



# Journal of Applied Sciences

ISSN 1812-5654

**science**  
alert

**ANSI***net*  
an open access publisher  
<http://ansinet.com>

## The Effect of Text and Picture Accompaniments to Audio Materials on Comprehension Level and Response Time of EFL Learners

<sup>1</sup>Samad Sajjadi, <sup>2</sup>Hooshang Yazdani, <sup>3</sup>Mohammad-Reza Sohrabi and <sup>4</sup>Ataollah Maleki

<sup>1</sup>Department of English Language, College of Paramedical Sciences,  
Shahid Beheshti University of Medical Sciences, Tehran, Iran

<sup>2</sup>Department of Language and Linguistics, Arak University, Arak, Iran

<sup>3</sup>Department of Community Medicine, Faculty of Medicine,  
Shahid Beheshti University of Medical Sciences, Iran

<sup>4</sup>Department of English Language, Faculty of Medicine,  
Zanjan Medical Sciences University, Zanjan, Iran

---

**Abstract:** Visual aids accompanying listening tasks are of special significance in second or foreign language teaching. This study examined the effects of text/picture accompaniments to audio recordings on language learners' performance. The main aim was to investigate the effect of texts and pictures, presented jointly, on audio-based listening scores and response time. A group of language learners took some listening comprehension tests, with or without text/picture aids, under three conditions of (1) text/picture-before listening to audio recordings, (2) text/picture-during listening to audio recordings and (3) text/picture-free listening. According to results from tests of within-subjects effects and pair-wise comparisons, listeners' scores varied significantly across the three test conditions ( $p = 0.002$ ), with 82, 73 and 72 for the text/picture-before, text/picture-during and text/picture-free listening, respectively. As for the response time, the text/picture-before and text/picture-during conditions consumed significantly more time than the text/picture-free condition ( $M = 40.97$  and  $40.12$  versus  $23.61$  sec). The response time was significantly affected by the presentation modes (texts, pictures or audio recording) used to convey audio information, especially in the text/picture-included conditions.

**Key words:** Visual support, listening skill, planning time, language learner

---

### INTRODUCTION

The study is intended to investigate the effect of visual aids-texts and pictures-on second language (L2) learners' comprehension scores and response time. Visual aids accompanying listening tasks are of special significance in foreign language teaching and research (Jones, 2003; Verdugo and Belmonte, 2007; Sydorenko, 2010; Kaivanpanah and Zandi, 2009) because they help listeners to comprehend the key points of audio materials (Kellerman, 1992; Kashani *et al.*, 2011). Audio materials alone are less intelligible than when they are accompanied by visual stimuli (Mayer, 2001). As such visual aids and auditory materials presented jointly as various modes or channels with similar information can noticeably improve listening comprehension (Kashani *et al.*, 2011).

The enhancing effect of visual aids on L2 learners has been referred to in different respects. In study of

Bird and Williams (2002), for example, the joint presentation of auditory and visual helped new word learning. They presented new vocabulary items to advanced learners of English under three conditions: (1) text with sound, (2) text without sound and (3) sound without text. They examined the effect of these conditions on recognition of new words. The authors concluded that the coupling of auditory and visuals enhanced language comprehension by facilitating mental processes related to oral speech. According to Sydorenko (2010), subjects seeing video with captions scored higher on written than on oral recognition of word forms while the reverse applied to subjects watching video with audio. According to Garza (1991), visuals served as important pedagogical tools by helping language learners connect auditory to visual input; the visuals apparently assisted L2 learners in recognizing what might otherwise have been an incomprehensible sequence of audio materials.

---

**Corresponding Author:** Ataollah Maleki, Department of English Language, Faculty of Medicine,  
Zanjan Medical Sciences University, Zanjan, Iran  
Tel: 982414256473, 0912 2481030

According to Elkhafaifi (2005) and Sajjadi *et al.* (2012), visuals accompanying audio materials helped language learners to improve their listening comprehension skills. Visual aids in caption forms facilitated language learning by helping learners visualize what they hear, especially if the audio input was beyond their linguistic ability (Doughty, 2004). According to Chung (1999), the use of a combination of techniques resulted in more effective comprehension rather than the use of any single technique.

