COMPREHENSION OF AN AUDIO VERSUS AN AUDIOVISUAL LECTURE AT 50% TIME-COMPRESSION

by

Nicole Perez

A Thesis Submitted to the Faculty of

Charles E. Schmidt College of Science

In Partial Fulfillment of the Requirements for the Degree of

Master of Arts

Florida Atlantic University

Boca Raton, FL

May 2017

COMPREHENSION OF AN AUDIO VERSUS AN AUDIOVISUAL LECTURE AT

50% TIME COMPRESSION

by

Nicole Perez

This thesis was prepared under the direction of the candidate's thesis advisor, Dr. Elan Barenholtz, Department of Psychology, and has been approved by the members of her supervisory committee. It was submitted to the faculty of the Charles E. Schmidt College of Science and was accepted in partial fulfillment of the requirements for the degree of Master of Arts.

SUPERVISORY COMMITTEE:

Elan Barenholtz, Ph.D.
Thesis Advisor

Sang Wook Hong, Ph.D.

Summer Sheremata, Ph.D.

Robert Stackman, Ph.D.
Interim Chair, Department of Psychology

Ata Sarajedini, Ph.D.
Dean, Charles E. Schmidt College of Science

Deborah L. Floyd, Ed.D.
Dean, Graduate College

April 14, 2017
Date

ACKNOWLEDGMENTS

Thanks to my advisor and committee for their feedback and recommendations. I want to thank the graduate students in the Visual Mind Lab, including Mike Kleiman and Lauren Mavica PhD for recording the videos used in this research. I also thank all the undergraduate students who ran participants for this study.

ABSTRACT

Author: Nicole Perez

Title: Comprehension of an audio versus an audiovisual lecture at 50% time-compression

Institution: Florida Atlantic University

Thesis Advisor: Dr. Elan Barenholtz

Degree: Master of Arts

Year: 2017

Since students can adjust the speed of online videos by time-compression which is available through common software (Pastore & Ritzhaupt, 2015), it is important to learn at which point compression impacts comprehension. The focus of the study is whether the speaker's face benefits comprehension during a 50% compressed lecture. Participants listened to a normal lecture or a 50% compressed lecture. Each participant saw an audio and audiovisual lecture, and were eye tracked during the audiovisual lecture. A comprehension test revealed that participants in the compressed lecture group performed better with the face. Eye fixations revealed that participants in the compressed lecture group looked less at the eyes and more at the nose when compared to eye fixations for those that viewed the normal lecture. This study demonstrates that 50% compression affects eye fixations and that the face benefits the listener, but this much compression will still lessen comprehension.

DEDICATION

To my mom and dad, for all their love and support.

LIST OF FIGURES

Figure 1. Average fixation proportion towards eyes, relative to the face......................... 15

Figure 2. Average fixation proportion towards mouth, relative to the face...................... 16

Figure 3. Average fixation proportion towards nose, relative to the face......................... 17

Figure 4. Average audiovisual comprehension scores: normal versus compressed lecture... 19

Figure 5. Average audio comprehension scores: normal versus compressed lecture....... 20

Figure 6. Normal lecture comprehension scores... 22

Figure 7. Compressed lecture comprehension scores. 23

INTRODUCTION

Online lectures may include multimedia presentations which can consist of audio, visual images, and slides (Homer, Plass, & Blake, 2008). The speed of the lectures' audio can be changed by the listener without affecting the pitch, and this is known as time-compression (Pastore & Ritzhaupt, 2015). Studies have shown that at 50% compression (two times the normal speed) or greater than 50% compression, comprehension of the material significantly worsens (Pastore 2012; Ritzhaupt & Barron 2008). If comprehension is affected at these speeds, is there anything that can benefit the learner, other than not compressing the lecture to these speeds? The face of the speaker present in the online video, provides a social cue and comprehension may be higher since the person can read the speaker's lips as well as hear his or her voice (Kizilcec, Bailenson, & Gomez, 2015). Therefore, this study examines if the face helps listeners during time-compressed lectures. This was accomplished by comparing scores on a test given after the participant views a compressed audiovisual lecture with the speaker's face or hears only the compressed audio. As participants view the audiovisual lecture of the speaker, whether or not they look more at the mouth was also addressed. Previous studies have shown that during a challenging learning condition, such as listening to a language people are unaccustomed to or hearing audio under noise, participants tend to look more towards the mouth (Barenholtz, Mavica, and Lewkowicz 2016; Vatikiotis-Bateson, Eigsti, Yano & Munhall, 1998).

To see if viewing the speaker's face allows for increased learning, Homer et al., (2008) had one group view a video in which they could hear and see the speaker and read the slides, and the other group could only hear the speaker and see the slides, and did not see the face of the speaker. Comprehension was tested with multiple choice questions. There was no difference in comprehension between the groups that saw the video with or without the face; however, participants who saw the face indicated a significantly higher amount of cognitive load. This was explained by participants' attention being split as they saw the video of the speaker and read the slides.

