The Effect of Repeated Reading Aloud on the Speaking Fluency of Russian Language Learners

Evgenia Nikolayevna Stroh
Brigham Young University - Provo

Follow this and additional works at: https://scholarsarchive.byu.edu/etd
Part of the Other Languages, Societies, and Cultures Commons

BYU ScholarsArchive Citation
Stroh, Evgenia Nikolayevna, "The Effect of Repeated Reading Aloud on the Speaking Fluency of Russian Language Learners" (2012). All Theses and Dissertations. 3325.
https://scholarsarchive.byu.edu/etd/3325

This Thesis is brought to you for free and open access by BYU ScholarsArchive. It has been accepted for inclusion in All Theses and Dissertations by an authorized administrator of BYU ScholarsArchive. For more information, please contact scholarsarchive@byu.edu, ellen_amatangelo@byu.edu.
The Effect of Repeated Reading Aloud on the Speaking Fluency of
Russian Language Learners

Evgenia Nikolayevna Stroh

A thesis submitted to the faculty of
Brigham Young University
in partial fulfillment of the requirements for the degree of
Master of Arts

Michael D. Bush, Chair
David K. Hart
Grant H. Lundberg

Center for Language Studies
Brigham Young University
August 2012

Copyright © 2012 Evgenia Stroh
All Rights Reserved
ABSTRACT

The Effect of Repeated Reading Aloud on the Speaking Fluency of Russian Language Learners

Evgenia Nikolayevna Stroh
Center for Language Studies, BYU
Master of Arts

The current study examines the effect of repeated reading aloud upon speaking fluency. Because there is little evidence in the literature that the practice of repeated reading aloud can have a positive effect upon speaking fluency, the primary goal of this study was to investigate this relationship further. For the purposes of the study, speaking fluency was defined as fluidity and smoothness of speech with little pausing and hesitation. It is measured by evaluating the following fluency features: speech rate, number of pauses, length of pauses, phonation/time ratio, and articulation rate. The repeated measures experimental design of the study involved current and former Brigham Young University students learning Russian as a foreign language. They were divided into two groups: control and experimental. The participants in the experimental group performed repeated reading aloud activities daily, while those in the control group read the same passages silently. All participants took weekly speaking tests consisting of simple speaking prompts. The final post-test included both reading aloud and speaking tests.

The speech samples collected from the tests were evaluated using computer-based analysis as well as scores from three raters who are native speakers of the Russian language. The statistical analysis and comparison of these scores revealed mixed results. The rater scores did not exhibit any statistically significant difference between the groups, which could be attributed to overall low inter-rater reliability and short duration of the experiment. On the other hand, the computer-generated scores for mean length of pauses, phonation/time ratio, and speech rate of the experimental group were better than those of the control group. This difference proved to be statistically significant based on the results of one-way and repeated measures ANOVA analyses.

Unfortunately because of the high attrition rate and short duration of the study, these results cannot be generalized. Therefore further research is necessary to confirm or reject these findings.

Keywords: speaking fluency, repeated reading aloud, repetition, practice, Russian language learners
ACKNOWLEDGEMENTS

I would like to express my deepest gratitude to my thesis chair and mentor Dr. Michael D. Bush. Without his direction and constant encouragement I would not have learned so much and would not have been able to complete this research project. I would also like to thank Dr. David K. Hart for his guidance and support throughout this study, and Dr. Grant H. Lundberg for his time and feedback. I am grateful to the ARCLITE Lab of the Center for Language Studies for assisting me with the design and development of the software used in the study, and Troy Cox for his help with data analysis. I greatly appreciate all the help and assistance I received throughout the years from the Center for Language Studies and especially Agnes Welch. I also want to extend my gratitude to the Fritz Burns Foundation, and Paul and Tonita Crookston for the scholarships and financial support that made my education at Brigham Young University possible.

Most of all I want to thank my parents, Nikolay and Yulia Kuznetsov, my sister, Alina Crookston, and my husband, Travis Stroh, for their unwavering support and faith in me.
# Table of Contents

ABSTRACT ........................................................................................................................................ iii

ACKNOWLEDGEMENTS .................................................................................................................. v

List of Figures ..................................................................................................................................... xi

List of tables ....................................................................................................................................... xiii

Chapter 1: Introduction ....................................................................................................................... 1
  Background ....................................................................................................................................... 1
  Significance of the Problem and its Solution ...................................................................................... 2
  Research Question and Overview of the Study ................................................................................. 3
  Definitions of Terms .......................................................................................................................... 5
  Assumptions and Limitations ............................................................................................................. 5

Chapter 2: Review of the Literature .................................................................................................... 7
  Fluency .............................................................................................................................................. 8
    Definition of fluency. .......................................................................................................................... 8
    Fluency and automaticity. .................................................................................................................. 9
    Measuring speaking fluency ............................................................................................................ 10
  The Role of Repetition in Language Learning ............................................................................... 15
  Repeated Reading as a Method ........................................................................................................ 18
    Repeated reading in L1 studies. ....................................................................................................... 18
    Repeated reading in SLA ................................................................................................................ 18
  Research Question ............................................................................................................................ 19
## Chapter 3: Research Design and Procedures

- Overview of the Study ........................................................................................................ 21
- Materials and Instruments .................................................................................................. 22
  - Pre-treatment survey. ........................................................................................................ 22
  - Extending Your Russian .................................................................................................. 23
  - Post-test. .......................................................................................................................... 25
  - Post-study survey. .......................................................................................................... 28
- Design of the Experiment .................................................................................................... 29
  - Experimental group. ....................................................................................................... 29
  - Control group. ............................................................................................................... 31
- Procedures ............................................................................................................................ 31
  - First iteration. .................................................................................................................. 32
  - Second iteration. ............................................................................................................. 32
  - Recruiting subjects for the final iterations. .................................................................. 33
  - Third iteration. ................................................................................................................. 33
  - Final iteration. ................................................................................................................ 33
- Subjects ................................................................................................................................ 34
- Data collection and analysis ............................................................................................... 36
  - Human raters. .................................................................................................................. 37
  - Computer generated analysis. ........................................................................................ 37
  - Statistical analysis. ......................................................................................................... 38

## Chapter 4: Results

- Statistical Analysis .............................................................................................................. 41
  - Rater scores. ..................................................................................................................... 41
ASR speaking scores results.................................44
Repeated reading aloud results.............................50
Relationships between scores...............................51
Summary of statistical analysis...............................52
Addressing Attrition in the Study...........................52
Post-study survey results.....................................53

Chapter 5: Discussion and Conclusions..........................55
Discussion..........................................................55
Raters...............................................................55
Computer based analysis.......................................56
Attrition.............................................................56
Limitations...........................................................57
Recommendations for Future Research......................57
Summary and Conclusions ........................................58

References..................................................................61

APPENDICES ................................................................68

Appendix A: Informed Consent Form..........................68
Appendix B: Pre-treatment Survey................................69
Appendix C: List of Reading Articles...........................72
Appendix D: Comprehension Questions........................73
Appendix E: Post-study Survey....................................77
List of Figures

Figure 1. Self-assessment questions ................................................................. 25
Figure 2. Speaking prompts for Week 1 and Week 2 tests ................................ 26
Figure 3. Reading passages for the reading aloud section of the post-test .......... 27
Figure 4. Speaking prompts for the final post-test ........................................... 28
Figure 5. Screen shot of the first reading tab of the modified software ............... 30
Figure 6. Screen shot of the second reading tab of the modified software .......... 31
Figure 7. The N of people participating in the various stages of the study ............ 35
Figure 8. Stages of attrition of the final iteration of the study .......................... 53
List of tables

Table 3.1. Subjects’ self-perception of overall proficiency in the Russian Language ..................... 36
Table 4.1. Intraclass Correlation Coefficient for rater scores .......................................................... 42
Table 4.2 Rater scores correlation ................................................................................................... 43
Table 4.3. Experimental and control group mean scores from raters ............................................. 43
Table 4.4. Independent samples t-test for human rater scores .......................................................... 44
Table 4.5. Abbreviations used in the statistical analysis of fluency features .................................. 44
Table 4.6. Means for speaking scores on fluency features ................................................................. 45
Table 4.7. One-way ANOVA comparison for speaking fluency scores .......................................... 46
Table 4.8. Between-groups comparison for repeated measures ANOVA: Week 1, 2, and Week 3 first test ........................................................................................................................................ 47
Table 4.9. Between-groups comparison for repeated measures ANOVA: Week 1, 2, and Week 3 mean scores ..................................................................................................................................... 49
Table 4.10. Means for reading aloud fluency features ...................................................................... 50
Table 4.11. One-way ANOVA comparison for reading aloud fluency scores .................................. 51
Table 4.12. Pearson correlation between scores ................................................................................. 52
Table 4.13. Reasons for abandoning the study ................................................................................ 54
Chapter 1: Introduction

The Russian language is perceived to be among the most difficult languages to learn. With its complex grammar and morphological systems, stress and intonation patterns, and pronunciation difficulty, many people struggle with achieving fluency in this language viewed as a challenge by its learners and beautiful by its speakers. Nevertheless, experience shows that with constant exposure and practice, developing fluency in a second language is made possible even with the most difficult languages. Individuals who spend time in the country of the target language come back with improved pronunciation and understanding of the mechanics of the language. On the other hand when individuals do not have constant access to native speakers and practice opportunities, achieving or maintaining fluency becomes problematic.

Background

An individual who is proficient in a second or a foreign language is often described as “fluent.” Even though language proficiency incorporates a wide array of skills from being able to produce language accurately in both spoken and written form to understanding written and spoken communication, fluency in a foreign language is often associated with one’s ability to speak smoothly and comprehensibly. In fact, in language acquisition literature fluency is often referred to as “flow, continuity, automaticity, or smoothness of speech” (Koponen & Riggenbach, 2000, p. 6).

Speaking a foreign language is just like any other skill and thus requires regular practice if a learner is to develop and maintain fluency (Ericsson, Krampe, & Tesch-Romer, 1993). Unfortunately, speech production is not as simple as riding a bicycle. In the words of MacKay (1981) it is “a sequentially organized output system par excellence, requiring sequential
organization of many different types of response components: phrases, words, syllables, and phonemes as well as muscle movements” (p. 485). Therefore, obtaining and maintaining this skill in a foreign language is contingent upon the amount of practice in which learners engage.

**Significance of the Problem and its Solution**

For Russian language learners who are not surrounded by native Russian speakers, there are very few opportunities to practice speaking outside of the classroom. As with any other skill, without practice learners’ ability to speak Russian fluently not only stops progressing, but also regresses with time. Learners, who do not have adequate opportunity to practice, experience difficulties with pronouncing foreign sounds and accessing complex and even basic vocabulary. Thus their speech becomes unnaturally slow and is filled with pauses, hesitations and repeats.

It has been hypothesized that one of the ways to solve this problem can perhaps be repeated reading aloud. In private conversations individuals who are learning a second or a foreign language admit that reading out loud repeatedly helps them work on their articulation and thus improves their speaking fluency. This idea has also been supported by Gibson (2008), who suggested that although this exercise cannot replace real life interpersonal communication, it could facilitate the maintenance of pronunciation and speed, and provide necessary articulatory warm-up for increased speaking fluency.

The method of repeated reading, which was presented by Samuels (1979), is most often used for improving reading fluency in either one’s first or second language. Students read short passages repeatedly until they reach the desired reading rate and then move on to the next passage. This activity can be carried out silently as well as aloud. Research in first and second language acquisition shows that the practice of repeated reading leads to improved reading rate
and shortened speech pauses (Herman 1985; Gunter, 1995; Steventon & Fredrick, 2003; Taguchi, Takayasu-Mass, & Grsuch, 2004; Wiley & Deno, 2005; Roundy & Roundy, 2009). Moreover, some studies show that performance in oral reading can predict speaking fluency (Coniam, 1991; Cucchiarini, Strik, & Boves, 2002).

The activity of repeated reading aloud for the purpose of practicing speaking was recently implemented in *Extending Your Russian* (EYR), a software package available online for learners of Russian as a foreign language (eyr.byu.edu). The program features articles from an authentic Russian news journal *Itogi* with word definitions, explanations of stress patterns, examples of native pronunciation and intonation, and finally, repeated reading activities. Overall, users read each article three times or more, and move on to the next article only after they reach a predetermined reading rate. The instructions for the reading section of the software package state that if users participate in these activities for at least 15 minutes a day, their overall language ability should improve, especially their speaking fluency.

**Research Question and Overview of the Study**

Although the practice of repeated reading has been researched in the past, the main focus of that research was on its effects upon reading abilities and not speaking fluency. The primary goal of the current study was to answer the following research question:

- What is the effect of repeated reading aloud upon speaking fluency of non-native speakers of Russian as measured by a repeated measures elicited response test?

A careful review of the literature produced only a few studies connecting the practice of repeated reading aloud and speaking fluency in a foreign language. Therefore, the working hypothesis for this study was the null hypothesis stating that repeated reading aloud has no effect
on speaking fluency. The study implemented experimental design with repeated measures involving learners of Russian as a foreign language.

In order to examine the effects of repeated reading aloud upon speaking fluency, the present study used a subsection of *Extending Your Russian*. The dependent variable for this study was speaking fluency, which was measured by computer-generated fluency features as well as overall perception by trained native speakers. The independent variable was the practice of repeated reading aloud, defined as reading passages repeatedly until the desired reading rate is reached. Some possible intervening variables that were taken into consideration at the beginning of the study were: various proficiency levels and other activities involving the Russian language at the time of the research.

A very important feature of the study was accessibility. Due to the fact that not all of the participants of this study were current BYU students, making the reading activities available online and easy to use was crucial. This eliminated the need for the participants to report to a testing center to read or take tests and allowed them to carry out the research activities at a place and time convenient for them. Administering the study online, however, presented limitations as well. More specifically, there was no control over the consistency of subject involvement in the study.

The participants in the study, current and former students of Brigham Young University (BYU) volunteered to take part in the research after receiving personalized invitations via e-mail. They were divided into two groups: control and experimental. The experimental group performed repeated reading aloud activities with the EYR content, whereas those in the control group read the same content silently. Each week all subjects recorded their responses to speaking
prompts. These recordings were later used to examine the effect of repeated reading aloud upon their speaking fluency by comparing rater-assigned and computer generated speaking fluency scores.

**Definitions of Terms**

The following terms appear frequently throughout the study and thus will be defined in this section.

The working definition of *speaking fluency* for the purposes of this study is adapted from Crystal and Varley (1993) where it is described as “ease and rapidity of speaking, a continuous flow with little hesitation, and a good command of grammar and vocabulary” (p. 189). The speaking fluency is further distinguished into two types: *higher order of fluency*, which refers to overall oral proficiency encompassing good command of grammar and vocabulary; and *lower order of fluency*, which refers to the ability to produce speech effortlessly and smoothly. The lower order of fluency can be characterized using various *fluency features*, quantitative measures of speaking fluency, such as speech rate, articulation rate, number and length of pauses, and repeats and hesitations. These features can also be referred to as *mechanics of speech*.

The exercise of *repeated reading* is defined as reading passages repeatedly until the desired rate is achieved. *Reading aloud* is defined in this study as the practice of reading passages out loud. Therefore *repeated reading aloud* refers to an activity of reading passages out loud repeatedly with the goal of reaching a specified reading speed.