The studies summarized above are typical examples of previous studies about the effect of visual aids such as pictures, graphs, captions, video clips or TV broadcasts accompanying listening activities on L2 performance. Though immensely informative, the studies have not investigated the effect of texts and pictures accompaniments on audio-based listening materials; nor have they examined the link between presentation mode and response time under different listening conditions in EFL settings. This is at odd with the reality that EFL practitioners and learners tend to use such aids for individual or classroom practice more than ever before as, with new developments in on-line facilities, they find it much easier to get access to or produce such materials nowadays. Such realities warrant further research to shed light on unresolved issues about multimedia aids such as: (1) the ordering effects of text/picture accompaniments (i.e., visuals shown before or during the audio materials), and (2) their effect on the response time in different listening conditions. In response to such unresolved issues, the following research questions are addressed in this study:

- Is participants' listening comprehension boosted if they are exposed to both texts and pictures before listening to the accompanying audio materials?
- Does participants' listening comprehension improve if they are presented with both texts and pictures and the accompanying audio materials simultaneously?
- Is participants' performance in text/picture-before listening better than their performance in text/picture-during listening?
- Does the presence of text/picture in different conditions of listening affect the response time similarly?

Based on the research questions, the following hypotheses are formulated: (1) participants' listening comprehension will improve if they are subjected to text/picture accompaniments before the audio materials, (2) participants' listening comprehension will improve if they are subjected to text/picture accompaniments

simultaneously with the audio material, (3) participants' performance in text/picture-before listening will be similar to their performance in text/picture-during audio materials and (4) the presence of text/picture in different conditions of listening will not affect participants' response time. The main objective of the hypotheses was to examine the effect of texts and pictures, presented jointly, on audio-based listening comprehension and to investigate the effect of such accompaniments on the amount of time used to process the incoming audio information.

## MATERIALS AND METHODS

**Participants:** Forty-two homogeneous participants, in upper intermediate level of English proficiency, were selected from 90 students attending Upper Intermediate English courses of a language institute in Tehran. The selection was made on the basis of an objective placement test. The students who were to take part in the study were relatively similar in their English proficiency as they all attended Upper Intermediate courses of an English language Institute which insisted on linguistically homogeneous class members. Nonetheless, the administration of the placement test was an additional measure to reassure the homogeneity of the participants as the test made it possible to exclude students with elementary knowledge and lower intermediate knowledge of English proficiency who, according a piloting scheme (Kashani *et al.*, 2011), were unable to follow the listening comprehension stories properly. Furthermore, the authors decided to exclude the advanced participants from the study because it was thought that these participants' higher command of English proficiency would provide them privilege with extra linguistic knowledge that might affect the results reversely. The test accordingly made it possible to choose those who were similar in English proficiency as the study's participants.

**Objective placement test:** Following Kashani *et al.* (2011), a placement test with 70 multiple-choice items, designed by Lesley *et al.* (2005), was used to select the participants. The test measured receptive skills of listening, reading and grammar recognition. The reason for using a test of general language skills rather than a listening comprehension test was that, to authors, the former, as a more conclusive test, could provide a more accurate picture of the target group. All parts of the test were based on the objectives, content and language of different levels of Interchange Third edition and passages (Lesley *et al.*, 2005). The main reasons for the selection of the Placement test Lesley *et al.* (2005) was that: (1) the institution, where the participants came from, used the

Interchange Third edition as their source materials to teach and practice English, (2) it was recommended by EFL teachers and test developers who were familiar with the research objectives and subject selection procedure and (3) it enabled the researchers to choose a homogeneous group.

**Audio recordings:** Six audio recordings with accompanying texts and pictures served as the study's data elicitation tools. The texts and audio materials had already been used to collect data in the first phase of the research (Kashani *et al.*, 2011). The listening texts used in the study were designed in such a way that they did not appear alongside pictures as taking the pictures away might in itself constitute an added level of difficulty. The elicitation materials (texts, picture and audio recordings) were taken from Harris and Rowe (1995). The illustrations accompanying the texts, according to the authors, were to encourage students to use their own experience to communicate in English and accordingly, their exclusion from the texts was not supposed to bring about an added level of difficulty to make them totally incomprehensible. Nonetheless, to obtain further reassurance that the aural texts serving as listening materials met the study's criteria, they were subjected to a piloting scheme to make sure they meet the requirements of the study (Kashani *et al.*, 2011; Sajjadi *et al.*, 2012).