Split attention was also seen on tests given after a presentation in Mayer and Moreno's (1998) study. They presented participants with an animation about lightning. One group heard the narration and saw the animation, while the other group read text on the screen and saw the animation. Attention was split when the participants had to read text, and those that listened to the narration instead of reading performed better on the matching, retention, and transfer tests following the presentations. Overall, a presentation with figures is learned better with an audio explanation rather than text because there is split attention when participants need to look at figures and read. To avoid split attention, the focus of this study will be on comparing the audiovisual recording of the face of the speaker and the audio recording only; therefore, there will be no slides, text, or figures in this study.

Learning is enhanced when students are presented with the learning style they prefer according to Plass, Chun, Mayer, and Leutner (1998). The study presented English speakers with words in German. Visual or verbal cues or both were present. If a participant preferred visual cues, they performed better at remembering words with visual cues, and if the participant preferred verbal cues, they performed better at remembering words with verbal cues. Since participant preferences may also have an effect, the current study will be within subjects, and each participant will see both lecture styles (audiovisual and audio). Those lecture style presentations will be counterbalanced as explained later in the methods section.

To determine if students prefer the presence of a face in an online video, Kizilcec et al., (2015) gave students the option to view an online video with or without a speaker's face. After the first week of watching both forms of video, students would be able to make the decision as to which lecture they would view. 57% of students viewed the lecture with the face and 35% viewed the lecture without the face, which means that more students have a preference for viewing a lecture with a face. Kizilcec, Papadopoulos, and Sritanyaratana (2014) also had participants watch videos with or without the instructor's face, to determine how this affects recall. The face of the instructor was in the lower right-hand corner next to the slides. A survey revealed that 15 of the 22 participants indicated their preference of the instructor's face in the lecture. In the lectures that had the speaker's face, eye tracking revealed that students looked at the face an average of 41% of the time, but participants still had no difference in recall right after watching the lecture with or without the face. Even though participants wanted to watch a video with a face, it was not necessarily helping them learn.

Time-compression

When a student sped up audio in the past, both tempo and pitch used to be affected, but this is no longer the case since pitch can be kept the same as the tempo changes (Pastore & Ritzhaupt, 2015). To examine whether performance is affected by compressed video, Ritzhaupt, Pastore, and Davis (2015) presented a video at normal speed, 1.25 times the normal speed, and 1.5 times the normal speed. No significant differences were found for the performance of the participants based on video speed. Ritzhaupt, Gomes, and Barron (2008) showed participants multimedia presentations at normal speed, 1.4 times the normal speed, and 1.8 times the normal speed, and speed did not significantly affect performance either.

At what point does compression affect comprehension? It seems that even up to 1.8 times the normal speed, participants' performance does not get affected. Ritzhaupt and Barron (2008) presented a narration at normal speed, 1.5 times, 2.0 times, and 2.5 times the normal speed. Recall and recognition was affected significantly at 2.5 times the normal speed. Pastore (2012) presented participants with either normal speed audio, 25% compressed audio, or 50% compressed audio each with figures of heart anatomy. An identification test and a test with 20 multiple choice questions and transfer questions were given after listening to the audio. There was no significant difference in scores between the normal audio group and the 25% compression group, and both groups had significantly higher scores when compared to those participants that listened to the 50% compressed audio, even though participants could replay the presentation. Learning was found to be negatively affected if audio is compressed by 50% or higher.

Pastore (2010) presented a normal lecture about the heart, a lecture at 25% compression, or a lecture at 50% compression. In addition to speed, the other independent variable was presence or absence of diagrams of the heart. Participants had the option to replay the lecture. Comprehension was measured with 20 multiple-choice problem solving questions. Those participants who listened to the 25% compressed lecture with the diagrams did just as well on the comprehension test as those who listened to the normal lecture. This did not occur with the 50% compressed lecture, though.

Eye tracking

Using eye tracking it is possible to determine which areas of the face a participant directs his or her gaze. Barenholtz et al., (2016) used this technique to demonstrate that when participants who only speak English are presented with a language they are not familiar with (Icelandic or Spanish) they fixate more at the mouth than the eyes. Sentences were presented in each language, and after the video presentation of those sentences, a third sentence was presented in audio format and participants had to match the audio sentence to one of the two videos in both the familiar and unfamiliar language. When the participants spoke both English and Spanish, no difference was seen in fixations towards the mouth and eyes. This demonstrates that when someone is familiar with a language, that person will not look at the mouth as much as someone who does not speak the language. However, these tests had the matching sentence task. Without this task, English speakers did not show a difference between looking at the eyes and mouth in their native language or in Icelandic. Therefore, when speech needs to be processed, like when being asked to match which was the sentence heard in the unfamiliar language, participants will fixate more at the mouth.

