

THE RELATIONSHIP BETWEEN EFL LEARNERS' LEARNING STYLES AND THEIR SCORES IN AUDIO AND VIDEO-MEDIATED L2 LISTENING TESTS

(İNGİLİZCEYİ YABANCI DİL OLARAK ÖĞRENEN ÖĞRENCİLERİN ÖĞRENME
BİÇİMLERİ İLE İŞİTSEL VE HEM GÖRSEL HEM İŞİTSEL DİNLEME
TESTLERİNDEKİ PUANLARININ İLİŞKİSİ)

Nazlınur GÖKTÜRK¹
Ayşe ALTAY²

ABSTRACT

The inclusion of visuals in second language (L2) listening tests has long been a matter of interest for researchers. Yet, studies have often yielded indecisive and conflicting results. Besides, there have been few, if any, studies exploring the relationship between learners' individual differences and their performance in audio and video-mediated listening tests. The aim of this study is thus to investigate the relationship between learners' perceptual learning styles and their scores in audio and video-mediated FL listening tests in a classroom based assessment context. It also investigates the impact of channel of input on EFL learners' listening comprehension test scores. 27 Turkish learners of English as a foreign language took part in this study. A listening test consisting of two short academic lectures and eighteen short-answer comprehension questions was designed to assess the listening comprehension of learners. To detect the learners' perceptual learning styles, the adapted version of Cohen, Oxford and Chi's (2001) Learning Style Survey was conducted. The listening comprehension test scores of learners and their responses to the survey were statistically analyzed through SPSS version 20.0. The results indicated that there was not a significant difference between learners' scores in audio and video-mediated listening tests. However, a one-way ANOVA revealed that kinesthetic learners scored significantly higher than the visual learners in audio listening test. As such, the results showed that audio listening tests favored or disfavored certain groups of learners, contrary to the commonly held assumption that they do not lead to test unfairness. It is hoped that the findings of this study will offer implications for designing valid and reliable listening tests to be used in classroom based assessment contexts.

Keywords: Learning styles, testing listening, video-mediated L2 listening test

ÖZET

Görsel öğelerin ikinci dilde yapılan dinleme testlerinde kullanılması uzun süredir araştırmacıların dikkatini çekmektedir. Fakat çalışmalarda çoğunlukla ortak bir sonuca ulaşılamamış ve çelişkili sonuçlar ortaya çıkmıştır. Bunun yanında öğrencilerin kişisel farklılıklar ile işitsel ve hem işitsel hem görsel dinleme testlerindeki performansları arasındaki ilişkiyi irdeleyen çalışmalara pek rastlanmamıştır. Bu nedenle bu çalışmada öğrencilerin algısal öğrenme biçimleri ile işitsel ve hem görsel hem işitsel dinleme testlerindeki puanlarının ilişkisi araştırılmaktadır. Ayrıca, girdi kanalının ikinci dil öğrenen öğrencilerin dinleme testlerindeki puanlarına etkisi de incelenmektedir. Bu çalışmada İngilizceyi yabancı dil olarak öğrenen 27 Türk öğrenci yer almıştır. İki kısa akademik konuşma ve onsekiz kısa cevaplı soru, öğrencilerin dinleme becerilerini ölçmek için kullanılmıştır. Öğrencilerin algısal öğrenme biçimlerini anlamak için, Cohen, Oxford ve Chi (2001)'nin Öğrenme Biçimi Anketi'nin uyarlanmış hali kullanılmıştır. Öğrencilerin dinleme testlerindeki puanları ve ankete verdikleri cevaplar SPSS 20.0 programı aracılığıyla istatistiksel olarak hesaplanmıştır. Araştırma sonuçları öğrencilerin işitsel ve hem işitsel hem görsel dinleme testlerindeki

¹Instructor, Hacettepe University, School of Foreign Languages, ngokturk@hacettepe.edu.tr

²Instructor, Hacettepe University, School of Foreign Languages, aysealtay@hacettepe.edu.tr

puanlarının arasında istatistiksel olarak anlamlı bir farklılık olmadığını göstermiştir. Fakat tek yönlü ANOVA analizi kinestetik öğrencilerin işitsel dinleme testlerinde görsel öğrencilerden istatistiksel olarak daha iyi puanlar aldıklarını ortaya çıkarmıştır. Sonuç olarak, işitsel dinleme testlerinin test adaletsizliğine yol açmadığı kanısının aksine işitsel dinleme testlerinin bazı öğrenci gruplarına test sırasında ayrıcalık sağlarken bazı öğrenci gruplarına ayrıcalık sağlamadığı ortaya çıkmıştır. Çalışma sonucunda sınıf içi değerlendirmeleri için geçerli ve güvenilir dinleme sınavı oluşturmaya yönelik çıkarımlarda bulunulmuş ve öneriler belirtilmiştir.

Anahtar sözcükler: Öğrenme biçimi, dinleme test etme, hem görsel hem işitsel ikinci dil dinleme testi.

INTRODUCTION

The use of visuals in L2 listening tests has received the attention of many researchers over past two decades (Progosh, 1996; Coniam, 2001; Wagner, 2007; Suvorov, 2009). However, the results of the studies have been inconclusive in that there are some researchers who suggested the incorporation of visuals into L2 listening tests (Ginther, 2002) with the purpose of increasing the construct validity, whereas others do not favor the inclusion of visual input in L2 listening tests as it may distract some learners and lead to test unfairness (Gruba, 1993; Coniam, 2001; Ockey, 2007).

