Effects of Audio Dilation and Listening Skill Ability for English as a Second Language

BY

Daniel Bunn
B.S., Calumet College of St. Joseph, 2015

THESIS

Submitted as partial fulfillment of the requirements
for the degree of Master of Science in Computer Science
in the Graduate College of the
University of Illinois at Chicago, 2018

Chicago, Illinois

Defense Committee:

Robert Kenyon, Chair and Advisor
Steven Jones, Communications
Ugo Buy

TABLE OF CONTENTS

Dedicated in loving memory to my friend, Greg. You have left an indelible mark on my life, and all I could ever hope to do is pay forward your kindness to honor your memory and legacy. Any amount of measurable effort to provide service to another person, in any capacity, is nothing more than a thank you for the time you took to help me.

Stay beautiful.

ACKNOWLEDGMENTS

I would like to thank my committee members, professors Robert Kenyon, Steven Jones, and Ugo Buy for their time and effort in helping me create and finalize this work. It would not have been possible without their help. I would also like to thank John Novak, whose work, as well as guidance, played an integral role in this thesis.

To my family, for putting up with me.

To the international students who took the time to participate in this experiment, thank you. I cannot possibly fully express my respect and amazement at your intellect and bravery to work so hard and, literally, go so far to pursue an education.

Ruggero, Sina, Yui, Vinnie, Vinit and Nick: you may have provided me with endless amounts of technical advice - but most importantly - it was your friendship that made this chapter of my life bearable.

DWB

“TWICE…two time”

SUMMARY

This project examines the effects of an audio dilation technique to slow the speech rate of audio for English as a Second language individuals, using listening skill assessment tests from the Test of English as a Foreign Language practice exam textbook. The goal of this research experiment is to measure the effect of audio dilation on subject performance on listening skill tasks.

These listening skill tests were loaded in to software designed to give the user the ability to dilate – or slow – the audio at the user's preference, to measure if this dilation treatment ability would result in an increased ability to comprehend the listening skill tests being administered, and subsequently result in an increased listening skill test score. These scores would be compared to listening skill tests delivered at unaltered, or natural rate, of speed.

The experiments were carried out A.) to see if subjects choose to dilate when given the opportunity, and B.) To see if there is a difference in test performance between dilation and natural rate listening test scores. After the experiment, test subjects were presented a Likert test questionnaire, as well as interviewed on their thoughts and opinions about the experiment.

The majority of test subjects used the dilation technique when given the opportunity, as well as rated their experience positively in both the questionnaire and post-experiment interview. The results indicate that, for our population that slowed their audio during the listening skill tests, difference in score was statistically insignificant.

INTRODUCTION

The ability to actively listen for meaning and context are critical aspects in linguistics, considered to be a cornerstone in language acquisition (Meskill, 1996). In the acquisition of a second language – in the case of this research, the English language, the listening skill is considered to be the most important aspect in helping an English as a Second Language (ESL) individual in learning the English language (Feyten, 1991). This is also coupled with the belief that the listening skill is the most difficult skill to learn of the four language skills (reading, writing, and speaking being the other three) (Vandergrift, 2004). A common metric to measure a person's English skill level is through their Test of English as a Foreign Language (TOEFL) score.

TOEFL is a standardized test created by the Educational Testing Services (ETS) group to measure the English language ability of non-native speakers, most commonly used by individuals wishing to enroll in English-speaking universities. The common testing method for TOEFL exams is through the internet, via the TOEFL internet based test (iBT), which is administered in over 50 countries, according to the ETS website. According to U.S Weekly News, the average Test of English as a Foreign Language score for a student to enter a university in the United States is 78. To enter the University of Illinois at Chicago (UIC), the score is 80.

This TOEFL requirement means that in many instances, regardless of intellectual or academic prowess in a field of study, a student's ability to participate at any given university could hinge directly on their English language skills, specifically measured through the TOEFL scoring system. Therefore, the desire to have a TOEFL score meeting or exceeding the scores required by any given university would be highly sought after by the prospective student. Indeed, while tutoring students at UIC and teaching English abroad in South East Asia, stories were relayed by ESL individuals about their experiences with preparing for, and taking, the TOEFL exam with emphasis on the perceived importance of passing this exam, or

scoring high enough, to allow them to attend not just a specific university they may desire, but in some cases *any* university at all in the United States.

According to the University of Illinois at Chicago's Office of Diversity, UIC is one of the most diverse universities in the nation. This diversity includes their international population, which currently accounts for just over 15% of all enrolled students, according to UIC's Office of International Services. It was because of this diverse population that initial exposure to this language acquisition problem was identified. While tutoring international students in English, a common request was made: "please slow down" rate of speech. Honoring this request and incorporating it into my tutoring seemed to be effective, from a strictly anecdotal perspective, but later became something worth exploring, scientifically.

Coincidentally, while attending a presentation by UIC faculty who introduced their research work, looking for students to join them, professor Robert Kenyon introduced the audience to PhD candidate John Novak's audio dilation technique (Novak et al, 2013) which was being used for a series of speech-in-noise experiments. The initial parameters of the experiment were laid out during the fall 2015 semester, where research was conducted through interviews with individuals from the School of Psychology, the Tutorium in Intensive English (TIE) program, a linguistics expert, and others in the field of computer science at UIC.

It was during this time that the notion that the listening skill, a critical aspect of learning, particularly in language acquisition, could benefit from the creation of specialized software that would dynamically dilate audio as it is being delivered, in real time, to an individual. The premise is rather simple: if slowing down speech rate (such as the common practice of using clear speech, when giving directions) can help someone to comprehend what they are hearing, then allowing someone to self-control a device to allow them to slow the rate of speech of a second language, where they may be in a constant state of desiring accentuated comprehension assistance, could be of great benefit in their ability to adequately

take in audio information while giving them additional time to process that information, resulting in better performance of the task related to that audio input.

The focus of this research was to implement an audio dilation technique developed by John Novak at the University of Illinois at Chicago (Novak et al, 2013), which time-stretches, slowing down audio tempo in real-time, as a tool for an ESL individual to use in order to help mitigate listening skill deficiency with a self-selection functionality that can be adjusted accordingly. The purpose of this research was to examine the effects of audio dilation on comprehension in listening skill ability of ESL individuals compared to natural rate speech, with TOEFL listening skills assessment tests as the metric.

The history of speech rate, whether it be speeding up speech rate or the slowing of speech rate, and the effect it has on non-native speakers (NNS) and comprehension has a rich history that is ever-evolving with technology. Research related to this specific experiment dates back to the late 1960's/early 1970's, where Foulke, Sticht, and Carver began trying to calculate words per minute (wpm) in relation to comprehension. According to Nichols and Stevens (as cited by Foulke and Sticht, 1969), the average words per minute in conversational speech was 125. Through increasing speech rate, it was discovered that listening skill comprehension declined after exposing the listener to speech rate at over 260 to 275 wpm (Foulke, 1968) (Foulke and Sticht, 1969) (Carver, 1973).

On the subject of slowing audio, Friedman and Johnson (1971) proposed the use of slower speech rates in language teaching by introducing a technique of inserting pauses in to recorded audio samples in order to facilitate dilation (this being their best option to facilitate slowing speech rate, due to technology constraints at the time). This concept was in response to their experiment results showing that increasing speech rate had no significant effect on Russian and Vietnamese L2 students. By the end of the 1970's, the concept of slowing speech rate iterated to attempting to match listener proficiency and optimal speech rates via dilation and compression (Pimsleur et al, 1977) with insertion of pauses in Spanish audio

clips. Pimsleur suggested that the average rate of speech for comprehension falls between 160 to 190 words per minute.