**Assumptions and Limitations**

For the study to be successful in establishing effects of repeated reading aloud upon speaking fluency, several assumptions and limitations must be identified. First of all, the
participants for the study were assumed to have at least intermediate knowledge of the Russian language. This level of proficiency was necessary for successful completion of the reading tasks for the experiment. In addition, the expectation was that the subjects had not been taking any Russian classes while participating in the study. Finally, it was anticipated that the subjects would carry out the reading activities five days a week for a minimum of three weeks.

Nevertheless, because no pre-test was administered, the proficiency level among the subjects could not be confirmed. Moreover, because the experiment took place online, there was little control over the research environment; therefore ensuring that the subjects consistently followed instructions was challenging. Finally attrition in the study was a foreseeable limitation.
Chapter 2: Review of the Literature

Acquiring a foreign language encompasses mastering the skills of reading, writing, listening, and speaking. Each of these skills is equally important for successful communication; however, speaking takes on a special role when it comes to interactions with native speakers of the target language. Developing speaking fluency is a process that requires much practice. This practice is abundant in a classroom setting or in any environment where the target language is predominant. However, developing or maintaining speaking fluency is a much more difficult task when opportunities for authentic practice are limited. Upon careful survey of the literature it became evident that repeated reading aloud perhaps can be a valuable tool in building up speech rate, reducing the number of repeats and hesitations, and improving overall flow of speech. Nevertheless, little research has been done to investigate whether there is a connection between the practice of repeated reading aloud and speaking fluency. There is sufficient evidence in the literature, however, to suggest that repeated reading aloud could, in fact, provide necessary exercise for improving some aspects of speaking fluency in a foreign language.

This chapter will discuss the role of fluency in the process of acquiring a second/foreign language as described in the literature. In addition, this review will demonstrate that even though some researchers in the field of second language acquisition have discounted the value of repetition, we will see where it is a viable choice of practice for developing fluency in a foreign language. It will then be shown that repeated reading aloud has already been used as a method for improvement of reading skills in both first and second language teaching, and that it indeed presents an opportunity for improving speaking fluency in a foreign language. Thus, this chapter will address three main points: fluency, repetition, and repeated reading aloud.
Fluency

Individuals who achieve success in mastering a second or a foreign language are often described as fluent. This description is usually connected to their ability to fully understand what they hear and read, as well as their ability to speak without lengthy pauses and with correct pronunciation and intonation. Essentially, becoming fluent in a target language is the main goal for every language learner. Fluency, therefore, has been a subject of many studies that focused on its definition, development, and assessment.

**Definition of fluency.** The investigation of this phenomenon holds significant importance, but unfortunately the term “fluency” does not have a standard definition in the literature. However, it is often associated with notions of smoothness, fluidity, and rapidity (Koponen & Riggenbach, 2000). In addition, fluency has been commonly viewed as a demonstration of an acquired skill in a foreign or a second language.

In the field of second language acquisition (SLA) fluency appears to be an essential component and is often used as an indicator of proficiency level. For example, ACTFL Proficiency Guidelines for speaking state that at a Superior level speakers “are able to communicate in the language with accuracy and fluency” (ACTFL, 1999, p. 3), where fluency is characterized as production of speech with ease and lack of unnatural hesitations. Similarly, Crystal and Varley (1993) suggest that “in relation to language, the term implies ease and rapidity of speaking, a continuous flow with little hesitation, and a good command of grammar and vocabulary” (p. 189).

Furthermore, Pawley and Syder (1983) distinguish between two types of speaking fluency: “nativelike [sic] selection” and “nativelike delivery.” “Nativelike selection” is the ability of a second language learner to accurately produce utterances with vocabulary and
phrases that a native speaker would use. Lennon (2000) refers to it as a higher order of fluency, or overall oral proficiency. On the other hand, “nativelike delivery” is the ability to produce speech as smoothly and effortlessly as a native speaker. Similar in definition to “nativelike delivery” is the lower order of fluency that can be measured through assessing the following characteristics: speech rate, length of pauses, hesitations, restarts, stretches of speech, intonation patterns, and stress-timed delivery (Oppenheim, 2000; Lennon, 2000). Despite its description as a “lower order of fluency,” the following sections will support the idea that these characteristics of speech are good predictors of overall speaking fluency.

**Fluency and automaticity.** The second language acquisition process has often been considered similar to acquiring other complex skills. For example, just like playing the piano, acquiring a second language requires practice. In the case of learning to play the piano, constant practice leads to development of “motor memory,” which is also commonly referred to as “muscle memory,” a type of movement with which the nervous system and muscles become familiar over time (Krakauer & Shadmehr, 2006). This movement then becomes automatic for the body as the organism decreases the level of effort and attention necessary to accomplish the task. With these realizations in mind, automaticity is now considered to play an essential role in the skill development process and thus has been presented in the SLA literature as an attribute of fluency.

Automaticity has been characterized as fast, unstoppable, effortless, and unconscious processing (Segalowitz, 2003). Anderson’s adaptive control of thought (ACT) theory provides good illustration of the role of automaticity in skill acquisition (Anderson 1983; Anderson & Lebriere, 1998). According to the ACT theory, a skill is acquired by transition from declarative knowledge, knowledge about the skill, to procedural knowledge, demonstration of the skill. This
process involves passing from explicit knowledge of the rules, to repeated application of these rules, to automaticity or “an autonomous stage where the rules are no longer explicit and are executed automatically, implicitly in a fast, coordinated fashion” (Segalowitz, 2003, p. 395). Thus automaticity is shown to be the end result of the process of transition from declarative to procedural knowledge. In SLA this principle is illustrated by transition from knowledge about how the language works to the development of the skill that enables its use.

Automaticity has also been associated with some aspects of fluency. More specifically, lower order processes are more likely to become automatic (Lennon, 2000). These processes include phonological articulation, lexical assessing, as well as morphological and syntactic processing. Segalowitz (2003) also suggests that “to the extent that fluency represents the ability to speak or read quickly, accurately, and without undue hesitation, then automatic execution of certain aspects of L2 performance such as pronunciation, grammatical processing, and word recognition would, by definition, promote fluency” (p. 401). In addition, achieving automaticity in lower order processes allows the speaker to pay more attention to semantic, pragmatic, and sociolinguistic aspects of language, which may ultimately lead to increased level of proficiency.

**Measuring speaking fluency.** Speaking fluency often serves as an indicator of achieving a high level of proficiency in a target language, and as such it has been an important part of classroom instruction. However, with classroom instruction comes the need for assessment. The survey of the literature revealed various approaches to assessing speaking fluency that can be divided into two main categories: in-person oral interviews and computer-based assessment.

Oral interviews have been widely used to measure one’s speaking fluency as well as overall oral proficiency. For instance, the American Council of the Teaching of Foreign Languages (ACTFL) developed Oral Proficiency Interviews (OPI) in 37 different languages,
assessments that are widely used as means of determining the level of “functioning speaking ability” for both academic and professional purposes (www.actfl.org). According to the ACTFL website, these interviews are conducted by highly trained certified testers who lead a “carefully structured conversation” to determine the proficiency level of the interviewee. Throughout the years after the launch of the OPI, the inter-rater reliability and validity concerns have been brought up repeatedly. However, further research and practice have shown that this method of testing is in fact reliable and valid when it comes to measuring oral proficiency (Malone, 2003).

In fact, Malone (2003) reports that the use of OPIs has been constantly increasing. Today academic institutions, government agencies, and private corporations use the OPI for assessment and placement purposes because this one-on-one testing approach provides an accurate assessment of one’s oral proficiency at all levels.

Despite the advantages of using an in-person oral interview as a means of assessing speaking fluency, this method is not efficient when it comes to testing large numbers of subjects. More specifically, because each interview takes between 20 and 30 minutes, one-on-one testing of large groups of individuals would not be feasible, due to the cumulative time requirements of such an undertaking. The solution to this issue was found in using technology. As a case in point, the Computerized Oral Proficiency Interview (COPI) was designed to be administered to larger numbers of subjects with more flexibility in the interview process. This assessment tool is similar in structure to the OPI, but instead of a face-to-face or a phone interview test takers use a computer program to respond to questions and demonstrate their abilities in a foreign language. Kenyon and Malabonga (2001) found that COPI was as effective in assessing oral proficiency as the OPI. Another example of computer-based assessment of oral proficiency is the speaking portion of the Internet based Test of English as a Foreign Language (TOEFL iBT). This test
consists of speaking prompts that draw on the subject’s personal knowledge and integrated tasks where test takers read or listen to a passage and then are asked to summarize or synthesize information from those passages. Several studies have demonstrated that this test is a valid measure of speaking ability for international students in academic settings (Butler, et al 2000; Rosenfeld, Leung, & Oltman, 2001; Xi, 2007).

All of the methods of speaking assessment discussed above are rated by one or several trained individuals who strictly follow the rubrics that are presented to them. The rubrics used in the assessments are usually designed to measure overall oral proficiency and include measures of content, accuracy, and speaking fluency. Unfortunately, speaking fluency as employed in these rubrics has a general description that lacks details. For instance, the speaking rubric for TOEFL iBT at level 3 out of 4 possible levels describes speech as “clear, with some fluidity of expression” (Educational Testing Service, 2004). Similarly, ACTFL (1999) speaking guidelines describe speech at an advanced-low level as “marked by substantial, albeit irregular flow,” and “somewhat strained and tentative” (p.4). Thus, it is evident that such methods of assessment are aimed at evaluating the higher order of fluency, or overall oral competency, and are not detailed enough for assessing the lower order of fluency, or mechanics of speech.

When it comes to measuring lower order of fluency, or mechanical aspects of speech production, the survey of the literature revealed that using automatic speech recognition systems (ASR) has proven to be highly effective. Although ASR come in different software packages, the following measures have been commonly used individually or in combinations and have been shown in various studies to be valid predictors of speaking fluency.

The first measure, the speech rate, is determined by the number of syllables articulated per minute. It has also been calculated by dividing the total number of syllables in the speech
sample by the total time taken to produce the sample including all pauses (Riggenbach, 1991).

The speech rate scores have been consistently shown to correlate with other fluency features and human rater scores (de Jong & Perfetti, 2011; Kormos & Denes, 2004; Cucchiarini, Strik, & Boves, 2002; Riggenbach, 1991).

Articulation rate is another common measure used in the studies of speaking fluency. It is determined by dividing the total number of syllables produced by the amount of time taken to produce them excluding the pause time. Although this measure provides valuable information about one’s speaking fluency, it has been inconsistent in correlating with rater generated scores (de Jong & Perfetti, 2011; Kormos & Denes, 2004; Cucchiarini, Strik, & Boves, 2002; van Gelderen, 1994; Riggenbach, 1991)

The mean length of fluent runs, which is the average number of syllables produced in between pauses, has also been used in measuring speaking fluency. The cut off point for the pauses has been suggested by Towell et al. (1996) at 0.25 seconds and has been used as a standard in other studies.

In addition, phonation/time ratio has been shown to be a predictor of speaking fluency. It is calculated as the percentage of actual time spent speaking as a proportion of the total time taken to produce the utterance. Like speech rate, this measure has been consistent in producing high correlation scores when compared to rater or reading scores (de Jong & Perfetti, 2011; Kormos & Denes, 2004; Cucchiarini, Strik, & Boves, 2002; Towell et al, 1996; van Gelderen, 1994; Riggenbach, 1991; Lennon, 1990).

And finally, the mean length of pauses has been reported in the studies as an additional fluency measure. It is calculated by dividing the total length of pauses above 0.2 seconds by the total number of pauses (Kormos & Denes, 2004). The scores obtained from this measure help

Studies have also been conducted to determine whether there is correlation between machine scores and human ratings. For instance, Neumeyer at al. (2000) used data from 100 American students of French who had read 30 sentences from newspapers and found that the machine scores served as good predictors of human ratings. In another study Cucchiarini, Strik and Boves (2002) used speech samples from an existing database for the Profieltoets, a language proficiency test used by the Dutch National Institute of Educational Measurement. The speech samples were evaluated by trained raters and analyzed by continuous speech recognizer (CSR), a type of ASR, for fluency features. They found that speech rate, phonation/time ratio, mean length of runs, and number of pauses correlated well with human scores with $r$ value ranging from 0.82 to 0.92. On the other hand, articulation rate and mean length of pauses did not exhibit statistically significant correlations. Kormos and Denes (2004) also found strong correlations between human rater scores and computer generated scores when they investigated the speaking fluency of Hungarian learners of English with various levels of proficiency. Both ASR and human raters were used to analyze the speech samples. The results were then compared, revealing that speech rate ($r = 0.87$), phonation/time ratio ($r = 0.80$), and mean length of fluent runs ($r = 0.91$) predicted the human scores well. In this study mean length of pauses and articulation rate scores did not correlate with the human rater scores.

In summary, fluency is a complex concept that involves more than just smoothness and fluidity. In SLA it can refer to both a higher order indicator of overall proficiency, and
underlying elements of language production, such as phonological articulation, speech rate, pauses, hesitations, etc. Speaking fluency can be effectively measured by both human raters and computer generated speech recognition technology. It has also been shown that fluency is closely related to automaticity. The following section will propose that automaticity can be achieved by repetition and frequent practice.

The Role of Repetition in Language Learning

The idea that practice is essential for developing a skill is a commonly held belief in perhaps all or at least most human endeavors. In SLA as well, many activities include deliberate practice aimed at developing automaticity. Referring to reading, Huey (1908/1968) suggests that practice “progressively frees the mind from attention to details and makes facile the total act, shortens the time, and reduces the extent to which consciousness must concern itself with the process” (p. 104).

Practice has often been connected to repetition. As a case in point, MacKay and Bowman (1969) found that after repeating normal or scrambled sentences twelve times, the subjects’ production time of those sentences was significantly reduced. However, today the role of repetition in learning a second or a foreign language has fallen into disfavor as a technique for the language learning experience. Following Chomsky’s takedown of Skinner and the connection of his behaviorist psychology to language, repetition lost its value, becoming collateral damage for the language acquisition field (Bush, Melby, & Lewis, 2010). The loss is related specifically to how the field considers the nature of language and the manner in which it is acquired (Bush, n.d.).
Repetition, mainly in the form of pattern and substitution drills, was widely popular in the 1950s and 1960s as the main means of language learning. It was the core feature of the audio-lingual method (ALM) that dominated academic and military language programs at that time. This method, originally created by Charles Fries in 1940s and later closely associated with Skinner’s behaviorist psychology, maintained that a language is learned through habit formation and pattern analogies. It was believed at the time that repetitive pattern practice would eventually lead to the ability to communicate freely through using analogies to those practiced patterns (Wong & VanPatten, 2003).

After behaviorism was replaced by a Chomskian theory of linguistic competency that rapidly became more popular, repetition was regarded as an ineffective way of practicing and learning a second or foreign language. In the words of Chomsky (1967), the behaviorist point of view “was largely mythology, and … its widespread acceptance is not the result of empirical support, persuasive reasoning, or the absence of a plausible alternative” (p. 142). For experts like Chomsky, analyzing language and developing grammatical knowledge became more important than building analogies and repeating patterns. The role of output was further downplayed by Krashen, whose Input Hypothesis became the foundation of language learning practice for many years after he had introduced the concept as part of his Monitor Theory. He believed that progress in language learning “comes from supplying communicative and comprehensible input, and not from forcing and correcting production” (Krashen, 1982, p.7). This way of thinking has long dominated language acquisition and teaching and created the commonly accepted framework for communicative language teaching (CLT).

