

ARE THEY WATCHING? TEST-TAKER VIEWING BEHAVIOR DURING AN L2 VIDEO LISTENING TEST

Elvis Wagner

Teachers College, Columbia University

ABSTRACT

The accessibility of video technology has made it possible to utilize both the auditory and visual channels to present listening texts in the second language (L2) classroom and on L2 listening tests. However, there has been little research investigating the extent to which L2 listeners actually watch the video monitor when presented with a listening video text. The current study investigated test-taker behavior on an L2 video listening test. Thirty-six test-takers were videotaped while taking a listening test composed of six separate video texts, and the amount of time test-takers made eye contact with the video monitor was computed. An analysis of the data indicated that the group of participants oriented to the video monitor 69% of the time while the video text was played. In addition, the study yielded valuable information concerning the consistency of the test-takers' viewing behavior.

INTRODUCTION

Traditionally, the aural input for second language (L2) listening tasks has been delivered by a teacher reading aloud a text for the students. Later, as audio technology developed, a text was recorded on audiotape and played for students. However, with the advent and dissemination in the 1980s of inexpensive, reliable, and high-quality video recording equipment, it became practical to deliver listening texts using video texts, which involve both the auditory and visual channels. Subsequently, the use of video to teach listening comprehension has become more common in the L2 classroom. As Nunan (2005) suggested, in many aspects technology has become as effective as humans in delivering content for L2 listening classrooms.

As the use of video to teach L2 listening increased, researchers became more cognizant of the role of non-verbal communication in listening ability. A general consensus seems to have emerged among L2 listening researchers that the non-verbal components of spoken communication are an important component of L2 listening ability, and that L2 listeners are able to more easily construct the meaning of a spoken text that includes non-verbal input than a spoken text that does not include non-verbal input (e.g., Allan, 1984; Altman, 1990; Baltova, 1994; Gruba, 1997; Hasan, 2000; Kellerman, 1990, 1992; Progosh, 1996; Shin, 1998). The use of video texts allows listeners to utilize the non-verbal components of communication that can assist them in processing and comprehending aural input. In the majority of L2 listening situations (excluding situations such as talking on the phone, listening to the radio, or listening to loudspeakers, etc.), the listener is able to see the speaker. Depending on the purpose of the test, the inclusion of the non-verbal components of spoken communication through the use of video texts on L2 listening test tasks might be advantageous, because not only would the tasks more closely simulate the characteristics of authentic spoken language, but the inclusion of the visual channel in presenting the spoken input might lead to more construct relevant variance in the assessments, allowing for more valid inferences to be made from the results of those assessments (Wagner, 2006).

While numerous researchers (Baltova, 1994; Brett, 1997; Dunkel, 1991; Gruba, 1993; Parry & Meredith, 1984; Progosh, 1996; Shin, 1998; Thompson & Rubin, 1996) have investigated how the use of

technology to deliver listening texts that included both the aural and visual channel affected performance on L2 listening tests, there does not seem to be any systematic research on L2 listener behavior in relation to this technology. When presented with a video text, the listeners are not forced to watch the monitor. They can look away from the monitor, they can focus on their test papers, or they can even close their eyes. Listeners cannot utilize the non-verbal components of spoken communication provided by the video text if they are not watching the video monitor. What listeners actually orient¹ and attend to when presented with a video listening text has not been researched, and little is known about how L2 listeners interact with a listening video text.

The intent of this study², therefore, was to investigate the extent to which L2 listeners oriented to non-verbal information while taking an L2 video listening test. A group of test-takers was videotaped in order to determine the amount of time they oriented to the video monitor and to examine the consistency of their test-taking behavior.

LITERATURE REVIEW

The Use of Video Texts for Testing L2 Listening Ability

As noted previously, there seems to be a general consensus that being able to see the speaker allows L2 listeners to utilize non-verbal components of spoken communication, which can assist them in processing and comprehending the spoken input, and this might lead to increased test-taker performance. Numerous researchers (e.g., Bejar, Douglas, Jamieson, Nissan, & Turner, 2000; Brown, 1995; Buck, 2001; Kellerman, 1990, 1992; Lynch, 1998; Samuels, 1984, 1987) have presented hypotheses about why being able to see the speaker is useful for the L2 listener. For example, listeners can utilize the information transmitted by the kinesic behavior of the speaker, including gestures, body movements, and facial expressions (Burgoon, 1994; Kellerman, 1992; Von Raffler-Engel, 1980). Segmentally connected speech (which includes phonological reductions) results in marked morphophonological changes by the speaker, and these morphophonological changes are paralleled by a visible change in articulation, which the listener can utilize in trying to comprehend the spoken input (Brown, 1995; Rost, 1990). A speaker's body movement and the stressed syllables of the spoken text are often linked, and this visual stress can be useful for the listener in segmenting and processing the spoken input (Brown, 1995; Kellerman, 1992). Being able to see the speaker also can allow the listener to make more accurate initial hypotheses about the roles of the speaker and the context of the speaking situation (Shin, 1998, Wagner, 2006).

The role of non-verbal information, however, has been neglected in construct definitions of L2 listening ability. This neglect may be accidental, or it may be that researchers have concluded that the listener's utilization of non-verbal information to create meaning from a spoken text is not actually part of the L2 listening process, and thus should be omitted from the construct definition of listening ability. In reviewing the use of video in listening assessments, Buck (2001) implied that it is best to avoid the use of video, arguing that test developers should focus on testing language ability, "rather than the ability to understand subtle visual information" (p. 172). In addition, Buck stated that because research suggests that people "differ quite considerably in their ability to utilize visual information" (p. 172), it is better to emphasize comprehension of the aural rather than visual information.

It seems that Buck sees the non-verbal aspects of interpersonal communication as somehow less important than or separate from the verbal aspects of spoken communication. Similarly, Gruba (1993) questioned the use of video in L2 listening testing because of construct validity concerns and asked if the presenting of visual information on an L2 listening test can have listening trait validity. Indeed, a traditional definition of listening ability might focus solely on the linguistic (verbal) aspects of spoken texts. However, if the goal of the test is to assess a person's L2 listening ability in a communicative language ability framework, as suggested by Bachman (1990) and Bachman and Palmer (1996), then it is important to include the non-verbal components of spoken communication in the construct definition of

L2 listening ability. Non-verbal information is often a vital component of interpersonal communication. Obviously, there are instances of spoken communication in which the listener cannot see the speaker, including listening to the radio, talking on the phone, or listening to loudspeakers, and there are instances when it would not be appropriate to use the visual channel when testing L2 listening ability (e.g., a test to measure a person's ability to work in a telephone call center). However, in the majority of communicative language use situations, a listener is able to see the speaker and is able to utilize and exploit the non-verbal information the speaker is projecting. Burgoon (1994), while writing about L1 listening, stated that the verbal and nonverbal channels are "inextricably intertwined in the communication of the total meaning of an interpersonal exchange" (p. 347). To preclude non-verbal information on listening tests could be seen as a threat to the validity of the inferences made about a person's L2 listening ability based on those tests. Exactly because some people are better than others at utilizing nonverbal information to infer meaning in spoken communication, it is important that this component be considered when creating assessments meant to measure L2 listening ability.