(Boyle, 1984) further elaborated on the adverse effects of fast speech rate relating to listening comprehension in a series of surveys of 30 Chinese teachers and 60 Chinese students, asking their opinions on native speakers speech rates and teaching methods related to speech rates. It was after analyzing these surveys that he postulated slower speech rates would be optimal in comprehension. In the 1990's, Zhao's study (1997) was instrumental in providing context for this experiment. Test subjects were presented with audio samples that were pre-treated by Zhao at various rates (again, this method was deemed the best at the time due to technology constraints), with the test subjects able to pick from the rates they preferred. Zhao found that using slowed speech rates resulted in improved comprehension. It should be noted that a similar experiment, conducted by Blau several years prior to Zhao, also using pre-modified speech rates, showed no statistical significance (Blau, 1990).

More recent experiments in the effects of slower speech rates include Khatib and Khodabaksh (2010), who tested a population of Iranian students on American television shows Oprah and Dr. Phil with modified speech rates. The results of their experiment did not show evidence of improved comprehension. The contradicting results in experiments, particularly in terms of finding optimal speech rates, may suggest that the speech rate should not be a static variable, and instead be able to be freely controlled by the user.

Numerous works have been insightful in not only the immediate subject at hand of ESL and the implementation of techniques to bolster the listening skill, but there has also been a wealth of knowledge exposed in the linguistics and psychological aspects of learning a second language, both in the mechanics of listening and the derivation of context and meaning involved in listening (Rivers 1981). The importance of the listening skill is described best when given the context of its use in the daily life of an individual,

where it is measured to be used nearly twice as much as speaking and a staggering four to five times as much as reading and writing (Morley, 1999).

Another aspect important to this research is the concept of pausology (Griffiths 1991), the silences added to communication, whether unintentional or for affect, and the role it can play in ESL listening skill. The pause, hesitation and the resulting affected speech rate may make comprehension difficult for someone unfamiliar with the intricacies involved in speaking the English language; coupled with the differences in what one demographic considers a "normal" speech rate compared to another (Zhao 1997) have shown to be important factors when considering protocols for experimentation within the realm of audio dilation related to listening skill tasks.

Technology has opened up a wide range of software solutions aiming at assisting acquisition of a second language. From popular software suites such as Rosetta Stone (Nielson 2011), the myriad language learning apps for mobile devices, to more nuanced applications like speech recognition software to aid in pronunciation (Kim 2006), as well as audio manipulating software, or features built within audio software, like VLC Media Player, or Windows Media Player. The approach of this work is to suggest that technology has reached the point to allow a more comprehensive self-selection rate for the user, which would allow an individual to adjust the rate of speech that would be optimal for their ability to comprehend any given speech rate they may encounter.

With technology continuing to evolve, it is not out of the realm of possibility that this work could be used in conjunction with the next evolution of listening aids, such as concepts similar to language-translating earbuds currently available, where one could have easy self-selection of preferred speech rate of a physical earpiece through a user interface on their smartphone. This could be useful for those interested in language acquisition -instead of outright translation-, but still seeking the assistance of technology to aid in language acquisition.

METHODS

Use of Test of English as a Foreign Language (TOEFL), a common standardized testing method to determine an individual's skill as an English language user, was agreed to be the primary metric used to reach a target demographic for testing. TOEFL, developed by the Education Testing Service (ETS) non-profit organization, is used by most United States universities in their selection process of international students. The TOEFL internet based test (iBT), is the most commonly used testing method to generate TOEFL scores, using four categories: listening, speaking, writing and reading, scored from 0 to 30, to create a composite score for the individual, for a maximum score of 120. In the case of University of Illinois at Chicago, according to the admissions office, a TOEFL iBT score of 80 is the minimum score required for a prospective student (these scores may be higher for other colleges or programs within UIC).

A. **Subject Population:**

The ideal pool for experimentation would consist of ESL individuals with a TOEFL score range between 75 and 85 to determine the usefulness of the dilation technique. This metric was refined to target individuals with no lower than a 60 TOEFL score, as any individual below this score may be too deficient at the English language to find the dilation tool as useful as those within the target score range. In compliance with university protocol for experimenting with human test subjects, Internal Review Board permission was attained.

An additional restriction was to limit the target test subject demographic to include only people who had lived in the USA for less than one year in order to gather a test subject population that could be reasonably assumed to still be within their TOEFL test result score range. The reasoning for this stems from the fact that students are not required to continue to prove their English skills once they begin their studies in the USA, and so it cannot be determined that the test subject populace could reasonably be expected to progress (or regress) at the same rate. Increased duration rates in the USA could result in

more variability in language skill as compared to their established TOEFL score, so procurement of subjects with more recent TOEFL scores appeared optimal.

B. **Test Environment:**

The test subject is placed in front of a computer with audio files played for them through over-the-ear headphones. After being presented with the listening test audio clip, the test subject answers five multiple choice questions related to the subject of the audio clip they just heard. These questions were taken directly from the book The Official TOEFL iBT Tests with Audio test preparation book. For this experiment, the blank sheet of paper and writing utensil provided to the user in a standard TOEFL listening skill test was removed. It was postulated that having this utility, coupled with the ability to slow the audio, could allow the user to slow the audio enough to allow the user to write a transcript of the audio while they were listening to it, thus giving an inaccurate test result on the subsequent comprehension test.

The biggest difference between the TOEFL test and this experiment is the incorporation of the dilation software, which allows the user to dynamically manipulate the audio as they are listening to the audio clip being presented to them. This dilation is achieved by moving the “dilation slider” (Figure 1A), which allows the user to listen from a range of natural (default speed) of the audio clip, to a slowed down rate of their choosing. Starting at the right-most tick on the slider bar plays the audio at natural rate. Each tick to the left on the slider bar applies one unit of Novak’s dilation algorithm treatment to the audio, up to a maximum of 60 ticks/units of dilation. Each tick is 1% dilation, or slowing of audio speed. The further the dilation slider is moved to the left, the slower the audio becomes. For example, setting the slider to .5 results in half speed. The dilation slider is manipulated by the test subject through the use of a standard PC mouse. The user can simply click and drag the slider to the desired position of preferred dilation.

C. **Testing Format:**

The test subject would alternate from being allowed to dilate an audio clip and then having to listen to an audio clip in its natural, default speed during the experiment. This provides two benefits: 1.) to establish a control parameter and 2.) to compare results against that control parameter within the demographic of test subjects. On tests that incorporated dilation, the dilation slider position would be randomly placed on the slider bar via random number generator. This functionality was implemented to show the user's intent to dilate, to strengthen the argument that that subject was purposely attempting to dilate the audio they were being exposed to.

D. **Complete Protocol:**

The experimental protocol is as follows: an experiment group of 26 English as a Second Language individuals between the ages of 18 and 30, in good physical health (related to hearing), with a target range of Test of English as a Foreign Language skill between 75-85, will be presented with six audio tests, chosen at random from a pool of nine, from The Official TOEFL iBT Tests with Audio test preparation book. The test subjects would alternate between being able to dilate and having to listen to the audio clips in natural form. After each audio clip, a related quiz would be presented to the test subject (Figure 2). Following the six audio tests, a Likert scale survey and interview would be conducted. It was initially hypothesized that test subjects using audio dilation would score higher on dilated TOEFL listening skill tasks than compared to natural rate TOEFL listening skill tasks.

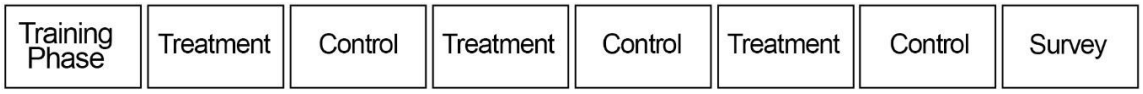


Fig. 2. Diagram of phases of the experiment, starting with the training phase and ending with the survey.