Despite the fact that repetition has become unpopular, some researchers maintain that it can be beneficial for development of various language skills. For instance, Leaver, Rifkin, and
Shekhtman (2004) argue that when it comes to acquiring different aspects of more complex languages, such as Russian, repetition is a valuable tool. Similarly, Gatbonton and Segalowitz (2005) point out that teachers of foreign languages generally see more value in structured activities, such as drills, than in games and role-plays that are usually associated with CLT. Other researchers have suggested that repetition and drills could be used for building “precommunicative” knowledge as a stepping stone to free communication (Littlewood, 1980; Rivers, 1981; Hammerly, 1991; Stevick, 1996). Moreover, R. Ellis (1993) and Nunan (1999) agreed that although output practice or drills do not generate enough knowledge and abilities for demonstrating communicative competence in a foreign language, they are essential for developing the underlying skills.

Another opinion in the literature holds that meaningful repetition is an effective technique for acquisition of a foreign or a second language. Hadley (2001) suggests that when a drill is put in context it “links form with meanings that language learners might genuinely want to convey in natural communicative situations” (p. 142). Gatbonton and Segalowitz (2005) expressed concern that CLT does not provide the environment for developing the level of automaticity that is necessary for fluent communication and proposed a methodology that integrates traditional ways of teaching with communicative tasks. In this methodology repetition plays an important role in promoting fluency in production of target utterances within a context that is meaningful to the learners and communicative in nature. No empirical data was shown, however, to support this methodology, and no later studies have been reported.
Repeated Reading as a Method

It has been shown that practice has a positive effect on skill development (Anderson, 1983; MacKay, 1981; Segalowitz, 2003). In one study Samuels (1979) introduced the method of repeated reading as a tool for increasing reading fluency of children who struggle with this skill in schools. According to this method, a student should repeatedly read the same passage until the desired reading rate is achieved, using either silent or oral reading. Samuels (1979) suggested that through repeated reading students may develop automaticity in word recognition and thus free their attention for comprehension.

Repeated reading in L1 studies. A number of studies have been conducted in order to investigate the repeated reading method (Herman 1985; Steventon & Fredrick, 2003; Roundy & Roundy, 2009). Herman (1985) found that students not only improved their reading rates, but also that the length of pauses was shortened as a result of repeated reading. In addition, the percentage of words read correctly increased significantly between first and last readings. Similarly, participants in Steventon and Frederick (2003) and Roundy and Roundy (2009) studies have shown improvement in their reading rates of practiced passages. Overall the studies have shown that the practice of repeated reading is a useful tool for fluency development.

Repeated reading in SLA. Several studies have also been conducted to research effects of repeated reading in SLA. Wiley and Deno (2005) found that repeated reading aloud positively affected performance of English Learners on a state standards test. Taguchi, Takayasu-Mass, and Gorsuch (2004) combined repeated reading and extensive reading in search of effective techniques for developing reading fluency. Their study, which involved native speakers of Japanese learning English, revealed that using repeated reading for practice correlates with reading fluency development. In another study, Gunter (1995) found that reading aloud lead to
improvement in vocabulary development of second grade ESL learners. All in all, the literature supports the idea that repeated reading aloud has been successfully used in SLA. All of the studies cited so far focused on the effects of repeated reading on reading fluency and vocabulary acquisition. Unfortunately there has been very little research done on the connection of repeated reading aloud and speaking fluency in a foreign language.

Research Question

Although there are few references in the literature connecting repeated reading aloud and speaking fluency, the hypothesis is that practicing reading aloud on a regular basis may allow the speaker to maintain or even increase speaking fluency. Some suggestions have been made in the literature that reading aloud may be beneficial for developing learners’ speaking ability. Coniam (1990) in his study demonstrates that there is a high correlation between the speed of reading aloud and speaking task scores. Gibson (2008) argues that reading aloud can be used for multiple purposes in language learning. These include focusing on prosodic features for practicing natural-sounding flow of speech, articulatory and cognitive warm-up, and practicing pronunciation of sounds and intonation (pp. 31-32). More research needs to be done, however, in order to provide empirical evidence of the effects of repeated reading aloud upon speaking fluency. The proposed study will therefore address the following research question:

- What is the effect of repeated reading aloud upon speaking fluency of non-native speakers of Russian as measured by a repeated measures elicited response test?
Chapter 3: Research Design and Procedures

Overview of the Study

This study was conducted to investigate whether there is a connection between repeated reading aloud and speaking fluency. In this context, speaking fluency is defined as “ease and rapidity of speaking, a continuous flow with little hesitation, and a good command of grammar and vocabulary” (Crystal & Varley, 1993, p. 189). More specifically, for the purposes of this study, speaking fluency was measured by calculating computer generated fluency features, as well as surveying overall perception of the speech by three native speakers with a linguistic background.

The repeated measures experimental design of this study involved learners of the Russian language who accessed the reading passages online. All participants were divided into two groups: control and experimental. Participants in the control group read given passages silently, whereas those in the experimental group practiced repeated reading aloud. Each week participants recorded their responses to the speaking prompts. These recordings were later processed and compared in order to answer the research question.

This chapter will introduce materials and instruments used in the study, including pre- and post-treatment surveys, the software package used for administering the treatment, and the post-tests. It will then review the design of the study, describing specific activities performed by subjects in the experimental and control groups. In addition, the chapter will include reports of all iterations of the study and provide a detailed description of the individuals who participated in the experiment. It will conclude with an overview of procedures for data collection and analysis.
Materials and Instruments

The present study was fully administered online and included three distinct phases. The first phase was an online survey that informed the subjects of the format of the study and also collected initial information about them. The treatment was administered via a modified version of the Extending Your Russian software package with separate versions designed for control and experimental groups. Finally, the repeated measures post-test for this study, including speaking prompts and reading passages, was also accessed and administered online. Upon completion of the study, a survey was conducted to learn about participants’ experiences with the study including the reasons for leaving the study.

Pre-treatment survey. The initial steps in this research were recruiting the subjects for the study and collecting information about them. This was done via a Qualtrics survey\(^1\), which was designed to serve several purposes. The first question of the survey was the informed consent agreement for the study (see Appendix A). If after reading through the informed consent, the respondents agreed to participate in the study, they were directed to the second part, which gathered information about their background in the Russian language. More specifically, this part of the survey revealed when and where the participants learned the Russian language, how much time they spent in a Russian speaking country, and what their level of involvement was with the Russian language at the time they began the study. Finally, the participants were asked to evaluate their own ability in the Russian language through self-assessment questions and a small reading comprehension task (see Appendix B).

\(^1\) Qualtrics is an online survey software that allows creating, administering and collecting data online. It is made available to all BYU faculty and students at no charge.
**Extending Your Russian.** The key component of this research was a slightly modified version of the *Extending Your Russian* software package that allowed administering the treatment and collecting data in an efficient way. *Extending Your Russian* was designed by professionals at the Center for Teaching and Learning at Brigham Young University (BYU) in conjunction with Professor David Hart. The purpose of this package is to assist the learners of Russian in maintaining their skills after completing formal education and thus having limited exposure and practice in the Russian language. The program consists of two parts: Building Vocabulary and Spoken Language. The purpose of the first part is not only to retain and build vocabulary knowledge, but also to maintain and develop speaking fluency through the practice of repeated reading aloud. The second part builds on the foundation of Building Vocabulary and is focused on furthering the development of listening and speaking skills.

In the original version of *Extending Your Russian* the first part, Building Vocabulary, is designed to assist the learners of the Russian language with maintaining their vocabulary knowledge and improving their speaking fluency. This segment consists of 25 articles from the Russian news journal *Itogi*. Several different activities are designed to accompany each article that appears 3 times. During the first reading, select words are highlighted and upon selecting them, the user can hear the pronunciation of these words, learn their definitions, and learn about the rules of stress patterns that apply to the specific word. The goal of this reading is for the learner to fully understand the meaning of the passage. Once users fully comprehend the meaning, they can move on to the second part where full sentences are highlighted. By clicking on the highlighted sentence, the learners can hear a native speaker read that particular segment and then practice pronouncing the same sentence. A recording tool allows the users to record themselves and later compare their pronunciation to the native speaker. The focus of the third
and final stage is on building the speed of reading aloud. In this section, each paragraph in the article is marked with the time that it would take a native speaker to read the passage out loud. The users are encouraged here to practice reading aloud repeatedly until their reading speed matches that of the native speaker.

In order to ensure simplicity of delivering and administering the treatment, only a subset of *Extending Your Russian*, which provides passages and additional exercises for repeated reading aloud, was used for the purposes of the present study. The user interface of the program underwent minor changes to facilitate research needs and data collection. For example, only the first part, Building Vocabulary, was made accessible to the participants. Furthermore, not all elements of the Building Vocabulary section of the software were included in this new version. As an additional modification, the second section of each article with highlighted sentences was deleted and the recording tool was moved to the repeated reading section. Tutorials in the complete version were replaced with a comprehensive set of instructions and three post-tests, each appearing only after the participants have completed at least five articles.

Making the activity of repeated reading meaningful and interesting to the participants was an important goal of this study. In addition to enhancing the existing software features, self-evaluation and comprehension questions were designed to meet that goal. When it comes to the existing features, first, the articles for the software package came from an authentic Russian source and the content for these articles was varied in order to meet diverse interests of the participants. Moreover, due to the fact, that the texts have not been simplified, each article was heavily glossed to ensure that the users understood the meaning. For the purposes of the study, the longer articles were broken down into smaller, more manageable segments (see Appendix C
for the full list of the articles used in the study). Finally, each article was followed by a set of self-assessment questions (see Figure 1), and selected articles had additional comprehension questions (see Appendix D).

![Self-Assessment Questions](image)

*Figure 1. Self-assessment questions*

**Post-test.** The post-test for the present study was designed to measure mechanical aspects of the oral fluency of individuals who participated in both control and experimental studies. As part of the need to conform to the repeated measures design, the test was administered on three different occasions. More specifically, at the end of the first and second weeks of the study, the participants responded to short speaking prompts, and at the end of the third week the main post-test was administered.

The post-tests for the first two weeks each included a simple speaking prompt and allowed 60 seconds for preparation and 90 seconds for speaking (see Figure 2). The prompts had a low difficulty level, allowing the subjects to produce speech without spending extra time on
accessing less frequently used vocabulary and grammar. The participants were instructed to record themselves at least once for each prompt. The responses were then stored as Flash media files on the server.

![WEEK 1 TEST](image)

**WEEK 1 TEST**

Connecting...

Imagine that you are in Russia meeting someone new. Tell about yourself in as much detail as possible. You may want to include your age, your occupation, information about your family, the place where you live.

**Preparation:** 60 sec

**Response:** 90 sec

![WEEK 2 TEST](image)

**WEEK 2 TEST**

Connecting...

What was your favorite story from these readings so far? What did you like about it?

**Preparation:** 60 sec

**Response:** 90 sec

**Figure 2.** Speaking prompts for Week 1 and Week 2 tests

As shown in the review of the literature in Chapter 2, the speed of reading aloud correlates well with speaking fluency scores (Coniam, 1990). Therefore, the first part of the final post-test consisted of three short passages meant to be read aloud. The passages for this section were selected from the same news journal, *Itogi* that was used during the experiment. This ensured the consistency of the language style to which the subjects were exposed during and after the study (see Figure 3). The participants were asked to read each passage, which appeared separately, only once. The readings were recorded and then stored on the server for analysis.
Figure 3. Reading passages for the reading aloud section of the post-test

The second part of the final post-test consisted of three speaking prompts that were designed to evaluate fluency in a less formal setting. Because the objective of this test was to measure speaking fluency in normal speech, the prompts included topics that were familiar to the participants and allowed them to draw on personal knowledge and experience. Each prompt, however, had limited times for preparation and response to exclude an opportunity for rehearsed recordings. (see Figure 4).
Week 3 Prompt 1

Do you prefer to live in a small town or in a big city? Explain your answer.
Preparation: 60 sec
Response: 90 sec

Week 3 Prompt 2

What kind of cultural differences have you noticed between Americans and people in the Russian speaking countries? Please include specific examples.
Preparation: 60 sec
Response: 2-3 min

Week 3 Prompt 3

How do you see the world in 15 years? Include ideas about international relations, technological development, moral values, etc.
Preparation: 90 sec
Response: 3-5 min

Figure 4. Speaking prompts for the final post-test

Post-study survey. The post-study survey was designed with the purpose of learning about the reasons why participants decided to leave the study and their overall experience with the research. The survey consisted of five questions (see Appendix E). It was administered via Internet using Qualtrics survey tools.
Design of the Experiment

This study implemented a repeated measures experimental design where all subjects were randomly divided into two groups: control and experimental. This was done by assigning each new participant a number, after which subjects with odd numbers were distributed to the control group and even numbers to the experimental. There was no pre-test in the study, however after completing each week of the treatment the subjects took a post-test consisting of speaking prompts and reading aloud. Both groups had access to the modified version of Extending Your Russian; however, the activities within the program varied depending on the group to which each subject had been previously assigned.

Experimental group. The main treatment for the experimental group was repeated reading aloud. Prior to moving to the repeated reading section, however, the participants in this group were asked to read each article for comprehension. This part was named Reading 1 in the modified software package (see Figure 3.5). To assist participants with comprehension, each article in the first reading was glossed. By clicking on highlighted words, the subjects could see the definition of those words, hear their pronunciation by a native speaker, and learn about specific stress patterns involved in pronouncing them.
Figure 5. Screen shot of the first reading tab of the modified software.

According to the instructions for the experimental group, only after they felt comfortable with the meaning of the article would the participants move to the second reading (see Appendix F). Between the first and second readings, subjects responded to self-evaluation and comprehension questions. During the second part, Reading 2, the subjects had to work on building up the rate of reading aloud (see Figure 6). As a working goal, they were given the time that it would take a native speaker to read the same passage. The participants were asked to re-read each article out loud minimum three times or until their reading time was no more than 10% slower than the standard. At least one of those readings was expected to be recorded. Upon completion of each article, the participants repeated the same activities with the next article until they had been reading regularly for at least three weeks. After each week, participants were instructed to take a short test that became available to them beneath the list of articles.
Figure 6. Screen shot of the second reading tab of the modified software

**Control group.** The subjects in the control group read the same articles as the experimental group; however, they experienced a slightly different interface to the program. There was only one reading section in which the article with highlighted vocabulary words was visible. The subjects were asked to read the article silently and work on their comprehension. There was no limitation on the number of times they could read each article nor were any additional activities that they were asked to perform. Before moving on to the next article, however, the subjects in the control group were also asked to respond to self-evaluation and comprehension questions (see Appendix G). Just like with the experimental group, the subjects in this group took a short speaking test at the end of each week.

**Procedures**

The present study was designed to be conducted online, and due to technical considerations and human factors, four separate iterations took place. The first two iterations
were pilot studies that served a purpose of testing the technical aspects of the experiment. The third iteration of the study did not produce any data for fluency analysis for various reasons. The fourth and final iteration of the study reflected all the changes that were prompted by previous iterations and provided the bulk of the data used in the analysis.