Video as a Distraction

Some researchers, however, have maintained that the use of video might actually inhibit comprehension, because of its supposed potential for distraction. MacWilliam (1986) argued that the visual aspects of a video text can distract learners' attention from the audio input and may actually impede comprehension. Bejar et al. (2000), in their working paper devoted to creating a listening framework for the TOEFL 2000 test, stated "There is no doubt that video offers the potential for enhanced face validity and authenticity, although there is a lot of concern about its potential for distraction" (p. 28). Bejar et al. do not state, however, what the potential distractions of video are. Alderson, Clapham, and Wall (1995) speculated that video texts may not be helpful for listening test-takers because the test-takers are too busy focusing on their (written) test materials to utilize the non-verbal information conveyed by the video texts.

While the points these researchers raise are valid, they are mostly theoretical in nature, and the claims made have not been subjected to empirical validation. Whether video is a distraction or not begs the more basic question of whether, and to what extent, test-takers actually watch the video text. There appear to be only two studies that have investigated the extent to which test-takers watch the video monitor while performing a listening task. Gruba (1994) videotaped a group of test-takers that were taking a video listening test. He described how many of the test-takers simply listened to the audio portion of the test, rather than watch the video. However, Gruba did not describe the extent of this behavior, saying only that many of the test-takers "were not continuously watching the video" (p. 34). Brett (1997) conducted a study comparing performance on an audio-only listening test, a video listening test, and a computer-based multi-media listening test. Brett described how "it was certainly noticeable" during the video listening test that "over half of the learners in each class never watched the video" (p. 48), but looked at their test books with the written test questions instead. But Brett neither describes how he analyzed the test-takers' behavior nor how he ascertained that over half of them did not watch the video.

The findings that Gruba and Brett present are certainly relevant to the issue of the usefulness of video texts in L2 listening tests, but the issues they raise require further and more systematic empirical research. The present study is an exploratory step in this direction. It would be premature to advocate for or argue against the use of video texts on L2 listening tests without first systematically examining what the test-takers actually do when taking a video listening test. In other words, studies are needed to investigate the extent to which test-takers utilize the video channel and orient to the video monitor while the textual input is being presented. Because test-takers would not be able to utilize the non-verbal information presented in the video text if they were not in fact watching the video text, I decided to examine the amount of time that the test-takers oriented to the video monitor while the video text was being presented. In addition, because two genres of texts were used in the test (academic lecture and dialogue), I hypothesized that one genre of texts might be inherently more interesting, attractive, or useful for the test-takers to watch, and for this reason I decided to compare the test-takers orientation behavior for the two different text types.

Finally, I also hypothesized that the rate at which test-takers oriented to the video monitor might decline as the test progressed, because the video might become less interesting to them as it progressed, or because the test-takers had determined that the nonverbal information found in the video was not helpful or useful for them in comprehending the aural text. Therefore, the test-takers' orienting rate to the video monitor as the video text progressed was also examined.

RESEARCH DESIGN

Research Questions

- 1) How do L2 listening test-takers interact with a video text? To what extent do they orient to the video monitor while the video text is playing?
- 2) Is the test-takers' orienting behavior affected by the type of text that is being played? In other words, do the test-takers orient to the video monitor at a higher rate for a particular text type (academic lecture vs. dialogue)?
- 3) Is the test-takers' orienting behavior consistent throughout the test? In other words, do the test-takers tend to orient to the video monitor at a lower or higher rate for the beginning of the test than they do for the end of the test?

Participants

The participants for this study were 36 students enrolled in the Community English Program (CEP) language program at a major university in the Eastern United States. The CEP is an adult language program with students ranging in age from 18 to 68, and from numerous and diverse language, national, and cultural backgrounds. The courses at the CEP are not credit-bearing. There are classes at six different ability levels: Levels 1 through 6. The 36 participants in this study were from eight different CEP classes representing Levels 2 through 6 (because of the difficulty level of the test, Level 1 classes were not included in this study).

Materials

I created the video listening test instrument that was used here as part of a larger study investigating the use of video texts on test-taker performance. The test was developed based on an operationalization of Buck's (2001) L2 listening ability construct, which includes the ability to listen for both explicitly-stated and implicitly-stated information, automatically and in real time. The test was developed for this population, and was piloted and revised a number of times. A more thorough account of the creation and validation of the test instrument is given in Wagner (2002).

The test consisted of six separate tasks that included three dialogue texts and three lecturette texts. All of the texts used an academic listening target language use domain. As used here, the term *dialogue* refers to a text consisting of two speakers in an interactive conversation. The dialogue texts used in this test involved two speakers discussing events that happened in their classes. In the first dialogue text, one speaker explains to the second person how he received a 'C' grade in his biology class. In the second dialogue text, the speaker relates an incident that happened in his class. In the third text, the speaker tells the second person about a group project he participated in. Each of the dialogue texts lasted from one and a half to three and a half minutes, and was followed by five or six comprehension items. These items were both limited-production and selected-response. The limited-production items required the participants to write short answers (SA) of 25 words or less. The selected-response items were multiple-choice (MC) with four response options.

The term *lecturette* refers to a description or explanation of a subject delivered before an audience or a class for the purpose of instruction (an academic lecture, but less than five minutes in duration). A lecturette usually has one speaker with an academic speaking style and includes language primarily for

transactional language purposes. The lecturette texts used in this study consisted of one speaker delivering a short lecture on history or biology subjects. Two pictures of the subjects were shown in each of the lecturettes. The first lecturette was about skunks, the second lecturette was about Wild Bill Hickok, and the third lecturette was about the Ivory-billed Woodpecker.³ Each of the lecturette texts lasted from three to four minutes, followed by eight items. Again, these items included both short answer and multiple-choice questions, with the multiple-choice items having four options.