Presenting recordings in an unfamiliar language is a challenging listening task and another challenging listening condition that can prompt people to look more at the mouth of the speaker is presenting audio under noise, which is what Vatikiotis-Bateson et al., (1998) did. When there was an increased amount of noise, the fixations of the participants moved from the eyes to the mouth. The time looking at the eyes went from over 65% with no noise, down to 45% with noise, and mouth fixations went from 35% with no noise or in a low noise condition to 55% under high noise. Buchan, Pare, and Munhall (2008) presented audiovisual recordings under noise as well. With noise, fixations on the eyes were not as long, and longer fixations were observed for the mouth and nose. Fixations on the nose were explained by the reasoning that these fixations might still allow participants to see the eyes and is a central location for fixation. Buchan, Pare, and Munhall (2007) also found that when sentences are presented under noise, eye fixations went towards the middle of the face, the nose. With the sentences being heard under noise, fixations to the eyes went down, mouth fixations did not change significantly, and fixations to the nose increased.

To test if audiovisual recordings offer a benefit, and where on the face the participants would look at, Lusk and Mitchell (2016) presented audiovisual and audio recordings of made up words to participants. Fixations on the eyes and mouth were longer than for the nose. As participants familiarized themselves with the words, fixations to the mouth were less than they were in the beginning. Fixations on the eyes and nose did not change. Since mouth fixations decreased with increased familiarity with words, this related to Barenholtz's et al., (2016) study, because with a familiar language, participants will not look at the mouth more whereas with an unfamiliar language and the need to process that language they will gaze at the mouth more.

Audio versus audiovisual

Audiovisual recordings seem to offer a benefit to learning, and to see if this benefit is present under accented speech and noise, Banks, Gowen, Munro, and Adank

If a student is viewing a lecture at 50% compressed speed and can also view the speaker's face, I expect that scores will be higher than for students who just listen to an audio lecture compressed by 50%, since an audiovisual presentation of the speaker allows for increased learning in challenging learning conditions (Banks et al., 2015; Sueyoshi and Harding 2005). For the lectures presented at normal speed, I expect that there will be no significant difference between audiovisual and audio lecture comprehension scores, since this has been the finding for normal multimedia lectures (Kizilcec et al., 2014; Homer et al., 2008).

Additionally, when someone is trying to comprehend what is being said under a difficult condition such as a new language (Barenholtz et al., 2016) or under noise (Vatikiotis-Bateson et al., 1998) they look more at the mouth. Based on these findings, I hypothesize that the participants will also look more at the mouth of the speaker than at the eyes at 50% compression.

METHODS

Participants

160 undergraduate students from Florida Atlantic University participated in the study to hear one lecture on monosodium glutamate (MSG) in food and another lecture on breakfast and grades. Participants would hear both lectures, and one would be audiovisual, with the face of the speaker, and the other would be just audio. The order of the recordings presented was counterbalanced. After adjusting for the minimum comprehension score required to be included in the analysis, 34 participants were in the audio MSG group, 38 in the audiovisual MSG group, 37 in the audio breakfast group, and 34 in the audiovisual breakfast group. The following is an example: 34 participants in the audio MSG group would listen to the audio MSG lecture first and the audiovisual breakfast lecture second, and the 34 in the audiovisual breakfast group would listen to the audiovisual breakfast lecture first and the audio MSG lecture second. Half of the participants in each group would listen to both lectures at normal speed and the other half would listen to both lectures at 50% time-compression. 98 (68.5%) of the participants were female and 45 (31.5%) were male. 123 (86%) of the participants were between the ages of 18-21, 12 (8.4%) were between the ages of 22-25, 7 (4.9%) were between the ages of 26-35, and 1 (0.7%) was between the ages of 36-50.

Materials

Two short lectures were recorded about Breakfast and Grades and about MSG in food. The Breakfast and Grades lecture was 3:05 minutes and the MSG lecture was 3:43

minutes. One of the recordings was an audiovisual recording in which the participant could see the face of the presenter and hear her speak and the other recording was just an audio recording of the presenter's voice. The recordings were time-compressed using a program called Adobe Premier and the time-compressed Breakfast and Grades lecture was shortened down to 1:32 minutes and the MSG lecture was shortened down to 1:51 minutes.

Procedures

Before listening to the recordings, participants were given five short answer questions to test their knowledge about what they knew about breakfast and grades and about MSG. For the audiovisual lecture, the participants' eye movements will be tracked using a Tobii T60 eye tracker. Fixations to the face, eyes, mouth, and nose are areas of interest (AOIs). The total fixations relative to the face will be considered, therefore proportions of the eye fixations to the eyes, mouth, and nose relative to the face are the focus. For the comprehension test, twelve short answer questions, one multiple choice question, and one fill in the blank questions were developed for these lectures, for a total of fourteen questions.

After taking the comprehension tests, participants took a survey which asked them to circle their learning style, which test they expected their highest score, and what lecture they preferred. To determine experience with online lectures and whether many participants have experience with time-compressed lectures they were asked to answer yes or no to the following questions: if they have taken online classes and if they speed up lectures online.