The listening texts and their visual accompaniments were composed of six rather long conversations, each with seven to nine multiple/choice questions, with four alternatives per question. The reason for choosing Multiple-Choice (M/C) questions rather than any other types of questions was firstly the objective measurement of the data that adds to the external validity of the study. The other reason was the familiarity of the subjects to M/C format. "In Iran, as in many parts of the world, M/C questions are the most common types of questions with which language learners are familiar. This familiarity could account for test format effect, which might, otherwise, act as a confounding factor" (Kashani *et al.*, 2011). All conversations included texts and pictures although in the text/picture-free condition texts or pictures did not accompany their relevant audio texts. The order of presenting the conversations to the subjects in text/picture-before, text/picture-during and text/picture-free conditions was counterbalanced. Thus, the six listening texts formed three sets of listening comprehension materials, each with two audio texts and their relevant pictures to be used in the three experimental conditions. As such, in each of the three test conditions, the participants heard a set of two stories with 15 or 16 questions (7 or 8 questions per story) as follows: 16

questions in set 1, 16 in set 2 and 15 in set 3, respectively. The listening texts and their text/picture accompaniments had already been subjected to a piloting scheme by Kashani *et al.* (2011) to make sure of: (1) the appropriateness of the listening texts and pictures, (2) to pretest the listening comprehension questions and (3) to determine the length of time necessary to understand the pictures. Piloting the materials gave credit to the suitability of joining these recordings with original texts and pictures in addition to the fact that taking the pictures away did not constitute an added level of difficulty for the text (Kashani *et al.*, 2011).

**Administration:** The data were collected in six sessions. Before each session, arrangements were made with seven students to attend. The reason for administering the procedure in more than one session was: (1) to offer more options to fit the participants' free time, (2) to exert better control over the administration procedure by limiting the number of participants per session and (3) to account for order effect by counterbalancing the administration order. In line with the balancing procedure, the 42 participants with homogeneous knowledge of English were randomly divided into six smaller groups each with 7 subjects, who were randomly assigned to one of the administration orders, appearing in Table 3 and 4. There were three dropouts in the data collection procedure. Each consecutive condition with two audio texts was separated by two minutes break, which was used to avoid making the experimental session boring. The need for two-minutes break was decided upon during the piloting phase carried out before the actual experiment (Kashani *et al.*, 2011).

**Procedure:** Participants had to perform the same task (listening with or without text/picture), under three different conditions. The audio materials for each condition were composed of two listening comprehension stories, altogether, making six stories read aloud by a native speaker of English. Having listened to the audio cassettes, the participants answered the accompanying questions in multiple choice formats. The conditions and the listening tests were counterbalanced in the presentation. All the participants sat for all the tasks as follows, though in a different order: (1) text/picture-before listening to audio cassettes, (2) text/picture-during listening to audio cassettes and (3) text/picture-free listening.

In the text/picture-before condition, subjects first read and looked at the accompanying texts and pictures taking their own time. They were given the texts and pictures to look at right before listening to the audio materials and

had to return the texts and pictures back to the researcher before listening to the audio texts. They were not allowed to discuss them with one another in or out of the class. The audio texts were then played out after all the participants indicated their readiness for listening to the audiotape, by returning the texts and pictures back to the researcher. Then, having returned the texts and pictures, they received the questions to answer. They had 15 sec (see the piloting section) to look at the questions before the listening began and this time was consistent across the three experimental conditions. Asking the students to be ready for the listening part, the researcher would then start up the audio player.

In the texts/pictures simultaneous condition, the subjects had the texts and pictures in front of them while they were listening to the audio broadcasting. In this condition also, like the other treatments, the participants received the questions before the listening component in order that they could have a look at the questions before the experimenter started up the audio cassette. The time length looking at the questions prior to the audio broadcasting was consistent across the six conditions. Then, 15 sec after the distribution of the questions, the experimenter told the participants to be ready for the audio materials while he started up the audio cassette. Then at the end of listening, like other conditions, the subjects were given sufficient time to answer the questions. In this condition, apart from the texts and pictures, which were left with the subjects during their listening activity, the rest of the procedure was similar to the previous condition.