Test-Taking Procedure

I went to the eight different CEP classes and administered the video listening test during class time. These eight different classes included: two Level 2 classes, three Level 3 classes, one Level 4 class, one Level 5 class, and one Level 6 class. This proportion is representative of the CEP class structure: there are more Level 2 and 3 class sections than other classes.

The test booklet (six pages long, with the set of test items for each test task on a separate page) was given to the test-takers, and then I played the videotape. The narrator on the video instructed the test-takers that they had one minute to preview the questions for the first task, and then the first video was played. At the end of the first video, the narrator instructed the test-takers that they had two minutes to answer the questions for task 1. At the end of this two-minute period, the narrator instructed the test-takers that they had one minute to preview the questions for task 2, and then the second video was played. When this video ended, the narrator instructed the test-takers that they had three minutes to answer the questions for task 2 (test-takers had two minutes to answer the questions for the dialogue texts, and three minutes to answer the questions for the lecturette texts). This procedure was repeated four more times through task six. With the one minute question preview period prior to each task, the six different texts, and the two or three minutes (depending on task type) for answering the questions for each task, the overall test took thirty-eight minutes for test-takers to complete. The characteristics of the tasks, and the order in which they are presented, are shown in [Table 1](#), and the test instrument can be seen in the [Appendix](#).

Table 1. Characteristics of the Test Tasks for the 40-item Assessment

Task Name	Text Type	Length (minutes)	# of Short Answer items	# of Multiple-Choice items	Time Given to Answer (minutes)
Skunk	Lecturette	2.5	2	6	3
Bob & Julie	Dialogue	1.5	3	2	2
WB Hickok	Lecturette	2.5	5	3	3
David & Amy	Dialogue	2	4	1	2
IB Woodpecker	Lecturette	4	4	4	3
Laura & Jimmy	Dialogue	3.5	4	2	2
		Total	22	18	

The participants were videotaped while they were taking the video listening test. A Sony digital video camera was mounted next to the video monitor and directed toward the test-takers while they were taking the test. In an attempt to keep the observation as objective as possible, test-takers were informed that I would analyze the videotape to examine what test-takers do when presented with a video listening test, but they were not explicitly told that the reason for the videotaping was to examine the percentage of time they oriented to the video monitor. Although the test-takers were aware that they were being videotaped, the video camera used for the taping was very small and inconspicuous, and it appeared that most of the test-takers were not consciously aware of the video camera while they were actually taking the test.

Two procedural issues related to the development and administration of this test are briefly discussed here. The first involves the number of times the text should be played for test-takers. Unfortunately, no

consensus was available in the literature as to how many times the text should be played. Sound arguments can be made for playing the text once, or for playing the text twice (e.g., Berne, 1995; Cervantes & Gainer, 1992; Chang & Read, 2006; Chiang & Dunkel, 1992). I decided to have each text played once, based partly on practicality considerations due to the overall length of the test.

The second issue relates to the idea of question preview. Again, no consensus was available in the literature regarding whether the questions should be given to test-takers before the texts were played or after, although there is evidence that a question preview period might lead to increased test performance (Berne, 1995, Chang & Read, 2006). For this test, it was decided to give the test-takers the test booklet containing all the items at the very beginning of the test to allow for question preview. Practicality constraints again necessitated this (because there were a set of test items for each of the six tasks, it would not have been feasible to give the test-takers each set of questions after the particular text was played), but theoretical considerations also influenced this decision.

Data and Method of Analysis

The eight classes in which the test was administered had a total of 71 test-takers present when the test was given. However, because of the inability of the camera to cover the entire classroom, because some test-takers elected to not be videotaped, and because some of the test-takers sat directly behind other test-takers and thus were unable to be seen on the videotape, only 36 of the 71 test-takers that were videotaped were actually analyzed for the amount of time spent orienting to the video monitor. This information, along with the class mean scores on the test, can be seen in [Table 2](#).

Table 2. Participants in the Study

CEP Class Level	Class Average on the Test (out of 40)	# of Test-takers in the Class	# of Test-takers Analyzed
Level 2	15.70	14	6
Level 2	17.11	9	5
Level 3	18.75	12	2
Level 3	19.11	9	5
Level 3	19.36	7	4
Level 4	20.71	7	4
Level 5	18.31	8	6
Level 6	33.07	7	4
		Total 71	36

The videotapes of the eight classes were then examined, and the amount of time test-takers oriented to the video monitor while the video was played was calculated. I did this by monitoring each individual test-taker on the videotape using a stopwatch timer. When the test-taker made eye contact with the video monitor, I started the timer, and when the test-taker looked away from the video monitor, the timer was stopped. This procedure was repeated 36 times⁴. This amount was converted into a percentage by dividing the amount of time the test-taker was watching the video monitor, by the total text time. In addition, the percentage of time the group of 36 test-takers oriented to the video monitor during the lecturette texts was computed and compared to the percentage of time they oriented to the monitor during the dialogue texts by conducting a repeated measures *t* test. Finally, the percentage of time the set of test-takers oriented to the video monitor during the initial two texts was computed and compared to the percentage of time they oriented to the monitor during the final two texts by conducting a repeated measures *t* test.

RESULTS

The first component of the data examined here focuses on the amount of time the test-takers oriented to the video monitor for the overall test. The six video texts had a combined length of 951 seconds. The mean amount of time spent orienting to the video monitor by the 36 participants was 655.86 seconds, with a standard deviation of 256.47, which means that the set of 36 participants were orienting to the video monitor 69% of the time. The amount of time each individual participant oriented to the monitor ranged from 165 seconds to 951 seconds. The test-taker who spent the least amount of time orienting to the video monitor oriented 17% of the time, while two of the test-takers oriented to the video monitor 100% of the time. In fact, six out of the 36 participants oriented to the video monitor more than 98% of the time. A histogram of the percentage the test-takers oriented to the video monitor can be seen in [Figure 1](#).