For listening is a complex process which involves the integration of multiple sources of information, a paper investigating the relationship between learners' perceptual learning styles and their comprehension test scores in audio and audiovisual listening tests is a worth of study as it may shed light on our understanding of the assessment of listening in a classroom based assessment context and provide new insights about how to design valid and reliable L2 listening tests, which, in turn, increases the test quality in the assessment of listening skill.

LITERATURE REVIEW

Given the prominence of listening skill as a way of obtaining comprehensible input in one's both L1 and any subsequent languages, it is not surprising that teaching and testing L2 listening skills has received a lot of attention among researchers (Thompson, 1995; Bejar, Douglas, Jamieson, Nissan, & Turner, 2000; Buck, 2001; Vandergrift, 2004; and Wagner, 2007). Though promising in nature, there have been a couple of practical constraints on the assessment of listening skills since listening is a multifaceted skill which requires the listener to have the ability to not only "extract basic linguistic information, but also interpret that in terms of some broader context" (Buck, 2001, p.59). In other words, the listener should be able to both process the linguistic features of a text (e.g., phonology, stress, intonation, word meanings, syntax and discourse) and incorporate that linguistic processing into the context of the situation to reach the intended meaning.

As listening processes interact in complex ways with different types of knowledge, the construct definition of L2 listening ability is not clearly well defined (Dunkel, Henning, & Chaudron, 1993; Buck, 2001; Field, 2003; Vandergrift, 2007).

Thus, various definitions of L2 listening ability have been suggested by researchers (Rubin, 1995; Brett, 1997). While early definitions focused on linguistic features as a way of expressing meaning in spoken discourse (Lado, 1961), later views of listening have primarily emphasized the inclusion of the both verbal and non-verbal stimuli such as lip movements of the speaker, body movement, gestures, and facial expression because in most real-life communicative situations, verbal information is often accompanied by visual information (Rubin, 1995; Buck, 2001, Wagner, 2007).

In order to determine the definition of the construct, Bachman and Palmer (1996) asserted that one should take the purpose of the test and target language use (TLU) domain which is defined as “situation or context in which the test taker will be using the language outside of the test itself” into account (p.18). As such, the characteristics of the test tasks should be representative of the types of tasks that learners are likely to encounter in real life target language use situations. Based on Messick's (1989, 1996) notion of construct validity, Wagner (2007) stated that if the TLU domain involved a listening event in which the listener could see the speaker and thus be able to utilize the nonverbal information transmitted by that speaker, then the construct definition of listening ability must include this aspect of listening as well. Hence, the authenticity of the tasks may increase, which will make it more likely to generalize the results of the test to a real life situation.

With regard to the rationale for the use of visuals in L2 listening tests, Progosh (1996) and Wagner (2007) argue for the inclusion of non-verbal components in the construct definition of L2 listening ability in listening tests based on the premise that non-verbal information is an essential part of communication in many real life situations. They maintain that the exclusion of non-verbal information from listening tests may threaten their validity and therefore, the listening construct in most cases needs to include the ability to obtain information from visual clues and even the ability to take notes (Ockey, 2007). Besides, they argue that the incorporation of visuals into L2 listening tests can assist the learners in understanding and processing the verbal input, which, in turn, may lead to increased listening comprehension. Contrary to Progosh (1996) and Wagner (2007), a number of researchers (Buck, 2001; Coniam, 2001) argue that though the inclusion of nonverbal features of language in listening tests appears to enhance its authenticity, the test's construct validity might be compromised, as this leads to assessment of something beyond listening ability.

Even though significant amount of research has been devoted to the use of visuals in both teaching and testing L2 listening ability over the past years (Progosh, 1996; Coniam, 2001; Ginther, 2002; Wagner, 2007; Suvorov, 2009), the findings of these studies were often conflicting. Some studies suggested that less proficient learners benefited most from the use of visuals (Rubin, 1995), while in other studies it was found that non-verbal clues have little, if any, facilitative effect on the understanding of verbal input (Lynch, 1998). The inclusion of video technology as a type of visual support in L2 listening tests has also been the focus of some research

on L2 listening (Coniam, 2001; Ginther, 2001, 2002; Ockey, 2007). However, there has been a dissonance among researchers with respect to the use of video technology in L2 listening tests (Progosh, 1996; Gruba, 1997; Shin, 1998; Wagner, 2002, 2008). While some studies showed that visuals can improve learners' performance on listening tests (Ginther, 2002), others provided no evidence for the facilitative effect of visuals on listening comprehension of test-takers (Gruba, 1993; Coniam, 2001; Ockey, 2007). One possible explanation for the overall lack of agreement among those studies may be concerned with the use of different types of visuals, that is to say, context and content visuals. Context visuals present the information about the context of the speech that takes place to set the scene, such as a picture of a lecturer talking to the learners in the classroom. However, content visuals provide information related to the content of the audio portion of the stimulus such as diagrams and/or drawings (Bejar et al., 2000; Ginther, 2002). While context visuals aid the comprehension of the text through setting the scene and activating the listeners' schema, content visuals give some helpful information about the content of the speech to help the listener comprehend the text. Since they provide different types of information, it is of paramount importance to decide what type of visual should be used in L2 listening tests.