In the text/pictures-free condition, a rather conventional form of listening comprehension was followed. That is, 15 sec after distributing the questions, the experimenter told the participants to be ready for the audio materials while he started up the audio cassette. Then, at the end of each listening, like other conditions, they were told to take their time to answer the questions. The rest of the procedure in this condition was similar to other conditions, except that there were no texts or pictures accompanying the listening comprehension materials in this condition.

In all the three conditions, participants were allowed to take their time to answer the questions at the end of each listening comprehension as there were no time restrictions in answering the questions. However, the experimenter recorded the time taken by each participant to answer the questions, upon receiving her answer sheet. Then, he added them up and divided the sum by total number of the participants to obtain the average time they spent to answer the questions across the three conditions of text/picture-before, text/picture-during and

text/picture-free conditions, which were 40.97, 40.13 and 23.61 sec, respectively. It is interesting to find that subjects spent significantly shorter time in the text/picture-free condition whereas they used a rather similar length of time in the two text/picture-included conditions.

**Statistical analysis:** A number of statistical computations were performed to analyze the data as follows: (1) descriptive statistics including means, standard deviations and standard Error of Mean for different conditions were calculated, (2) One-sample Kolmogorov-Smirnov test was then used to check the normality of variables, (3) the effect of presentation order, using comprehension scores, was calculated to make sure the administration order of the test formats had not biased the results, (4) the administration order based on the answering time for different test conditions was also performed to find out if the answering time had affected participants' performance, (5) repeated measure ANOVA (test of between and within-subjects effects) was used to test the significance of averages between tests, (6) pair-wise comparisons (*post hoc* tests) were administered to identify significant contrasts across different conditions. The p-value <0.05 was considered statistically significant.

## RESULTS

The Mean±SD scores for variables of text/picture-before, text/picture-during and pictures-free were 82.26 (SD = 13.81), 72.97 (SD = 13.75) and 71.89 (SD = 14.83) (out of 100), respectively (Table 1). One-sample Kolmogorov-Smirnov test showed that the assumption of normality of variables was met (Table 2). According to this test, the p-values for variables of text/picture-before, text/picture-during and pictures-free were 0.229, 0.293 and 0.225, respectively, justifying the application of further computations (Table 2). Then, the

Table 1: Descriptive statistics for comprehension scores and response time under different conditions of text/picture listening

Conditions	N	Mean	SD	Standard error mean
TPB scores	39	82.26	13.81	2.21
TPD scores	39	72.97	13.75	2.20
U scores	39	71.89	14.83	2.37
TPB time	39	40.97	42.99	6.88
TPD time	39	40.13	49.50	7.92
U time	39	23.61	25.22	4.03

TPB: Text/picture-before audio materials, TPD: Text/picture-during audio materials, U: Usual or text/picture-free audio materials

Table 2: One-sample Kolmogorov-Smirnov test

Condition	N	Significance (2-tailed)
TPB	39	0.229
TPD	39	0.293
U	39	0.225

Table 3: Scores based on the order of administering listening activities to participants

First		Second		Third		Total*
Group	Mean±SD	Group	Mean±SD	Group	Mean±SD	Mean±SD
U	83.93±8.73	TPD	72.81±9.29	TPB	66.60±1.84	74.45±14.31
TPB	89.98±9.14	U	59.82±1.39	TPD	76.11±1.43	75.30±17.44
TPD	79.06±1.95	TPB	72.12±1.72	U	68.50±1.52	73.22±17.10
U	71.43±7.95	TPB	75.89±1.42	TPD	68.50±1.32	71.94±11.87
TPD	79.46±1.60	U	82.14±9.83	TPB	72.31±1.36	77.97±13.39
TPB	95.31±3.13	TPD	76.56±3.13	U	84.92±1.14	85.60±10.24

TPB: Text/picture-before audio materials, TPD: Text/picture-during audio materials, U: Usual or text/picture-free audio materials, \*p = 0.09

Table 4: Answering time based on the order of administering listening activities to participants