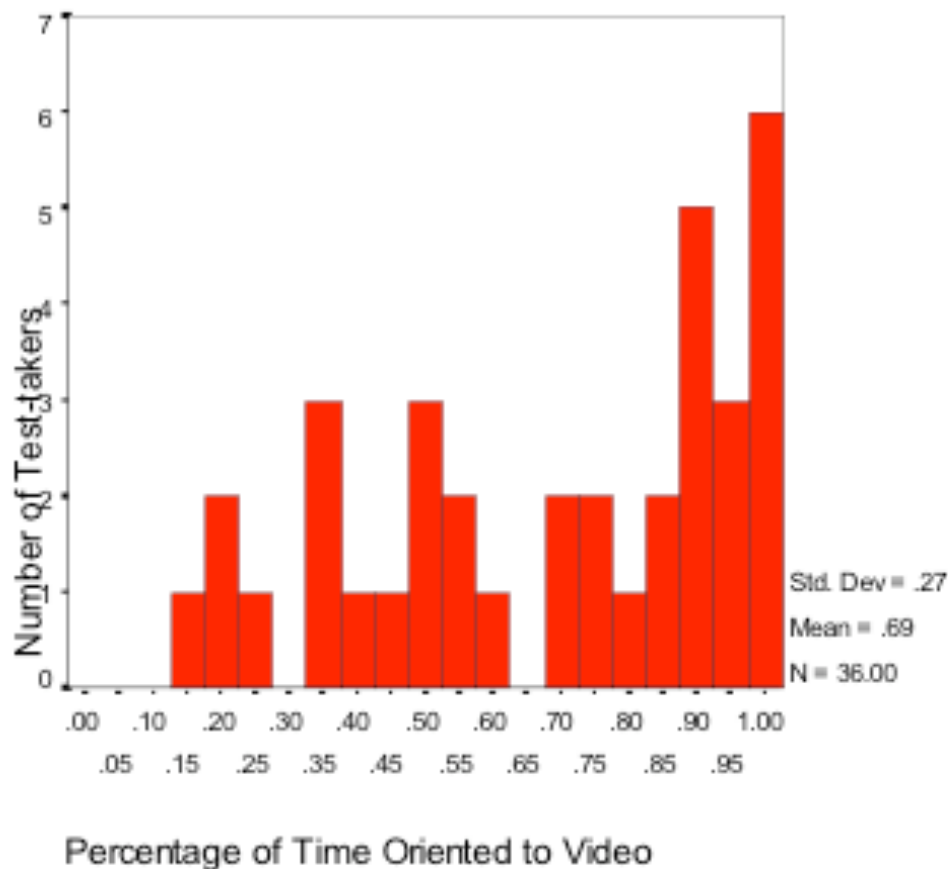


Figure 1. Distribution of the percentage of time participants oriented to the video monitor

The mean percentage of time the set of participants oriented to the video monitor for each task was also computed. The percentages ranged from as low as 62% for the "Skunk" task (task 1), to a high of 77% for the "David and Amy" task (task 4). In addition, the mean percentage of time participants oriented to the monitor for the three dialogue tasks (72%) and for the three lecturette tasks (67%) was computed. Finally, the mean percentage of time the participants oriented to the monitor for the first two tasks combined (65%) and the final two tasks combined (68%) was computed. These figures can be seen in [Table 3](#).

Table 3. Participants' Orienting to the Video Monitor

Task (in order they appear on test)	Text Type	Overall Video Text Length (in seconds)	Average Time Oriented to Video (in seconds)	Average Time Oriented to Video (%)	SD
Skunk	Lecturette	138	85.61	62	32
Bob & Julie	Dialogue	77	54.89	71	31
WB Hickok	Lecturette	160	116.22	73	27
David & Amy	Dialogue	101	77.61	77	28
IB Woodpecker	Lecturette	254	166.86	66	31
Laura & Jimmy	Dialogue	221	154.67	70	29
Overall Test (6 Texts)	3 Dialogues and 3 Lecturettes	951	655.86	69	27
Dialogue Texts	3 Dialogues	399	287.17	72	28
Lecturette Texts	3 Lecturettes	552	368.69	67	28
Initial Two Texts	1 Lecturette and 1 Dialogue	205	140.50	65	29
Final Two Texts	1 Lecturette and 1 Dialogue	475	321.53	68	29

To test whether there was a difference in the percentage of time the participants oriented to the video monitor for the different types of texts (lecturette versus dialogue), a repeated measures *t* test was conducted. It was found that the set of participants oriented to the video monitor at a higher rate for the dialogue texts (72% of the time) than for the lecturette texts (67% of the time), and this difference was statistically significant. These figures are summarized in Table 4.

Table 4. Comparison of Participants' Orienting to the Video Monitor by Text Type

Text Type	N	M (%)	SD	<i>t</i> -value
Dialogue (Bob and Julie, David & Amy, Laura & Jimmy)	36	72	28	2.272* (<i>p</i> =.029)
Lecturette (Skunk, Wild Bill, IB Woodpecker)	36	67	28	

To test whether there was a difference in the percentage of time the participants oriented to the video monitor for the beginning of the test (the initial two tasks) and for the final portion of the test (the final two tasks), another repeated measures *t* test was conducted. It was found that the participants oriented to the video monitor a higher percentage of the time for the final two tasks (68% of the time) than for the initial two tasks (65% of the time), although this difference was not statistically significant. These figures are summarized in Table 5.

Table 5. Comparison of Participants' Orienting to the Video Monitor by Text Chronology

Text Type	N	M (%)	SD	<i>t</i> -value
Initial Two Texts (Skunk, Bob and Julie)	36	65	29	-.823 (<i>p</i> =.416)
Final Two Texts (IB Woodpecker, Laura & Jimmy)	36	68	29	

DISCUSSION

Research Question 1

1) How do L2 listening test-takers interact with a video text? To what extent do they orient to the video monitor while the video text is playing?

The set of 36 participants in this part of the study oriented to the video monitor 69% of the time while the video text was playing. These results suggest that the test-takers interacted extensively with the video text, orienting to the video monitor over two thirds of the time. This finding itself seems important, because there has been no systematic and quantitative investigation of whether test-takers even watch the video when taking an L2 listening test. Because there is no figure to compare the results of this with, it is difficult to say if this figure of 69% is more or less than anticipated. However, in light of anecdotal accounts of observations of test-takers by researchers (e.g., Alderson, Clapham, & Wall, 1995; Brett, 1997; Gruba, 1994) who have commented that the video might not be useful for test-takers because they do not even watch the video monitor, this finding does seem to be important in that it indicates that the test-takers in this study did indeed orient to the video monitor for a large portion of the time while the text was being played.

It must also be noted that while the mean orienting rate for the set of 36 test-takers was 69%, there was a great range of orienting behavior, as was seen in [Figure 1](#). In fact, 14 of the 36 test-takers oriented to the video monitor at least 90% of the time, while 10 of the 36 test-takers oriented to the video monitor less than 50% of the time. Thus, rather than interpreting the mean as indicative of the orienting rate for each of the individual test-takers, it can be interpreted as indicating that while the individual orienting rate varied widely, as a group the test-takers did tend to orient to the video while the text was played.

