



2016-12-01

A Longitudinal Analysis of Adult ESL Speakers' Oral Fluency Gains

Kostiantyn Fesenko
Brigham Young University

Follow this and additional works at: <https://scholarsarchive.byu.edu/etd>

 Part of the [Linguistics Commons](#)

BYU ScholarsArchive Citation

Fesenko, Kostiantyn, "A Longitudinal Analysis of Adult ESL Speakers' Oral Fluency Gains" (2016). *All Theses and Dissertations*. 6130.
<https://scholarsarchive.byu.edu/etd/6130>

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.

A Longitudinal Analysis of Adult ESL Speakers' Oral Fluency Gains

Kostiantyn Fesenko

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

Mark W. Tanner, Chair
Lynn E. Henrichsen
K. James Hartshorn

Department of Linguistics and English Language
Brigham Young University

Copyright © 2016 Kostiantyn Fesenko

All Rights Reserved

ABSTRACT

A Longitudinal Analysis of Adult ESL Speakers' Oral Fluency Gains

Kostiantyn Fesenko
Department of Linguistics and English Language, BYU
Master of Arts

While a number of studies have sought to investigate ESL speakers' fluency gains over the course of one 15-week semester, few if any studies have investigated these changes over a longer developmental period. A critical factor in researching longitudinal change is that students do not often remain in an intensive English program (IEP) for more than two semesters before moving to a new school, applying to an American university, or returning to their home country. Longitudinal research, therefore, is necessary as program administrators, teachers, and learners all seek to understand points where change in oral fluency actually occurs.

For this study data were collected from students in a large intensive English program over a 45-week period. For 39 ESL learners audio files from speaking tasks that were part of placement and end-of-semester level achievement tests were collected and analyzed. Specific oral fluency features such as speech rate, articulation rate, and pause frequency were investigated. This thesis will share the results of the analysis while also discussing the implications of the data for program administrators, teachers, and learners. Particular focus will be given to helping stakeholders understand specific changes that occurred in learners' fluency over the time period of three semesters.

Keywords: oral fluency, speaking, ESL, speech rate, articulation rate, phonation time

ACKNOWLEDGMENTS

This thesis would not have been possible without my committee and chair Dr. Mark Tanner. His guidance and patience have helped me to successfully complete this project. He has shown me on multiple occasions that there is always room for improvement. Dr. James Hartshorn has been another wonderful resource who kindly and patiently helped with some of the hardest parts of this process. Finally, Dr. Lynn Henrichsen provided some very valuable feedback and always asked questions that helped me analyze my writing from different perspectives.

I am also incredibly grateful for my incredible parents Tatiana and Vasiliy Fesenko who encouraged me throughout this journey. Their kind words of encouragement have been especially helpful in times of discouragement and disappointment.

Last but not least I would like to express my gratitude to my friends who have been extremely supportive and encouraging: Skyler Crouch, Jacob Newman, Matthew Gilliland, Alisha Chugg, Paul Lee, Christina Mahler, and Jerilyn Brunson.

TABLE OF CONTENTS

LIST OF TABLES	vi
LIST OF FIGURES	vii
PREFACE	viii
Introduction.....	1
Review of Literature	2
Definition of Fluency	2
Fluency Features	4
Definition of Fluency Features	6
Short-term vs. Longitudinal Studies	8
Research Questions.....	11
Methodology.....	12
Participants.....	12
Setting.....	13
Instruments.....	14
Procedure	17
Analysis	18
Results.....	18
Changes in Fluency Over 45 Weeks of Instruction	19
Speech Rate.....	19
Articulation rate	22
Number of Syllables	24

Number of Pauses	26
Phonation Time	28
Discussion	30
Gains Based on Fluency Features	31
Gains Based on Proficiency Level	33
Implications	36
Limitations and Directions for Future Research	37
Conclusion	39
References	41
APPENDIX: Sample Placement/LAT Prompts at ACTFL Proficiency Levels	46