E. **Experiment - User Logging and Training Phase:**

The experiment begins with the logging of the test subject, the researcher, and designating the test subject's status as either a TOEFL (an ESL individual) or a non-TOEFL (native English speaker). After logging the users with a non-identifying string of characters, the next phase of the software introduces the test subject to the audio dilation process, where they are allowed to experiment with the dilation slider in what is known as the Training Phase (Figure 1B). The user is encouraged to move the slider to the left-most position (full dilation) and to the right-most position (unaltered audio file, or natural speech rate) as well as randomly in between to fully understand the dilation process. The test subject is presented with an audio clip of a re-telling of the Gettysburg Address, which gives the user several minutes to experiment with the dilation technique. After the training phase, the user is presented with the option to either repeat the Training Phase or move on to the Testing Phase.

F. **Experiment - Testing Phase:**

When the test subject is ready, the Testing Phase can begin (Figure 1C). The testing phase consists of playing an audio clip taken from The Official TOEFL iBT Tests with Audio test preparation book with material one would expect to encounter in the listening skill section of a TOEFL test. As the test subject manipulates the dilation slider, the movement is logged in to a text file, including time stamps in the milliseconds to later be used to help in creating graphs and determine the amount of time spent at any particular dilation setting. It should be noted, in order to determine intent to use the treatment, the dilation slider bar position is set at a random setting at the start of the dilation-enabled test phases.

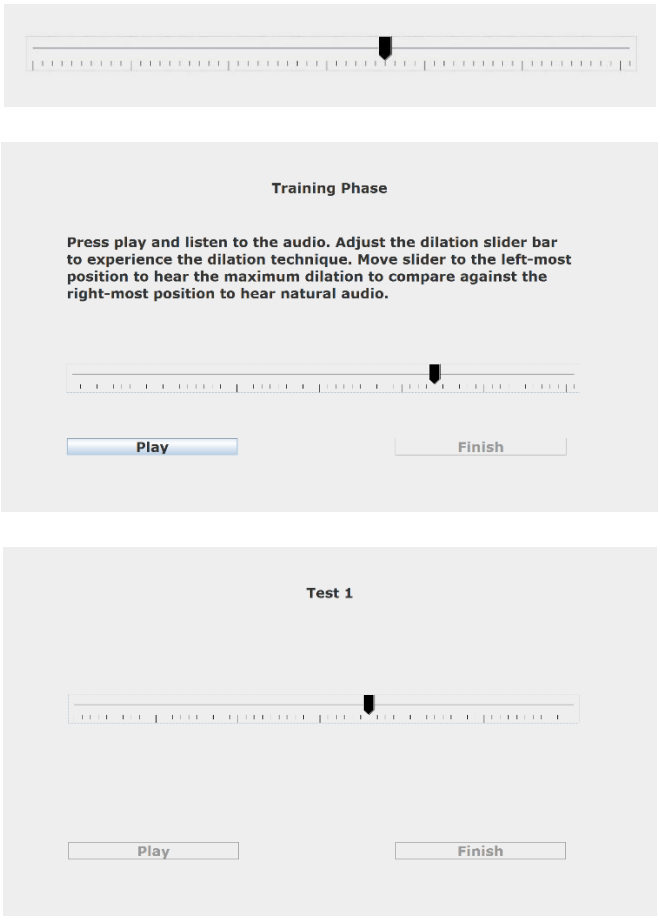


Fig. 1A, 1B, 1C. (1A - top) The dilation slider, when positioned at the right-most position, will result in the audio file being played at the natural, default speed of the audio clip. Every tick movement to the left results in an additional 1% increase of Novak's dilation algorithm being applied to the audio clip, up to a maximum of 60%. (1B - middle) The Training Phase displays the Play button and the Finish button (available to be clicked only after the Gettysburg Address audio clip finishes playing). The Training Phase allows the test subject to experiment with dilation. (1C – bottom) The Testing Phase, featuring an audio clip that can be dilated, as evidenced by the slider bar being present on the screen. The Testing Phase functions similarly to the Training Phase.

Upon the completion of the audio clip presentation, the test subject clicks the Finish button and is taken to the associated quiz for the audio clip. The quiz consists of five multiple-choice questions that each have four answers to choose from. An example from one of the quizzes:

What is the main purpose of the lecture?

A: To introduce a method that can help students remember new information.

B: To introduce a way to study how information passes from one person to another.

C: To explain the differences between biological information and cultural information.

D: To explain the differences between stories, songs, and other pieces of information.

Why does the professor tell the story about alligators?

A: To explain the difference between true and false stories.

B: To draw an analogy between alligator reproduction and cultural transmission.

C: To give an example of a piece of information that functions as a meme.

D: To show how a story can gradually change into a song.

What example does the professor give of a meme's longevity?

A: A story has been changing since it first appeared in the 1930s.

B: A person remembers a story for many years.

C: A gene is passed on through many generations without changing.

D: A song quickly becomes popular all over the world.

What does the professor compare to a housefly laying many eggs?

A: A child learning many different ideas from his or her parents.

B: Alligators reproducing in New York sewers.

C: Different people remembering different versions of a story.

D: A person singing the "Twinkle, twinkle" song many times.

According to the professor, which of the following are examples of meme transfer? (Choose two answers.)

A: Telling familiar stories.

B: Sharing feelings.

C: Composing original music.

D: Learning scientific theory.

After the listening test clip featuring the ability to dilate is played, the following test clip will not allow the test subject to dilate the clip, as evidenced by the entire slider bar being removed from the screen. The test subject listens to an audio clip at natural rate, answering the associated quiz accordingly. The test subject alternates between being able to use treatment to dilate their audio test and having to listen to an audio test in natural state. This is done three times for dilation and three times for control, for a total of six audio tests. The order of testing phases is as follows: dilation, natural, dilation, natural, dilation, natural (Figure 2).

G. **Experiment – Likert Scale Phase:**

A Likert Scale questionnaire (Likert 1932) was also added to the end of the experiment to survey the opinions of the test subjects on the usefulness of the dilation technique. A Likert Scale is a five to seven-point scale (in this case five) which is used to allow the individual to express how much they agree or disagree with a particular statement. The subject answers a question with the options of Strongly Agree, Agree, Neutral, Disagree, and Strongly Disagree. The three questions are as follows:

- "In general, audio dilation for listening skill was useful"
- "Slowing speech for the audio files I listened to was useful"
- "The ability to control audio dilation was easy to use"

The reasoning behind these questions was to differentiate between the user feeling that dilation technique, as a whole, was useful for listening skill rather than focusing specifically on just acquiring an opinion on dilation of the specific audio files, as the latter option leaves possible bias if the user feels they did poorly, or didn't understand, the audio files they just experienced. The third question, on the ability to control audio dilation being easy to use, was more trivial, with implications on possible future user interface design changes.

H. **Experiment – Post Interview:**

An interview was then conducted after the experiment concluded to gather opinions of the test subjects in order to better understand their experience of the material, the test environment, and the software. While there was initial structure to the questions, user responses and attitudes were heavily factored in the follow-up questions, or the continuation of the interview at all. Sensitivity to the test subjects comfort and feelings was of paramount importance throughout the entire experiment phase, from recruitment, testing, and posttest interviews.

Questions such as "what were you thoughts on the experiment?"; "Did you notice a difference in difficulty when comparing natural rate to slowed rate?"; "Do you think dilation is useful?" were asked. Depending on the subject's answers, follow up questions were asked accordingly. Test subjects were encouraged to speak freely about their experiences with the experiment, as well as provide any suggestions or insights they may have to the topic of ESL and living in America (these comments generally came about organically over the course of the interview, and could be considered small talk). Sensitivity to any discomfort or unwillingness to talk during the interview phase were given priority, as to prevent any stress or discomfort to the test subject.