Test Scores

Two graders, who did not know the condition that the participants were in, scored the comprehension responses. There were 14 questions, but a total of 22 points since some questions had more than one answer and points were awarded for those answers as well. If the scores of the two graders did not match, another grader scored the responses. If the third score matched one of the first two scores, then that was the score that was used for data analysis. If all three scores were different, the average of the three scores was taken.

Of the fourteen questions that were on both comprehension tests, one was multiple choice and could have been guessed correctly by chance. Another question on each test was hinted about in the test about their prior knowledge of the subjects. If students paid attention to the test of prior knowledge and to the lectures, these questions should have been answered correctly. Therefore, only students that scored higher than 2 points were included in the analysis.

Data analysis: Eye tracking

The proportion of the total fixation duration towards the eyes, mouth, and nose in relation to the face was calculated. Three independent samples t-tests were performed to determine whether eye fixations differ when a lecture is time-compressed.

Data analysis: Comprehension scores

An independent samples t-test was conducted to determine whether comprehension scores differ between normal and time-compressed lectures. A paired samples t-test would determine if a video of a time-compressed lecture would help participants achieve higher comprehension scores when compared to just an audio time-compressed lecture.

RESULTS

Survey responses:

A survey was given for participants to choose their learning style, and 43 (30.1%) of the participants indicated that they are visual learners, 8 (5.6%) of the participants indicated that they are auditory learners, 88 (61.5%) of the participants indicated they are a combination of these learning styles, and 3 (2.1%) indicated that they were none of these learning styles. Table 1 shows the learning style of participants separated by condition, either compressed or normal.

Table 1. *Learning style for participants in compressed and normal conditions*

Learning Style Compressed Condition			Learning Style Normal Condition	
	Frequency	Percent	Frequency	Percent
Visual	21	31.8	22	28.6
Auditory	4	6.1	4	5.2
Combination	39	59.1	49	63.6
None	1	1.5	2	2.6
Total	65	98.5	77	100

The next survey question asked participants on which test they expected their highest score to verify that one lecture would not be perceived as easier. There were 79 (55.2%) of the participants who indicated that their highest score would be on the MSG lecture and 64 (44.8%) participants indicated that their highest score would be on the Breakfast lecture. Table 2 shows the exact number of participants based on condition. Participants were also asked to indicate which lecture they preferred, the video lecture,

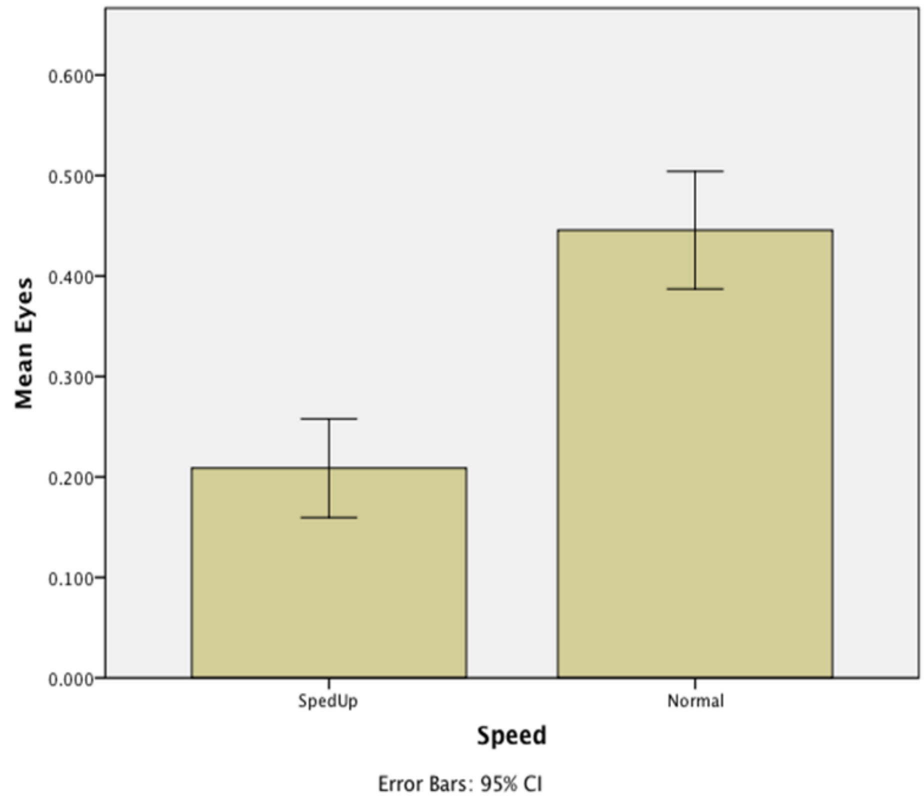

Figure 1. Average fixation proportion towards eyes, relative to the face

During the sped up (compressed) lecture, participants gazed significantly more towards the eyes when compared to participants who saw the normal lecture.