In his study, Gruba (1993) compared the performance of 91 advanced level ESL students who took an audio listening comprehension test with those who took a video-mediated version of the same test which is a type of context visual for it does not convey any content-related information. The findings indicated that there was no significant difference between the learners' scores in audio-only and video-mediated listening tests. Gruba (1993) hypothesized that no differences between groups could be attributed to the fact that advanced second language learners are not "medium-dependent" (p.87). He also noted that some test-takers were apparently distracted by the visual stimuli and chose not to look at the video input. In a similar vein, Londe (2009) found no difference on short-answer listening comprehension test scores of undergraduate and graduate university students who heard a lecture in one of the three conditions: (a) with an accompanying video of the head of the speaker, (b) with a video of whole body of the speaker, (c) with the audio only. Londe (2009) concluded that different modes of delivery do not necessarily contribute to or take away from performance in L2 listening tests. Likewise, Coniam (2001) carried out a study to examine the impact of the two modes of a listening test (i.e., video-mediated and audio-only) on the test-takers' performance. Besides, the test-takers were asked to fill out a questionnaire on the advantages and disadvantages of video and audio as modes of a listening test and their preference for either mode. The participants of the study were 104 English language teachers in Hong Kong. The analysis of the test-takers' scores indicated no significant difference between scores of test takers who either watched a video or listened to audio only, even though the test takers who listened to audio only scored slightly higher on the multiple-choice test aimed to measure their listening comprehension. However, the difference was not statistically significant. In addition, the test-takers who watched a video reported that the video did not help, but distracted them

during the exam. Thus, Coniam (2011) concluded that L2 listening tests should be delivered by audio, rather than video mode.

The effect of context visuals on test-takers' listening performance on a computer-based test was also the focus of Suvurov's (2009) study which involved thirty four international students. In the study, the listening test was divided into three parts: one part presented with a single photograph, the other one presented with video-mediated input and the last one presented with audio-only input. Suvurov (2009) found that although there was not a significant difference between the test-takers' scores on the audio-only and photo-mediated parts of the listening test, participants' scores on video-mediated part of the listening were lower. Like Suvurov (2009), Ginther (2002) studied the effects of the presence and absence of two types of visuals, namely content and context visuals, with different types of stimulus (i.e., Dialogues/Short conversations, Academic discussions and Mini-talks) on the test-takers' performance on TOEFL CBT listening comprehension section. The results of the study revealed that effective content visuals increased the comprehension of the listening texts, while context visuals did not have a facilitative effect on comprehension.

On the other hand, Wagner (2013) investigated the impact of the channel of input on the listening performance of L2 learners and he found that test-takers who received audio-only input scored lower than test-takers who received audiovisual input, which he attributed to test takers' use of nonverbal clues given by the speaker in the video. However, it is worth noting that in Wagner's (2013) study, the participants may have been more familiar with the video-based listening tests, which, in turn, might have provided them with a chance to effectively make use of the nonverbal clues presented in the video.

Of relevance to the present study is also the research on learners' perceptual learning styles (Myers, 1962; Larsen, 1992; Riding & Sadler-Smith, 1992; Biggs, 1993; Jonassen & Grabowski, 1993; Vermunt, 1996; Karns, 2006; Morrison, Sweeney, & Heffernan, 2006). Ehrman and Oxford (1990) define learning styles as "preferred or habitual patterns of mental functioning and dealing with new information" (p. 311). Even though most educators are well aware of the fact that they should take the learners' preferred learning styles into consideration in order to increase the effectiveness of learning and teaching, they still rely on the same, often standardized tests to assess learners (Swain, 2004; Butler & Roediger, 2008; Nichols and Berliner, 2008). As stated by Leithner (2011), even when the educators utilize different testing formats, the variety of students' learning styles are likely to always disadvantage certain students whenever an assessment strategy is used excessively.

As there have been few, if any, studies conducted to investigate the relationship between learners' perceptual learning styles and their performance in a listening test presented through different channels of input (i.e., audio and video-mediated), the current study seeks to investigate how the inclusion of video in L2 listening tests affects the test performance of learners with different learning styles. Furthermore, there are only few studies conducted on Turkish speaking learners of

English as a foreign language (EFL) with respect to their performance in audio and video-mediated listening tests in a classroom based assessment context. By relying on the statistical analyses of the learners' scores on two listening tests and the data collected through a learning style questionnaire, this study aims to contribute to our understanding of L2 listening ability, and further to offer implications for designing valid L2 listening tests.

The research questions addressed in this study are the following:

1. What is the relationship between Turkish EFL learners' perceptual learning styles and their scores in audio and video-mediated listening tests?
2. What is the effect of the channel of input (i.e., audio and video-mediated) on EFL learners' listening test scores?

METHODOLOGY

In order to examine the relationship between EFL learners' perceptual learning styles and their performance in a listening test presented through different channels of input (i.e., audio and video-mediated), this study employed a within-subjects design rather than a between-subject design to avoid error variance associated with the individual differences of the learners. The independent variables measured throughout the experiment were the channel of input (i.e. video-mediated and audio-only) and the perceptual learning styles of learners (i.e. visual, auditory and kinesthetic). The dependent variable consisted of participants' scores on each of the two parts of the listening test.

Context of the Study

The study was conducted in the School of Foreign Languages at Hacettepe University where each student enrolled is required to certify a certain level of English proficiency to be eligible for the freshman year. The minimum score required for exemption from the preparatory school program is 65 on the proficiency exam for those learners enrolled in a program where the medium of instruction is completely (100%) English, whereas it is 55 for the learners enrolled in a program where the medium of instruction is partially (30%) English. The learners who fail in the proficiency exam are required to take the Placement Test in order to determine the level of classes in which they will be enrolled. Placement Test consists of multiple choice questions with a focus on grammar, vocabulary and reading skills, whereas proficiency exam is composed of Listening Comprehension and Note-Taking, Reading Comprehension, Language Use and Writing sections. Each of the four sections is worth 25 points and the composite score of the exam is 100 points.