Results

A Shapiro-Wilk test was applied and confirmed that the listening skill test scores were normally distributed. A paired-samples T-test was conducted to compare listening skill test results between treatment and control conditions. There was no significant difference in the listening skill scores for treatment (M = .56, SD =.17) compared to control (M =.61, SD =.18) conditions; t (.98) = 25, P = 0.34.

On average, 88% of candidates dilated, suggesting that members of the subject population will slow speech when given the chance. There was no significant difference in scores between the test and control conditions, however, the dilation scores were lower than scores at natural rate, with average dilation quiz scores of 56% (SD 17%), while natural-rate scores were at 61% (SD 18%) (Figure 3); the percentage change being an 8.19% decrease in score with treatment, for a percentage difference of 8.54%.

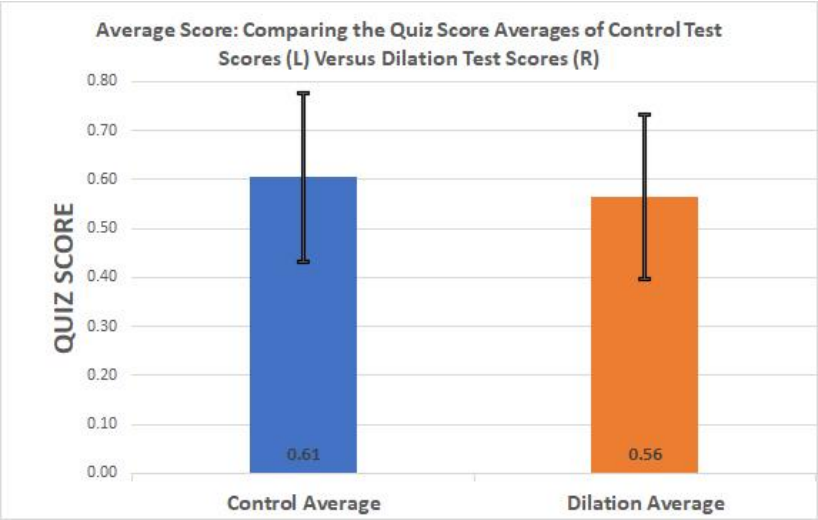


Fig. 3. Average quiz scores: test subject's listening skill assessment with control tests vs treatment (using dilation) tests. Results show that control scores averaged 61% (SD 18%) compared to dilation scores of 56% (SD 17%) There was no significant difference in the listening skill scores for treatment (M = .56, SD =.17) compared to control (M =.61, SD =.18) conditions; t (.98) = 25, P = .34.

The experiment had two different types of audio tests: Conversation style, where a student talks to an advisor in a “real world” setting about problems they are having within the academic environment; and lecture style, where a professor is giving a lecture to students about various topics, replicating a “real world” lecture. 57% of experiment participants said that they felt they understood conversation better (or that conversation was easier), but, on average they scored lower using dilation on conversation style tests, with an average score of 57% (SD 30%) with dilation, compared to 63% (SD 24%) at natural rate (Figure 4).

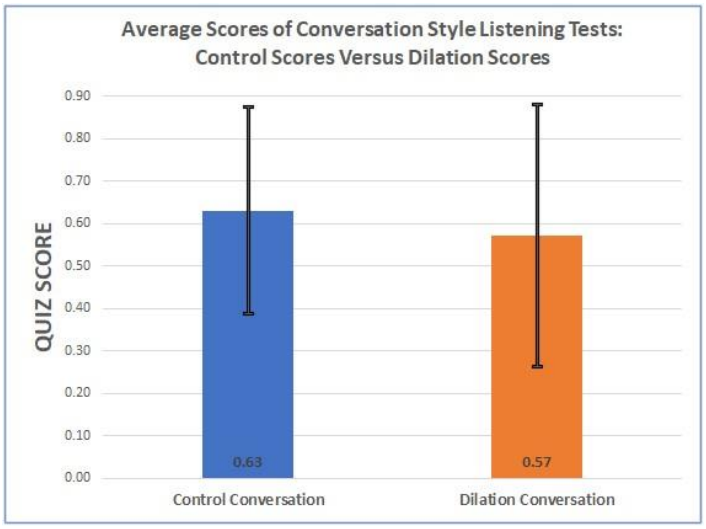


Fig. 4. Average quiz scores of listening skill assessment tests in conversation style. Results show that control scores averaged 63% (SD 24%) compared to dilation scores of 57% (SD 30%)

On Lecture style tests, subjects had an average score of 56% (SD 17%) with dilation compared to a similar average score of 59% (SD 22%) at natural rate (Figure 5).

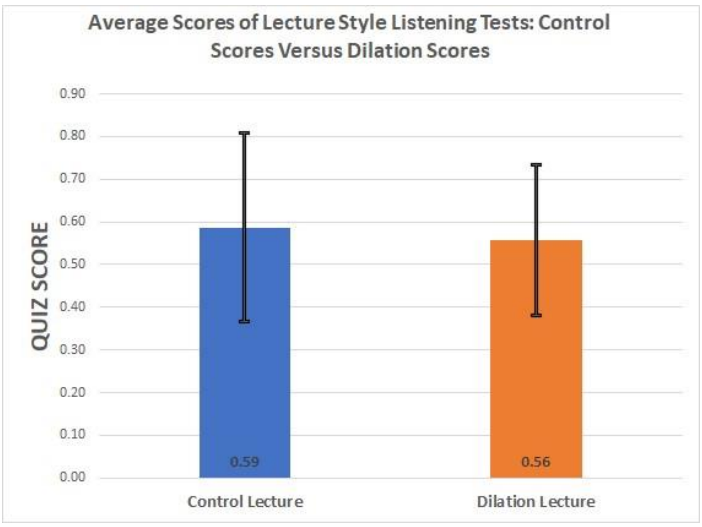


Fig. 5. Average quiz scores of listening skill assessment tests in lecture style. Results show that control scores averaged 59% (SD 22%) compared to dilation scores of 56% (SD 17%)

Overall, combining both control condition scores and treatment condition scores were virtually the same. The average scores of lecture style compared to conversation style showed a difference of one percent, 59% (SD 27%) average score on conversation compared to 58% (SD 19%) for lecture (Figure 6).

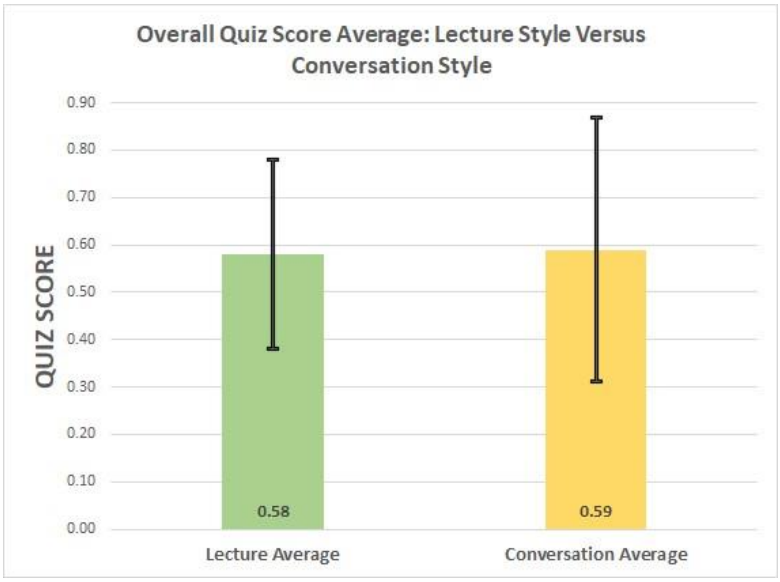


Fig. 6 Overall scores of lecture style versus conversation style audio, combined scores of treatment and control. The average scores of lecture style compared to conversation style showed a difference of one percent: lecture score average of 58% (SD 19%) compared to conversation score of 59% (SD 27%).