LIST OF TABLES

Table 1: Placement Test Level Assignment and Tasks	17
Table 2: Data Collection Time Periods.....	18
Table 3: Descriptive Statistics for Speech Rate Across the 45-Week Period.....	20
Table 4: Descriptive Statistics for Articulation Rate Across the 45-Week Period	22
Table 5: Descriptive Statistics for Number of Syllables Across the 45-Week Period.....	25
Table 6: Descriptive Statistics of Number of Pauses Across the 45-Week Period.....	27
Table 7: Descriptive Statistics of Phonation Time Across the 45-Week Period	29

LIST OF FIGURES

Figure 1: Speech Rate changes over time by level of proficiency.....	20
Figure 2: Articulation Rate changes over time by level of proficiency.....	23
Figure 3: Average number of syllabus per response over time by level of proficiency... ..	25
Figure 4: Average number of pauses per response over time by level of proficiency.	27
Figure 5: Phonation Time changes over time by level of proficiency.....	29

PREFACE

This thesis is intended to be a manuscript length document that will be submitted to the journal *Applied Language Learning* for publication. The purpose of this journal is to inform second language researchers and educators about empirical research that focuses on the longitudinal development of oral fluency.

Manuscripts submitted to this journal need to be in APA format according to the Publication Manual of the American Psychological Association (6th edition). Furthermore, manuscripts should not exceed 6,000 words (excluding charts, appendix, references, or/and notes, etc.) or 15 to 20 pages. At present, this manuscript is over that allotted amount due to having additional pages that are required by the university for submission as a master's thesis.

There are two other journals that are being considered. One of them is *The Modern Language Journal*, which is primarily focused on teaching second and foreign languages. Among the audience of this publication are language teachers who would be interested in the topic of oral fluency and its development over a long period of time. The length of manuscripts submitted to *The Modern Language Journal* should be between 8,000 and 10,000 words, which will not require a reduction in the manuscript length.

The other alternative journal is *English Teaching Research*. This particular journal includes articles investigating second or foreign language teaching. The content of this article appears to be relevant for language teachers and researches interested in the oral proficiency development of second language learners in instructed context. Articles submitted to *English Language Teaching* must not exceed 8,000 words, which is why this journal is another appropriate venue which to consider for this manuscript.

Introduction

Imagine a foreign student who comes to the U.S. to improve his English so that he can enter an American university. He chooses to study in an intensive English program (IEP) with the objective of improving his language skills. Questions he may ask himself include, “How quickly will I be able to improve? What changes will I actually be able to see in my speaking skills? How many semesters do I need to make the kind of change that will help me be successful in college?”

This student is like many who enroll in intensive English programs (IEP) across the United States every year. Many have already spent years studying English, but they stress about how long it will take to obtain the knowledge and skills necessary to achieve a high score on the Test of English as a Foreign Language (TOEFL). A high score is necessary to be admitted to an American university. One way students know they have made improvement in their language skills is if they are promoted to the next program level. Noticeable improvement is needed in all language skills: vocabulary, grammar, writing, reading, listening, and speaking. While overall language proficiency improvement is the goal, learners often sometimes express concern with their oral skills as this is a means by which they are regularly judged. Proficiency is often evaluated by listeners through the speakers’ oral fluency skills. The focus of this particular study was on evaluating oral fluency of English as a second language (ESL) learners. Successful and effective speaking means intelligible and comprehensible communication (Derwing, Munro, & Thomson, 2008). Rossiter, Derwing, Manimtim, & Thomson (2010) argue that “because the primary goal of most communicative L2 programs is to foster communicative competence, oral fluency is an important outcome criterion” (p. 599).