In the preparatory program, the levels of classes, curriculum and teaching objectives are aligned with the Common European Framework of Reference (2001). Accordingly, learners are divided into six levels: A1, A2, B1, B1+, B2, B2+. Learners with A1, A2, B1 and B1+ levels of proficiency receive 25 hours of

instruction per week while learners with higher levels of proficiency (i.e., B2, B2+) take 20 hours of instruction each week. At all levels, eight hours a week is devoted to skill based courses (i.e., listening and speaking, reading and writing) and the rest of the weekly hours of instruction includes integrated courses aimed at improving general language knowledge of learners. Of eight hours, four hours is allocated to listening and speaking classes. At A1 and A2 levels, listening and speaking classes focus on tasks where learners listen to short authentic texts and engage in speaking activities to improve their communication skills. At B1, B1+, B2 and B2+ levels, the classes aim to develop not only academic listening and note-taking skills of learners through several authentic lectures but also their communication skills by engaging them in group discussions and negotiations. During the classes, listening texts are commonly accompanied by pre-listening activities such as predicting the content, discussing the topic and activating the relevant vocabulary. At while listening stage, the texts are played two times: on the first listening the learners are required to listen for main ideas and on the second one they are asked to focus on details. Besides, the audio-visual texts are frequently utilized in teaching listening skills at all levels not only to increase the motivation of learners, but also to familiarize them with authentic texts.

As for the assessment of listening skills, at B1, B1+, B2, B2+ levels the learners take two listening quizzes with the purpose of measuring both their note-taking skills and listening comprehension in each quarter. The contribution of listening comprehension quiz scores of learners to their overall grade is 10%. In addition to the quizzes, the level achievement tests administered to the learners at the end of each quarter contain a listening section similar to the ones given in the quizzes. The contribution of scores of learners obtained from listening section of the achievement test to their overall grade is approximately 15%. Although a sizeable portion of the listening and speaking classes involve audiovisual texts as part of teaching procedure, audio texts are only used in the listening exams conducted in the school.

Participants

A total of 27 Turkish-speaking adult EFL learners of English (female = 17; male = 10) participated in this study (age range 18 to 21). The English proficiency level of the learners is B1 based on their scores on the proficiency exam and placement test conducted by the school. The learners are supposed to obtain at least 65 points out of 100 points on the proficiency exam as they are enrolled in the departments where the medium of instruction is completely (100%) English. The departments of the learners are Medicine, Computer, Physics and Chemical Engineering. The numbers of learners enrolled in the departments of Medicine, Computer, Physics and Chemical Engineering are respectively 10, 8, 5 and 4. All of the learners have been learning English as a foreign language for about 8 years.

Instruments

A listening test consisting of two short academic lectures given by a university professor, each of which lasted for 4.5 minutes and was followed by nine short-answer comprehension questions, was designed to assess the listening comprehension of learners (Appendix-A). The channel of input for one of the lectures in the test was video-mediated, while the other one was audio-only. In the video-mediated part of the test, the lecturer was standing behind the lectern and could be seen from the waist up. He did not use the board or any special equipment, but did use natural body language and gestures. According to the classification of visuals proposed by Bejar et al. (2000) and Ginther (2002), only context visuals were used in the listening test.

To ensure the content equivalence between the texts, both lectures covered topics in Business. The video-mediated lecture was on the effects of high and low context communication styles on international business, while the audio-only lecture was on the effects of neutral and affective cultures on international business.

For the two lectures, two sets of short answer comprehension questions were developed by one of the researchers, who was also the teacher of the learners, as she was familiar with the teaching and learning processes in class. These sets of questions were composed of limited response (i.e., short-answer) comprehension items for two reasons. First, because short-answer items are largely used in classroom based assessment contexts, it is important to examine how the channel of input in listening tests affects the performance of learners on those types of comprehension items. Second, because multiple-choice items may lend themselves to test-taking strategies, which do not evaluate the learners' understanding of the question (Hearst, 2000); it is reasonable to include short-answer items in the listening test to reduce the risk of random guessing. Additionally, it should be noted that the items in the test were not constructed in a way that could only be answered by viewing the videotext.

With the purpose of ascertaining that two sets of questions and lectures in the test were of equal difficulty, a pilot study was undertaken on a different group of learners with the same proficiency level (i.e., B1). 15 adult Turkish learners of English took place in the pilot study. During the piloting, the learners were asked to evaluate the level of difficulty of both texts and to indicate any points which were not clear enough in the questions.

In the pilot study, the internal consistency of the items in the test was checked. The Cronbach alpha coefficient of the whole test was estimated ($=.70$), which can be regarded as acceptable. The internal consistency reliability (KR-20) of the audio part of the listening test (9 items) was $.54$ and the KR-20 of the video-mediated part of the listening test (9 items) was $.69$.

After piloting, one of the items in the audio part of the listening test was reformulated to make the meaning clear and one of them was replaced with a different item. Due to constraints of time, the changed items were not piloted again.

Next, the edited version of the listening test was conducted in the original study and the internal consistency reliability of the listening test as a whole was found to be good (Cronbach alpha=.81). KR-20 of the audio part of the listening test (9 items) was .62 and the KR-20 of the video-mediated part of the listening test (9 items) was .77, which can be considered as reasonable given the relatively small number of participants.