**First iteration.** The first iteration was an initial pilot study that was designed to test the technical aspects of the modified software package. During the Spring Term 2011 students enrolled in Russian 321 course were invited to participate in a limited version of the modified software for the purpose of determining whether the programming of the software package was functioning properly. Even though the students were offered extra credit for participation, three out of seventeen students participated and only one completed the study. This initial pilot study did not reveal any technical issues with the instrument.

**Second iteration.** Because the first iteration of the study did not attract a sufficient number of participants, a new pilot study was launched in June 2011. The second group consisted of the author’s personal acquaintances who had previously served missions for the Church of Jesus Christ of Latter-Day Saints (LDS) in Voronezh and other cities of Russia. These individuals were initially contacted via Facebook, and invited to participate in the study. Out of 117 people that were contacted, 17 individuals agreed to participate in the experiment. The individuals who completed the online survey received a link to the instrument by e-mail. They were instructed to log in using the nickname they created during the survey. Upon logging in, each participant was randomly assigned to either the experimental or the control group. Over a period of five weeks, the activities of the participants were logged and recorded. However, with time it became evident that fewer and fewer people were logging in to perform the reading activities. In the end, only one post-test was collected from this iteration of the study.
**Recruiting subjects for the final iterations.** To help in locating participants for the final iteration and after approval of the Institutional Review Board the Brigham Young University (BYU) Institutional Assessment and Analysis provided a database export with the information of current and former BYU students who listed Russian as their second language and who served a Russian speaking mission. This export included information on 3,327 individuals. However, in order to generate the final list of prospective subjects, this database had to undergo several rounds of elimination. First of all, people who listed Russian as their native language, or who grew up in a Russian-speaking country were removed from the list. Second, people who listed missions in places where Russian is not the dominant language or is not spoken at all were also eliminated. Finally, records with missing or outdated information could not be used for the study. As a result of this process of elimination 2,475 records remained in the database.

**Third iteration.** Recruiting for the third wave of the study began in July 2011 and continued through August 2011. The database that was obtained from the BYU Institutional Assessment and Analysis was divided into 14 manageable contact lists organized in alphabetical order. Every week an e-mail message with the invitation to participate in the research study was sent to 180-200 people from the first half of the lists. As in previous iterations, despite the high level of interest in the beginning of this study, only two post-tests were collected in the end of this third iteration. This was attributed to two factors at the time. First, it became evident that the server that was used for data collection was unstable and some crucial data was never recorded. Second, the participants who did agree to participate quickly lost interest in the activities and did not stay with the study long enough to complete the post-test.

**Final iteration.** The proceedings of the second and third iterations of the experiment not only revealed technical issues with the software package, but also prompted changes in the
design of the study. The first three iterations were conducted as post-test only studies over the course of five weeks and did not produce sufficient data for analysis. The final iteration maintained the original experimental design; however, the repeated measures tests administered at the end of each week replaced the final post-test. In addition, the length of the study was reduced from five weeks down to three weeks to increase the likelihood that participants would complete the study. At the beginning of the Winter Semester 2012, after all necessary changes were made to the instruments and all technical issues were eliminated from the server, individuals from the second half of the database list received invitations to participate in the study. Once the study began, the activities of the participants were logged and carefully followed. Participants who had not logged on to the website for more than two days received reminder e-mail messages. Additionally, several participants agreed to receive daily reminders for the study regardless of their recorded activity. Reminder messages were also sent for the weekly tests. Despite these efforts to improve the study and to provide timely and repetitive reminders only 15 people completed all three weeks of the study.

Subjects

The process of selecting the subjects for the study followed several important criteria. More specifically, this group consisted of individuals who had previously studied the Russian language and spent between 18 months and two years in a Russian speaking country, and who were assumed to be at intermediate to low advanced level of proficiency on the ACTFL scale. However, taking into consideration the fact that some of these participants have not had regular language practice in several years, their actual level of proficiency may have been lower at the time of the study. In addition, it was assumed that the participants of this study were not enrolled
in any Russian classes during the duration of the experiment. Applying these criteria to the sample population significantly lowered the possibility of intervening variables that could have skewed the results of the experiment.

After eliminating irrelevant and unusable records from the database, 2,475 e-mail messages with invitations to participate in the study were sent. The chart in Figure 5 represents the response level for this study. Of 265 respondents to the initial survey, 227 respondents agreed to participate in the study, and 158 created logins to the instrument and read at least one article. However, only 16 individuals ultimately completed the full experiment, amounting to 6.6% of the total number of people who signed the informed consent and agreed to participate in the study.

![Figure 7](chart.png)

*Figure 7. The N of people participating in the various stages of the study*

All of the participants who completed the study reported that they spent between 15 and 25 months in Russia and had been back in the United States on average for about five years. Eight were male and seven were female. One participant in the experimental group was taking a Russian class while the rest did not have any formal training at the time of the experiment.
However, 7 people reported that they were using the Russian language at least two-to-three times a month or more. These people were evenly distributed in the control and experimental groups. The rest admitted that they rarely use the language. Overall most participants evaluated their ability in Russian as fair or good, and only two perceived their speaking ability as poor or bad (see Table 3.1).

Table 3.1

| Subjects’ self-perception of overall proficiency in the Russian Language |
|-----------------|-----|-----|-----|-----|-----|-----|
|                 | Very Bad | Bad | Poor | Fair | Good | Very Good |
| Speaking        | 0     | 1   | 1    | 8    | 5    | 0          |
| Reading         | 0     | 0   | 0    | 7    | 8    | 0          |
| Listening       | 0     | 1   | 1    | 6    | 6    | 1          |
| Writing         | 0     | 1   | 3    | 9    | 2    | 0          |
| Grammar         | 0     | 0   | 3    | 7    | 4    | 1          |

*Note:* The results come from 15 subjects who completed the study.

**Data collection and analysis**

The purpose of the experiment was to collect speech samples in order to determine whether there is a connection between the practice of repeated reading aloud and speaking fluency. The data collection was done by instructing participants to record their responses to the speaking prompts and read aloud paragraphs in the post-tests. Voice recognition software, or ASR, and three trained raters provided the data used in the analysis. The voice recognition software provided data about mechanical measures of lower order of fluency, such as speech and
articulation rate, length and number of pauses, and phonation/time ratio, such as discussed below. On the other hand, the raters gave a holistic score for the use of grammar and vocabulary, as well as overall flow of speech from the perspective of a native speaker.

**Human raters.** Three individuals rated the same selected speech samples in order to provide information on the speaking fluency of the participants from the perspective of a native speaker. All three raters are native speakers of the Russian language with linguistic backgrounds and experience in rating speech samples. One graduated from Brigham Young University with a Master’s degree in Teaching English as a Second Language. The other two raters are current students in the Second Language Acquisition and Teaching graduate program at BYU. The raters assigned scores ranging from one to four to selected speech samples for each participant relying on an adaptation of the speaking rubric that is used by the BYU English Language Center for grading Level Achievement Tests (see Appendix H).

**Computer generated analysis.** In addition to human ratings, the samples were analyzed using Praat, an ASR software package that has been previously used in fluency research. It is useful because it provides numerical data on various features of speech that have been linked to speaking fluency and that correlate well with human ratings (de Jong & Wempe, 2009; de Jong & Perfetti, 2011). For the purposes of this study, the following fluency features were extracted and analyzed from the post-test recordings.

- Phonation/time ratio (PTR) shows the percentage of time spent speaking as a proportion of the total time taken to produce the speech sample.
- Speech rate (SR) is calculated in Praat by dividing the total number of syllables uttered by the total duration of the speech sample.
• Articulation rate (AR) is the number of syllables divided by the phonation time, or the time spent speaking minus the time spent on pauses.

• Average Syllable Duration (ASD) is calculated by dividing the phonation time, or actual time spent speaking, by the total number of syllables in the speech sample.

• Mean length of pauses (MLP) is calculated by dividing the silent time by the number of pauses in the speech sample.

• Number of syllables (NSYL) and number of pauses (NP) were also used in analysis.

The Praat software is capable of extracting these features because it is scripted to recognize syllable nuclei and silent runs (de Jong & Wempe, 2009). In order for the program to work properly, the sound quality of the speech files had to be as high as possible. In fact, it became evident that when recorded files had excess background noise, Praat was not be able to register any pausing and thus gave an inaccurate reading. For this reason, several files were edited using Audacity software for sound amplification and noise reduction. This software also allowed for trimming the tracks that had lengthy silence at the end, which allowed for a more accurate reading.

Statistical analysis. The data from raters and Praat ASR were further analyzed using several statistical procedures. First, the inter-rater reliability was established using Intraclass Correlation Coefficient (ICC) and Pearson correlation. Second, the independent sample t-test allowed comparing the means of rater scores between the control and experimental groups. Furthermore, computer-generated scores were analyzed using one-way ANOVA. In addition, ANOVA with repeated measures provided information on the difference in scores between the control and experimental groups that took place over time. Finally, Pearson correlation measures
were used to establish relationships between rater scores, speaking test scores, and reading test scores.
Chapter 4: Results

This study was conducted to determine whether there is a connection between repeated reading aloud and speaking fluency. As part of the weekly tests administered during each of the three weeks in the study, participants recorded themselves responding to speaking prompts and reading three passages out loud. These speech samples were rated by three trained individuals and processed using automatic speech recognition (ASR) software and techniques. The results were then analyzed using various statistical measures in the Statistical Package for the Social Sciences (SPSS). This chapter will describe the results of the statistical analysis of speaking scores and address the issue of attrition that arose during the study.

Statistical Analysis

The final sample size for the statistical analysis included data from 16 subjects. Seven subjects were in the experimental group, and nine were from the control group. To determine whether there was a difference in speaking fluency between the groups from the perspective of the human raters, an independent sample $t$-test was used where the mean of scores assigned by raters was the dependent variable. Human rater scores were first analyzed for inter-rater reliability, however. The scores identified as best reflecting speaking fluency and available from the Praat analysis were compared using a one-way ANOVA to compare the mean of all scores, and a repeated-measures ANOVA to determine whether there was a difference in scores over time. A series of correlation tests were also performed to explore relationships between different measures of speaking fluency.

**Rater scores.** Three native speakers of the Russian language with graduate level education in language related fields evaluated the speech samples from both control and
experimental groups. They then assigned a score ranging from one to four for grammar, pronunciation, content, and overall perception of the speech for each of the subjects in the final group. For the purposes of the statistical analysis the means of all scores assigned by each rater for each participant were used. To ensure inter-rater reliability, the scores from the three raters were compared using interclass correlation coefficient (ICC), a procedure commonly used to measure correlation when the number of raters exceeds two. As shown in Table 4.1, the differences in the results from ICC analysis, for the scores of all three raters were statistically significant ($p = .005$). The low ICC (.270) indicates that the rater scores were inconsistent with each other and therefore have low inter-rater reliability.

Table 4.1

| Intraclass Correlation Coefficient for rater scores |
|-----------------------------------------------|-----------------|------------------|------------------|------------------|------------------|------------------|------------------|
|                                               | Intraclass Correlation | 95% Confidence Interval | F Test with True Value 0 |                      |                  |                  |                  |
|                                               |                      | Lower Bound | Upper Bound | Value | df1 | df2 | Sig  |
| Single Measures                               | .270                | .008        | .584       | 3.037 | 15 | 30 | .005 |

To better understand the relationship between rater scores, Pearson correlation analysis was performed on the results of the three raters. The results, as shown in Table 4.2, reveal that scores from Rater 1 were inconsistent with scores from the other two raters: Rater 2 ($r = .390$), Rater 3 ($r = .075$). At the same time, Raters 2 and 3 assigned similar scores exhibiting strong correlation of $r = .685$, which was also statistically significant ($p = 0.003$). Based on these results, only the scores from Raters 2 and 3 were used for further analysis and comparison.
Table 4.2

*Rater scores correlation*

<table>
<thead>
<tr>
<th>Pearson Correlation</th>
<th>Rater 1</th>
<th>Rater 2</th>
<th>Rater 3</th>
</tr>
</thead>
<tbody>
<tr>
<td>Rater 1</td>
<td>1</td>
<td>.390</td>
<td>.075</td>
</tr>
<tr>
<td>Rater 2</td>
<td>.390</td>
<td>1</td>
<td>.685**</td>
</tr>
<tr>
<td>Rater 3</td>
<td>.075</td>
<td>.685**</td>
<td>1</td>
</tr>
</tbody>
</table>

*Note.* (***) Correlation is significant at the $p < 0.01$ level (2-tailed).

Upon establishing inter-rater reliability the mean of scores from Raters 2 and 3 was used to compare experimental and control groups. Although Table 4.3 indicates that the scores for participants in the experimental group were slightly higher than those of the control group, it is not possible to assume that this difference in scores was a result of treatment and did not occur by chance due to the low number of the participants in the groups and the small difference in scores between the groups (0.23). Therefore, an independent samples $t$-test would reveal whether this difference has statistical significance.

Table 4.3

*Experimental and control group mean scores from raters*

<table>
<thead>
<tr>
<th>Group</th>
<th>N</th>
<th>Mean</th>
<th>Std. Deviation</th>
<th>Std. Error Mean</th>
</tr>
</thead>
<tbody>
<tr>
<td>Experimental</td>
<td>7</td>
<td>3.0657</td>
<td>.59761</td>
<td>.22588</td>
</tr>
<tr>
<td>Control</td>
<td>9</td>
<td>2.8289</td>
<td>.60675</td>
<td>.20225</td>
</tr>
</tbody>
</table>

The independent samples $t$-test results indicate that the difference between experimental and control group rater scores is not statistically significant. In fact, the $t$ value is only .780 with significance level at $p = 0.449$ (see Table 4.4). Therefore, human rater scores support the null hypothesis and indicate that there is no difference in speaking fluency between the control and experimental groups.
Table 4.4

Independent samples t-test for human rater scores

<table>
<thead>
<tr>
<th>T</th>
<th>df</th>
<th>Sig. (2-tailed)</th>
<th>Mean Difference</th>
<th>Std. Error Difference</th>
</tr>
</thead>
<tbody>
<tr>
<td>.780</td>
<td>14</td>
<td>.449</td>
<td>.2363</td>
<td>.3038</td>
</tr>
</tbody>
</table>

Notes. The Levene’s test for equality of variances shows significance at .759, which means that data had equal variances and was normally distributed.

**ASR speaking scores results.** In contrast with the rater scores, the analysis of speaking scores obtained from Praat revealed statistically significant difference for some of the fluency features. These results were obtained by comparing the means of all speaking scores between the groups, and performing a repeated measures ANOVA. Descriptions for fluency features used for this part of analysis are outlined in Table 4.5

Table 4.5

Abbreviations used in the statistical analysis of fluency features

<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Definition</th>
</tr>
</thead>
<tbody>
<tr>
<td>NSYL</td>
<td>Number of syllables in the speech sample</td>
</tr>
<tr>
<td>NP</td>
<td>Number of pauses</td>
</tr>
<tr>
<td>MLP</td>
<td>Mean length of pauses</td>
</tr>
<tr>
<td>PTR</td>
<td>Phonation/time ratio</td>
</tr>
<tr>
<td>SR</td>
<td>Speech rate</td>
</tr>
<tr>
<td>AR</td>
<td>Articulation rate</td>
</tr>
<tr>
<td>ASD</td>
<td>Average syllable duration</td>
</tr>
</tbody>
</table>

For the first part of this analysis the means of fluency scores from all five speaking tests were compared. Table 4.6 displays the mean scores for speaking tests for experimental and control groups. It is evident here that the subjects in the control group had longer pauses than those in the experimental group. Both speech rate and articulation rate means also show that
subjects in the control group were slower. The average syllable duration is slightly shorter for the subjects in the experimental group.