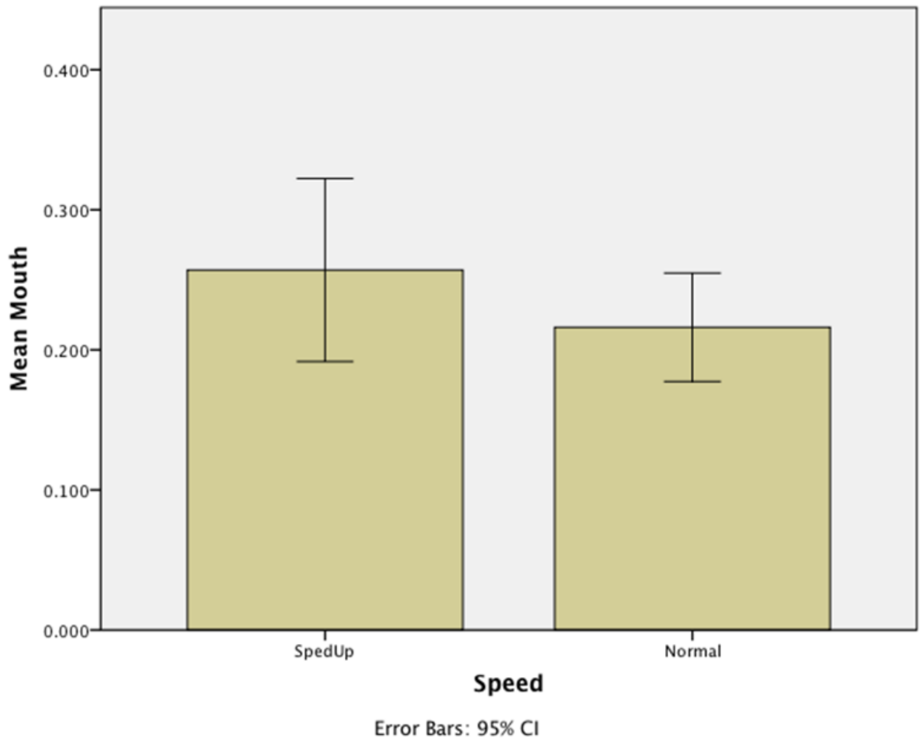

Figure 2. Average fixation proportion towards mouth, relative to the face

There were no significant differences in mouth fixations based on speed of the lecture.

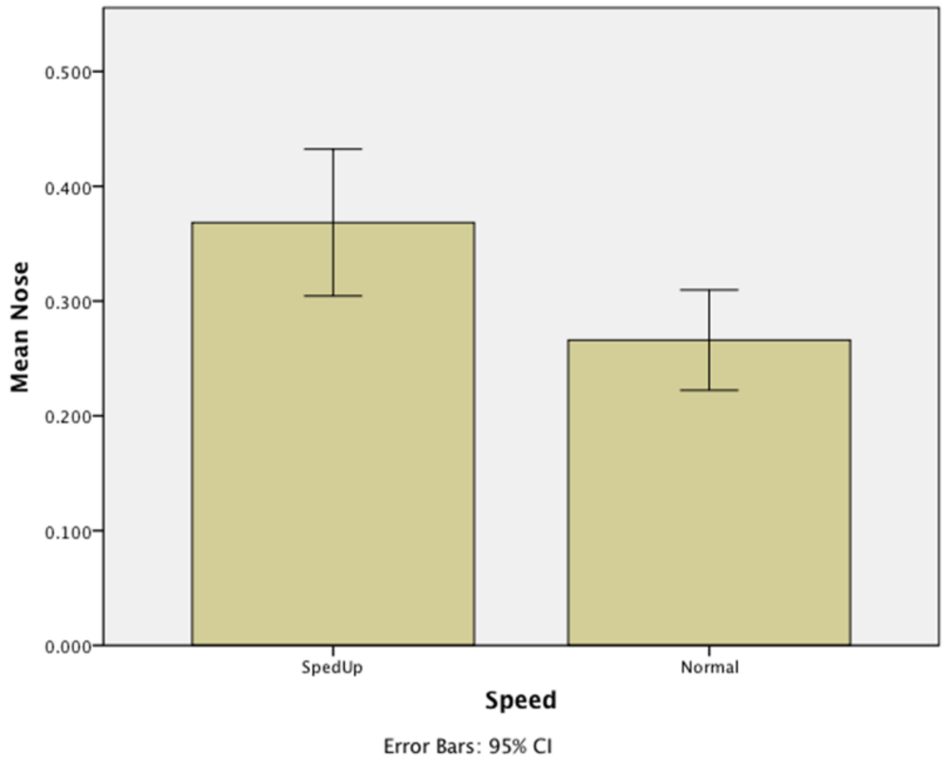

Figure 3. Average fixation proportion towards nose, relative to the face

During the sped up (compressed) lecture, participants gazed significantly more towards the nose when compared to participants who saw the normal lecture.

Comprehension Scores

For the comprehension score analysis, 2 participants were also excluded because of their detailed answers to the prior knowledge questions about MSG. There were 66 participants that met the requirement of having over 2 points in the time compressed lecture condition and there were 77 participants that met this requirement in the normal lecture. The outliers from the eye tracking data were included in the analysis of comprehension scores since these participants still received over 2 points, and were assumed to be paying attention to the recordings. Therefore, a total of 143 participants were included in the data analysis.