In researching change in oral proficiency as a result of instruction, many studies have focused on gains that learners have made with respect to the area of fluency as one way in which progress has been evaluated. Many of the current studies (Blake, 2009; Christensen, 2012; Cucchiarini, Strik, & Boves, 2000; Ushigusa, 2009; de Wet, Van der Walt, & Niesler, 2009), however, have investigated fluency change over the course of only semester. In contrast, few studies have investigated the longitudinal development of fluency in ESL students enrolled in an intensive English program over an extended time period. Possible reasons for this may be that learners tend to transition between schools after one or two semesters, or they have completed admission requirements enabling them to matriculate to an American university. The purpose of this research was to investigate key features of oral fluency development made by ESL learners enrolled for 45 weeks of instruction in an Intensive English Program (IEP), which focuses on preparing students in English for academic purposes.

Review of Literature

This section will begin by providing various definitions of fluency as well as a description of fluency features that were used in this study. The last section will give an overview of various studies that have been conducted on oral fluency, both short-term and long-term.

Definition of Fluency

There are several major reasons why international ESL students in the United States strive to speak fluently in English: some need to pass important proficiency tests such as the Test of English as a Foreign Language (TOEFL) which are required by US colleges for admission; others want to use English as a language of wider communication

to better communicate for specific purposes such as English for business, English for law, or English for medicine. Still others simply desire to sound more confident in their everyday speech. Defining the term *fluency*, however, is difficult as students, educators, and administrators may have different concepts of this term. Learners of English may consider fluent speech to be error free utterances, while teachers and administrators may understand fluency as the ability to speak comprehensibly and intelligibly. These definitions of fluency, however, are very broad. In the following section, the definition of fluency will be further explored.

A generic definition for fluency, according to Crystal (1987), is “smooth, rapid, effortless use of language” (p. 427). Fluency can also be described as “successful performance in task-based contexts...containing the capacity to produce speech at a normal rate and without interruption” (Skehan, 2009, p. 510). Another definition of oral fluency has been proposed by Derwing & Munro (2013) who state that fluency “refers to listeners’ perceptions of the flow of the speakers’ language output, for example, whether there are frequent pauses, false starts, or other dysfluencies” (p. 197). When focusing on oral fluency, though, Fillmore (1979) proposes four definitions of fluency. The first way in which he defines fluency is “the ability to talk at length with few pauses; the ability to fill time with talk” (p. 93). The second description of fluency is “the ability to talk in coherent, reasoned, and “semantically dense” sentences” (p. 93). The third definition is “the ability to have appropriate things to say in a wide range of contexts” (p. 93). And finally, a fourth classification of fluency is “the ability some people have to be creative and imaginative in their language use, to express their ideas in novel ways, to pun, to make up jokes, to attend to the sound independently of the sense, to vary styles, to create

and build on metaphors” (p. 93). One may infer from this variety of definitions that it is challenging to describe oral fluency in a single definition. Neither is it sufficient to combine all of them in one definition.

Other definitions proposed by Baker-Smemoe, Dewey, Bown, & Martinsen (2014) indicate that “fluency can refer to language that is produced fluidly and smoothly as one combines words and sentences in speech” (p. 708). Besides giving a general definition of fluency, the authors also suggest a number of different types of fluency that demonstrate how hard it is to simply choose one definition. Clarifications on the types of fluency that exist also provide those researching fluency with clearer ways to measure this complex aspect of language production. These types of measures then help to provide a workable definition of fluency that other scholars can use in subsequent research.

With clearer definitions and agreed upon units of measurement, it should also be easier to detect ways in which learners’ language changes and improvement can be seen empirically. The next section will discuss various fluency variables identified by researchers and how these variables are measured in order to identify fluency gains in a speaker’s oral language.

Fluency Features

As mentioned in the previous paragraph, in order to study or improve fluency, researchers track certain features to measure progress and change. Götz (2013) used the following features in her research on fluency: speech rate, mean length of run, number of unfilled pauses, and phonation time ratio. In another study, Zechner, Higgins, Xi, &

Williamson (2009) investigated several language features such as response duration, articulation rate, number of silences and pauses, duration of silences per word.