First		Second		Third		Total*
Group	Mean±SD	Group	Mean±SD	Group	Mean±SD	Mean±SD
U	52.86±44.89	TPD	63.57±56.47	TPB	42.14±35.81	52.886±45.01
TPB	30.71±40.77	U	32.14±47.16	TPD	16.00±28.23	26.29±38.22
TPD	45.00±53.46	TPB	49.29±69.24	U	21.29±22.47	38.52±51.06
U	31.42±45.16	TPB	17.14±24.47	TPD	16.71±20.99	21.76±31.19
TPD	41.86±51.84	U	47.86±53.61	TPB	23.57±20.35	37.76±43.62
TPB	46.25±47.87	TPD	23.75±17.96	U	20.75±10.28	30.25±16.26

TPB: Text/picture-before audio materials, TPD: Text/picture-during audio materials, U: Usual or Text/picture-free audio materials, \*p = 0.21

Table 5: Test of within-subjects effects, including mean scores and F-value in the three conditions across all subjects

SOV	Sum of squares	df	Mean square	F-value	p-value
Between groups	2536.213	2	1268.107	6.344	0.002
Within groups	22788.403	114	199.898		
Total	25324.616	116			

Table 6: Pairwise comparison results (*post-hoc* test) for comparing score among groups of participants

Groups	Mean difference	Standard error	Sig.	95% CI
TPB: TPD	9.29	3.02	0.008	1.93-16.66
TPB: Usual	10.37	3.02	0.003	3.01-17.73
TPD: Usual	1.08	3.02	1.000	-6.28-8.44

order effect was also tested to find out whether or not the administration of the tests had affected the scores. According to the results, the order of administering listening activities showed no significant difference across different treatment conditions ( $p = 0.09$ ; Table 3), indicating that the six different presentation orders were well balanced and there were no order effect with the sequence of administering text/picture-before, text/picture-during and text/picture-free conditions; likewise, the answering time in different administration conditions revealed no order effect ( $p = 0.21$ ; Table 4), indicating that the time order in different test formats did not affect subjects' performance.

The computations proceeded with test of Within-subjects effects which showed a significant overall difference across the three test conditions ( $p = 0.002$ , Table 5). As such, because the p-value for within-subjects effects test was statistically highly significant ( $p = 0.002$ ), among the three groups of tests, the mean scores for at least two groups should have been

meaningfully different; so, to find out which specific pairs of means were significantly different, pair-wise comparisons were performed (Table 6). Among the pairs compared, the mean scores for text/picture-before condition were significantly higher than the mean scores for text/picture-during condition ( $p = 0.008$ ). Likewise, the mean scores for the text/picture-before condition were significantly higher than the mean scores for text/picture-free condition ( $p = 0.003$ ). However, the mean difference between text/picture-during and text/picture-free conditions failed to be statistically significant.

## DISCUSSION

As indicated in the result section, participants' performance varied significantly and differently across the three conditions ( $p = 0.002$ ), with mean scores of 82.26, 72.97 and 71.89 (Table 1) for the text/picture-before, text/picture-during and text/picture-free conditions in order. As for the response time, the mean times in seconds were  $M = 40.97$  and  $40.13$  versus  $23.61$  sec (Table 1) for text/picture-before, text/picture-during and text/picture-free conditions, respectively.

The discussion addresses the results in two lines of reasoning referred to as comprehension level and response time. It first elaborates on the role of text/picture accompaniments on participants' comprehension level and then proceeds to explain the relationship between such accompaniments and the time spent to process audio materials in each of the three conditions.

Regarding the superiority of comprehension in text/picture-before condition over the others, the results

seem to be in line with previous reports (Teichert, 1996; Kashani *et al.*, 2011; Ur, 1984) that pre-listening activities in visual, textual or audio forms can help L2 learners activate their prior knowledge, resulting in better listening comprehension. In study of Teichert (1996) pre-listening activities such as illustrations and videotapes, serving as advance organizers, allegedly activated subjects' prior knowledge, resulting in better listening comprehension at the intermediate level of German. According to Ur (1984) a short pre-listening activity helped listeners to construct a "mental model" and make predictions about what they were going to hear and these all made them motivated and anxious to listen.

Jones (2003) and Kashani *et al.* (2011) attributed the pre-listening superiority to subjects' exposure to the visual, verbal and oral stimuli, serving as three different mental representations. The participants had possibly integrated the organized representations of different modes into their mental model of the aural passage. In the current work also comprehension of the aural input would have been more difficult for the participants without proper access to the texts and pictures which might have served as complementary mental representations, helping the participants to achieve better understanding of the incoming verbal stimuli.