An independent samples t-test revealed that the time-compressed condition caused significantly lower scores in the time-compressed video lecture (t (141) = -2.808, p=0.006) and in the time-compressed audio lecture (t (141) = -3.206, p=0.002) when compared to the normal video lecture and normal audio lecture, respectively (Figure 4 and Figure 5).

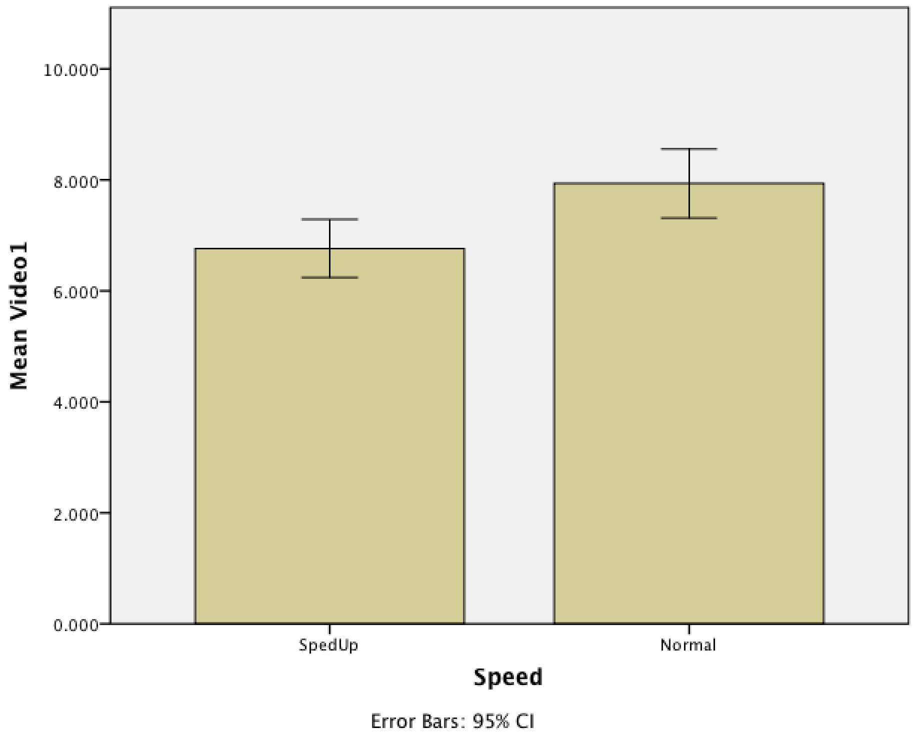

Figure 4. Average audiovisual comprehension scores: normal versus compressed lecture

This bar graph depicts significantly higher scores which were observed for the normal video (audiovisual) lecture group when compared to the sped up (compressed) lecture group.

Table 5. *Average comprehension scores for the normal lectures*

	N	Mean	Std. Deviation	Std. Error Mean
Audiovisual Normal Lecture	77	7.93939	2.75019	0.313413
Audio Normal Lecture	77	7.47618	2.797988	0.31886

Table 6. *Average comprehension scores for the compressed lectures*

	N	Mean	Std. Deviation	Std. Error Mean
Audiovisual Compressed Lecture	66	6.76765	2.141169	0.26356
Audio Compressed Lecture	66	6.11615	2.171684	0.267316