Some researchers have focused on only one particular fluency feature. For instance, Sato (2014) focused on different types of speech rate, which she found to be one of the most important features of oral fluency. On the other hand, other researchers have investigated a variety of fluency features. In a study of oral fluency by Christensen (2012), he suggested tracking the following fluency measures: speech rate, articulation rate, number of syllables, number of pauses, phonation time, and average syllable duration.

Cucchiarini, Strik, & Boves (2000) and Kormos & Dénes (2004) found the previously mentioned fluency features to be a good descriptor of oral fluency and its measurements. These studies have determined that “fluency ratings are strongly affected by rate of speech, articulation rate, phonation/time ratio, number of silent pauses, total duration of pauses, and mean length of run” (Cucchiarini et al., 2000, p. 996).

To simplify the data analysis process and measure particular features, more researchers (Boersma & van Heuven, 2001) suggest using a software program called Praat. This program can be used to “analyze, synthesize, and manipulate speech” (p. 341). They state, “It is a remarkable instrument that allows scholars and linguists to “label and segment their speech recordings” (p. 342). According to Christensen (2012), scripts in Praat enable the program to measure the following features: (1) speech rate, (2) articulation rate, (3) number of syllables, (4) number of pauses, and (5) phonation time. Given that these fluency features are commonly recognized by researchers as viable measures of general oral fluency, these features will be used as variables in this research

study. In all of these studies on fluency, a variety of features were analyzed; however, not all of them seem equally valuable or even necessary. Therefore, the following section will provide more detailed descriptions of each of these fluency features.

Definition of Fluency Features

Current research indicates that there are numerous features and software programs that researchers have been using to measure oral fluency. Speech rate, number of syllables, articulation rate, mean length of run, number of pauses and other features appear to be the most common in measuring oral fluency. This section is intended to provide a more complete description of five key features that Praat can measure with regards to fluency.

Speech rate. Speech rate is defined as “the number of syllables uttered per second” (Chambers, 1997, p. 538). Speaking rate is “an overall measure which includes *articulation rate* and *pause time*” (p. 538). It is “one of the most frequently used measures of fluency” (Lambert & Kormos, 2014, p. 610). As Ginther, Dimova, & Yang (2010) indicate, it is “the most salient parameter of language fluency” (p. 385) and one of the “rate-related variables most frequently found in the literature” (p. 382). They also suggest that this fluency feature has the strongest relationship to the Oral English Proficiency Test (OEPT). In her research, Ushigusa (2009) found that “speaking rate had the strongest correlation and mean length of run and phonation time ratio (or silent pause time rate) had the second strongest correlation with the proficiency scores among all temporal variables” (p. 169). In addition to the previous studies, de Wet, et al. (2009) found speech rate to be “one of the best indicators of fluency” (p. 871). In order to measure speech rate in a given speech sample, one must take the total number of syllables

produced in a speech sample and divide them by the total response time, which includes both filled and silent pauses. Finally, this number (in seconds) is multiplied by 60 in order to produce speech rate per minute.

Articulation Rate. This variable is defined as the total number of syllables divided by the sum of speech time and total filled pause time multiplied by 60, or in other words, the speed with which participants produce words and sentences. Ushigusa (2009) defines articulation rate as a fluency variable that “measures the speed of delivery based on the number of syllables produced” (p. 118) or “how many syllables examinees produced per 60 seconds of utterances” (p. 120). In their research on spontaneous speech, Xi, Higgins, Zechner, & Williamson (2008) used a variety of variables related to fluency, grammar, accuracy, pronunciation, and vocabulary diversity. They came to the conclusion that out of the many fluency features often measured: articulation rate, number of pauses, number of silences, response length, and other features, articulation rate was the most significant, which could indicate that the higher the articulation rate is, the higher the overall oral proficiency becomes. Finally, articulation rate “focuses on the amount of time required for a speaker to physically produce speech and is argued to represent the efficiency of the articulator” (Ginther, Dimova, & Yang, 2010, p. 382).