The post-experiment interview statistic of 57% of participants stating that conversation style audio was easier, or preferred over, lecture style audio inspired the exploration in to seeing if subjects would dilate more for lecture style than conversation style. Analysis of the data shows that dilation rates of conversation style audio were the same as lecture style audio, a rate of 10% dilation, or slowdown in speech rate.

Graphing average dilation rate by users shows that while as TOEFL score increases (the test subject population contained a maximum TOEFL score of 110), the user tends to listen to audio closer to the natural rate of speed (Figure 7).

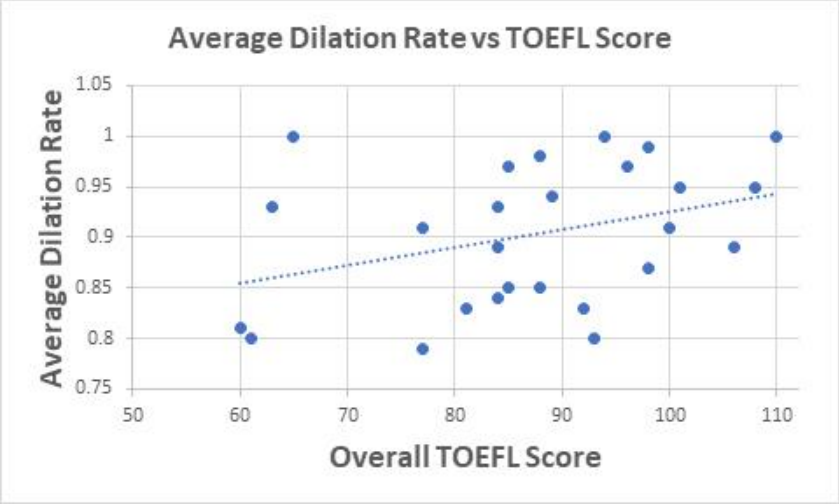


Fig. 7. Graph showing overall TOEFL scores on the X axis, ranging from TOEFL scores of 60 through 110, and their related dilation rate average. Trend line shows as TOEFL score increases, dilation rate decreases, meaning suggesting that the higher the person's TOEFL score, the more likely they are to listen to audio at a natural rate.

Figure 8 shows test subject's TOEFL scores on the X axis, with that test subject's results on the Y axis. The square denotes control score, while the circle represents dilation score, conditions: t (.98) = 25, P = 0.34.

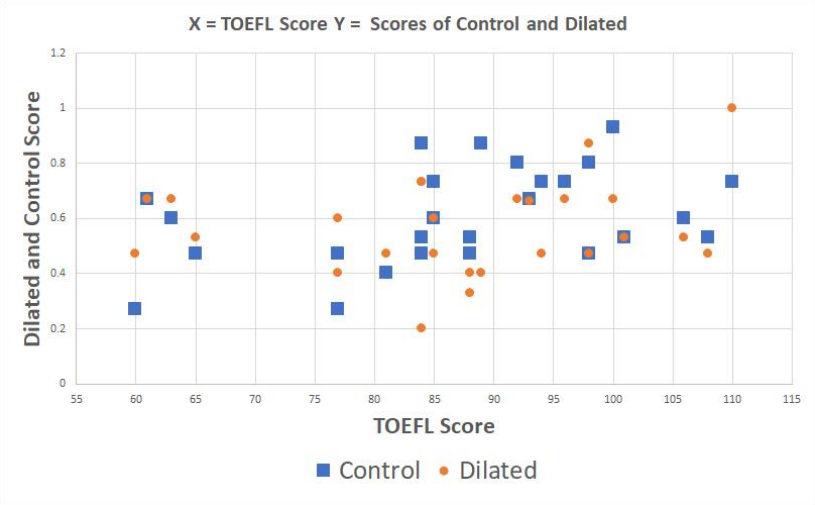


Fig. 8. Graph showing overall TOEFL scores on the X axis, ranging from TOEFL scores of 60 through 110, and their related test score average: treatment vs control. The graph illustrates high variability of score ranges, regardless of TOEFL score, illustrating the lack of statistical significance between treatment and control scores; conditions: t (.98) = 25, P = 0.34.

Of the 26 subjects, six (23%) were below a TOEFL score of 80. Five of those six (83%) dilated the audio, when given the opportunity, at an average dilation rate of 15% (dilation point of .85 on the slider). The average score of those who dilated was 56% (SD 10%), compared with their natural rate score of 46% (SD 16%) with no statistical significance of condition t(-1.7) = 5, P = 0.16. The population average quiz score was 57% (SD 18%) when dilation was used and 64% (SD 15%) with the natural rate (Figure 9).

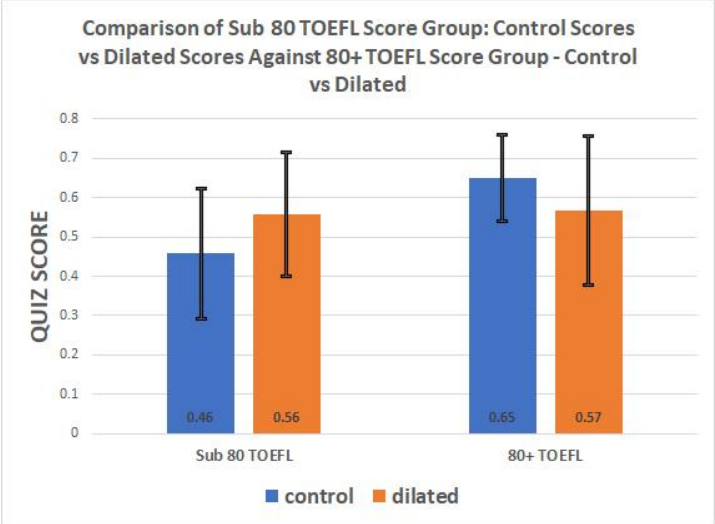


Fig. 9. Graph comparing test subjects with a TOEFL score below 80 against test subjects with a TOEFL score 81 and higher. Results show that the sub 80 TOEFL score group had a higher average score using dilation, 56% (SD 10%), than at natural rate, .46% (SD 16%). By contrast, the 81 and higher TOEFL score group scored higher with natural rate 65% (SD 15%) than with dilation, 57% (SD 18%).

A. **Likert Scale Questionnaire and Interview Responses:**

Results of the Likert scale questionnaire shows favorable opinions of the dilation technique from the test subject population, with 80.7% of respondents giving a positive response to the question "In general, audio dilation for listening skill was useful (Figure 10)." Negative opinions by respondents was at 11.5%, with 7.7% having a neutral opinion; with a median 4, IQR of 1, suggesting consensus.

In general, audio dilation for listening skill was useful

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	Strongly Disagree	1	3.8	3.8	3.8
	Disagree	2	7.7	7.7	11.5
	Neutral	2	7.7	7.7	19.2
	Agree	14	53.8	53.8	73.1
	Strongly Agree	7	26.9	26.9	100.0
	Total	26	100.0	100.0	

Fig. 10. Showing 80.7% of respondents had a favorable opinion to the question "In general, audio dilation for listening skill was useful." 11.5% had a negative opinion to the question, with 7.7% of respondents having a neutral opinion to the question. Median 4; IQR of 1, suggesting consensus

Results for the question "Slowing speech for the audio files I listened to was useful" resulted in in 61.6% of respondents selecting a favorable opinion, 19.2% with a negative opinion, and 7.7% with a neutral opinion (Figure 11). Median 4; IQR of 1.25, suggesting consensus.