Table 4.6

Means for speaking scores on fluency features

<table>
<thead>
<tr>
<th>Group</th>
<th>NSYL</th>
<th>NP</th>
<th>MLP</th>
<th>PTR</th>
<th>SR</th>
<th>AR</th>
<th>ASD</th>
</tr>
</thead>
<tbody>
<tr>
<td>Experimental</td>
<td>322.628</td>
<td>45.866</td>
<td>0.768</td>
<td>71.675</td>
<td>2.828</td>
<td>3.890</td>
<td>0.261</td>
</tr>
<tr>
<td>Control</td>
<td>275.052</td>
<td>44.659</td>
<td>1.124</td>
<td>59.015</td>
<td>2.190</td>
<td>3.780</td>
<td>0.272</td>
</tr>
<tr>
<td>Total</td>
<td>295.867</td>
<td>45.187</td>
<td>0.968</td>
<td>64.554</td>
<td>2.469</td>
<td>3.828</td>
<td>0.267</td>
</tr>
</tbody>
</table>

Notes. Experimental group N = 7, Control group N = 9, Total N = 16

The means of speaking fluency scores were then compared using one-way ANOVA. The results, as shown in Table 4.7, reveal that the difference in scores that was observed between the control and experimental groups did not occur by chance for some of the fluency features. Under the assumption of homogeneity of variances, three of the fluency features exhibit statistical significance at $p < 0.05$: mean length of pauses ($p = 0.013$), speech rate ($p = 0.018$), and phonation/time ratio ($p = 0.042$).

One of the participants in the experimental group reported using the Russian language on a daily basis and taking a Russian class during the study. With such a small group of participants the scores of this individual could have skewed the overall analysis. For this reason a separate comparison of the speaking scores of the experimental and control groups without the scores of the outlier was performed. The results of this analysis are consistent with the findings reported for the whole group. As shown in Appendix I there is a statistically significant difference between the two groups in mean length of pauses ($p = 0.030$), phonation/time ratio ($p = 0.08$), and speech rate ($p = 0.43$) with the experimental group outperforming the control group.
Table 4.7

One-way ANOVA comparison for speaking fluency scores

<table>
<thead>
<tr>
<th>Variable</th>
<th>Sum of Squares</th>
<th>f</th>
<th>Mean Square</th>
<th>F</th>
<th>Sig.</th>
</tr>
</thead>
<tbody>
<tr>
<td>NSYL</td>
<td>Between Groups</td>
<td>8912.692</td>
<td>1</td>
<td>8912.692</td>
<td>.523</td>
</tr>
<tr>
<td></td>
<td>Within Groups</td>
<td>81923.569</td>
<td>4</td>
<td>5851.683</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Total</td>
<td>90836.260</td>
<td>5</td>
<td></td>
<td></td>
</tr>
<tr>
<td>NP</td>
<td>Between Groups</td>
<td>5.740</td>
<td>1</td>
<td>5.740</td>
<td>0.026</td>
</tr>
<tr>
<td></td>
<td>Within Groups</td>
<td>3135.583</td>
<td>4</td>
<td>223.970</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Total</td>
<td>3141.323</td>
<td>5</td>
<td></td>
<td></td>
</tr>
<tr>
<td>MLP</td>
<td>Between Groups</td>
<td>0.497</td>
<td>1</td>
<td>0.497</td>
<td>8.047</td>
</tr>
<tr>
<td></td>
<td>Within Groups</td>
<td>0.865</td>
<td>4</td>
<td>0.062</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Total</td>
<td>1.362</td>
<td>5</td>
<td></td>
<td></td>
</tr>
<tr>
<td>PTR</td>
<td>Between Groups</td>
<td>631.099</td>
<td>1</td>
<td>631.099</td>
<td>4.986</td>
</tr>
<tr>
<td></td>
<td>Within Groups</td>
<td>1772.064</td>
<td>4</td>
<td>126.576</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Total</td>
<td>2403.163</td>
<td>5</td>
<td></td>
<td></td>
</tr>
<tr>
<td>SR</td>
<td>Between Groups</td>
<td>1.605</td>
<td>1</td>
<td>1.605</td>
<td>7.251</td>
</tr>
<tr>
<td></td>
<td>Within Groups</td>
<td>3.099</td>
<td>14</td>
<td>0.221</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Total</td>
<td>4.704</td>
<td>15</td>
<td></td>
<td></td>
</tr>
<tr>
<td>AR</td>
<td>Between Groups</td>
<td>0.048</td>
<td>1</td>
<td>0.048</td>
<td>0.250</td>
</tr>
<tr>
<td></td>
<td>Within Groups</td>
<td>2.669</td>
<td>14</td>
<td>0.191</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Total</td>
<td>2.717</td>
<td>15</td>
<td></td>
<td></td>
</tr>
<tr>
<td>ASD</td>
<td>Between Groups</td>
<td>0.000</td>
<td>1</td>
<td>0.000</td>
<td>0.288</td>
</tr>
<tr>
<td></td>
<td>Within Groups</td>
<td>0.022</td>
<td>14</td>
<td>0.002</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Total</td>
<td>0.023</td>
<td>15</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

An ANOVA with repeated measures was necessary to perform because the participants took three different tests during the study. This analysis provided information about the changes in the speaking fluency that could have occurred over the course of the experiment. Tests from Week 1 and Week 2 consisted of one simple speaking prompt each. However, the final test included three speaking prompts varying in the levels of difficulty from low to moderately high. Listing the results of all five speaking prompts as separate measures in the repeated measures ANOVA would have produced inaccurate results because the last three tests occurred on the same day and not over time. Therefore, two separate analyses were performed with this data set.
The first analysis included test scores from Week 1 and Week 2 as well as scores from the first speaking prompt from Week 3 test that matched the level of difficulty of the first two tests. The results for this analysis can be found under Group section in Table 4.8. Even with the low sample size, statistical significance ($p < 0.05$) in the difference between group scores was observed for mean length of pauses at $p = 0.020$. Phonation/time ratio $p$ value was only slightly higher than 0.05 at $p = 0.058$. Speech rate scores also approached statistical significance at 0.063. Scores from other fluency features did not produce any statistically significant results.

Table 4.8

_Between-groups comparison for repeated measures ANOVA: Week 1, 2, and Week 3 first test_

<table>
<thead>
<tr>
<th>Measure</th>
<th>Type III Sum of Squares</th>
<th>df</th>
<th>Mean Square</th>
<th>F</th>
<th>Sig.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Intercept</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>NSYL</td>
<td>1805097.299</td>
<td>1</td>
<td>1805097.299</td>
<td>242.404</td>
<td>.000</td>
</tr>
<tr>
<td>NP</td>
<td>33902.505</td>
<td>1</td>
<td>33902.505</td>
<td>128.689</td>
<td>.000</td>
</tr>
<tr>
<td>MLP</td>
<td>28.962</td>
<td>1</td>
<td>28.962</td>
<td>345.833</td>
<td>.000</td>
</tr>
<tr>
<td>PTR</td>
<td>183026.062</td>
<td>1</td>
<td>183026.062</td>
<td>576.703</td>
<td>.000</td>
</tr>
<tr>
<td>SR</td>
<td>269.248</td>
<td>1</td>
<td>269.248</td>
<td>339.246</td>
<td>.000</td>
</tr>
<tr>
<td>AR</td>
<td>569.256</td>
<td>1</td>
<td>569.256</td>
<td>698.483</td>
<td>.000</td>
</tr>
<tr>
<td>ASD</td>
<td>2.796</td>
<td>1</td>
<td>2.796</td>
<td>420.231</td>
<td>.000</td>
</tr>
<tr>
<td>Group</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>NSYL</td>
<td>11797.812</td>
<td>1</td>
<td>11797.812</td>
<td>1.584</td>
<td>.234</td>
</tr>
<tr>
<td>NP</td>
<td>12.659</td>
<td>1</td>
<td>12.659</td>
<td>.048</td>
<td>.831</td>
</tr>
<tr>
<td>MLP</td>
<td>.620</td>
<td>1</td>
<td>.620</td>
<td>7.400</td>
<td>.020</td>
</tr>
<tr>
<td>PTR</td>
<td>1419.272</td>
<td>1</td>
<td>1419.272</td>
<td>4.472</td>
<td>.058</td>
</tr>
<tr>
<td>SR</td>
<td>3.392</td>
<td>1</td>
<td>3.392</td>
<td>4.274</td>
<td>.063</td>
</tr>
<tr>
<td>AR</td>
<td>.017</td>
<td>1</td>
<td>.017</td>
<td>.021</td>
<td>.888</td>
</tr>
<tr>
<td>ASD</td>
<td>.000</td>
<td>1</td>
<td>.000</td>
<td>.036</td>
<td>.854</td>
</tr>
<tr>
<td>Error</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>NSYL</td>
<td>81913.111</td>
<td>11</td>
<td>7446.646</td>
<td></td>
<td></td>
</tr>
<tr>
<td>NP</td>
<td>2897.905</td>
<td>11</td>
<td>263.446</td>
<td></td>
<td></td>
</tr>
<tr>
<td>MLP</td>
<td>.921</td>
<td>11</td>
<td>.084</td>
<td></td>
<td></td>
</tr>
<tr>
<td>PTR</td>
<td>3491.029</td>
<td>11</td>
<td>317.366</td>
<td></td>
<td></td>
</tr>
<tr>
<td>SR</td>
<td>8.730</td>
<td>11</td>
<td>.794</td>
<td></td>
<td></td>
</tr>
<tr>
<td>AR</td>
<td>8.965</td>
<td>11</td>
<td>.815</td>
<td></td>
<td></td>
</tr>
<tr>
<td>ASD</td>
<td>.073</td>
<td>11</td>
<td>.007</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Data set for the second analysis included scores from Week 1 and Week 2 tests and the mean of scores from all three tests in the Week 3 test. The findings for this analysis were somewhat different from findings in previous analyses performed for this study. In this case the scores for mean length of pauses ($p = 0.056$), phonation/time ratio ($p = 0.084$), and speech rate ($p = 0.063$) are only approaching statistical significance (see Table 4.9). Explanation of this phenomenon can be found in the low sample size or difference in the levels of difficulty of the speaking tasks. However, these findings are consistent with findings from previously performed analyses in this study in identifying mean length of pauses, phonation/time ratio, and speech rate as fluency features that are most affected by the treatment.
Table 4.9

*Between-groups comparison for repeated measures ANOVA: Week 1, 2, and Week 3 mean scores*

<table>
<thead>
<tr>
<th>Measure</th>
<th>Type III Sum of Squares</th>
<th>df</th>
<th>Mean Square</th>
<th>F</th>
<th>Sig.</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Intercept</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>NSYL</td>
<td>2741191.736</td>
<td>1</td>
<td>2741191.736</td>
<td>246.281</td>
<td>.000</td>
</tr>
<tr>
<td>NP</td>
<td>57280.252</td>
<td>1</td>
<td>57280.252</td>
<td>111.501</td>
<td>.000</td>
</tr>
<tr>
<td>MLP</td>
<td>30.310</td>
<td>1</td>
<td>30.310</td>
<td>305.866</td>
<td>.000</td>
</tr>
<tr>
<td>PTR</td>
<td>177534.292</td>
<td>1</td>
<td>177534.292</td>
<td>540.233</td>
<td>.000</td>
</tr>
<tr>
<td>SR</td>
<td>259.760</td>
<td>1</td>
<td>259.760</td>
<td>341.849</td>
<td>.000</td>
</tr>
<tr>
<td>AR</td>
<td>463.059</td>
<td>1</td>
<td>463.059</td>
<td>657.931</td>
<td>.000</td>
</tr>
<tr>
<td>ASD</td>
<td>2.812</td>
<td>1</td>
<td>2.812</td>
<td>398.840</td>
<td>.000</td>
</tr>
<tr>
<td><strong>Group</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>NSYL</td>
<td>9768.010</td>
<td>1</td>
<td>9768.010</td>
<td>.878</td>
<td>.369</td>
</tr>
<tr>
<td>NP</td>
<td>21.962</td>
<td>1</td>
<td>21.962</td>
<td>.043</td>
<td>.840</td>
</tr>
<tr>
<td>MLP</td>
<td>.450</td>
<td>1</td>
<td>.450</td>
<td>4.541</td>
<td>.056</td>
</tr>
<tr>
<td>PTR</td>
<td>1187.236</td>
<td>1</td>
<td>1187.236</td>
<td>3.613</td>
<td>.084</td>
</tr>
<tr>
<td>AR</td>
<td>.103</td>
<td>1</td>
<td>.103</td>
<td>.146</td>
<td>.710</td>
</tr>
<tr>
<td>ASD</td>
<td>.001</td>
<td>1</td>
<td>.001</td>
<td>.103</td>
<td>.755</td>
</tr>
<tr>
<td><strong>Error</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>NSYL</td>
<td>122433.75</td>
<td>11</td>
<td>11130.342</td>
<td></td>
<td></td>
</tr>
<tr>
<td>NP</td>
<td>5650.899</td>
<td>11</td>
<td>513.718</td>
<td></td>
<td></td>
</tr>
<tr>
<td>MLP</td>
<td>1.090</td>
<td>11</td>
<td>.099</td>
<td></td>
<td></td>
</tr>
<tr>
<td>PTR</td>
<td>3614.882</td>
<td>11</td>
<td>328.626</td>
<td></td>
<td></td>
</tr>
<tr>
<td>SR</td>
<td>8.359</td>
<td>11</td>
<td>.760</td>
<td></td>
<td></td>
</tr>
<tr>
<td>AR</td>
<td>7.742</td>
<td>11</td>
<td>.704</td>
<td></td>
<td></td>
</tr>
<tr>
<td>ASD</td>
<td>.078</td>
<td>11</td>
<td>.007</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Repeated reading aloud results. As discussed in Chapter 2 reading aloud scores can predict the scores for speaking fluency. In the present study the subjects were asked to read three passages out loud as part of the final test. The scores from these passages were then processed using Praat and results for control and experimental groups were compared. Table 4.10 displays the means of fluency feature scores for the two groups, again on the objectively derived scores obtained from Praat. The simple comparison of means reveals that there is a difference in performance between the experimental and control groups. It is evident that the mean length of pauses was shorter, and the phonation/time ratio and speech rate were higher for the subjects in the experimental group.