In addition to speech rate and articulation rate, three other fluency features were measured in this study as a means of providing a clearer understanding of how fluency changed over time. As the review of literature suggests, these five features appear to those which are most studied among linguists in recent years (Christensen, 2012; Ginther, Dimova, & Yang, 2010; Götz, 2013).

Short-term vs. Longitudinal Studies

Studies investigating oral fluency have been both short-term (several weeks or months) and long-term (several years) in nature. Studies, which have been short-term, will be discussed first and insights obtained from this research will be shared.

Short-term studies. Several studies have investigated fluency in short-term studies that have ranged in length from four weeks to several months (Baker-Smemoe et al., 2014; Baró & Serrano, 2011; Cucchiarini et al., 2000; De Jong, Steinel, Florijn, Schoonen, & Hulstijn, 2013; Delaney, 2012; Nakano, Kondo, & Tsutsui, 2008; O'brien, Segalowitz, Freed, & Collentine, 2007; Pinget, Bosker, Quené, & de Jong, 2014; Sato, 2014; de Wet et al., 2009). For instance, Llanes & Muñoz (2009) conducted a four-week study on learners' oral fluency gains while students studied abroad. A total of 24 Spanish speakers aged 13-22 years old spent almost a month in countries in which English is used as the primary medium of communication. The fluency features for this study included articulation rate, number of syllables, number of pauses, other language word ratio, and the longest fluent run. The results of the study suggest that “differences between pre-test and post-test scores turned out to be statistically significant in four of the six measures of oral fluency analyzed: syllables per minute, other language word ratio, articulation rate, and the longest fluent run” (p. 361).

Blake (2009) conducted a six-week study on oral fluency development studying 34 ESL learners in order to compare their use of Internet chats and face-to-face communication, and how both of these types of communication helped learners of English improve their speaking skills. The fluency variables that were used in this study were rate of speech (speaking rate), articulation rate, average length of pauses, and

phonation time ratio. The results of the study suggest that phonation time was significantly higher than the rest of the fluency measures used in this research.

Based on these short-term studies, it is evident that certain fluency features improve much faster than others. Some learners may even think that four weeks of formal instruction might be sufficient to see a pattern of oral fluency and proficiency development. However, even though the above-mentioned studies definitely provide some meaningful statistics to show how fluency changes over a short period of time, learning usually takes much longer periods of time, and therefore studying these features over longer periods of time might reveal a bigger and more in-depth picture of the development of oral fluency. In order to see the importance of longitudinal studies and the purpose of this particular research, the following section will describe some of the studies that were conducted longitudinally.

Longitudinal studies. While the majority of studies on oral fluency have been short-term in nature, a few studies have been longitudinal (Miller & Schwanenflugel, 2008; Polat & Kim, 2014), but the number of these studies is rather small, or unlike the present study, they have focused on examining oral fluency in children rather than adult learners of English. In their study on fluency, accuracy, and complexity, Mora & Valls-Ferrer (2012) collected oral data from 30 adult ESL learners and 10 native speakers of English to investigate a number of fluency variables: speech rate, articulation rate, phonation rate, mean length of run, dysfluency ratio, pause frequency, and pause time ratio. The results of this study were very interesting. In terms of fluency, after two years, the participants' speech rate "increased substantially towards [native speakers'] performance...suggesting that [nonnative speakers'] oral production became faster" (p.

622). In addition to improving their speech rate, the participants in this study improved their pause duration and frequency. This suggests that speech rate, pause duration and frequency are critical features in measuring oral fluency gains longitudinally.