Slowing speech for the audio files I listened to was useful

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	Strongly Disagree	3	11.5	11.5	11.5
	Disagree	2	7.7	7.7	19.2
	Neutral	5	19.2	19.2	38.5
	Agree	10	38.5	38.5	76.9
	Strongly Agree	6	23.1	23.1	100.0
	Total	26	100.0	100.0	

Fig. 11. Showing 61.6% of respondents had a favorable opinion to the question "Slowing speech for the audio files I listened to was useful." 19.2% had a negative opinion to the question, with 7.7% of respondents having a neutral opinion to the question. Median 4; IQR of 1.25, suggesting consensus

Results for the question “The ability to control audio dilation was easy to use” resulted in in 92.3% of respondents selecting a favorable opinion, 3.8% with a negative opinion, and 3.8% with a neutral opinion (Figure 12). Median 5; IQR of 1, suggesting consensus.

The ability to control audio dilation was easy to use

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	Strongly Disagree	.00	.00	.00	.00
	Disagree	1	3.8	3.8	3.8
	Neutral	1	3.8	3.8	7.7
	Agree	5	19.2	19.2	26.9
	Strongly Agree	19	73.1	73.1	100.0
	Total	26	100.0	100.0	

Fig. 12. Showing 92.3% of respondents had a favorable opinion to the question “The ability to control audio dilation was easy to use.” 3.8% had a negative opinion to the question, with 3.8% of respondents having a neutral opinion to the question. Median 5; IQR of 1, suggesting consensus

When asked their opinion of the experiment, 15 respondents stated that they felt conversation style audio tests were easier, or preferred over, lecture style audio tests. Those who chose to elaborate on the response typically relayed the opinion that the vocabulary used in lecture style audio tests was more difficult than the vocabulary used in conversation style audio tests.

Discussion

The intent of this work was to find an ideal speech rate for a demographic of ESL individuals in order to improve comprehension of listening skill-related tasks. The target demographic for this research was ESL individuals with a TOEFL score between 75-85. While the results of the experiment showed no statistical significance, this study has shown a scientific contribution to the body of L2 listening comprehension with the aid of computer science. These results illustrate several observations and questions to consider, as well as areas to explore for future work.

This experiment didn't show a statistical significance with this particular test group (N = 26) on this particular set of testing materials. It is in my opinion that this could suggest that the test subject group could be expanded to a larger number of subjects, over a wider range of TOEFL scores (with emphasis on gathering more subjects below a TOEFL score of 80), and by possibly altering the subject material they are tested on. In terms of sample size for future work, for an effect size 0.5, alpha 0.05, and power of 0.80, a sample size of at least 34 should be used (Faul, et al, 2007).

While the TOEFL testing material comes directly from retired material used in previous TOEFL iBT tests, the pool of materials used for testing could be expanded to include International English Language Testing System material (IELTS being another international standardized test of English language proficiency for non-native English language speakers, as well as another metric used by universities) as well as television shows, or similar "natural conversation" media considered to be examples of common English conversation topics and speech patterns, similar to the experiments of Khatib and Khodabaksh (2010), who tested a population of Iranian students on American television shows Oprah and Dr. Phil with modified speech rates. This method of incorporating several different testing materials could deliver more insights into similarities and differences in results, possibly providing more clarity into the effects of dilation.

From a linguistics perspective, it is advised that future work attempts to take into consideration the effects of pausology when altering speech rate, particularly with dilation/slowing of speech rate. Pausology is best defined by O'Connell and Kowal as "the behavioral investigation of temporal dimensions of human speech" (1980). The critical importance of processing and deriving meaning through the listening skill lies in the ability to recognize words in connected speech (Matthews and O'Toole, 2013). Because of the difficulty early-stage L2 learners exhibit in language acquisition (Goh, 2000), explorations of methods to compensate for this deficiency arise, such as in the case of this thesis experiment: audio dilation. Something that should be considered is the role pausology plays in comprehension and the possible adverse effects that dilating these temporal variables may have on ESL comprehension.

For example, the role of pauses, particularly deliberate pauses for effect, might be best left intact over the course of listening to an audio clip. It would also be worth exploring the possibility of leaving the filler "ah" and "um" sounds intact, as well. This is merely speculative, coming from the perspective of someone who has only briefly been exposed to the field of linguistics. But the fact that so much work has been put in to understanding the role pauses play in the topic of pausology, the possibility that pauses might be a separate entity from typical delivery of messages in the context of speech, it could be a worthwhile endeavor to attempt to isolate these phenomena and preserve them.

It could be worth considering the exploration of the effects of workload on ESL individuals during participation of skill assessments like those in this experiment. Specifically, workload measurements to test for a change in physical state which could be associated with increased workload, such as by the use of heart rate monitors. There could be insights attained from attempting to measure workload of the test subjects and postulating that if there are workload changes in test subjects, what those workload changes might mean.

The workload theory is inspired by experiments conducted by Roscoe in 1987 on pilots in various flight conditions. His work showed that depending on the flight conditions, the pilots heart rate would increase, and that this increase could suggest that the pilots were becoming neurologically aroused, resulting in increased performance in order to compensate for the demands of piloting the plane (Roscoe 1987). It could be theorized that if one showed higher heart rate/higher workload during the natural rate listening skill tasks while showing a decreased heart rate/workload of treatment listening skill tasks, that the subject may exhibit less workload to process listening skill tasks, which may be the result of less arousal of the subject and possibly a reason for a lower test score. If Roscoe's theory of arousal playing a role in increased ability can be seen in ESL subjects during listening skill tasks, measuring arousal during control and treatment tests could be worth exploring.

It would also be interesting to compare how a group of native English speakers score compared to a group of ESL subjects. By including a native English-speaking population considered to be similar – that is, a UIC student between the ages of 18 and 30 without hearing impairment – it may be possible to use the native speakers to establish a metric, in terms of average test score, with which to compare the ESL subject group against, possibly providing more insight into expected outcomes of treatment and control test scores. This was initially part of our early concepts for this experiment, but logistics and time constraints made the incorporation of native speakers infeasible.

Allowing the test subject to self-select dilation rates could also pose an issue worth exploring. While providing self-selection is a desired element to this experiment, it is worth considering if the rate chosen by the test subject is accurate. Simply put, the test subject may not know exactly how deficient (if they are deficient at all) they truly are. Self-selecting a rate may be a first step of fine tuning to find their optimal dilation setting or speech rate. I propose that during the training phase, the test subject attempts to find their ideal dilation rate, and then the ensuing tests could automatically add a range of additional

dilation settings per each test, to see if perhaps the test subject over -or under- estimated their optimal dilation rate.

The logistics of executing this experiment resulted in a sample size that ran a broad TOEFL spectrum, from scores of 66 through 110. The preferred range of people with a TOEFL score between 75-85 was limited to merely 8 people. Our initial assumption, when creating a target TOEFL score range, was that it would be rare to find people with a score lower than 80, because the UIC cut rate for admittance is 80. Through a sort of blind luck, the bulk of volunteers for this experiment came from a special Electrical and Computer Engineering program within UIC that had a relaxed TOEFL score requirement compared to the normal university requirements. And so, while many participants from the class fell within the target range, four did not. This presented us with a group of six people with sub 80 TOEFL scores, allowing us to explore the possibility that our target range could be adjusted to a sub 80 TOEFL score range.

Looking at the results from the sub 80 TOEFL group, though a small sample size, could be promising. Compared to the rest of the pool's TOEFL range (81 through 110), the sub 80 group was the only group to score higher with dilation than with control; those who dilated averaged a score of 56%, compared with their natural rate score of 45%. Focusing on TOEFL scores below 80, or at the very least, actively pursuing them in an experiment's protocol, could be insightful. This also includes more rigorous research of all international-focused student programs on campus, to uncover those that may have relaxed English language proficiency scores required for participation with which to recruit participants.

