press.
- [2] Bachman, L. F., & Palmer, A. S. (1996). *Language testing in practice*. Oxford: Oxford University Press.
- [3] Balizet, S., Treder, D., & Parshall, C. (1999). The development of an audio computer based classroom test of ESL listening skills. Paper presented at the Annual Meeting of the American Educational Research Association, Montreal, Quebec, Canada.
- [4] Blake, R., Wilson, N. L., Cetto, M., & Pardo Ballester, C. (2008). Measuring oral proficiency in distance, face to face, and blended classrooms. *Language Learning & Technology*, 12 (3), 114-127.
- [5] Brindley, G. (2001). Assessment. In R. Carter & D. Nunan (Eds.), *The Cambridge guide to teaching English to speakers of other languages* (pp. 137-143). Cambridge: Cambridge University Press.
- [6] Brindley, G., & Slatyer, H. (2002). Exploring task difficulty in ESL listening assessment. *Language Testing*, 19 (4), 369-394.
- [7] Brown, H. D. (2003). *Language assessment: Principles and classroom practices*. New York: Longman.
- [8] Bubel, C. M. (2008). Film audiences as overhearers. *Journal of Pragmatics*, 40, 55-71.
- [9] Buck, G. (1991). The testing of listening comprehension: an introspective study. *Language Testing*, 8 (1), 67-91.
- [10] Buck, G. (2001). *Assessing listening*. Cambridge: Cambridge University Press.
- [11] Chalhoub Deville, M. (2001). Language testing and technology: past and future. *Language Learning & Technology*, 5 (2), 95-98.
- [12] Chang, A. C-S., & Read, J. (2007). Support for foreign language listeners: its effectiveness and limitations. *RELC Journal*, 38 (3), 375-395.
- [13] Chastain, K. (1998). *Developing second language skills: Theory and practice* (3rd ed.). FL: Harcourt Brace Jovanovich.
- [14] Chen, L., Chen, G., Xu, C., March, J., & Benford, S. (2007). EmoPlayer: A media player for video clips with affective annotations. *Interacting with Computers*, 20, 17-28.
- [15] Chun, D. M., & Plass, J. L. (1997). Research on text comprehension in multimedia environments. *Language Learning & Technology*, 1 (1), 60-81.
- [16] Field, J. (2002). The changing face of listening. In J. C. Richards & W. A. Renandya (Eds.), *Methodology in language teaching: An anthology of current practice* (pp. 242-247). Cambridge: Cambridge University Press.
- [17] Fulcher, G., & Davidson, F. (2007). *Language testing and assessment: An advanced resource book* (C. N. Candlin & R. Carter, Eds.). New York: Routledge Applied Linguistics.
- [18] Ginther, A. (2001). Effects of the presence and absence of visuals on performance on TOEFL CBT listening comprehension stimuli (TOEFL Research Rep. No. 66). Princeton, NJ: Educational Testing Service.
- [19] Ginther, A. (2002). Context and content visuals and performance on listening comprehension stimuli. *Language Testing*, 19 (2), 133-167.
- [20] Grgurović, M., & Hegelheimer, V. (2007). Help options and multimedia listening: Students' use of subtitles and the transcript. *Language Learning & Technology*, 11 (1), 45-66.
- [21] Griffiths, R. (1990). Facilitating listening comprehension through rate control. *RELC Journal*, 21 (1), 55-65.
- [22] Gruba, P. (2006). Playing the videotext: A media literacy perspective on video mediated L2 listening. *Language Learning & Technology*, 10 (2), 77-92.
- [23] Harrison, M., & Whitehead, R. (2005). *Thomson exam essentials: IELTS practice tests*. London: Thomson ELT.
- [24] Hoven, D. (1999). A model for listening and viewing comprehension in multimedia environments. *Language Learning & Technology*, 3 (1), 88-103.
- [25] Hughes, A. (2003). *Testing for language teachers* (2nd ed.). Cambridge: Cambridge University Press.
- [26] Kramsch, C., & Andersen, R. W. (1999). Teaching text and context through multimedia. *Language Learning & Technology*, 2 (2), 31-42.
- [27] Lam, W. Y. K. (2002). Raising students' awareness of the features of real world listening input. In J. C. Richards & W. A. Renandya (Eds.), *Methodology in language teaching: An anthology of current practice* (pp. 248-253). Cambridge: Cambridge University Press.
- [28] Lynch, T. (2002). Listening: Questions of level. In R. B. Kaplan (Ed.), *The Oxford handbook of applied linguistics* (pp. 39-49). Oxford University Press.
- [29] Morley, J. (2001). Aural comprehension instruction: Principles and practices. In M. Celce Murcia (Ed.) *Teaching English as a second or foreign language* (3rd ed.) (pp. 69-84). US: Thomson Learning; Heinle & Heinle.
- [30] Nunan, D. (2002). Listening in language learning. In J. C. Richards & W. A. Renandya (Eds.), *Methodology in language teaching: An anthology of current practice* (pp. 238-242). Cambridge: Cambridge University Press.
- [31] Ockey, G. J. (2007). Construct implications of including still image or video in computer based listening tests. *Language Testing*, 24 (4), 517-537.
- [32] Peterson, P. W. (2001). Skills and strategies for proficient listening. In M. Celce Murcia (Ed.) *Teaching English as a second or foreign language* (3rd ed.) (pp. 87-99). US: Thomson Learning; Heinle & Heinle.
- [33] Richards, J. C. (2005). Second thoughts on teaching listening. *RELC Journal*, 36 (1), 85-92.
- [34] Robin, R. (2007). Commentary: Learner based listening and technological authenticity. *Language Learning & Technology*, 11 (1), 109-115.
- [35] Rost, M. (2001). Listening. In R. Carter & D. Nunan (Eds.), *The Cambridge guide to teaching English to speakers of other languages* (pp. 7-13). Cambridge: Cambridge University Press.
- [36] Rost, M. (2007). Commentary: I'm only trying to help: A role for interventions in teaching listening. *Language Learning &*