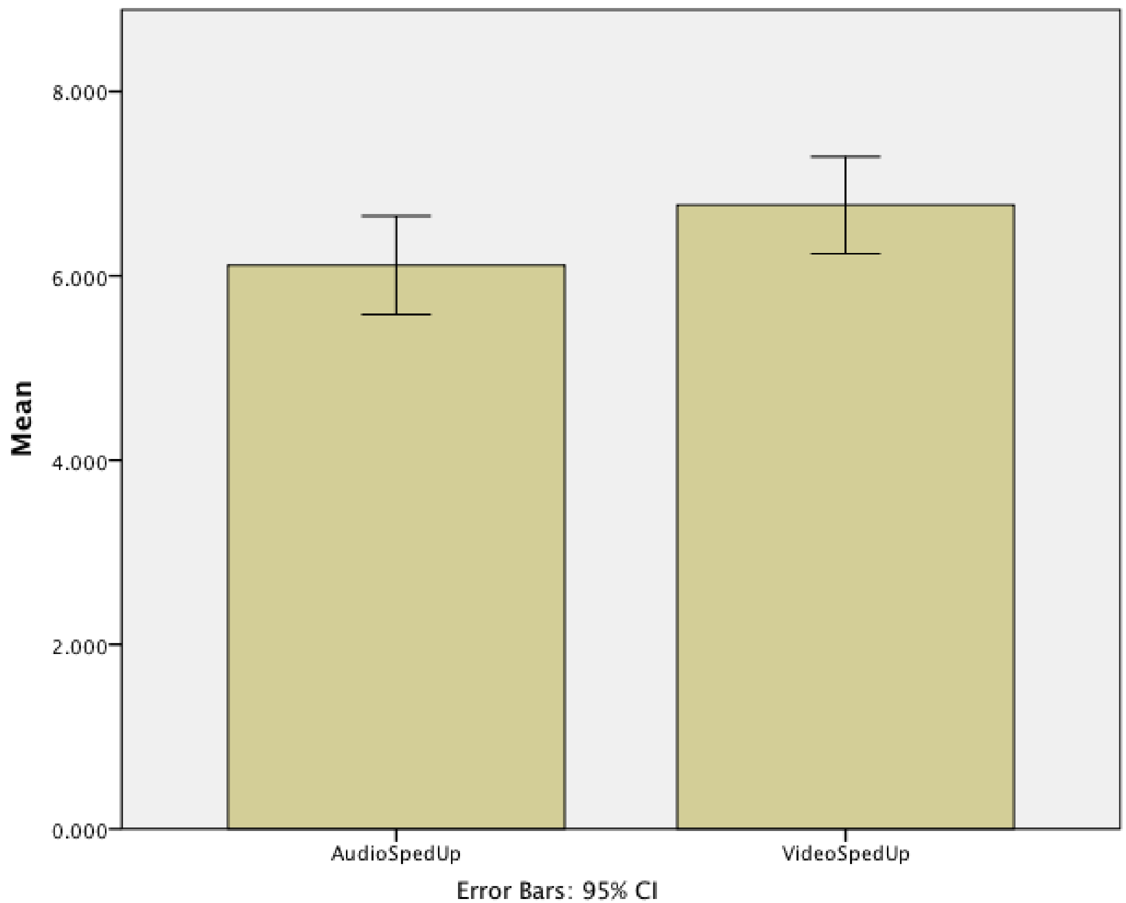

Figure 7. Compressed lecture comprehension scores.

This bar graph depicts that the video (audiovisual) lecture led to significantly higher scores when compared to an audio sped up (compressed) lecture.

DISCUSSION

This study was done to determine if there is a benefit from watching an audiovisual compressed lecture as opposed to an audio compressed lecture. Compressing online lectures can be easily done by the student and if a student chooses to compress a lecture, they might save time, but affect comprehension. The survey questions about online learning and learning preferences were asked to determine how many people in the sample have taken at least one online class and to see if speeding up lectures was common practice among people in the sample. Learning preferences were of interest since previous studies have shown that there is a preference for the face. Therefore, the participants were asked if they preferred the audiovisual lecture (with the face) or the audio lecture (without the face). Over both normal and compressed conditions, 81.1% of participants indicated that they preferred the audiovisual lecture. This is consistent with previous studies in which most participants prefer lectures with the speaker's face (Kizilcec et al., 2015; Kizilcec et al., 2014).

However, is the speaker's face benefitting comprehension? This seems to be the case for the compressed condition and not for the normal condition, which was expected. The compressed audiovisual lecture allowed participants to get higher comprehension scores than the compressed audio lecture. This effect was significant only when including participants who got higher than 2 points on the comprehension test. If participants got less than 2 points, it is assumed that they were not paying attention to the recordings. With the exclusion of these participants, it was clear that the presence of the face was providing some benefit to comprehension. This is consistent with studies that have presented challenging learning conditions like audio under noise and/or accents and have found that the face of the speaker helps the recognition of speech (Banks et al., 2015; Sueyoshi & Hardison, 2005). A 50% compressed lecture appears to be a challenging listening condition because comprehension was significantly lower in both compressed audio and audiovisual lectures when compared to the normal audio and audiovisual lectures which is similar to Pastore (2012). When the normal audio and normal audiovisual lecture comprehension scores were compared to each other, no significant differences were observed, unlike the compressed lecture. This is related to Kizlilcec et al., (2014) and Homer et al., (2008) which did not manipulate speed and presented lectures with or without the face, yet there was no difference between comprehension or recall scores. The present study was not a multimedia presentation because the focus was solely on the face of the speaker, though.

Eye tracking data did not prove the hypothesis that a compressed lecture would cause participants to look more towards the mouth. Interestingly, participants looked significantly less at the eyes and more at the nose in the compressed lecture which is consistent with Buchan's et al., (2008) and Buchan's et al., (2007) findings. Participants in the normal condition did not gaze more towards the nose, and there were higher proportions of fixations towards the eyes. Therefore, in addition to significantly lower comprehension scores, the compressed lecture caused eye fixations to change as well. They shifted towards the fixation trend seen in previous studies that presented recordings under noise. Future research about time compression and eye tracking could examine more speeds and at which point participants begin to fixate more at the nose. This was observed at 50% compression, but how about 25% compression and other speeds? Although, the presentation of the face helped participants get higher comprehension scores in the 50% compressed condition, this amount of compression still caused significantly lower scores than listening to the recordings at normal speed. Refraining from compressing lectures to 2 times the normal speed would benefit the learner because even with the face of the speaker, comprehension at this speed would still be affected.