Another interpretation for the participants' better performance in the text/picture-before condition can be that the text/picture accompaniments helped EFL learners use their prior knowledge more effectively. Pre-listening activities apparently served as advance organizers activating learners' background knowledge, in line with the principles of schema theory in reconstructing meaning resulting in better comprehension of the incoming aural stimuli (Bonnell and Hafner, 1998; Teichert, 1996; Ur, 1984). The implication might be that learners can be helped to comprehend listening materials better by activating their schematic knowledge or background information. New materials may not be comprehensible without prior knowledge that could be made available to a conscious level by activating the relevant schemata. (Bransford, 1979; Mueller, 1980). To Bransford (1979), the most effective learning is meaningful learning, which can be enhanced by prior knowledge that can be presented in the form of advance organizers.

Another reason for text/picture-before to be the sole condition to significantly enhance comprehension, in Jones (2003), can be that students had possibly developed three different mental representations of the incoming stimuli-textual, pictorial and verbal-, in addition to that of the aural input, to process the verbal stimuli. This helped them to integrate "...the organized representations of the different modes into their mental

model of the aural passage" (Jones, 2003). Indeed, without text/picture aids, the processing of the listening information turned out to be difficult for the participants, as manifested in the text/picture-free condition. Lack of visual accompaniments such as text and picture and low knowledge of L2 prevented subjects from effectively processing the aural input under a text/picture-free condition. They could not construct meaning in a text/picture-free condition, with no annotations, as effectively as a text/picture-included condition with pictures serving as visual annotations which, compared to other annotations, according to Plass *et al.* (1998), can be more effective in facilitating listening comprehension since texts and pictures can help subjects choose the type of input that best suits their needs. With proper access to aural and visual modes, subjects could select the mode that facilitated listening comprehension (Kashani *et al.*, 2011).

In this study, the text/picture additions improved subjects' comprehension as subjects in text/picture-before condition achieved score gains significantly greater than pictures-during condition. In other words, subjects had significant gains in the former condition but failed to do so in the latter i.e., text/picture-during condition. The question is why in the text/picture-during condition the participants did not perform as well as the text/picture-before condition although they were subjected to the same materials in both conditions. One implication might be that the addition of visual annotations *per se* may not be sufficient to bring about a noticeable change in L2 performance; rather, the order of presenting the text/picture accompaniments- i.e., prior to audio materials- is of special significance in achieving the desired impact. In other words, the mere presence of pictorial, textual or verbal accompaniments to aural stimuli may not be sufficient to bring about the desired outcome; rather, their presentation order should also be taken into account in order to achieve the intended outcome.

Another reason for the superiority of text/picture-before condition might be the split-attention effect (Tarmizi and Sweller, 1988). In the text/picture-during condition, unlike the text/picture-before, the participants were operating under some information processing pressure. They had to focus on three different activities-looking at pictures, reading and listening-simultaneously which apparently were less conducive to comprehension. In the text/picture-before, nonetheless, they had the opportunity to first look at the texts and pictures and then, putting them aside, they solely focused on the audio broadcasting. This may be due to the fact that "... attention is limited in capacity and that its use to achieve one goal can reduce the capacity

remaining for the achievement of other goals” (Foster and Skehan, 1996). To put it differently, “... our mind has the ability to focus attention on just one task (as in the text/picture-before condition) as spreading our attention across different types of tasks (in the text/picture-before condition) can result in attention deficits” (Bonnell and Hafter, 1998).