- Technology*, 11 (1), 102-108.
- [37] Shohamy, E., & Inbar, O. (1991). Validation of listening comprehension tests: The effect of text and question type. *Language Testing*, 8 (1), 23-40.
  - [38] Wagner, E. (2002). Video listening tests: A pilot study. Teachers College, Columbia University Working Papers in TESOL and Applied Linguistics, 2, 1.
  - [39] Wagner, E. (2007). Are they watching? Test taker viewing behavior during an L2 video listening test. *Language Learning & Technology*, 11 (1), 67-86.
  - [40] Weir, C. J. (1990). Communicative language testing. New York: Prentice Hall.
  - [41] Willmington, S. C., & Steinbrecher, M. M. (1993, April). Assessing listening in the basic course: The university of Wisconsin Oshkosh listening test. Paper presented at the Joint Meeting of the Southern States Communication Association and the Central States Communication Association, Lexington, KY.
  - [42] Winke, P., Gass, S., & Sydorenko, T. (2010). The effects of captioning videos used for foreign language listening activities. *Language Learning & Technology*, 14 (1), 65-86.
  - [43] Yanagawa, K., & Green, A. (2008). To show or not to show: The effects of item stems and answer options on performance on a multiple choice listening comprehension test. *System*, 36, 107-122.
  - [44] Yi'an, W. (1998). What do tests of listening comprehension test? A retrospection study of EFL test takers performing a multiple choice task. *Language Testing*, 15 (1), 21-44.
  - [45] Zielinski, B. W. (2008). The listener: No longer the silent partner in reduced intelligibility. *System*, 36, 69-84.



**Ataollah Maleki** is currently Associate Professor of TEFL at Zanjan Medical Sciences University, IRAN. His main interest of research is communication strategies, EFL teaching, teaching and testing, ESP, discourse analysis, and phonetics. He has been engaged in teaching English and applied linguistics for more than twenty years. He holds a Ph. D. degree in TEFL, an M. A. in TEFL, and a postgraduate diploma in applied linguistics and phonetics. He has published a number of books and articles in his areas of interest in academic journals.



**Masoud Safaee Rad** is an instructor at Azad University of Hamedan, Iran. His main area of research is CALL. He holds a B. A. degree in English translation and an M. A. degree in TEFL.