Table 4.10

Means for reading aloud fluency features

<table>
<thead>
<tr>
<th>Group</th>
<th>NSYL</th>
<th>NP</th>
<th>MLP</th>
<th>PTR</th>
<th>SR</th>
<th>AR</th>
<th>ASD</th>
</tr>
</thead>
<tbody>
<tr>
<td>Experimental</td>
<td>445.285</td>
<td>58.095</td>
<td>0.708</td>
<td>78.816</td>
<td>3.064</td>
<td>3.888</td>
<td>0.260</td>
</tr>
<tr>
<td>Control</td>
<td>357.407</td>
<td>49.296</td>
<td>0.766</td>
<td>71.172</td>
<td>2.818</td>
<td>3.954</td>
<td>0.254</td>
</tr>
<tr>
<td>Total</td>
<td>395.854</td>
<td>53.146</td>
<td>0.741</td>
<td>74.516</td>
<td>2.926</td>
<td>3.925</td>
<td>0.257</td>
</tr>
</tbody>
</table>

Notes. Experimental group N = 7, Control group N = 9, Total N = 16

These findings were further investigated using one-way ANOVA. The results for this analysis, as shown in Table 4.11, revealed that the differences described above were not statistically significant. In this set of data, phonation/time ratio scores were the only ones approaching significance at \( p = 0.080 \). However, when each set of scores was analyzed separately, the differences in mean length of pauses scores (\( p_1 = 0.041; p_2 = 0.044 \)) for readings one and two were statistically significant at \( p < 0.05 \), and phonation time ratio scores (\( p_1 = 0.052; p_2 = 0.060 \)) were near the significance level (see Appendix J).
Table 4.11

*One-way ANOVA comparison for reading aloud fluency scores*

<table>
<thead>
<tr>
<th>Group</th>
<th>Sum of Squares</th>
<th>f</th>
<th>Mean Square</th>
<th>F</th>
<th>Sig.</th>
</tr>
</thead>
<tbody>
<tr>
<td>NSYL</td>
<td>Between Groups</td>
<td>30407.725</td>
<td>1</td>
<td>30407.725</td>
<td>.089</td>
</tr>
<tr>
<td></td>
<td>Within Groups</td>
<td>203780.046</td>
<td>4</td>
<td>14555.718</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Total</td>
<td>234187.771</td>
<td>15</td>
<td></td>
<td></td>
</tr>
<tr>
<td>NP</td>
<td>Between Groups</td>
<td>304.847</td>
<td>1</td>
<td>304.847</td>
<td>0.342</td>
</tr>
<tr>
<td></td>
<td>Within Groups</td>
<td>12474.480</td>
<td>14</td>
<td>891.034</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Total</td>
<td>12779.326</td>
<td>15</td>
<td></td>
<td></td>
</tr>
<tr>
<td>MLP</td>
<td>Between Groups</td>
<td>0.013</td>
<td>1</td>
<td>0.013</td>
<td>.371</td>
</tr>
<tr>
<td></td>
<td>Within Groups</td>
<td>0.496</td>
<td>14</td>
<td>0.035</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Total</td>
<td>0.509</td>
<td>15</td>
<td></td>
<td></td>
</tr>
<tr>
<td>PTR</td>
<td>Between Groups</td>
<td>230.123</td>
<td>1</td>
<td>230.123</td>
<td>3.553</td>
</tr>
<tr>
<td></td>
<td>Within Groups</td>
<td>906.828</td>
<td>14</td>
<td>64.773</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Total</td>
<td>1136.952</td>
<td>15</td>
<td></td>
<td></td>
</tr>
<tr>
<td>SR</td>
<td>Between Groups</td>
<td>0.238</td>
<td>1</td>
<td>0.238</td>
<td>1.346</td>
</tr>
<tr>
<td></td>
<td>Within Groups</td>
<td>2.481</td>
<td>14</td>
<td>0.177</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Total</td>
<td>2.719</td>
<td>15</td>
<td></td>
<td></td>
</tr>
<tr>
<td>AR</td>
<td>Between Groups</td>
<td>0.017</td>
<td>1</td>
<td>0.017</td>
<td>0.177</td>
</tr>
<tr>
<td></td>
<td>Within Groups</td>
<td>1.371</td>
<td>14</td>
<td>0.098</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Total</td>
<td>1.389</td>
<td>15</td>
<td></td>
<td></td>
</tr>
<tr>
<td>ASR</td>
<td>Between Groups</td>
<td>0.000</td>
<td>1</td>
<td>0.000</td>
<td>0.311</td>
</tr>
<tr>
<td></td>
<td>Within Groups</td>
<td>0.007</td>
<td>4</td>
<td>0.000</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Total</td>
<td>0.007</td>
<td>15</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**Relationships between scores.** For the final stage of the analysis the scores from human raters, ASR generated speaking tests, and reading aloud tests were compared to determine whether there is any relationship between these scores. Based on the results from the previous analysis, only mean length of pauses, phonation/time ratio, and speech rate scores were used for this comparison. The results, as shown in Table 4.12, reveal that there is no correlation between the rater and computer generated scores. However, there is a strong relationship between scores obtained from speaking tests and reading aloud for phonation/time ratio ($r = 0.707, p = 0.003$) and speech rate ($r = 0.732, p = 0.002$) as predicted by Coniam (1991).
Table 4.1

*Pearson correlation between scores*

<table>
<thead>
<tr>
<th></th>
<th>Raters</th>
<th>Reading MLP</th>
<th>Reading PTR</th>
<th>Reading SR</th>
</tr>
</thead>
<tbody>
<tr>
<td>Raters</td>
<td>Pearson Correlation</td>
<td>1</td>
<td>.168</td>
<td>.162</td>
</tr>
<tr>
<td></td>
<td>Sig. (2-tailed)</td>
<td>.535</td>
<td>.549</td>
<td>.197</td>
</tr>
<tr>
<td>Speaking MLP</td>
<td>Pearson Correlation</td>
<td>-.122</td>
<td>.261</td>
<td>-.416</td>
</tr>
<tr>
<td></td>
<td>Sig. (2-tailed)</td>
<td>.664</td>
<td>.348</td>
<td>.123</td>
</tr>
<tr>
<td>Speaking PTR</td>
<td>Pearson Correlation</td>
<td>.145</td>
<td>-.378</td>
<td>.707**</td>
</tr>
<tr>
<td></td>
<td>Sig. (2-tailed)</td>
<td>.606</td>
<td>.165</td>
<td>.003</td>
</tr>
<tr>
<td>Speaking SR</td>
<td>Pearson Correlation</td>
<td>.452</td>
<td>-.340</td>
<td>.698**</td>
</tr>
<tr>
<td></td>
<td>Sig. (2-tailed)</td>
<td>.091</td>
<td>.215</td>
<td>.004</td>
</tr>
</tbody>
</table>

**. Correlation is significant at the 0.01 level (2-tailed).

*. Correlation is significant at the 0.05 level (2-tailed).

**Summary of statistical analysis.** The null hypothesis for this study states that repeated reading aloud does not have an effect upon speaking fluency, therefore there is no statistically significant difference between speaking fluency scores for control and experimental groups. While analysis of the human rater scores fails to reject the null hypothesis, some of the results of ASR analysis ran counter to these findings. In fact, the findings in the statistical comparison of the groups’ scores for mean length of pauses, phonation/time ratio, and speech rate directly reject the null hypothesis. Unfortunately, the outcomes of this study cannot be generalized due to the low sample size.

**Addressing Attrition in the Study**

The final data set used for the statistical analysis included scores from 16 individuals who completed the study. Initially, however, 158 people indicated interest in participating in the study and logged on to the system to the reading activities at least once. Thus, the level of attrition was
approximately 90% for the third and the final iterations combined, and 80% for the final iteration alone.

Figure 4.1 represents the stages at which participants left the study during the final and most successful iteration. Nearly half of the subjects discontinued participating in the study after their first or second log in, and an additional 18% logged in less than 5 times. Therefore, 65% of all participants left the study within one week. It is also evident from this chart that the longer participants stayed with the study the more likely they were to complete it.

![Pie chart showing stages of attrition](image)

*Figure 8. Stages of attrition of the final iteration of the study*

**Post-study survey results.** As shown in Table 4.13, which reports the results from the post-study online survey, most participants who left the study before completing it admitted that they had done so because of time constraints (60%) or lack of commitment to participate every day (52%). Additionally, 32% of the subjects experienced technical difficulties with the website.
or did not have the necessary equipment, and only five people revealed that they did not understand the instructions or the reading articles.

Table 4.13

*Reasons for abandoning the study*

<table>
<thead>
<tr>
<th>Answer</th>
<th>N</th>
<th>%</th>
</tr>
</thead>
<tbody>
<tr>
<td>I did not understand the instructions</td>
<td>3</td>
<td>12%</td>
</tr>
<tr>
<td>The articles were too difficult for me</td>
<td>2</td>
<td>8%</td>
</tr>
<tr>
<td>I did not have enough time to read</td>
<td>15</td>
<td>60%</td>
</tr>
<tr>
<td>I experienced technical difficulties with the website</td>
<td>8</td>
<td>32%</td>
</tr>
<tr>
<td>I lost interest in the activities</td>
<td>3</td>
<td>12%</td>
</tr>
<tr>
<td>I forgot to read every day</td>
<td>13</td>
<td>52%</td>
</tr>
<tr>
<td>Other (please specify)</td>
<td>4</td>
<td>16%</td>
</tr>
</tbody>
</table>

*Note:* Total N of people who took the survey = 53

In the comments section of the survey many participants, stated that they were spending more than 15 minutes a day on the reading activities and could not allocate so much time from their busy schedules. Others admitted that although they found the reading activities helpful, they could not dedicate the necessary time to practice daily. The following comment from one of the participants demonstrates the difficulties most subjects encountered with the study:

I had other things come up and was not able to keep up with the reading schedule. Finally, I realized that my haphazard participation would likely not provide much value to the study, so I gave up on it. I apologize for committing to the study and having to back out of it (see Appendix K for all comments).

Students who completed the study, however, felt that the practice of repeated reading was beneficial to their ability to speak in the Russian language and many indicated that they would like to continue practicing reading.
Chapter 5: Discussion and Conclusions

This study investigated effects of repeated reading aloud upon speaking fluency. The treatment was administered online using modified parts of an existing software package, *Extending Your Russian*. Even though over two thousand invitations were sent out, only 16 individuals completed the study. The speech recordings, obtained from weekly tests, were reviewed by raters and processed through Praat. The results provided enough evidence to reject the null hypothesis within the scope of the study, although further investigation is necessary to confirm these findings.

Discussion

The outcome of the statistical analysis suggests that repeated reading aloud does have an effect on mechanical aspects of speaking fluency. After performing repeated reading activities for three weeks, the subjects in the experimental group performed better on the speaking scores. ANOVA output suggests that these differences can be attributed to the effect of treatment. Rater scores, and repeated reading results, however, did not produce similar results.

**Raters.** The raters assigned subjective scores to the speech recordings based on their perception of participants’ fluency. Comparing these scores for control and experimental groups was challenging because inter-rater reliability among the raters was low. Comparison of means of the most consistent scores from Raters 2 and 3 did not produce any statistically significant results indicating that there was no difference in speaking fluency between the control and experimental groups from the perspective of a native speaker. However, participants from the experimental group did receive slightly higher scores than those from the control group. It is
possible, that the outcome of the analysis could be different with a larger sample size and raters having had the benefit of calibration training.

**Computer based analysis.** The analysis of computer generated scores produced interesting results. Although there was no statistically significant difference in the reading aloud scores between the control and experimental groups, it was confirmed that these scores can in fact predict the speaking fluency scores for phonation/time ratio and speech rate. The analysis also confirmed that mean length of pauses, phonation/time ratio, and speech rate were most affected by the practice of repeated reading aloud. In fact, even with a small group size, participants in the experimental group outperformed those in the control group. This supports Gibson’s (2008) suggestion that reading aloud may facilitate development of mechanical aspects of speaking fluency. In the end, computer generated scores proved to be more precise in measuring speaking fluency. However, these findings cannot be generalized due to the small sample size and various levels of exposure to the Russian language within the groups outside of the study environment.

**Attrition.** Although the attrition rate for the study was high, it is not uncommon. Overall online studies and distance learning institutions report attrition rates ranging from 20%-30% to 70-80% (Tyler-Smith, 2006). In the case of this study, lack of time and inconsistency in participation were the two main reasons for discontinuing participation in the study. Those who stayed with the study confirmed that the activities were engaging and beneficial for development of fluency in the Russian language.
Limitations

The present study had several limitations. The most serious limitation was the steady level of attrition in the participation. Such a high level can be attributed to the lack of commitment to the study. The participants who reported back admitted that their decision to leave the study was motivated by the lack of free time or necessity to focus on their education or work. Another explanation to the high level of attrition could be found in the instructions for the participants. Perhaps reading one article each day was an overwhelming task for those participants who had not previously practiced reading in Russian on regular basis. Related to attrition is the relative inability to control the research environment. Because the study was done online, no controls were in place to govern how instructions were followed, including when and for how long the subjects performed the reading activities, contributing to the high attrition rate. Data collection was also done online and due to instability in server performance, vital information was lost or not recorded. Another issue with data collection was the equipment. Some of the participants, who did not complete the study, reported that they did not have access to a microphone or their Internet connection was not stable, indicating a lack of bandwidth may have adversely affected the study’s integrity. The quality of sound for those people who recorded the tests was sometimes poor and required editing, which could have resulted in inaccurate readings in Praat. Finally, the duration of the study can also be viewed as a limitation. A longer exposure to the reading activities would have likely produced more conclusive results.

Recommendations for Future Research

The results of this study indicate that there is a strong possibility that the practice of repeated reading aloud can in fact improve speaking fluency in foreign language. However, due
to the limitations in the study, further research is necessary to confirm these findings. Should the study be replicated, several enhancements could potentially mitigate the limitations referenced above. First of all, the success of the study is dependent on enlisting larger, more motivated group of participants. Perhaps current BYU students who can receive extra credit and other incentives for participating in the study would be more motivated to complete the experiment. Another possible solution for the issue of attrition and data collection is performing the study in a more controlled environment, such as a computer lab on BYU campus. Doing so could provide participants with access to the necessary equipment and facilitate more reliable data collection. A longer duration of the study would allow for collecting a larger sample of data, which could in return provide more accurate, reliable, and generalizable results. In addition, limiting the time of the reading activities to 15 minutes a day and shortening the length of the passages may result in higher level of involvement in the study. Finally, calibration training and scoring benchmarks could increase the level of inter-rater reliability in the study.

The study also raised a few questions for additional research. First, the long-term effects of the practice of repeated reading aloud should be investigated. Second, exploring the role of pronunciation and intonation feedback for repeated reading aloud in developing and maintaining speaking fluency would likely provide actionable feedback. Finally, the effectiveness and reliability of computer-based speaking fluency assessment requires further examination.

**Summary and Conclusions**

This study had the purpose of investigating effects of repeated reading upon speaking fluency. A review of the literature established that although there was little research in connecting repeated reading aloud and speaking fluency, the theory and research on fluency,
repetition, and skill development provided the theoretical rationale for this study. Russian language learners volunteered to participate in the research that had experimental design with repeated measures. The study had the null hypothesis stating that repeated reading aloud does not have an effect upon speaking fluency.

Taking into consideration the small group size, a conclusion can be made that within the scope of the study the practice of repeated reading aloud improved some of the aspects of lower order speaking fluency, thus rejecting the null hypothesis. Those who read out loud every day with the purpose of increasing their reading speed hesitated less and spoke faster while responding to speaking prompts. However, the native speakers who reviewed these responses did not observe the differences indicated by computer analysis. Perhaps a longer duration of the experiment and higher level of participation is necessary for the difference to be more evident.
References


APPENDICES

Appendix A: Informed Consent Form

This research study is being conducted by Evgenia Stroh, a graduate student in the Master's Program in Second Language Acquisition and Teaching at Brigham Young University, working under the direction of Dr. Michael Bush and Dr. David Hart. The purpose of the study is to determine how various reading practices affect development of speaking fluency. You are invited to participate because you have learned the Russian language.