Let us now turn to the next line of discussion, i.e., the relationship between the text/picture accompaniments to the audio materials and the time spent to process them in each of the three conditions. The fact is that in terms of semantic load and question type (i.e., comprehension questions) the three test conditions were similar, but in terms of presentation modes they differed—the text/picture-free condition with a single mode-aural stimuli-used an average of 23.61 sec, while the two others with two additional modes—texts and pictures—used nearly twice this figure, i.e., 40.97 and 40.13 sec. One implication may be that additional modes resulted in greater amount of time, regardless of conceptual similarities. Another implication can be that, in terms of the amount of response time, the presentation modes themselves are more important than the presentation orders, as text/picture-before and text/picture-during (with similar presentation modes) both consumed nearly the same amount of time (40.97 versus 40.13 sec). This contradicts the comprehension scores (i.e., 82.26 versus 72.97) which were significantly different across the same two conditions. The implication can be that: the presentation modes of audio materials can influence response time, even if the materials in multi-presentation modes are semantically similar to materials in their single mode counterpart. Hence, we may very tentatively postulate that diversity in presentation modes per se, independent of semantic load, would involve more time to process the incoming information. The reason may be that, in L2 acquisition, to Wigglesworth (1997), response time is affected by the overall length of the test and that the increase in length is justified by language outcome. In this study, the justification can be that the addition of texts and pictures, as complementary materials, resulted in greater lengths of test materials which, in turn, affected the amount of time spent to process the audio information.

The effect of planning time on issues like speech accuracy, complexity and authenticity has already been well established in previous studies (Hulstijn and Hulstijn, 1984; Crookes, 1989; Tannen, 1982). Nonetheless, the studies have seldom referred to presentation mode as a factor affecting the length of time and, accordingly, the time issue, as indicated in the current study, to the authors’ knowledge, is left unattended. Further work is essential to delve further into such issue to shed light on classroom instructions and

testing. A noticeable point here is that previous studies have artificially manipulated the time variable to seek subjects’ performance under time limited versus time-free conditions (Wigglesworth, 1997; Hulstijn and Hulstijn, 1984). In this study, however, there was no interference in the response time across different conditions. Instead, the researchers just had to jot down the amount of time each participant spent to listen to the audio materials (with or without text/picture accompaniments) and answer the questions, upon the completion and return of her answer sheet. The results are therefore, comparable to authentic test conditions with learners sitting for exam questions on aural stimuli with or without visual aids.

This study had its own limitations. The participants were upper intermediate in their level of English; the results, accordingly, may not be applicable to learners at other proficiency levels. The research was performed in a foreign language setting with participants learning English as a foreign language; so, the findings may not be satisfactorily applicable to second language learners or those practicing their first language. Multiple/choice questions, serving as the study’s elicitation technique to test participant’s comprehension, might not have been as authentic as other techniques in testing subjects’ comprehension. The reason for choosing multiple-choice questions rather than any other types of questions was the objective nature of the test and familiarity of the subjects with such tests. The visual aids were confined to texts and pictures which may not be sufficient for a solid generalization on the way visual aids influence comprehension score and response time.

Future studies can account for such limitations by devising more conclusive designs with fewer pitfalls. They need to examine the impact of other forms of aids (video clips with transcripts, charts and graphs) on language learners at different proficiency levels. Such studies can go beyond listening, the main focus of current work, to examine the effect of visual aids on other major skills like speaking, reading and writing. They had better use open ended questions or other forms of elicitation techniques to test subjects’ performance in visually assisted listening or reading comprehension. Also, the issue of presentation mode as a factor significantly affecting response time needs to be investigated further as it can have a significant effect on language teaching and testing. The current study adopted a cross-sectional design while similar research with longitudinal designs is inevitable.

## **CONCLUSION**

Channels of receiving information (pictorial, textual, aural) and their proper order of presentation (prior to or during audio activities) can play an important role in



enhancing L2 comprehension. In line with such role, future teaching and learning are expected to plan for new materials with proper order of presentation. Future plans need to give due attention to visual aids as advance organizers preceding audio materials to facilitate comprehension. In all such activities, the role played by response time should not be underestimated. As such, classroom practitioners may need to consider the link between presentation mode and the length of time as, in the current work, the time spent in the single mode task, i.e., the text/picture-free condition, was significantly less than the time spent in the multi mode tasks-included conditions.

### ACKNOWLEDGMENTS

The authors would like to thank the students who took part in data collection procedure.