The issue of *attention* is relevant here. L2 listeners, because of their lack of linguistic knowledge and limited processing capacity, quite often can experience information overload (VanPatten, 1996). The test instrument used in the current study was probably perceived as a cognitively demanding task by many of the test-takers. They had to listen to six fairly extended texts in a foreign language that were spoken at normal speeds with natural language. The test-takers had to read the comprehension questions (in English) and remember these questions while listening to the text. They had to remember the answers to the questions and write the correct answers. However, the results of this study indicated that even when faced with test tasks that might be causing the test-takers to reach the point of cognitive overload, test-takers still tended to orient to an additional input source, the video monitor.

With the video texts, the test-takers had the choice of where to devote their attentional resources. If they felt that it would be more efficient and worthwhile to allocate their attentional resources to process both the visual and the aural input, they were able to do this. If they felt that it was more efficient and worthwhile to ignore the visual input completely, they were able to do this. The test-takers were able to focus their attentional resources in whichever manner they felt was most advantageous for them to comprehend the text. This is similar to the target language use domain, where listeners are generally able to focus their attentional resources in the manner most advantageous to comprehension.

Research Question 2

2) Is the test-takers' orienting behavior affected by the type of text that is being played? In other words, do the test-takers orient to the video monitor at a higher rate for a particular text type (lecturette vs. dialogue)?

The statistical analysis indicated that the percentage of time test-takers oriented to the video monitor for the three dialogue texts (72%) was higher than the percentage of time they oriented to the three lecturette texts (67%). While this difference was statistically significant at the $p < .05$ level, the practical significance of this finding is less clear. It was anticipated that the test-takers would watch the dialogue texts a larger percentage of the time because the test-takers would find these texts inherently more interesting. While the lecturette texts are on fairly academic subjects, the dialogue texts are more oriented

to interpersonal communication, with two speakers discussing events that I had hypothesized would mirror the interests of the test-takers. The "Bob and Julie" text has one of the speakers complaining about his grade in a class. The "David and Amy" text has one of the speakers recounting a funny incident that happened in class. The "Laura and Jimmy" text has one of the speakers complaining about another student in his group project. The test-takers were observed laughing and smiling a number of times while they were watching and listening to these texts, which was not the case with the lecturette texts. In addition, participants from the same population that participated in a different component of the larger study reported verbally that the dialogue texts were more interesting for them than the lecturette texts. If indeed the test-takers found the dialogue texts more interesting or amusing to watch than the lecturette texts, this might have contributed to the statistically significant higher orienting rate for the dialogue texts.

Another reason I conjectured that test-takers would orient to the dialogue tasks a higher percentage of the time was because of the context-embedded nature of the dialogues. It was believed that the contextual cues and non-verbal components of communication would be more prevalent and salient for the participants while listening to and viewing the texts involving two speakers engaging in interpersonal communication, and thus might be more useful for the participants. In contrast, the lecturette tasks were more context-reduced, and contained more transactional language with only one speaker, thus leading to the conjecture that the non-verbal components of these texts would be less salient and beneficial to the participants for processing the incoming text. Burgoon (1994) described how adults rely more on nonverbal cues than verbal cues in trying to determine social meaning in communicative situations and how verbal cues seem to be more important for factual and persuasive messages, while nonverbal cues are more important for interpersonal relationships. Because the dialogue texts were more interpersonal in nature than the transactional and academic tone of the lecturette texts, it was predicted that the more emotional nature of the dialogue texts would lend itself to the speakers displaying that emotion through body language including facial expressions, hand gestures, and other types of non-verbal communication. In addition, these texts had a number of comprehension items that attempted to assess the listener's ability to comprehend the more emotional and interpersonal content of the dialogue texts.

Related to this notion of the type of language found in the texts (interpersonal or transactional) are the findings described by Nissan, Devnicensi, and Tang (1996), who found that the variable *Role of Speakers* had a significant effect on item difficulty. In their study, they found that when there was more than one person speaking on the audio-only text, the test-takers sometimes had difficulty determining who was speaking, as well as the speakers' roles. Nissan et al. suggested that this might be due to the test-takers not having the sociolinguistic competence to make the necessary inferences about situations and speaker roles. For this study, the test-takers might have been able to utilize visual information in the text to assist them in determining speaker roles, and thus it was conjectured that the idea of speaker roles would be more salient in the two-person dialogues than in the single speaker lecturette texts.

The statistical analysis indicated that the test-takers did indeed orient to the video at a higher rate during the dialogue texts. Although the difference in the rate of orienting to the video monitor between the two types of tasks is small, it is statistically significant. While the reason for this higher rate can only be conjectured, the hypothesis that test-takers would orient to the video at a higher rate during the dialogue texts than the lecturette texts was confirmed by the data.

Research Question 3

3) Is the test-takers' orienting behavior consistent throughout the test? In other words, do the test-takers tend to orient to the video monitor at a lower or higher rate for the beginning of the test than they do for the end of the test?

The analysis of the percentage of time test-takers oriented to the video for each of the six texts indicates that the test-takers' orienting patterns were fairly consistent. For the first text, test-takers watched 62% of the time, and this percentage increased slightly for each of the next three tasks, reaching a high of 77% in

task four (David and Amy). The rate then decreased to 66% in task five, and then increased again to 70% in the sixth and final task (Laura and Jimmy), as was summarized in Table 3. Although there was some variation, the test-takers' orienting rate was fairly consistent across the six tasks (ranging from 62% to 77%), and was fairly consistent as the test progressed.

I had conjectured that the rate of orienting to the video might decrease as the test progressed. Specifically, I reasoned that the orienting rate might be lower for the final two tasks than the rate for the initial two tasks. This was conjectured for a number of reasons, including the idea that the novelty of watching the videos might wear off. I thought that the test-takers might orient to the initial video texts because the characters and images shown in the first two videos would be novel. The same actor performed the lecturette texts one, three, and five, and the same two actors performed the dialogue texts two, four, and six (the actor in the lecturette texts was also one of the speakers in the dialogue tasks). Because the actors were the same, it was conjectured that by task three the test-takers might be less interested in the images shown on the video text, and thus might orient to the video monitor at a lower rate in the final portion of the test.