Procedures
Prior to the research you will be asked to complete a 10 minute survey that includes a short reading comprehension task. During the research you will be asked to read provided passages in the Russian language 15 minutes a day, at least 5 days a week for 3 weeks. Upon completion of the experimental portion of this study you will be asked to complete a test consisting of two parts: reading and responding to speaking prompts.

Risks/Discomforts
There are no known risks for participation in this study.

Benefits
By participating in this study you may improve your Russian reading and speaking ability.

Confidentiality
All information provided will remain confidential and will only be reported as group data with no information that can be used to identify individuals. All data will be stored electronically solely for the purposes of the research. Only the small number of researchers directly involved with the study will have access to any data collected.

Participation
Participation in this research study is voluntary.

Questions about the research
Evgenia Stroh: jenya.kuz@gmail.com; 801-386-2231
Prof. Michael Bush: michael_bush@byu.edu; 801-422-4515

Questions about your rights:
BYU IRB Administraton
A-285 ASB
Brigham Young University
Provo, UT 84602
Phone: 801-422-1461
Email: irb@byu.edu

I have read, and understood, the above consent and desire of my own free will to participate in this study.

- Yes
- No
Appendix B: Pre-treatment Survey

Survey

1. Please create a nickname
   
   

2. Please indicate your gender
   - Male
   - Female

3. Are you currently a BYU student?
   - Yes
   - No

4. For how long have you studied the Russian Language?
   - 1 year
   - 2 years
   - 3 years
   - 4 years
   - 5 years
   - 6 years

5. Have you lived in a Russian speaking country?
   - Yes
   - No

6. Which country did you live in?
   
   

7. How many months did you spend in this country?
   - 1
   - 2
   - 3
   - 4
   - 5
   - 6
   - 7
   - 8
   - 9
   - 10
   - 11
8. How long ago did you return from this country?
   - Less than 1 year
   - 1-2 years
   - 2-3 years
   - 3-4 years
   - 4-5 years
   - More than 5 years

9. How much do you use the Russian language right now?
   - Never
   - Less than once a month
   - Once a month
   - 2-3 times a month
   - Once a week
   - 2-3 times a week
   - Daily

10. Are you taking any Russian courses right now
    - Yes
    - No

11. How would you describe your proficiency in Russian right now?

<table>
<thead>
<tr>
<th>Speaking</th>
<th>Very Bad</th>
<th>Bad</th>
<th>Poor</th>
<th>Fair</th>
<th>Good</th>
<th>Very Good</th>
</tr>
</thead>
<tbody>
<tr>
<td>Reading</td>
<td>Very Bad</td>
<td>Bad</td>
<td>Poor</td>
<td>Fair</td>
<td>Good</td>
<td>Very Good</td>
</tr>
<tr>
<td>Listening</td>
<td>Very Bad</td>
<td>Bad</td>
<td>Poor</td>
<td>Fair</td>
<td>Good</td>
<td>Very Good</td>
</tr>
<tr>
<td>Writing</td>
<td>Very Bad</td>
<td>Bad</td>
<td>Poor</td>
<td>Fair</td>
<td>Good</td>
<td>Very Good</td>
</tr>
<tr>
<td>Grammar</td>
<td>Very Bad</td>
<td>Bad</td>
<td>Poor</td>
<td>Fair</td>
<td>Good</td>
<td>Very Good</td>
</tr>
</tbody>
</table>
Reading Comprehension Test

Please read the Russian text below and answer the following questions.

А к кому у нас вообще хорошее отношение? Порой к собственному населению отношение далеко не дружественное. Если у нас граждан различают по цвету волос, а милиция проверяет таких людей в метро, так о чем тут говорить? Вообще когда из-за рубежа идет большой иммиграционный поток, тем более представляющий различный социокультурный уровень, то это чревато взрывом националистических настроений. Тем не менее, для того чтобы удержать экономический рост и рост благосостояния населения, нам до середины века необходимо принять как минимум 20 миллионов иммигрантов. В противном случае, к примеру, оправдаются опасения, что реформа образования выльется в массовое закрытие вузов. Даже если все нынешние ученики 3-х или 4-х классов в дальнейшем пойдут в университеты и институты, то вузы все равно придется сократить наполовину: учиться в них будет некому. А ведь есть еще ПГУ и, наконец, армия. Да и начальные школы ожидает дефицит учащихся. То есть имеющегося демографического "пирога", как его ни дели, на все структуры, необходимые государству, не хватит.

1. What is the main idea of the passage?
   - Attitude towards immigrants in Russia
   - Current demographical situation in Russia
   - The education reform in Russia
   - I don’t know

2. According to this passage, what is a possible consequence of increased flow of immigrants into Russia?
   - Immigrants will be arrested
   - Economy will decline
   - There may be increase in nationalist movement
   - I don’t know

3. What will happen with Russian universities if the population doesn't increase?
   - They will be reformed
   - Nothing will happen
   - Many universities will be closed down
   - I don’t know

Email

Enter your E-mail address:

Re-enter your E-mail address:
Appendix C: List of Reading Articles

1. Осторожно жабы
2. Линия бедра
3. Марсианские микробы
4. Контольный выстрел
   4.1 Оптимизм – это не лечится 1
   4.2 Оптимизм – это не лечится 2
   4.3 Оптимизм – это не лечится 3
   4.4 Оптимизм – это не лечится 4
   4.5 Тайная геометрия 1
   4.6 Тайная геометрия 2
   4.7 Тайная геометрия 3
   4.8 Тайная геометрия 4
   4.9 Космическое мужество 1
   4.10 Космическое мужество 2
   4.11 Космическое мужество 3
   4.12 Космическое мужество 4
   4.13 Похваля лении 1
   4.14 Похваля лении 2
   4.15 Похваля лении 3
   4.16 Похваля лении 4
   4.17 Какие люди! 1
   4.18 Какие люди! 2
   4.19 Какие люди! 3
   4.20 Какие люди! 4
   4.21 Предки недорого1
   4.22 Предки недорого 2
   4.23 Предки недорого 3
   4.24 Предки недорого 4

The full list of articles and contents can be found at:

http://arclite.byu.edu/Evgenia/EYR/Index.html?test=g1&id=evgenia
Appendix D: Comprehension Questions

Article #1

1.1 What is the main idea of the passage? (Correct Answer: C)
   a. Denmark is invaded by aliens
   b. There is an ozone hole in Denmark
   c. Toads are exploding for no reason in Denmark
   d. I don’t know

1.2 What is the reason for the death of toads? (Correct Answer: A)
   a. The reason is unknown
   b. Cosmic plague
   c. An unknown disease
   d. I don’t know

Article #2

2.1 What is the main idea of the passage? (Correct Answer: B)
   a. A recent discovery of a dinosaur soft tissue may give a chance to clone these ancient animals
   b. A recent discovery of a dinosaur soft tissue may shed some light on genetics of these ancient animals
   c. A recent discovery of a dinosaur soft tissue proved that their DNA is similar to the one of ostrich
   d. I don’t know

2.2 What did Dr. Ыпайшеп bring from her last trip to Montana? (Correct Answer: C)
   a. A new type of stone
   b. A mammoth
   c. A hip bone of a tyrannosaurus
   d. I don’t know

2.3 Why is Dr. Ыпайшеп’s discovery important? (Correct Answer: A)
   a. The scientists hope to restore the DNA from the molecules that may be left in the sample
   b. The scientists found that its DNA is similar to the one of an ostrich
   c. The scientists discovered a new type of fossil
   d. I don’t know
Article #3

3.1 What is the main idea of this passage? (Correct Answer: B)
   a. The possibility of life on Mars
   b. The possibility of bringing new viruses or bacteria from Mars
   c. The possibility of curing diseases with microelements from Mars
   d. I don’t know

3.2 What is considered to be the best biomarker? (Correct Answer: B)
   a. Ice
   b. Methane
   c. Soil
   d. I don’t know

3.3 What were the findings of the Russian scientists? (Correct Answer: C)
   a. All elements and viruses survived the freezing temperatures
   b. Nothing survived the freezing temperatures
   c. Some parts of viruses survived the freezing temperatures
   d. I don’t know

Article #4

4.1 What is this article about? (Correct Answer: B)
   a. The life of a British spy
   b. An alternate explanation of the assassination of Rasputin
   c. Relationships between Russia and Great Britain
   d. I don’t know

4.2 Why did the British want to kill Rasputin? (Correct Answer: A)
   a. They were afraid that he would convince the tsar to sign a treaty with Germany
   b. Rasputin wanted to take over the British throne
   c. He had a love affair with the British spy
   d. I don’t know

4.3 How many shooters fired shots at Rasputin according to Andrew Cook? (Correct Answer: C)
   a. One
   b. Two
   c. Three
   d. I don’t know

4.4 What did Рейнеп’s nephew have in his possession? (Correct Answer: B)
   a. A ring made from the bullet that killed Rasputin
   b. Journal entries proving that Рейнеп was involved
   c. A picture of Rasputin
   d. I don’t know
Article #7

7.1 What happened to the information from the first space ship? (Correct Answer: C)
   a. The ship fell down in a foreign territory
   b. The information was not useful
   c. The information was lost in the explosion of the ship
   d. I don’t know

7.2 What choice did the astronauts make about the departure? (Correct Answer: B)
   a. They decided to put it off until more information is obtained
   b. They decided to proceed with their mission without the information
   c. They decided to cancel the mission
   d. I don’t know

7.3 Why did Леонов have a hard time getting back into the ship? (Correct Answer: B)
   a. The door was automatically shut behind him after he exited the ship
   b. His gear became bigger in size and he could not fit through the door
   c. He could not move in the suit
   d. I don’t know

7.4 What kind of problem did Леонов encounter when it was time to close the ship door? (Correct Answer: C)
   a. His suit was enlarged because of the pressure
   b. A piece of cord got stuck in the doorway
   c. He had a hard time turning around to close the door
   d. I don’t know

7.5 How much time did Леонов spend trouble shooting all together? (Correct Answer: C)
   a.
   b. 1 ½ min
   c. 2 min
   d. 3 ½ min
   e. I don’t know

7.6 Why was the spaceship in danger of exploding? (Correct Answer: A)
   a. The oxygen pressure grew to extremely dangerous levels
   b. The temperature in the ship was too high
   c. There was a spark in the wiring
   d. I don’t know

7.7 There was a secret order to the team in case of a serious emergency, what was it according to Леонов? (Correct Answer: B)
   a. Both pilots should commit suicide
   b. Бегин should kill Леонов and then himself
   c. They should do their best to stay alive until help arrives
   d. I don’t know
7.8 What was Беляев expected to do in case something went wrong with Леонов’s mission in outer space according to Богдашевский? (Correct Answer: A)
   a. Cut off the cord and leave Леонов in outer space
   b. Come out to the outer space and try to save Леонов
   c. He was not trained for such situation
   d. I don’t know

Article #8

8.1 What does Peter Akst claim about laziness? (Correct Answer: B)
   a. It is harmful for health, happiness, and longevity of life
   b. It is the key to health, happiness, and longevity of life
   c. It leads to loss of memory
   d. I don’t know

8.2 What is one of the things that Akst recommends to do in order to prolong life? (Correct Answer: B)
   a. Go on a special diet
   b. Run regularly
   c. Avoid stress
   d. I don’t know

8.3 What is the “effect of frozen emotions”? (Correct Answer: A)
   a. Constant muscle tension as a result of continuous stress
   b. Result of repeated meditation
   c. Inability to have any kind of emotions
   d. I don’t know

8.4 What does the following phrase characterize: “Если же после тренажерного зала ноги еле ведут домой, вы падаете и наутро не можете проснуться...”? (Correct Answer: C)
   a. You went to the wrong gym
   b. You had a great work out
   c. You overworked yourself at the gym
   d. I don’t know

8.5 Which of these statements is NOT listed as an explanation why alcohol is not good for people who try to lose weight? (Correct Answer: B)
   a. It has a lot of empty calories
   b. It makes people grow old faster
   c. It increases appetite
   d. I don’t know

8.6 What do you think phrase “Худая корова еще не газель!” means? (Correct Answer: B)
   a. Being skinny is very attractive
   b. Being thin and being attractive is not same
   c. A cow can never be a gazelle
   d. I don’t know
Appendix E: Post-study Survey

1. After taking the initial survey for the Russian Research, did you log on to the reading activities website?
   - Yes
   - No

2. (If, Q1 answer is NO) Why did you not log in to the reading activities website?
   - I never received the link to the reading activities website
   - I did not have time to participate in the study
   - I did not know what I was supposed to do
   - I lost interest in the study
   - Other (please specify)

3. (If Q1 answer is YES) Did you complete the study (completing the study means reading for at least 3 weeks and taking 3 tests including the final test)
   - Yes
   - No

4. (If Q3 answer is NO) Why did you stop participating in the study?
   - I did not understand the instructions
   - The articles were too difficult for me
   - I did not have enough time to read
   - I experienced technical difficulties with the website
   - I lost interest in the activities
   - I forgot to read every day
   - Other (please specify)

5. (ALL) Please share any additional comments about this study.
Appendix F: Experimental Group Instructions

- For this study to be successful it is important that you read at least one article every day from the list that is provided.
- For each article, you will see two tabs (Reading 1 and Reading 2) at the top. Under the “Reading 1” tab, there will be text on the left, a set of pictures in the top right corner, and the dictionary section or a timer in the bottom right corner.
- You should read each article out loud.
- For the first time through, read the text slowly; be sure you can understand the meaning of each word.
  - You can click on a highlighted word to see its definition, hear a native speaker pronounce it, and learn about the stress pattern that applies to that word.
  - As you read, pay close attention to your pronunciation.
- Once you are confident in the meaning of the text and the pronunciation of those selected words, you may move forward by clicking the “Reading 2” tab. A short quiz will appear automatically, and after completing the quiz, you will move on to reading out loud, while increasing speed.
  - You should still pay attention to pronunciation and proper pausing. Each time you read, time yourself by using the timer button in the bottom right corner. You can compare your time to that of a native speaker (located at the bottom of each paragraph). If your time is over 10% slower than the native speaker’s time you should repeat the reading until you reach a higher speed. Once you’ve reached this goal you should move to the next article.
- Record yourself at least once for each text that you read.
- If you complete the article in less than 15 minutes, move on to the next one on the list.
- At the end of each week, you will take a short speaking test that is located at the bottom of the table of contents.
- You do not need to go through the entire list of articles. At the end of week 3 you may proceed to the final test.
Appendix G: Control Group Instructions

- For this study to be successful, it is important that you read at least one article every day from the list that is provided.
- For each article you will see a text on the left, a set of pictures in the top right corner, and the dictionary section in the bottom right corner.
- You should read each text at least three times.
- For the first time through, read the text slowly; be sure you can **understand** the meaning of each word.
  - You can click on a highlighted word to see its definition, hear a native speaker pronounce it, and learn about the stress pattern that applies to that word.
- Once you are confident in the meaning of the text take a short quiz and move on to the next article on the list.
- **At the end of each week, you will take a short speaking test that is located at the bottom of the table of contents**
- You do not need to go through the entire list of articles. At the end of **week 3** you may proceed to the final test.
## Appendix H: Speaking Rubric for Raters

<table>
<thead>
<tr>
<th>Score Key:</th>
<th>Overall Score:</th>
</tr>
</thead>
<tbody>
<tr>
<td>1 Needs a lot of improvement</td>
<td>☐ 4</td>
</tr>
<tr>
<td>2 Satisfactory</td>
<td>☐ 3</td>
</tr>
<tr>
<td>3 Good</td>
<td>☐ 2</td>
</tr>
<tr>
<td>4 Excellent</td>
<td>☐ 1</td>
</tr>
</tbody>
</table>

### Appropriateness & Pragmatics: 1 2 3 4
- Understood the purpose of your speaking.
- Aware of the audience in your speaking.