With regard to the content validity of the test, a group of expert judges, namely the head of the testing unit of the department and two coordinators, were asked to evaluate the appropriateness of the listening texts, clarity of instructions and questions, and the quality of audio and video recordings. Thanks to the researchers' and experts' judgments, the content validity of the test was tried to be established.

In addition to the listening test, a survey was given to the students to explore the learners' perceptual learning styles. Cohen, Oxford and Chi's (2001) *Learning Style Survey* was adapted for the purposes of the study. As the focus of the study was on the sensory style preferences (visual, auditory and kinesthetic) of learners, the other sections in the original survey were excluded. Moreover, the constituent items were mixed up randomly in order to eliminate the impact of repetitive content on learners since repetition may lead to frustration among respondents when answering the questions (Dörnyei, 2002). As such, the adapted version of the perceptual learning styles survey was composed of two sections (Appendix-B). In the first section, personal information about the participants' gender, age, and proficiency level was sought. The second section of the survey included 30 items categorized under three scales with 10 items each. The items of the first scale were related to the visual learners, the second scale was related to the auditory learners, and the third one was related to the kinesthetic learners.

To identify the possible problems to be encountered in the actual study with respect to the survey and to test whether each subscale is measuring a single idea and whether the items that compose the survey are internally consistent, a pilot study was carried out with another group of 40 adult Turkish learners of English at the same institution. During the piloting, the learners were also asked to point out any questions that were not clear.

In the pilot study, internal reliability data was estimated through Cronbach alpha coefficient. It was found that the internal consistency of 30 items in the questionnaire was .70. The alpha values estimated for visual, auditory and kinesthetic subscales were respectively .58, .65 and .68, which can be taken as satisfactory considering the small number of respondents.

Later, the adapted version of the survey was carried out in the original study. The total internal consistency reliability of 30 items in the survey was 0.65. The alpha values for visual, auditory and kinesthetic subscales were 0.51, 0.20, and 0.65 respectively. Even though the values found in this study are lower than the ones found in Cesur and Fer (2009) and the pilot study, those alpha scores were regarded

as satisfactory in the current study given the homogeneous profile of the participants as freshman students and small sample size as Schmitt (1996) argues that alpha values of 0.5 would not attenuate validity. In addition, the low reliability score in the auditory subscale can be attributed to the fact that there were few learners with auditory learning style.

Procedure

Before taking the listening comprehension quiz in class, the learners filled out the Perceptual Learning Styles Survey. After completing the survey, they were given blank sheets for taking notes to be later used for answering the questions during the test. There are two reasons for asking the learners to take notes while listening to lectures in the test. First, listeners, unlike readers, do not have a chance of reviewing the information presented to them (Thompson, 1995). Furthermore, taking notes can reduce the cognitive load and therefore increase the validity of a listening test (Vandergrift, 2007).

After the audio-only text was played twice using a CD on a Samsung PC with external speakers, the sheets with the short-answer comprehension questions written for the audio-only text were distributed to the learners. The learners were instructed to answer all the questions using their notes and the time allotted for the learners to answer the questions was 8 minutes.

Subsequently, the videotext was played twice from a DVD on a Samsung PC connected to a wall mounted LCD projector and external speakers. The procedure followed in the audio-only part of the listening test was undertaken again. At the end of the listening test, the answer sheets of the learners were gathered for evaluation.

Data Analysis

Quantitative data obtained in this study were analyzed through SPSS Version 20.0 and the reliability of the participants' scores on each part of the listening test and their responses on each subscale of the survey was calculated. To answer the first research question, descriptive statistics for each part of the test were calculated and a paired-sample t-test was conducted. To answer the second research question, the ANOVA procedure was used, which was followed by the Tukey HSD test for post-hoc comparison, to determine whether there was a significant difference among the learners' mean scores on audio and video-mediated parts of the listening test with respect to their perceptual learning styles (i.e., visual, kinesthetic, auditory/visual and auditory/kinesthetic). In this analysis, the independent variable was the participants' perceptual learning styles (i.e., visual, kinesthetic, auditory/visual and auditory/kinesthetic), and the dependent variable was their scores on audio-only and video-mediated parts of the listening test.

RESULTS

The Comparison of Overall Scores of Learners on Audio and Video-Mediated Listening Tests

In order to answer the first research question addressing the effect of the channel of input (i.e., video and audio-mediated) on listening comprehension test scores of learners, learners' scores on audio and video-mediated listening tests were compared. Table 1 presents the descriptive statistics related to the learners' scores on both audio and video-mediated listening tests.

Table 1. The Descriptive Statistics of Learners' Overall Scores

	N	Minimum	Maximum	Mean	Std. Deviation
Video Overall	27	.00	9.00	3.57	2.55
Audio Overall	27	.00	8.00	4.25	2.02

A paired samples t-test was performed to determine if the mean difference in learners' scores was statistically significant. The results revealed that there was not a statistically significant difference between the audio ($M=4.25$, $SD=2.02$) and video-mediated ($M=3.57$, $SD=2.55$) listening test scores of the participants, ($t(26) = -1.64$, $p > .05$). These results suggest that the channel of input in listening tests does not affect the listening comprehension test scores of learners. Specifically, the use of video technology in listening tests did not have any effect on the learners' scores, providing a negative answer to the first research question.