I also predicted that the test-takers would stop orienting to the video monitor if they found that the non-verbal components of the video were not helpful for them in comprehending the aural text and answering the comprehension questions. In other words, if the test-taker watched the first couple of video texts, and determined (consciously or unconsciously) that the video was not providing them with any useful information in interpreting the aural text, or in answering the comprehension questions, they might decide to stop watching the video. Brett (1997) described how most of the test-takers in his experiment using a video listening comprehension test did not watch the video, but instead focused on the true/false questions on their test papers while the video was being played. Similarly, other researchers (including Bejar et al., 2000), speculated that the video might actually be distracting for test-takers, and thus test-takers might avoid watching the video if they found that it was not useful or beneficial for them. If test-takers found that the video was not helpful for them in comprehending the text, or if they found the video distracting, it was hypothesized that their orienting behavior would change; the test-takers would orient less and less to the video monitor as the test progressed. However, the statistical analysis of the percentage of time test-takers oriented to the video in the first two tasks in comparison to the final two tasks showed that the orienting rate actually increased (65% for the initial two tasks, compared with 68% for the final two tasks), although this difference was not statistically significant. The test-takers' behavior in orienting to the video was fairly consistent through the course of the test, and there was not a drop-off in their orienting to the video as the test progressed. I interpreted this as indicating that the test-takers did not conclude that the video was unhelpful in comprehending the text. It would also seem to indicate that test-takers did not find the video distracting.

Study Limitations

There are a number of limitations to this study that must be acknowledged. First, the methodology required that I observe the eye contact that test-takers made with the video monitor. Although the test-takers oriented to the video almost 69% of the time, this does not necessarily imply that they were attending to the non-verbal components of the video text the video. Even though they were physically oriented to the video monitor, they might not have been attending to the visual information on the monitor.

As mentioned earlier, this study was part of a larger study. For the other components of the study, it was necessary to present the six different test tasks in the same order. For this study, however, it might have been more useful to present the six texts in random order. Doing so might have presented useful information about the test-takers' viewing behavior, especially in examining if the test-takers tended to orient to the video at a lower rate as the test progressed.

Another possible limitation to the study involves the observer's paradox. Although efforts were made to minimize this issue by using an unobtrusive video camera to record the test-takers, and by not explicitly informing the test-takers that the reason for the videotaping was to examine the amount of time they oriented to the monitor, the fact that they were videotaped may have influenced their orienting behavior.

It is also important to acknowledge the limited generalizability of these findings. While this set of 36 test-takers oriented to the video monitor 69% of the time while the text was played, as noted above, the individual test-takers varied greatly in their orienting behavior. It seems logical to believe that orienting behavior is influenced by a great number of variables, including the background characteristics of the listeners and the characteristics of the test tasks. For this test, the test-takers came from a number of different language and cultural backgrounds, and varied extensively in their L2 listening ability. It could be that test-takers from a particular cultural group might be more inclined to orient to the video monitor compared with test-takers from another cultural group. The ability level of the test-takers might also affect their orienting rates; it might be that lower ability test-takers orient to the video monitor at higher or lower rates than higher ability test-takers. Similarly, it seems likely that the characteristics of the test tasks used in the study influenced orienting behavior. The test-takers had a question preview period in which they could read the questions before the texts were played, and this might have made them feel freer to orient to the video monitor rather than looking at their test booklets while the video text was played.

Finally, this study did not correlate orienting rates and test performance. It might have been informative to do so. However, because this study involved videotaping the test-takers, in order to get institutional review board (IRB) approval for the study, I was not able to identify who the individual test-takers were on the video tape. Therefore, it was impossible to correlate the individual test-takers' orienting rate with their test scores.

CONCLUSION

Implications

This study was a preliminary investigation into how L2 listening test-takers interact with video listening texts, and the extent to which test-takers orient to the video texts. Clearly, these results are not conclusive, and more research is warranted. Even after acknowledging its weaknesses and limitations, however, this study does provide some theoretical and practical insight into L2 video listening test-taking behavior. This population of test-takers did orient to the video extensively—69% of the time. Their orienting to the video monitor was fairly consistent over time, and the percentage of time they oriented to the video did not drop off as the test progressed. This would seem to present evidence indicating that the test-takers did not decide that the video text was not useful for them in interpreting the aural text or in answering the items. It would also seem to present evidence dispelling the notion that L2 listening test-takers are too busy attending to the written material in their test booklets to attend to the video monitor during L2 listening tests. The data examined here provided evidence that the test-takers did not find the video "distracting" (as a number of researchers have conjectured). Rather, the test-takers in this study seemed willing and eager to watch the video texts. In addition, the data provided limited evidence that test-takers tended to orient to the video at a higher rate during the dialogue texts than the lecture texts, possibly because they found the dialogue texts more interesting to watch, and possibly because they found the non-verbal information in these texts more useful in comprehending the aural text and in answering the comprehension items.

Suggestions for Future Research

This analysis of the data from the present study has made it possible to provide some recommendations for future research regarding test-taker behavior during a test of L2 listening ability using video texts. A useful future study would be one which investigated the correlation between test scores and test-takers' orientation to the video monitor, in order to address whether test-takers who score higher on a video listening test tend to orient to the video monitor at a higher rate than test-takers who score lower on the test. Another useful study would be one that assesses test-takers' awareness of their own orienting behavior to the video monitor. Asking test-takers to estimate what percentage of the time they orient to the video monitor, and then comparing these self-assessments with an analysis of the actual time oriented to the video, might provide information about listeners' awareness and consciousness of their viewing behavior and preferences. Finally, a study that investigated orienting behavior with different test task characteristics would be informative. The study conducted here utilized a test instrument in which the test-takers saw the video texts one time, and which had a question preview period in which they could read the multiple-choice and short answer comprehension items before the video texts were played. It would be informative to investigate how an individual test-taker's orienting behavior might change if the video texts were presented two times, or if the test-takers did not have a question preview period, or if different item types were used.

APPENDIX

Video Listening Post-Test Instrument

Part 1 SKUNKS

Watch the video, and then answer the following questions. For the multiple-choice questions, circle the letter of the correct answer. For the other questions, answer them as completely as you can in 25 words or less. You will see the video one time.