This study also considered how many people in the sample of participants have ever sped up a lecture. The words sped up were used as opposed to time-compressed, just in case participants were not familiar with the term. Based on the survey responses, only 19.6% of participants indicated that they speed up lectures online, but 75.5% of participants indicated that they have taken an online class. With so many participants taking at least one online class, it would be interesting to see how prevalent speeding up lectures might become.

REFERENCES

Banks, B., Gowen, E., Munro, K.J., & Adank, P. (2015). Audiovisual cues benefit recognition of accented speech in noise but not perceptual adaptation. *Frontiers in Human Neuroscience*, *9*, 422. doi: 10.3389.fnhum.2015.00422.

Barenholtz, E., Mavica, L., & Lewkowicz, D.J. (2016). Language familiarity modulates relative attention to the eyes and mouth of a talker. *Cognition, 147*, 100-105. doi: 10.1016/j.cognition.2015.11.013.

Buchan, J.N., Pare, M., & Munhall, K.G. (2008). The effect of varying talker identity and listening conditions on gaze behavior during audiovisual speech perception. *Brain Research, 1242*, 162-171. doi: 10.1016/j.brainres.2008.06.083.

Buchan, J.N., Pare, M., & Munhall, K.G. (2007). Spatial statistics of gaze fixations during dynamic face processing. *Social Neuroscience, 2,* 1-13. doi: 10.1080/17470910601043644.

Homer, B.D., Plass, J.L., & Blake, L. (2008). The effects of video on cognitive load and social presence in multimedia- learning *Computers in Human Behavior*, *24*, 786-797. doi:10.1016/j.chb.2007.02.009.

Kizilcec, R.F., Bailenson, J.N., & Gomez, C.J. (2015). The instructor's face in video instruction: Evidence from two large-scale field studies. *Journal of Educational Psychology*, *107*,724-739. doi: 10.1037/edu0000013.

Kizilcec, R.F., Papadopoulos, K., & Sritanyaratana, L. (2014). Showing face in video instruction: Effects on information retention, visual attention, and affect.

Proceedings of SIGCHI Conference on Human Factors in Computing Systems, 2095-2102. doi: 10.1145/2556288.2557207.

Lusk, L.G., & Mitchell, A.D. (2016). Differential gaze patterns on eyes and mouth during audiovisual speech segmentation. *Frontiers in Psychology, 7,* 52. doi: 10.3389/fpsyg.2016.00052.

Mayer, R.E., & Moreno, R. (1998). A split-attention effect in multimedia learning: Evidence for dual processing systems in working memory. *Journal of Educational Psychology, 90,* 312-320. doi: 10.1037//0022-0663.90.2.312.

Pastore, R., & Ritzhaupt, A.D. (2015). Using time-compression to make multimedia learning more efficient: Current research and practice. *TechTrends, 59,* 66-74. doi: 10.1007/s11528-015-0841-2.

Pastore, R. (2012). The effects of time-compressed instruction and redundancy on learning and learners' perceptions of cognitive load. *Computers and Education, 58,* 641-651. doi: 10.1016/j.compedu..2011.09.018.

Pastore, R. (2010). The effects of diagrams and time-compressed instruction on learning and learner's perceptions of cognitive load. *Educational Technology Research and Development, 58,* 485-505. doi: 10.1007/s11423-009-9145-8.

Plass, J.L., Chun, D.M., Mayer, R.E., & Leutner, D. (1998). Supporting visual and verbal learning preferences in a second-language multimedia learning environment. *Journal of Educational Psychology, 90,* 25-36. doi: 10.1037//0022 0663.90.1.25.

Ritzhaupt, A.D., Pastore, R., & Davis, R. (2015). Effects of captions and time compressed video on learner performance and satisfaction. *Computers in Human Behavior, 45,* 222-227. doi: 10.1016/j.chb.2014.12.020.

Ritzhaupt, A.D., & Barron, A.E. (2008). Effects of time-compressed narration and representational adjunct images on cued-recall, content recognition, and learner satisfaction. *Journal of Educational Computing Research, 39,* 161-184. doi: 10.2190/EC.39.2.d.

Ritzhaupt, A.D., Gomes, N.D., & Barron, A.E. (2008). The effects of time-compressed audio and verbal redundancy on learner performance and satisfaction. *Computers in Human Behavior, 24,* 2434-2445. doi: 10.1016/j.chb.2008.02.017.

Sueyoshi, A., & Hardison, D.M. (2005). The role of gestures and facial cues in second language listening comprehension language learning. *Language Learning, 55,* 661-699. doi: 10.1111/j.0023-8333.2005.00320.x.

Vatikiotis-Bateson, E., Eigsti, I.M., Yano, S., & Munhall, K.G. (1998). Eye movement of perceivers during audiovisual speech perception. *Perception and Psychophysics, 60,* 926-940. doi: 10.3758/BF03211929.