Learning Styles of the Learners

Using the results of the adapted version of the *Learning Style Survey* (Cohen, Oxford & Chi, 2001), learners' learning style preferences (i.e., visual, auditory and kinesthetic) were identified. As shown in Table 2, the most preferred learning style was visual (44.4%), which means that almost half of the learners rely more on the sense of sight and learn best through visual means such as video, charts and pictures. However, there were not any participants who only preferred the auditory learning style. Instead, there were learners with amalgam learning styles (i.e., favoring both auditory and visual, or auditory and kinesthetic learning styles). In other words, learners with amalgam learning styles either employed visual means or implemented hands-on activities alongside the listening skill as a way of learning.

Table 2. The Distribution of Learners Through the Learning Styles

	N	%	
Valid	Visual	12	44.4
	Kinesthetic	4	14.8
	Auditory + Visual	5	18.5
	Auditory + Kinesthetic	5	18.5
	Total	26	96.3

Missing	System	1	3.7
Total		27	100.0

Relationship of Scores to Learning Styles

The second research question involved the relationship between learners' learning styles (i.e., visual, auditory and kinesthetic) and their scores on audio and video-mediated listening tests. The goal was to see whether learners with specific learning style preferences score higher or lower in audio and video-mediated listening tests than learners with other learning preferences. Table 3 shows the descriptive statistics of the scores of learners with different learning styles on audio and video-mediated listening tests.

Table 3. The Descriptive Statistics of the Scores of Learners with Different Learning Styles

		N	Mean	Std. Deviation
Video Overall	Visual	12	2.91	2.53
	Kinesthetic	4	4.25	3.01
	Auditory + Visual	5	3.00	1.96
	Auditory + Kinesthetic	5	4.10	2.01
	Total	26	3.36	2.35
Audio Overall	Visual	12	3.25	1.42
	Kinesthetic	4	6.87	1.31
	Auditory + Visual	5	4.10	1.19
	Auditory + Kinesthetic	5	4.10	2.40
	Total	26	4.13	1.95

To find out whether there was a significant difference among the learners' mean scores on audio and video-mediated parts of the listening test with respect to their preferred learning styles, a one-way ANOVA was run on the data. A one-way between subjects ANOVA was conducted to compare the effect of different types learning styles on audio and video mediated listening test scores. When the scores of learners on audio and video-mediated listening tests were dependent variables, there was not a significant effect of learning style on video-mediated listening test scores at the $p < .05$ level for four groups of learners with different learning styles [$F(3, 22) = .503, p > .05$]. However, there was a significant effect of learning style on audio listening test scores at the $p < .05$ level for the four groups of learners with different learning styles [$F(3, 22) = 5.13, p = .008$]. To determine the differences among groups, a post-hoc comparison test was undertaken. The Tukey HSD test indicated that the mean score of the learners with visual learning style ($M=3.25, SD=1.42$) was significantly different from the mean score of the learners with kinesthetic learning style ($M=6.87, SD=1.31$) in the audio listening test, which

indicates a meaningful relationship between learning styles and listening test scores, and thus providing an affirmative answer to the second research question. However, the mean score of learners with balancing learning style of auditory and visual ($M=4.10$, $SD=1.19$) did not significantly differ from the mean scores of learners with visual, kinesthetic and auditory/kinesthetic learning styles. Likewise, the mean score of learners with balancing learning style of auditory and kinesthetic ($M=4.10$, $SD=2.40$) did not significantly differ from the mean scores of visual, kinesthetic and auditory/visual learning styles. Taken together, the results of the one-way ANOVA showed that the kinesthetic learners performed better than visual learners in the audio listening test. As such, no significant difference was observed among the other groups of learners with different learning styles, namely visual, kinesthetic, auditory/visual and auditory/kinesthetic.

Besides, a Pearson correlation coefficient was computed to assess the relationship between the learning styles (i.e. visual, auditory and kinesthetic) and learners' scores on video-mediated and audio parts of the listening test. There was a negative correlation between the learners' scores on audio listening test and visual learning style ($r=-.447$, $n=26$, $p=.022$). Overall, there was a relatively strong, negative correlation between the learners' scores on audio listening test and visual learning style. In other words, the more a learner prefers visual learning style, the lower s/he scores in the audio listening test.

DISCUSSION AND CONCLUSION

The comparison of learners' overall scores on video-mediated and audio listening tests indicated that the channel of input in listening tests does not affect the listening comprehension test scores of learners in a classroom based assessment context. Although the learners who received audio only input scored slightly higher than the ones receiving video input on the short-answer listening comprehension test, no significant difference was found between learners' mean scores on audio and video-mediated listening tests, which accords with the findings of the research by Coniam (2001), Gruba (1993) and Londe (2009) in that the inclusion of video material does not necessarily lead to increased test scores. This lack of difference between the exam scores may stem from the fact that the participants in this study were already exposed to video materials during teaching procedure and thus had no difficulty in comprehending the audiovisual text and the visual stimuli did not unfairly advantage a certain group of learners.

Concerning the relationship between learners' learning styles and their scores on video-mediated and audio listening tests, the findings revealed that kinesthetic learners scored higher than the visual learners in audio listening test and the video material did not augment the visual. In fact, the audio text favored the kinesthetic learners and lead to a significant increase in kinesthetic learners' short answer comprehension test scores. The reason behind the outstanding performance of kinesthetic learners might be that these learners are usually good at dealing with more than one task simultaneously (Oxford, 1990, 1995) and thus managed to take

notes while they were listening to the text. Besides, they learn best by doing, testing and trying things out themselves (Kinsella, 1995). Hence, they took advantage of note-taking while listening to the lectures as they had already been engaged with the task.