There is a possible bias in data analysis related to this experiment, and possibly any experiment, that evaluates listening skill through questions administered via text. Text based questions then become reliant on the reading skill of the test subject, and in the scenario of the subject having a low reading score, there may be the possibility of the subject having difficulty answering the questions. This situation

appeared in a post interview with one of the test subjects, who commented that while they felt they understood what they were listening to, they could not understand the questions on the test.

It should also be noted that a single metric for average words per minute was not established in this paper. This was primarily due to time constraints, as this experiment was focused on whether dilation was being used and if dilation proved beneficial. Interpreting results in the context of words per minute may appear as an attractive metric for evaluation, but the average words per minute can be highly varied depending on region and topic (i.e. formal or informal), and so I was uncomfortable to attempt to quantify this data into a standard word per minute format.

Conclusion

An experiment testing the use of an audio dilation technique for speech rate manipulation on a population of ESL test subjects was shown to have no statistical significance compared to natural rate speech. The results of lack of statistical significance suggested that while there was no measured benefit to the technique, there was no detrimental effects from the use of the technique, either. It is my belief that slowing of speech rate can provide benefit to those trying to learn a second language, in this case English. I believe further research must still be conducted to find the optimal speech rate to provide benefit.

While my initial exposure to teaching English at UIC consisted of a small demographic of ESL individuals, during the course of this experiment I was able to participate in a Fulbright English Teaching Assistant program to teach English in Laos People Democratic Republic, where I was able to teach for 12 months, in two different locations across the country, as well as participate in a national conference on Teaching English to Speakers of Other Languages (TESOL). On numerous occasions, either my trainers or students would request that I slow my rate of speech. It was observed during conversations that slowing my rate of speech had a positive effect on the conversation.

During these events in Laos, and at the conference, I was able discuss the topic of audio dilation and the slowing of speech rate with professionals from across Asia. The countries these English teachers represented range from Cambodia, Vietnam, Thailand, Laos, China, Korea, Malaysia, Singapore, and Japan. Of the approximately twenty individuals I spoke with, every one of them affirmed the importance of slowing speech rate to help adequately address the deficiencies of their pupils. This doesn't suggest that slowing of speech is actually better than normal speech rate, but it further reinforces that slowing of speech rate is worth exploring, whether it be to find the optimal range of speech rate best for comprehension, or to dispel the notion the slowed speech rate is beneficial.

While explaining the process of the experiment to the test subjects for this experiment, there were several occasions where I was asked to slow down so the test subject could better understand the directions. Test subjects also stated their preference and perceived benefit of slowing speech rate during post experiment interviews. It is clear that there exists a population of people who prefer a slowed rate of speech from natural rate and exhibit an increased ability to participate in a conversation when presented with speech at a reduced speed. This could be seen as another example of need to further explore the topic of speech rate and comprehension.

Bibliography

Blau, E. K.: The effect of syntax, speed, and pauses on listening comprehension. TESOL Quarterly 24 746-53, 1990

Boyle, J.: Factors affecting listening comprehension. ELT Journal, 38,34-38, 1984

Carver R.: Effect of increasing the rate of speech presentation upon comprehension. Journal of Educational Psychology,65, 118-126, 1973

Faul, F., Erdfelder, E., Lang, A.-G., & Buchner, A.: G*Power 3: A flexible statistical power analysis program for the social, behavioral, and biomedical sciences. *Behavior Research Methods*, *39*, 175-191, 2007

Feyten, C. M.: The power of listening ability: An overlooked dimension in language acquisition. The Modern Language Journal, 75(2), 173-180. 1991

Foulke, E.: Listening comprehension as a function of word rate. Journal of Communication, 18, 198-206, 1968

Foulke, E. and Sticht, T.G.: Review of research on the intelligibility and comprehension of accelerated speech. Pausological Bulletin 72: 50-62, 1969

Friedman, H.L. & Johnson, R.L.: Some actual and potential uses of rate controlled speech in second language learning. In P.Pimsleur & T. Quinn (Eds.). The Psychology of Second Language Learning (pp. 167-164), 1971

Goh, C.: A cognitive perspective on language learners' listening comprehension problems. System, 28, 55-75, 2000

Griffiths, R. T.: Pausological research in an L2 context: A rationale and review of selected studies. Applied Linguistics, 12,345-361, 1991

Khatib, M and Khodabaksh, M.: The effect of modified speech on listening to authentic speech. Journal of Language Teaching and Research, 1, 5, 685-693, 2010

Kim, I.-S.: Automatic speech recognition: reliability and pedagogical implications for teaching pronunciation. Educational Technology & Society, 9 (1), 322-334 , 2006

Likert, Rensis.: A technique for the measurement of attitudes. Archives of Psychology, 140, 1–55. 1932

Matthews, J. and O'Toole, J.M.: Investigating an innovative computer application to improve L2 word recognition from speech, Computer Assisted Language Learning, 28:4, 364-382, 2013

Meskill, C.: Listening skills development through multimedia. Journal of Educational Multimedia and Hypermedia, 5(2), 179-201. Charlottesville, VA, 1996

Morley, J.: Current perspectives on improving aural comprehension. In ESL Magazine. January. 1-30, 1999

Nielson, Katharine B.: Self-study with language learning software in the workplace: what happens. Language Learning & Technology,V15 N3 110-129 2011

Novak, J. Archer, J. Shafiro, V. Kenyon, R. Leigh, J.: On-line audio dilation for human interaction. Interspeech 2013: Show & Tell Contribution, 2013.

O'Connell and Kowal.: Prospectus for a science of pausology, in temporal variables in speech: Studies in Honour of Frieda Goldman-Eisler, 3-10, 1980

Pimsleur, P.C. Hanckok, C. and Furey, P.: Speech rate and listening comprehension. In Viewpoints on English as a Second Language. New York: Regents, 1977

Rivers, W.M.: Teaching foreign-language skills. Chicago, University of Chicago Press, 1968

Roscoe, A.H.: The practical assessment of pilot workload (AGARD-AG-282, pp. 78–82). Neuilly-sur-Seine, France: Advisory Group for Aerospace Research and Development, 1987.

Vandergrift, L.: Listening to Learn or Learning to Listen? Annual Review of Applied Linguistics, 24,3-26,2004

Zhao, Y.: The effects of listeners' control of speech rate on second language comprehension. Applied Linguistics, 18,49-68, 1997

APPENDIX

Example transcript of audio test used in experiment:

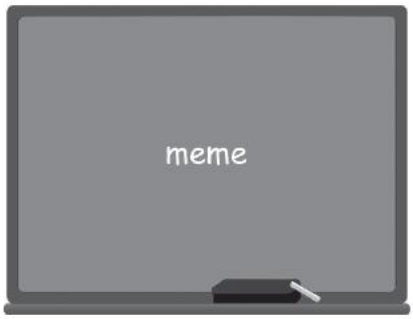
TRACK 23 TRANSCRIPT

Narrator
Listen to part of a lecture in a sociology class.

Professor
Have you ever heard the one about alligators living in New York sewers? The story goes like this: a family went on vacation in Florida, and bought a couple of baby alligators as presents for their children, then returned from vacation to New York, bringing the alligators home with them as pets. But the alligators would escape and find their way into the New York sewer system where they started reproducing, grew to huge sizes and now strike fear into sewer workers. Have you heard this story? Well, it isn't true and it never happened, but despite that, the story's been around since the 1930s.