1. A skunk is a type of _____.
 - a. bird
 - b. insect
 - c. reptile
 - d. mammal

2. What physical characteristic makes it easy to see skunks in nature?

3. The thick, yellow fluid that causes a skunk's odor is called _____.
 - a. musk
 - b. gland
 - c. species
 - d. omnivorous

4. Why is the skunk's spray an effective defense against most predators?

5. The skunk's spray is not effective against birds like hawks and owls because these birds _____.
 - a. eat agricultural crops
 - b. can't smell very well
 - c. like the taste of skunks

- d. are able to see skunks at night
6. According to the speaker, skunks have benefited from the presence of humans because humans have _____.
- not eaten skunks
 - always avoided skunks
 - cleared trees and forests
 - killed many of the skunks' predators
7. Grasshoppers and larvae are examples of _____.
- different types of skunks
 - agricultural pests that skunks eat
 - chemicals that cause a skunk's odor
 - predators that sometimes eat skunks
8. What is the best title for this video?
- "The Extinction of Skunks"
 - "The Story of Smelly Skunks"
 - "Man's Best Friend: The Skunk"
 - "Skunks: Food for Hawks and Owls"

Part 2 BOB AND JULIE

Watch the video, and then answer the following questions. For the multiple-choice questions, circle the letter of the correct answer. For the other questions, answer them as completely as you can in 25 words or less. You will see the video one time.

- How does Julie react to Bob's grade?
 - She is mad.
 - She is jealous.
 - She is excited.
 - She is surprised.
- Did Bob have a good reason for missing class? Why or why not?
- How did Bob do on his final exam?
 - He got a "B".
 - He did very well.
 - He doesn't know yet.
 - He hasn't taken it yet.
- What does Julie advise Bob to do?
- Will Bob do what Julie advises? Why or why not?

Part 3 WILD BILL HICKOK

Watch the video, and then answer the following questions. For the multiple-choice questions, circle the letter of the correct answer. For the other questions, answer them as completely as you can in 25 words or less. You will see the video one time.

1. What did Wild Bill look like?
2. How or why was Wild Bill's "twist" draw an unusual draw?
3. How many people were killed at Rock Creek Station? Who killed them?
4. Why did newspapers and magazines write exaggerated stories about Wild Bill?
5. Wild Bill was killed in _____.
 - a. Abilene, Kansas
 - b. Kansas City, Missouri
 - c. Deadwood, South Dakota
 - d. Rock Creek Station, Nebraska
6. Who was Jack McCall?
7. What's the best title for this passage?
 - a. "The Legend of Wild Bill"
 - b. "The Rock Creek Station Killer"
 - c. "Wild Bill: Best Poker Player in the West"
 - d. "Wild Bill: The World's Wildest Cowboy"
8. The speaker's tone in this video is _____.
 - a. sarcastic
 - b. humorous
 - c. persuasive
 - d. informative

Part 4 DAVID AND AMY

Watch the video, and then answer the following questions. For the multiple-choice questions, circle the letter of the correct answer. For the other questions, answer them as completely as you can in 25 words or less. You will see the video one time.

1. What happened to Tina in class, and why?
2. How did the professor respond to the incident? What did he say and do?
3. What is Amy's attitude toward Tina after she hears the story?
4. What is David's attitude towards Tina?
5. What is Amy's attitude toward David when he tells the story?

- a. She is upset with him for laughing.
- b. She likes him, so she laughs at his story.
- c. She feels sorry for him because he was hurt.
- d. She is jealous of him because he got a good grade in the class.

Part 5 THE IVORY-BILLED WOODPECKER

Watch the video, and then answer the following questions. For the multiple-choice questions, circle the letter of the correct answer. For the other questions, answer them as completely as you can in 25 words or less. You will see the video one time.

1. What did the Ivory-Billed Woodpecker look like?
2. Why was the Ivory-Billed Woodpecker sometimes called the "Lord God" bird?
3. The Ivory-Billed Woodpecker ate by _____.
 - a. getting nectar from flowers
 - b. catching insects flying in the air
 - c. digging out bugs from the wood in trees
 - d. collecting acorns and other nuts from trees
4. What did environmentalists do to try and save these birds?
5. The Ivory-Billed Woodpecker is (probably) extinct due to _____.
 - a. disease
 - b. hunting
 - c. logging
 - d. starvation
6. Who is David Kulivan?
7. Some people believe that the Ivory-Billed woodpecker is not extinct because _____.
 - a. it is human nature to maintain hope
 - b. there is a population of the birds alive in Cuba
 - c. scientists working with DNA hope to be able to clone the birds
 - d. there have been numerous confirmed sightings of the birds in recent years
8. Which word best describes the speaker's tone in this passage?
 - a. sarcastic
 - b. humorous
 - c. thoughtful
 - d. argumentative

Part 6 LAURA AND JIMMY

Watch the video, and then answer the following questions. For the multiple-choice questions, circle the letter of the correct answer. For the other questions, answer them as completely as you can in 25 words or less. You will see the video one time.

1. Why is Jimmy worried about working with Dave Rice?
 - a. Dave has skipped some classes.
 - b. Jimmy believes that Dave isn't very smart.
 - c. Jimmy doesn't think he can depend on Dave.
 - d. Dave is very bossy and always wants to be in charge.
 2. What does Dave Rice look like?
 3. Why doesn't Laura like Dave Rice?
 4. What two things does Laura advise Jimmy to do (not including "Run!")?
 5. Why doesn't Jimmy take Laura's advice?
 6. Does Laura believe that Jimmy's talk with Dave will work?
 - a. Yes, because Jimmy can be very forceful and persuasive.
 - b. No, because she believes that Dave won't listen to Jimmy.
 - c. Yes, because she knows that Dave will do what people tell him to do.
 - d. No, because she knows that Jimmy will not have the courage to confront Dave.
-

NOTES

1. The role of *attention* is an important construct in SLA research (Schmidt, 1990). For this paper, the term *orient to* is used according to Tomlin and Villa's (1994) definition, in which they divided attention into three separate but interrelated networks: alertness, orientation, and detection. *Orienting* is the directing of attentional resources to a certain type or class of sensory information, while excluding other types or classes of information. This directing of attentional resources usually serves to facilitate detection of the sensory information that is focused on. For this paper, then, *orienting to* the video monitor signifies that the listeners directed their visual attentional resources to the video monitor by looking at and making eye contact with the monitor while the video texts were being played.
 2. The study reported here is part of a larger study that investigated the impact of the use of video texts on L2 listening test-taker performance, and how L2 listeners process video texts.
 3. I sought to choose topics that were representative of an academic target language use domain. I also sought to choose topics that the test-takers would have limited background knowledge in. It is possible that some of the test-takers did have background knowledge of some of these topics, and thus might have been able to answer some of the comprehension items even if they were unable to understand the aural text. However, the results from the larger study which this test instrument was created for (Wagner, 2006) suggested that this was not an issue.
 4. This coding process was relatively straightforward and did not seem to require subjective judgment. However, in order to investigate the consistency of the coding, a subset of the data was reviewed by a second coder. This second coder followed the same procedure in independently reviewing six of the 36 participants' viewing behavior. A correlation analysis indicated that the two coders were remarkably consistent in their coding, with correlation coefficients ranging from a .991 to 1.000 for these six participants.
-

ACKNOWLEDGMENTS

I am grateful to the anonymous reviewers and the editors at Language Learning and Technology for their useful and insightful comments and suggestions for this article.