However, as highlighted by Kunnan (2004), test fairness is an important factor which may threaten the validity of a particular test. In this study the audio-only listening test either favored or disfavored some groups of learners unfairly, which lowers the validity of the scores obtained from this test. On the other hand, the audiovisual input did not lead to increased or decreased performance of particular learners, which, in turn, enhanced the construct validity of the video listening test.

The fact that the inclusion of visuals in L2 listening tests does not unfairly advantage a certain group of learners and the audio only L2 listening tests favor kinesthetic learners appears to cast doubt on the construct validity of audio-only L2 listening tests. Wagner (2006) argued, based on Messick's (1989, 1996) notion of construct validity, that target language use domain should dictate the test task characteristics. That is to say, if the TLU domain consists of situations where the listener can see the speaker and consequently utilize both verbal and nonverbal information conveyed by the speaker, then the characteristics of the task must also include such an aspect to contribute to the validity of the inferences made from the results of the test. Given that the TLU domain in the context of a university includes listening situations where learners can both see and hear the professor, it can be argued that the visual information is to be included not only in the construct definition of L2 listening ability, but also in the characteristics of the test tasks to enhance the validity of the test scores. Likewise, Progoosh (1996) and Wagner (2007) assert that the exclusion of visual input from listening tests might threaten their validity due to the construct-underrepresentation. Besides, to generalize the results of a test to non-test language situations, the test tasks must also be authentic. Bachman and Palmer (1996) define authenticity as "the degree of correspondence of the characteristics of a given language test task to the features of a TLU task" (p. 23). Thus, the inclusion of visual input in L2 listening tests may contribute to the authenticity of the test tasks, which are employed to measure the learners' listening comprehension in a university setting, provided that it does not lead some learners to score lower or higher/ perform worse or better.

In addition to increasing the construct validity and authenticity of L2 listening test, the incorporation of video technology into L2 listening tests has some pedagogical reasons as highlighted by Gruba (1997). Since the video materials are commonly employed to teach L2 listening skills in most EFL classes, they should also be included in the testing of L2 listening skills with the purpose of aligning the characteristics of tests tasks with the instructional practices (Bachman & Palmer, 1996).

Limitations

Owing to the constraints of time, this study was conducted on low proficiency EFL learners; however, it would yield more informative and contributing scores if the participants were both low and high proficiency learners. In addition, the current study involved only a small number of students as participants, yet having more participants would provide more reliable and precise results. Finally, in this study the impact of only context visuals has been explored; however, it would be interesting to see the relationship between learners' perceptual learning styles and their test scores in L2 listening tests presented with both content and context visuals.

Implications

The findings of this study have implications for practitioners and test developers. Since real-life language tasks usually include the listener being able to see the speaker, the video mediated listening tasks should be incorporated into classroom listening tests whenever possible and practical. In addition, teachers can teach some listening and note-taking strategies to their learners by providing classroom tasks where students can experiment with watching the videotext and taking notes the entire time. As the inclusion of video technology in L2 listening tests does not lead to test unfairness among learners, it is also safe to suggest that the video-mediated listening tests may be used for the assessment of listening skills in a classroom based assessment context.

REFERENCES

- Bachman, L. F., & Palmer, A. S. (1996). *Language testing in practice*. Oxford: Oxford University Press.
- Bejar, I., Douglas, D., Jamieson, J., Nissan, S., & Turner, J. (2000). *TOEFL 2000 listening framework: A working paper* (TOEFL Monograph Series, Report No. 19). Princeton, NJ: Educational Testing Service.
- Biggs, J. B. (1993). What do inventories of students' learning processes really measure? A theoretical review and clarification. *Educational Psychology*, 63, 3–19.
- Brett, P. (1997). A comparative study of the effects of the use of multimedia on listening comprehension. *System*, 25(1), 39–53.
- Buck, G. (2001). *Assessing listening*. Cambridge: Cambridge University Press.
- Butler, A. C., & Roediger, H. (2008). 'Feedback enhances the positive effects and reduces the negative effects of multiple-choice testing. *Memory & Cognition*, 36(4), 604–623.
- Cesur, M. O., & Fer, S. (2011). Öğrenme stratejileri, stilleri ve yabancı dilde okuma-anlama başarısı arasındaki ilişkiler örüntüsü. *Hacettepe Üniversitesi Eğitim Fakültesi Dergisi*, 41, 83-93.
- Cohen, A. D., Oxford, R. L., & Chi, J. C. (2001). Learning Style Survey. Retrieved from http://www.carla.umn.edu/maxsa/documents/LearningStyleSurvey_MAXSA_IG.pdf