### REFERENCES

- Bird, S.A. and J.N. Williams, 2002. The effect of bimodal input on implicit and explicit memory: An investigation into the benefits of within-language subtitling. *Applied Psycholinguistics*, 23: 509-533.
- Bonnel, A.M. and E.R. Hafter, 1998. Divided attention between simultaneous auditory and visual signals. *Percept. Psychophys.*, 60: 179-190.
- Bransford, J., 1979. *Human Cognition: Learning, Understanding and Remembering*. Wadsworth, Belmont, CA.
- Chung, J.M., 1999. The effects of using video texts supported with advance organizers and captions on Chinese college students listening comprehension: An empirical study. *Foreign Lang. Ann.*, 32: 295-308.
- Crookes, G., 1989. Planning and interlanguage variation. *Stud. Second Lang. Acquisit.*, 11: 367-383.
- Doughty, C.J., 2004. Effects of Instruction on Learning a Second Language: A Critique of Instructed SLA Research. In: *Form-Meaning Connections in Second Language Acquisition*, Van Patten, B., J. Williams, S. Rott and M. Overstreet (Eds.). Routledge, London, pp: 181-202.
- Elkhafafi, H., 2005. The effect of prelistening activities on listening comprehension in Arabic learners. *Foreign Lang. Ann.*, 38: 505-513.
- Foster, P. and P. Skehan, 1996. The influence of planning and task type on second language performance. *Stud. Second Lang. Acquisit.*, 18: 299-323.
- Garza, T.J., 1991. Evaluating the use of captioned video materials in advanced foreign language learning. *Foreign Lang. Ann.*, 24: 239-258.
- Harris, T. and A. Rowe, 1995. *Exploring English*. Vol. 4, Longman, New York, USA.
- Hulstijn, J.H. and W. Hulstijn, 1984. Grammatical errors as a function of processing constraints and explicit knowledge. *Lang. Learn.*, 34: 23-43.
- Jones, L.C., 2003. Supporting listening comprehension and vocabulary acquisition with multimedia annotations: The students voice. *CALICO J.*, 21: 41-65.
- Kaivanpanah, S. and H. Zandi, 2009. The role of depth of vocabulary knowledge in reading comprehension in EFL contexts. *J. Applied Sci.*, 9: 698-706.
- Kashani, A.S., S. Sajjadi, M.R. Sohrabi and S. Younespour, 2011. Optimizing visually-assisted listening comprehension. *Lang. Learn. J.*, 39: 75-84.
- Kellerman, S., 1992. I see what you mean: The role of kinesic behavior in listening and implications for foreign and second language learning. *Applied Ling.*, 13: 239-258.
- Lesley, T., C. Hansen and J.J. Zukowski/Faust, 2005. *Interchange/Passages: Placement and Evaluation Package*. 3rd Edn., Cambridge University Press, Cambridge, UK., Pages: 292.
- Mayer, R.E., 2001. *Multimedia Learning*. Oxford University Press, Oxford, UK.
- Mueller, G.A., 1980. Visual contextualized cues and listening comprehension: An experiment. *Modern Lang. J.*, 64: 335-340.
- Plass, J.L., D.M. Chun, R.E. Mayer and D. Leutner, 1998. Supporting visual and verbal learning preferences in a second language multimedia learning environment. *J. Educational Psychol.*, 90: 25-36.
- Sajjadi, S., M. Ahmadi, M. Heidarpour, A. Salahiyakta, N. Khadembashi and M. Rafatbakhsh, 2012. The effect of in-service English education on medical professionals language proficiency. *J. Res. Med. Sci.*, 17: 190-216.
- Sydorenko, T., 2010. Modality of input and vocabulary acquisition. *Lang. Learn. Teach.*, 14: 50-73.
- Tannen, D., 1982. Oral and literate strategies in spoken and written narratives. *Language*, 58: 1-21.
- Tarmizi, R.A. and J. Sweller, 1988. Guidance during mathematical problem solving. *J. Educ. Psychol.*, 80: 424-436.

- Teichert, H.U., 1996. A comparative study using illustrations, brainstorming, and questions as advance organizers in intermediate college German conversation classes. *Modern Lang. J.*, 80: 509-517.
- Ur, P., 1984. *Teaching Listening Comprehension*. Cambridge University Press, Cambridge, UK., ISBN-13: 978-0521287814, Pages: 184.
- Verdugo, D.R. and I.A. Belmonte, 2007. Using digital stories to improve listening comprehension with Spanish young learners of English. *Lang. Learn. Technol.*, 11: 87-101.
- Wigglesworth, G., 1997. An investigation of planning time and proficiency level on oral test discourse. *Lang. Testing*, 14: 85-106.