Title of Project:
Video-Mediated Listening Passages and Typed Notetaking: Investigating Their Impact on Comprehension, Test Structure, and Item Performance

Researcher:
Justin Cubilo
University of Hawaii at Manoa
cubiloju@gmail.com

Research Supervisor:
James Dean Brown
University of Hawaii at Manoa
brownj@hawaii.edu

Project Summary

Motivation for the Research
Technology has created many implications for second language (L2) listening assessment, particularly as it relates to the role of visuals and typed note-taking. However, while previous studies have investigated the effects of visuals and typed note-taking on listening test performance, the results of these studies have been contradictory at best, with research indicating that visuals and typed note-taking both help (Kim, Turner, & Perez-Quinones, 2009; Wagner, 2010) and hinder (Muller & Oppenheimer, 2014; Suvorov, 2009) performance. Additionally, while many of these studies have tried to investigate the role of visuals and note-taking on overall test performance, few studies have attempted to investigate the effect that these variables have on item characteristics or on different comprehension subskills. Therefore, the present study attempts to address this gap in the research by investigating the effects of these variables by on item performance and examinee performance in different comprehension skill areas.

Research Questions
1. To what degree do participants’ performances on listening tests vary when presented with listening input (i.e., video-based versus audio-only material) and note-taking (i.e., handwritten versus typed) conditions? To what degree do these conditions interact with each other?
2. How do item characteristics differ between the video-based and audio-only conditions? How do they differ between handwritten and typed note-taking conditions?
3. What is the extent to which visual support and note-taking conditions influence examinee’s abilities to answer items testing them on different listening comprehension skills?
4. What perceptions and opinions do examinees have of the different conditions to which they are exposed and how do these perceptions and opinions and explanations shed light on the results from the previous questions?

Research Methodology

I employed a quantitative-dominant mixed methods design (Brown, 2014), which incorporated data from a test of academic listening developed for this study as well as qualitative data from a post-test questionnaire that participants were asked to complete. Two hundred participants were asked to take two forms of a listening comprehension task that was designed to measure an academic listening construct focused on comprehension of introductory level university course content. Each test form consisted of three lectures that targeted content from the humanities, social sciences, and natural sciences, and each lecture was followed by 10 multiple-choice questions. Each question set targeted four different subskills representative of the academic domain (identifying the main idea, identifying details, making inferences, and determining speaker attitude/bias). Participants were randomly placed into one of eight experimental groups in which one form of the test was presented in video-mediated format, which included listener access to lecturer gestures and PowerPoint slides, and the other was presented in an audio-only listening format. Likewise, one form required only handwritten note-taking while the other required only typed note-taking. The test was piloted prior to experimental use and was determined to have acceptable reliability, item facility, and item discrimination values, with only three of the thirty items showing need for revision. Examinees finished both forms and then proceeded to complete the post-test questionnaire, which asked them to state their preference for visual and note-taking conditions and to provide reasons for their preference. They also answered questions asking what they found themselves focusing on most in the videos.

Test data were subjected to an ANOVA to determine the overall effect of visual and note-taking on test scores while Rasch analysis and path analysis were used to investigate their effects on item performance and comprehension subskill performance, respectively. Survey data were analyzed inductively to determine codes and themes associated with each response unit and these themes and their associated responses were later used to further explain results obtained from the quantitative analysis.

Summary of Findings

The ANOVA results showed that there was no significant effect of video or typed note-taking on overall test performance. However, low statistical power indicates that a larger sample size may reveal different results in respect to this analysis. Rasch path analysis showed item performance and examinee comprehension subskill performance in different visual and note-taking conditions. The results of the Rasch analysis indicated that items were made slightly easier through the presence of video and typed note-taking. However, this effect did not reach levels of significance. The results of the path analysis indicated that note-taking conditions did not have significant predictive power for examinees’ performance in any of the listening comprehension subskills targeted by the items on the listening test. However, it was found that video had significant predictive power on questions that asked examinees to identify details and that performance on these detail-focused questions served as an intermediary variable through which video also had predictive power on examinees’ performance on questions asking them to…
understand main ideas, make inferences, and determine the speaker’s attitude. In all cases, this relationship was positive, indicating that the presence of video significantly predicted better performance on different comprehension subskills, most directly the subskill targeting the identification of details.