- Coniam, D. (2001). The use of audio or video comprehension as an assessment instrument in the certification of English language teachers: A case study. *System*, 29(1), 1-14.
- Dörnyei, Z. (2002). *Questionnaires in second language research: Construction, administration, and processing*. New Jersey: Lawrence Erlbaum Associates, Publishers.
- Dunkel, P., Henning, G., & Chaudron, C. (1993). The assessment of an L2 listening comprehension construct: Tentative model for test specification and development. *The Modern Language Journal*, 77(2), 180-191.
- Ehrman, M., & Oxford, R. (1990). Adult language learning styles and strategies in an intensive training setting. *The Modern Language Journal*, 74(3), 311-327.
- Field, J. (2003). Promoting Perception: Lexical segmentation in L2 listening. *ELT Journal*, 57, 325-334.
- Ginther, A. (2002). Context and content visuals and performance on listening comprehension stimuli. *Language Testing*, 19(2), 133-167. doi:10.1191/0265532202lt225oa
- Gruba, P. (1993). A comparison study of audio and video in language testing. *JALT Journal*, 15(1), 85-88.
- Gruba, P. (1997). The role of video media in listening assessment. *System*, 25(3), 335-345.
- Gruba, P. (2006). Playing the videotext: A media literacy perspective on video-mediated L2 listening. *Language Learning & Technology*, 10(2), 77-92.
- Hearst, M. (2000). The debate on automated essay grading. *IEEE Intelligent Systems*, 15(5), 22-37.
- Karns, G. (2006). Learning style differences in the perceived effectiveness of learning activities. *Journal of Marketing Education*, 28, 56-63.
- Kinsella, K. (1995). Understanding and empowering diverse learners. In J. Reid (Ed.), *Learning styles in the ESL/EFL classroom* (pp. 170-194). Boston: Heinle & Heinle.
- Kunnan, A. J. (2010). Test fairness and Toulmin's argument structure. *Language Testing*, 27(2), 183-189.
- Jonassen, D. H., & Grabowski, B. (1993). *Handbook of individual differences, learning, and Instruction*. Hillsdale, NJ: Lawrence Erlbaum Associates.
- Lado, R. (1961). *Language testing: The construction and use of foreign language tests*. London: Longman.
- Larsen, R. E. (1992). Relationship of learning style to the effectiveness and acceptance of interactive video instruction. *Journal of Computer-Based Instruction*, 19, 17-21.
- Leithner, A. (2011). Do student learning styles translate to different testing styles?. *Journal of Political Science Education*, 7(4), 416-433.
- Londe, Z. C. (2009). The effects of video media in English as a second language listening comprehension tests. *Issues in Applied Linguistics*, 17(1), 41-50.
- Messick, S. (1989). Validity. In R. L. Linn (Ed.), *Educational measurement* (pp. 13-103). New York: Macmillan.

- Messick, S. (1996). Validity and washback in language testing. *Language Testing* 13(3): 241-256.
- Morrison, M., Sweeney, A., & Heffernan, T. (2006). Karns's learning styles and learning effectiveness: A rejoinder. *Journal of Marketing Education*, 28(1), 64- 68.
- Myers, I. B. (1962). *Manual: The Myers-Briggs Type Indicator*. Princeton, NJ: Education Testing Service.
- Nichols, S. L., & Berliner, D. C. (2008). Why has high-stakes testing so easily slipped into contemporary American life?. *Education Digest*, 74(1), 41-47.
- Ockey, G. J. (2007). Construct implications of including still image or vide in computer-based listening tests. *Language Testing*, 24(4), 517-537.
- Oxford, R. (1990). *Language learning strategies: What every teacher should know*. Newyork: Newbury House.
- Oxford, R. L. (1995). Gender differences in language learning styles: What do they mean?. In J. M. Reid (Ed.), *Learning styles in the ESL/EFL classroom* (pp. 34-46). Boston: Heinle & Heinle.
- Progosh, D. (1996). Using video for listening assessment: Opinions of test-takers. *TESL Canada Journal*, 14(1), 34-44.
- Riding, R., & Smith, E.S. (1992). Type of instructional material, cognitive style, and learning performance. *Educational Studies*, 18, 323-340.
- Rubin, J. (1995). The contribution of video to the development of competence in listening. In D. Mendelsohn, & J. Rubin (Eds.), *A guide for the teaching of second language listening* (pp. 151-165). San Diego: Dominie Press.
- Shin, D. (1998). Using videotaped lectures for testing academic listening proficiency. *International Journal of Listening*, 12(1), 57-80. doi:10.1080/10904018.1998.10499019
- Schmitt, N. (1996). Uses and abuses of coefficient alpha. *Psychological Assessment*, 8(4), 3 50-353.
- Suvorov, R. (2009). Context visuals in L2 listening tests: The effects of photographs and video vs. audio-only format. In C. A. Chapelle, H.G. Jun, & I. Katz (Eds.), *Developing and evaluating language learning materials* (pp. 53-68). Ames, IA: Iowa State University.
- Swain, H. (2004). It's as easy as A, B or C. *Times Higher Education Supplement*, 1648(7), 21-29.
- Thompson, I. (1995). Assessment of second/foreign language listening comprehension. In D. J. Mendelsohn, & J. Rubin (Eds.). *A Guide for the teaching of second language listening* (pp. 31-58). San Diego, CA: Dominie Press.
- Vandergrift, L. (2004). Listening to Learn or Learning to Listen? *Annual Review of Applied Linguistics*, 24, 3-25.
- Vandergrift, L. (2007). Recent developments in second and foreign language listening comprehension research. *Language Teaching*, 40, 191-210.
- Vermunt, J. D. (1996). Metacognitive, cognitive and affective aspects of learning styles and strategies. *Higher Education*, 31, 25-50.

- Wagner, E. (2002). Video listening tests: A pilot study. *Teachers College, Columbia University Working Papers in TESOL & Applied Linguistics*, 2(1).
- Wagner, E. (2006). *Utilizing the visual channel: An investigation of the use of video texts on tests of second language listening ability* (Unpublished doctoral dissertation). Teachers College, Columbia University, New York.
- Wagner, E. (2007). Video listening tests: What are they measuring?. *Language Assessment Quarterly*, 5(3), 218-243.
- Wagner, E. (2013). An investigation of how the channel of input and access to test questions affect L2 listening test performance. *Language Assessment Quarterly*, 10(2), 178-195.