### Content: 1 2 3 4
- Completed the task successfully as defined by the criteria.

### Detail: 1 2 3 4
- Provided sufficient detail for the task.

### Discourse Level: 1 2 3 4
- Used simple but complete sentences.
- Mostly fluent when using memorized phrases and expressions to complete tasks.

### Grammar: 1 2 3 4
- Errors were minimal and not distracting.

### Pronunciation: 1 2 3 4
- Speed (not too fast nor too slow) did not affect comprehensibility.
- Very little unnatural pausing or hesitation for basic survival tasks.
- Errors in linking and sentence stress were not distracting.
- Other pronunciation errors were not distracting.

### Vocabulary & Phrases: 1 2 3 4
- Used a variety of words and phrases.
- Used appropriate word forms.
- Used phrases appropriately.
Appendix I: Means and One-way ANOVA for Speaking Scores Less the Outlier

**Means of speaking scores**

<table>
<thead>
<tr>
<th>Group</th>
<th>NSYL</th>
<th>NP</th>
<th>MLP</th>
<th>PTR</th>
<th>SR</th>
<th>AR</th>
<th>ASD</th>
</tr>
</thead>
<tbody>
<tr>
<td>Experimental</td>
<td>331.67783</td>
<td>49.28883</td>
<td>.79983</td>
<td>69.88317</td>
<td>2.72000</td>
<td>3.89683</td>
<td>.26167</td>
</tr>
<tr>
<td>Control</td>
<td>275.05189</td>
<td>44.65922</td>
<td>1.12378</td>
<td>59.01500</td>
<td>2.19011</td>
<td>3.77967</td>
<td>.27167</td>
</tr>
<tr>
<td>Total</td>
<td>297.70227</td>
<td>46.51107</td>
<td>.99420</td>
<td>63.36227</td>
<td>2.40207</td>
<td>3.82653</td>
<td>.26767</td>
</tr>
</tbody>
</table>

**One-way ANOVA scores**

<table>
<thead>
<tr>
<th></th>
<th>Sum of Squares</th>
<th>df</th>
<th>Mean Square</th>
<th>F</th>
<th>Sig.</th>
</tr>
</thead>
<tbody>
<tr>
<td>NSYL</td>
<td>Between Groups</td>
<td>11543.391</td>
<td>11543.391</td>
<td>1.912</td>
<td>.190</td>
</tr>
<tr>
<td></td>
<td>Within Groups</td>
<td>78484.225</td>
<td>6037.248</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Total</td>
<td>90027.616</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>NP</td>
<td>Between Groups</td>
<td>77.160</td>
<td>77.160</td>
<td>.379</td>
<td>.549</td>
</tr>
<tr>
<td></td>
<td>Within Groups</td>
<td>2643.684</td>
<td>203.360</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Total</td>
<td>2720.844</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>MLP</td>
<td>Between Groups</td>
<td>.378</td>
<td>.378</td>
<td>5.963</td>
<td>.030</td>
</tr>
<tr>
<td></td>
<td>Within Groups</td>
<td>.824</td>
<td>.063</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Total</td>
<td>1.201</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>PTR</td>
<td>Between Groups</td>
<td>425.221</td>
<td>425.221</td>
<td>3.376</td>
<td>.089</td>
</tr>
<tr>
<td></td>
<td>Within Groups</td>
<td>1637.194</td>
<td>125.938</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Total</td>
<td>2062.416</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>SR</td>
<td>Between Groups</td>
<td>1.011</td>
<td>1.011</td>
<td>5.047</td>
<td>.043</td>
</tr>
<tr>
<td></td>
<td>Within Groups</td>
<td>2.604</td>
<td>.200</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Total</td>
<td>3.615</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>AR</td>
<td>Between Groups</td>
<td>.049</td>
<td>.049</td>
<td>.241</td>
<td>.632</td>
</tr>
<tr>
<td></td>
<td>Within Groups</td>
<td>2.667</td>
<td>.205</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Total</td>
<td>2.716</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ASD</td>
<td>Between Groups</td>
<td>.000</td>
<td>.000</td>
<td>.210</td>
<td>.655</td>
</tr>
<tr>
<td></td>
<td>Within Groups</td>
<td>.022</td>
<td>.002</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Total</td>
<td>.023</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
### Appendix J: One-way ANOVA scores for All Reading Aloud Scores

<table>
<thead>
<tr>
<th></th>
<th>Sum of Squares</th>
<th>df</th>
<th>Mean Square</th>
<th>F</th>
<th>Sig.</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>NSYL1</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Between Groups</td>
<td>2371.600</td>
<td>1</td>
<td>2371.600</td>
<td>1.465</td>
<td>.248</td>
</tr>
<tr>
<td>Within Groups</td>
<td>21045.333</td>
<td>13</td>
<td>1618.872</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>23416.933</td>
<td>14</td>
<td>176.670</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>NP1</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Between Groups</td>
<td>136.752</td>
<td>1</td>
<td>136.752</td>
<td>.774</td>
<td>.395</td>
</tr>
<tr>
<td>Within Groups</td>
<td>2296.715</td>
<td>13</td>
<td>176.670</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>2433.467</td>
<td>14</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>MLP1</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Between Groups</td>
<td>.077</td>
<td>1</td>
<td>.077</td>
<td>5.125</td>
<td>.041</td>
</tr>
<tr>
<td>Within Groups</td>
<td>.195</td>
<td>13</td>
<td>.015</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>.271</td>
<td>14</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>PTR1</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Between Groups</td>
<td>281.762</td>
<td>1</td>
<td>281.762</td>
<td>4.562</td>
<td>.052</td>
</tr>
<tr>
<td>Within Groups</td>
<td>802.901</td>
<td>13</td>
<td>61.762</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>1084.663</td>
<td>14</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>SR1</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Between Groups</td>
<td>.426</td>
<td>1</td>
<td>.426</td>
<td>1.847</td>
<td>.197</td>
</tr>
<tr>
<td>Within Groups</td>
<td>2.996</td>
<td>13</td>
<td>.230</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>3.422</td>
<td>14</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>AR1</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Between Groups</td>
<td>.01</td>
<td>1</td>
<td>.01</td>
<td>.011</td>
<td>.918</td>
</tr>
<tr>
<td>Within Groups</td>
<td>1.519</td>
<td>13</td>
<td>.117</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>1.521</td>
<td>14</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>ASD1</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Between Groups</td>
<td>.01</td>
<td>1</td>
<td>.01</td>
<td>.013</td>
<td>.910</td>
</tr>
<tr>
<td>Within Groups</td>
<td>.007</td>
<td>13</td>
<td>.001</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>.007</td>
<td>14</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>NSYL2</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Between Groups</td>
<td>472.500</td>
<td>1</td>
<td>472.500</td>
<td>.126</td>
<td>.728</td>
</tr>
<tr>
<td>Within Groups</td>
<td>48569.500</td>
<td>13</td>
<td>3736.115</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>49042.000</td>
<td>14</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>NP2</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Between Groups</td>
<td>26.448</td>
<td>1</td>
<td>26.448</td>
<td>.111</td>
<td>.744</td>
</tr>
<tr>
<td>Within Groups</td>
<td>3092.755</td>
<td>13</td>
<td>237.904</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>3119.203</td>
<td>14</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>MLP2</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Between Groups</td>
<td>.11</td>
<td>1</td>
<td>.11</td>
<td>4.972</td>
<td>.044</td>
</tr>
<tr>
<td>Within Groups</td>
<td>.295</td>
<td>13</td>
<td>.023</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>.407</td>
<td>14</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>PTR2</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Between Groups</td>
<td>294.959</td>
<td>1</td>
<td>294.959</td>
<td>4.238</td>
<td>.060</td>
</tr>
<tr>
<td>Within Groups</td>
<td>904.743</td>
<td>13</td>
<td>69.596</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>1199.703</td>
<td>14</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>SR2</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Between Groups</td>
<td>.417</td>
<td>1</td>
<td>.417</td>
<td>1.945</td>
<td>.187</td>
</tr>
<tr>
<td>Within Groups</td>
<td>2.789</td>
<td>13</td>
<td>.215</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>3.206</td>
<td>14</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>AR2</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Between Groups</td>
<td>.006</td>
<td>1</td>
<td>.006</td>
<td>.066</td>
<td>.801</td>
</tr>
<tr>
<td>Within Groups</td>
<td>1.197</td>
<td>13</td>
<td>.092</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>1.204</td>
<td>14</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>ASD2</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Between Groups</td>
<td>.005</td>
<td>1</td>
<td>.005</td>
<td>.119</td>
<td>.735</td>
</tr>
<tr>
<td>Within Groups</td>
<td>.005</td>
<td>13</td>
<td>.000</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>.005</td>
<td>14</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>NSYL3</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Between Groups</td>
<td>339974.349</td>
<td>1</td>
<td>339974.349</td>
<td>2.452</td>
<td>.140</td>
</tr>
<tr>
<td>Within Groups</td>
<td>1940897.651</td>
<td>13</td>
<td>138635.546</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>2280872.000</td>
<td>15</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>NP3</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Between Groups</td>
<td>32656.165</td>
<td>1</td>
<td>32656.165</td>
<td>3.181</td>
<td>.096</td>
</tr>
<tr>
<td>Within Groups</td>
<td>143712.455</td>
<td>14</td>
<td>10265.175</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>176368.620</td>
<td>15</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>MLP3</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Between Groups</td>
<td>.088</td>
<td>1</td>
<td>.088</td>
<td>.366</td>
<td>.555</td>
</tr>
<tr>
<td>Within Groups</td>
<td>3.377</td>
<td>14</td>
<td>.241</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>3.466</td>
<td>15</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>PTR3</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Between Groups</td>
<td>15.827</td>
<td>1</td>
<td>15.827</td>
<td>.086</td>
<td>.733</td>
</tr>
<tr>
<td>Within Groups</td>
<td>2567.591</td>
<td>14</td>
<td>183.399</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>2583.418</td>
<td>15</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>SR3</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Between Groups</td>
<td>.035</td>
<td>1</td>
<td>.035</td>
<td>.116</td>
<td>.738</td>
</tr>
<tr>
<td>Within Groups</td>
<td>4.268</td>
<td>14</td>
<td>.305</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>4.303</td>
<td>15</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>AR3</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Between Groups</td>
<td>.186</td>
<td>1</td>
<td>.186</td>
<td>1.244</td>
<td>.284</td>
</tr>
<tr>
<td>Within Groups</td>
<td>2.090</td>
<td>14</td>
<td>.149</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>2.276</td>
<td>15</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>ASD3</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Between Groups</td>
<td>.001</td>
<td>1</td>
<td>.001</td>
<td>1.575</td>
<td>.230</td>
</tr>
<tr>
<td>Within Groups</td>
<td>.013</td>
<td>14</td>
<td>.001</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>.015</td>
<td>15</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Appendix K: Comments about the Study from Participants

- I think mine was actually more like 4 weeks or 3 1/2...I got off a little but I took all the tests.

- I accumulated 15 days of reading and took a test after every five articles, but it took me about 4 and a half weeks to do so. I couldn't find the time to read every day. I thought the articles were interesting. I was surprised how hard it was to talk for a few minutes in Russian, and how rusty my speaking skills were even after listening a lot. Good luck with your project and writing your thesis, Evgenia!

- The readings were way above my ability, but I really would have liked to try them and see if I could have shown improvement. I truly am swamped for time (full-time work, doctoral program, family, demanding church calling, etc.

- there were technical difficulties, and i don't have much time

- I enjoyed reading and trying to understand in Russian again. I really do want to improve my Russian. I also found it helpful that I could listen to pronunciation as I read and see how long it would take a native Russian speaker to read that passage. The questions about the text were very good for measuring understanding. It was technical difficulties that I got hung up on with the microphone and not always remembering which article I was on. Then the time commitment was probably more than I could give because even though it was SAID to be 15 minutes, it would take me 20 minutes just to get through the passage the first time and understand what I had read. Then further readings would add to the time commitment and I really needed 60-90 minutes a day to complete one article's requirements. Perhaps it's for someone with a higher Russian ability than I, but it HAS been a while since I was in that class.

- I enjoyed participating and felt that my ability to speak and pronounce was enhanced by participation. Further, I feel that committed participation would have produced even better results. Unfortunately, I was not able to fully complete the study.

- I greatly apologize. Just bad timing. I had a great desire to help you but things came up in my personal life and I couldn't add any commitments at that time, though I knew it would only be a few minutes a day. I think the amount of time was necessary in order to see results and learn. Again, just bad timing on my part. So sorry I couldn't help. And so sorry I didn't do it as I know it would have increased my vocabulary, skills, and confidence with the Russian language. I regret not completing it.
It might be easier to complete on a busy schedule if participants were emailed the text of the reading assignment, with a link to the system and a short survey/quiz. Basically, when I would get a few free minutes it didn't seem like long enough to complete the study.

The passages were technically too difficult for me. I had to really study them for a long time and practice reading meticulously before I could read them quickly. I got discouraged because I didn't feel like I was progressing quickly enough to make a difference in the study. When it came time to take the tests (recordings of my speaking) it was intimidating for me. I think I speak better and with more confidence when there is not the pressure of being formally evaluated for language abilities. The entire experience was very good for me. I like to know what my strengths and weaknesses are in the language so that I can know what and how to improve. I'm glad I got invited to be part of the study and do my best to study and read out loud in Russian. *I would have benefitted from having the entire passages read aloud by a native Russian speaker and not just select vocabulary words. It probably would've helped my overall speed and understanding. But, you probably have your own reasons for not letting us hear the entire passage read out loud by a native Russian speaker... just saying it would've helped me. Thanks for the experience!

I think if I had had time and remembered to read every day it would have worked

I'm currently working full-time and completing my executive MBA. While I would have like to participate in the study, I simply did not have the time.

I had other things come up and was not able to keep up with the reading schedule. Finally, I realized that my haphazard participation would likely not provide much value to the study, so I gave up on it. I apologize for committing to the study and having to back out of it.

I wish it were still open to me- I miss Russian a lot and loved the practice. I would love to have the option to listen to the paragraph in Russian, and then read aloud and record myself. That way I could tune my ear more. Would you please give those who finished the study access to do it again? The email reminders helped a lot; with four small children, I often forgot to read.

I completed one reading and answered the questions. The next time I wasn't able to get in with the same browser. Sorry I didn't try again later on a different one.