Qualitative data were also analyzed from the post-test survey using inductive methods to develop a list of thematic categories. Results from this analysis indicated that most participants preferred the video-mediated listening passages because they not only made it easier to understand the material, but also because the visuals from the PowerPoint slides and the lecturer’s gestures aided them in getting back on track if they got lost in the middle of the passage, and they made the listening seem more authentic. Those who said they preferred audio-only listening passages primarily gave the reason that the video distracted them and made it difficult to pay attention to the information delivered by the lecturer.

When asked about preferences regarding note-taking conditions, participants overwhelmingly preferred handwritten notes for the primary reason that handwriting made it possible to write notes in their native languages and to show connections between ideas by using circles and arrows. Those who stated that they preferred typing stated that speed and neatness of notes made it the best method for them. Regardless of the preference, most participants stated that they chose their preference based on what they were most comfortable and had the most experience with.

**Implications**

The findings in this study have a number of implications for the way in which the listening construct is defined. While previous studies have primarily looked at the effect that video-based input has on overall scores for a listening test (Ginther, 2002; Suvorov, 2009; Wagner, 2008), few, if any, have examined the effects it has on different comprehension subskill types. The fact that this study found that visuals significantly contribute to detail comprehension in some situations signals that the impact of visuals may be more refined than previously envisioned. If the listening construct is going to be defined in such a way as to make it so that listening passages are only associated with still pictures of the lecturer or of an object that the lecturer is talking about, then it is necessary to limit this type of test only to subjects that do not see any underlying affects from video-based input. Rather, what seems the safer option is to redefine listening tasks by including the video-based component. In this case, those subjects who do not experience effects from video will produce similar output regardless, while those topics that lead to benefits from its inclusion are allowed to produce these benefits. To remove video-based listening from topics that are affected by it only serves to threaten the construct validity of the test through construct under-representation.

Video-based input also led participants to state a number of other opinions related to the helpfulness of the visuals and the authenticity of the lecture. Even though the slides were constructed to actually contain less information than what one would generally see on slides in a lecture in the real-world domain, participants overwhelmingly stated that they were helpful. In addition, several participants were also clear in stating that they felt that the lectures were more realistic with the video and that it reminded them of attending a real lecture. These findings have several implications. The first of these is that it is clear from the comments, once again, that visuals do play a role in comprehension. Participants made note of both the information on the slides and the lecturer’s gestures, body language, and lip movements. This indicates, as previous
studies have also done (Sueyoshi & Hardison, 2005), that visual and aural processing are interconnected and that divorcing the two would lead, once again to construct underrepresentation within a listening task. Additionally, comments made by participants related to the visuals providing greater authenticity not only help to provide confirmation that the target language use (TLU) domain is being adequately represented, but also help to establish face validity of the test.

While results related to video-based input seem to provide the greatest number of implications related to the listening construct, note-taking conditions also had an implication worth examining. While the findings for his study indicated that there were no significant differences between scores across note-taking conditions, participants were quite opinionated regarding their actual preferences for note-taking. On the one hand, many participants preferred to take notes by hand, making it so that the current means by which most listening tests require learners to take notes is fair enough for most of the population. On the other hand, however, there were still those who were more accustomed to typing their notes and who stated that they preferred to take notes in this manner because they were better and faster at doing so. This result could call into question the way in which the academic domain is currently conceptualized within academic listening tasks. Even though there were no differences in performance, the fact that many participants stated that one method led to greater comfort than another suggests the need to address this issue. Tests already present themselves as a stress-inducing event. To deprive examinees of the comfort of doing something in a way they would do it in the TLU domain poses potential challenges to the current definition of the TLU domain and, by extension, the construct validity. It may be worth considering whether future test development will take this issue into account and how it can be addressed, though such initiatives may be difficult because they could potentially require that students have constant access to the ability to switch between language typefaces quickly during the test, which may be a difficult feat to accomplish given the number of L1 backgrounds test takers come from.
References


NVivo qualitative data analysis software (Version 10) [Computer software]. (2014). Doncaster, Australia: QSR International Pty Ltd.