Or how about the song "Twinkle, twinkle, little star"? You know "Twinkle, twinkle, little star, how I wonder what you are . . ." Well, we've all heard this song. Where am I going with this? Well, both the song and the story are examples of memes, and that's what we'll talk about, the theory of memes.

This ebook was issued to **sherice nelson**, order **#10898554238**. Unlawful distribution of this ebook is prohibited.



A meme is defined as a piece of information copied from person to person. By this definition, most of what you know . . . ideas, skills, stories, songs . . . are memes. All the words you know, all the scientific theories you've learned, the rules your parents taught you to observe . . . all are memes that have been passed on from person to person.

So what? . . . you may say. Passing on ideas from one person to another is nothing new . . . Well, the whole point of defining this familiar process as transmission of memes is so that we can explore its analogy with the transmission of *genes*.

As you know, all living organisms pass on biological information through the genes. What's a gene? A gene is a piece of biological information that gets copied, or replicated, and the copy, or replica, is passed on to the new generation. So genes are defined as replicators . . .

This ebook was issued to **sherice nelson**, order **#10898554238**. Unlawful distribution of this ebook is prohibited.

Genes are replicators that pass on information about properties and characteristics of organisms. By analogy, *memes* also get replicated and in the process pass on cultural information from person to person, generation to generation. So memes are also replicators. To be a successful replicator, there are three key characteristics: longevity, fecundity, and fidelity. Let's take a closer look . . .

First, longevity. A replicator must exist long enough to be able to get copied and transfer its information. Clearly, the longer a replicator survives, the better its chances of getting its message copied and passed on. So longevity is a key characteristic of a replicator. If you take the alligator story, it can exist for a long time in individual memory—let's say my memory. I can tell you the story now, or ten years from now. The same with the "Twinkle, twinkle" song. So these memes have longevity, because they're memorable, for one reason or another.

This ebook was issued to **sherice nelson**, order **#10898554238**. Unlawful distribution of this ebook is prohibited.

Next, fecundity. Fecundity is the ability to reproduce in large numbers. For example, the common housefly reproduces by laying several thousand eggs. So each fly gene gets copied thousands of times. Memes? Well, they can be reproduced in large numbers as well. How many times have you sung the "Twinkle, twinkle" song to someone? Each time you replicated the song—and maybe passed it along to someone who didn't know it yet, a small child maybe.

And finally, fidelity. Fidelity means accuracy of the copying process. We know fidelity is an essential principle of genetic transmission. If a copy of a gene is a bit different from the original, that's called a *genetic* mutation, and mutations are usually bad news. An organism often cannot survive with a mutated gene—and so a gene usually cannot be passed on unless it's an exact copy. For *memes*, however, fidelity is not always so important. For example, if you tell someone the alligator story I told you today, it probably won't be word for word exactly as I said it. Still, it will be basically the same story, and the person who hears the story will be able to pass it along. Other memes are replicated with higher fidelity, though—like the "Twinkle, twinkle" song? It had the exact same words twenty years ago as it does now. Well, that's because we see songs as something that has to be performed accurately each time. If you change a word, the others will usually bring you in line. They'll say, "That's not how you sing it," right?

This ebook was issued to **sherice nelson**, order **#10898554238**. Unlawful distribution of this ebook is prohibited.

So, you can see how looking at pieces of cultural information as replicators, as memes, and analyzing them in terms of longevity, fecundity, and fidelity, we can gain some insight about how they spread, persist, or change.

TRACK 24 TRANSCRIPT

Narrator
Why does the professor say this:

Professor
If you change a word, the others will usually bring you in line. They'll say, "That's not how you sing it," right?

TRACK 25 TRANSCRIPT

Note: The actual lecture contains color images. The colors from one image are discussed by the professor. You do not need to see the colors to understand the lecture or to answer the questions.

Narrator
Listen to part of a lecture in an astronomy class.

This ebook was issued to **sherice nelson**, order **#10898554238**. Unlawful distribution of this ebook is prohibited.

Directions: Now answer the questions.

6. What is the main purpose of the lecture?
   - Ⓐ To introduce a method that can help students remember new information
   - Ⓑ To introduce a way to study how information passes from one person to another
   - Ⓒ To explain the differences between biological information and cultural information
   - Ⓓ To explain the differences between stories, songs, and other pieces of information

7. Why does the professor tell the story about alligators?
   - Ⓐ To explain the difference between true and false stories
   - Ⓑ To draw an analogy between alligator reproduction and cultural transmission
   - Ⓒ To give an example of a piece of information that functions as a meme
   - Ⓓ To show how a story can gradually change into a song

This ebook was issued to **sherice nelson**, order **#10898554238**. Unlawful distribution of this ebook is prohibited.

8. According to the professor, which of the following are examples of meme transfer? *Choose 2 answers.*

   A Telling familiar stories
   B Sharing feelings
   C Composing original music
   D Learning a scientific theory

9. What example does the professor give of a meme's longevity?

   (A) A story has been changing since it first appeared in the 1930s.
   (B) A person remembers a story for many years.
   (C) A gene is passed on through many generations without changing.
   (D) A song quickly becomes popular all over the world.

10. What does the professor compare to a housefly laying many eggs?

    (A) A child learning many different ideas from his or her parents
    (B) Alligators reproducing in New York sewers
    (C) Different people remembering different versions of a story
    (D) A person singing the "Twinkle, twinkle" song many times

11. Listen to Track 24.

    (A) To explain why some memes do not change much
    (B) To ask the students for their opinion about songs as memes
    (C) To acknowledge a problem with the meme theory
    (D) To ask the students to test an idea about memes

This ebook was issued to **sherice nelson**, order **#10898554238**. Unlawful distribution of this ebook is prohibited.

Professor
Last week, we covered some arguments *against* going back to the Moon. But there are compelling reasons *in favor of* another Moon landing, too, um, not the least of which is trying to pinpoint the Moon's age. We could do this, in theory, by studying an enormous impact crater known as the South Pole–Aitken Basin. Ah, it's located in the Moon's south polar region. But, since it's on the *far* side of the Moon, it can be seen only from space. Here's an image of . . . we'll call it the SPA Basin.


South Pole—Aitken Basin

This color-coded image of the SPA Basin—ahh, those aren't its actual colors, obviously—uh, this image is from the mid-nineties, from an American spacecraft called Clementine. Um, unlike earlier lunar missions, Clementine *didn't* orbit *only* around the Moon's *equator*. Its orbits enabled it to send back data to create *this* topographical map of . . . well, the gray-and-white area toward the bottom is the *South Pole*. The purples and blues in the middle correspond to low elevations—the SPA Basin itself. Uh, the oranges and reds around it are higher elevations. The Basin measures an amazing 2,500 kilometers in diameter, and its average depth is 12 kilometers. That makes it the biggest known crater in our *solar system*. And it may well be the *oldest*.

Y'know, planetary researchers *love* studying deep craters to learn about the impacts that created them, um, how they redistributed pieces of the planet's crust. And, in *this*

This ebook was issued to **sherice nelson**, order **#10898554238**. Unlawful distribution of this ebook is prohibited.

VITA

NAME:	Daniel Wayne Bunn
EDUCATION:	B.S., Computer Information Systems, Calumet College of St. Joseph, Whiting, Indiana, 2014
TEACHING:	Information Technology, Seton Academy, 2009-2011
	Information Communications Technology and English, Vientiane and Pakse, Laos People's Democratic Republic, 2016-2017
WORK EXPERIENCE:	Graduate Assistant, College of Engineering, University of Illinois at Chicago 2015-2016, 2017-2018
	Technology Resource Coordinator, Seton Academy, 2009-2011
	Internship Marketing and Web Design, Calumet College of St. Joseph, 2013
HONORS:	Fulbright Scholar, English Teaching Assistant 2016-2017