ABOUT THE AUTHOR

Dr. Elvis Wagner is currently a lecturer in the TESOL and Applied Linguistics programs at Teachers College, Columbia University. His research interests include language assessment (especially the assessment of second language listening), and the teaching and testing of second language pragmatic ability. The study presented here was part of his dissertation research.

REFERENCES

- Alderson, J., Clapham, C., & Wall, D. (1995). *Language test construction and evaluation*. Cambridge: Cambridge University Press.
- Allan, M. (1984). Viewing comprehension in the EFL classroom. *ELT Journal*, 38, 21-26.
- Altman, R. (1990). Toward a new video pedagogy: The role of schema theory and discourse analysis. *IALL Journal of Language Learning Technologies*, 23, 9-16.
- Bachman, L. (1990). *Fundamental considerations in language testing*. Oxford: Oxford University Press.
- Bachman, L., & Palmer, A. (1996). *Language testing in practice*. Oxford: Oxford University Press.
- Baltova, I. (1994). The impact of video on comprehension skills of core French students. *Canadian Modern Language Review*, 50, 507-531.
- 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.
- Berne, J. (1995). How does varying pre-listening activities affect second language listening comprehension? *Hispania*, 78, 316-329.
- Brett, P. (1997). A comparative study of the effects of the use of multimedia on listening comprehension. *System*, 25, 39-53.
- Brown, G. (1995). Dimensions of difficulty in listening comprehension. In D. Mendelshon & J. Rubin (Eds.), *A guide for the teaching of second language listening* (pp. 59-73). San Diego: Dominic Press.
- Buck, G. (2001). *Assessing Listening*. Cambridge: Cambridge University Press.
- Burgoon, J. (1994). Non-verbal signals. In M. Knapp & G. Miller (Eds.), *Handbook of interpersonal communication* (pp. 344-393). London: Routledge.
- Cervantes, R., & Gainer G. (1992). The effects of syntactic simplification and repetition on listening comprehension. *TESOL Quarterly*, 26, 767-770.
- Chiang, C., & Dunkel, P. (1992). The effect of speech modification, prior knowledge and listening proficiency on EFL lecture learning. *TESOL Quarterly*, 26, 345-374.
- Chang, A., & Read, J. (2006). The effects of listening support on the listening performance of EFL learners. *TESOL Quarterly*, 40, 375-397.

- Dunkel, P. (1991). Computerized testing of nonparticipatory L2 listening comprehension proficiency: An ESL prototype development effort. *Modern Language Journal*, 75, 64-73.
- Gruba, P. (1993). A comparison study of audio and video in language testing. *JALT Journal*, 15, 85-88.
- Gruba, P. (1994). Design and development of a video-mediated test of communicative proficiency. *JALT Journal*, 16, 25-40.
- Gruba, P. (1997). The role of video media in listening assessment. *System*, 25, 335-345.
- Hasan, A. (2000). Learners' perceptions of listening comprehension problems. *Language, Culture and Curriculum*, 13, 137-153.
- Kellerman, S. (1990). Lip service: The contribution of the visual modality to speech perception and its relevance to the teaching and testing of foreign language listening comprehension. *Applied Linguistics*, 11, 272-280.
- Kellerman, S. (1992). 'I see what you mean': The role of kinesic behaviour in listening and implications for foreign and second language learning. *Applied Linguistics*, 13, 239-258.
- Lynch, T. (1998). Theoretical perspective on listening. *Annual Review of Applied Linguistics*, 18, 3-19.
- MacWilliam, I. (1986). Video and language comprehension. *ELT Journal*, 40, 131-135.
- Nunan, D. (2005). From the special issues editor. *Language Learning & Technology*, 9(3), 2-3. Retrieved October 28, 2005 from <http://llt.msu.edu/vol9num3/pdf/speced.pdf>.
- Nissan, S., Devnicenzi, F., & Tang, K. (1996). *An analysis of factors affecting the difficulty of dialogue items in TOEFL listening comprehension* (TOEFL Research Report No. 51). Princeton, NJ: Educational Testing Service.
- Parry, T., & Meredith, R. (1984). Videotape vs. audiotape for listening comprehension tests: An experiment. (ERIC Document Reproduction Services ED 254 107).
- Progosh, D. (1996). Using video for listening assessment opinions of test-takers. *TESL Canada Journal*, 14, 34-43.
- Rost, M. (1990). *Listening in language learning*. New York: Longman.
- Samuels, S. (1984). Factors influencing listening: Inside and outside the head. *Theory into Practice*, 23, 183-189.
- Samuels, S. (1987). Factors that influence listening and reading comprehension. In R. Horowitz & S. Samuels (Eds.), *Comprehending oral and written language* (pp. 183-189). San Diego: Academic Press.
- Schmidt, R. (1990). The role of consciousness in second language learning. *Applied Linguistics*, 4, 179-191.
- Shin, D. (1998). Using video-taped lectures for testing academic language. *International Journal of Listening*, 12, 56-79.
- Thompson, I., & Rubin, J. (1996). Can strategy instruction improve listening comprehension? *Foreign Language Annals*, 29, 331-342.
- Tomlin, R., & Villa, V. (1994). Attention in cognitive science and second language acquisition. *Studies in Second Language Acquisition*, 16, 183-203.
- VanPatten, B. (1996). Can learners attend to form and content while processing input? *Hispania*, 72, 409-417.

von Raffler-Engel, W. (1980). Kinesics and paralinguistics: A neglected factor in second language research and teaching. *Canadian Modern Language Review*, 36, 225-237.

Wagner, E. (2002). Video listening tests: A pilot study. *Working Papers in TESOL & Applied Linguistics, Teachers College, Columbia University*, 2(1), Retrieved March 31, 2006, from <http://www.tc.edu/tesolalwebjournal>.

Wagner, M. (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.