Another longitudinal study was conducted by Derwing, Munro, & Thompson in 2008. The participants of the study were 16 Mandarin and 16 Slavic adults aged 19 to 49 years old who studied English in Canada for two years. Data were collected and compared during seven time periods (the first six measurements were taken every two months with one more a year later). The results of the study showed that “the Slavic language speakers showed a small but significant improvement in both fluency and comprehensibility” (p. 359), while the Mandarin speakers’ oral proficiency “did not change over 2 years” (p. 359). One of the strongest factors that could have influenced the results of this study was exposure to English outside of class, which is why those students who practice their English outside of their classes will probably achieve better results regarding their oral fluency and proficiency.

As Rossiter et al. (2010) put it, longitudinal studies together with long-term instruction seem to be more successful since “learner improvement is unlikely to be [effective] over the duration of a single ESL course” (p. 600). As shown, previous longitudinal studies have been done with adult immigrant learners, but no longitudinal studies have been done with IEP students coming to an English-speaking environment for the purpose of learning English as an additional language. As Derwing, Munro, and Thompson (2008) claim that “there are no systematic longitudinal studies of adult immigrants’ development of oral fluency in their second language (L2) environment” (p. 359). From this statement one can infer that even though a small number of longitudinal

studies have been conducted, more studies on oral fluency need to be conducted over long periods of time. Consequently, this particular study focuses on students studying English in an IEP for three semesters or a total of 45 weeks of instruction.

Native English Speakers' Oral Fluency

While some research has investigated oral fluency factors of native English speakers, most of this research has studied the speech rate that native English speaker listeners prefer when hearing English spoken by native English speakers and those of other languages. In a study by Derwing & Munro (2001) on speech rate, they found that the optimal speech rate for native English speakers was 4.7 syllables/second. This rate varied for speakers of Mandarin Chinese. For the Mandarin speakers, listeners identified that the optimal rate was 4.5 syllables/second. While more research needs to be done in this area, it is helpful to have a target speech rate goal for second language learners striving to achieve this target level fluency in English.

Research Questions

Previous studies have measured fluency gains in learners over different time periods. This research has been informative in identifying how different fluency features have changed in programs where the duration of instruction has been as little as four weeks to as long as two years. These studies have also targeted different learner populations. The focus of this study is to investigate adult ESL learners enrolled in an IEP whose focus is to teach English for academic purposes (EAP). The majority of the students in the program plan to attend an American university. In carrying out this research, the following questions were studied:

1. How do the fluency features of speech rate, articulation rate, number of syllables, number of pauses, and phonation time change over the course of three semesters (or 45 weeks of instruction) for students in an EAP program?
2. Which of the fluency features exhibits the most change?
3. During which time period do these EAP learners make the most gains?

Methodology

The purpose of this research was to examine changes in fluency features made by adult EAP students over a 45-week period of instruction. Particular fluency features (speech rate, articulation rate, number of syllables, number of pauses, and phonation time) were measured in learners' speech at the beginning of the instructional period and again after each subsequent semester over the course of three semesters of instruction. This section will describe the participants, instruments used, procedures followed to collect data, and the analyses performed.

Participants

In order to conduct this research, permission was received from Brigham Young University's (BYU's) Institutional Review Board for Human Subjects. The participants in this study were 39 international students who studied English as a second language at an intensive English language program attached to a university in the United States. Participants received 20 hours of formal classroom instruction a week and studied concurrently in the program for three semesters. Their native languages included the following: Spanish (21), Chinese (6), Korean (5), Russian (3), Portuguese (2), Thai (1), and Japanese (1). Participants included a total of 19 males and 20 females between the ages of 17 and 49 years. Speech samples for all 39 subjects were gathered as part of the

regular program language assessment activities that students participated in at the beginning and end of each semester of study. Participants' proficiency ranged from Novice Low to Advanced Mid as measured by the ACTFL OPI (Swender, Conrad, & Vicars, 2012) scale. The participants were organized into two groups, each described below.

Group # 1. This group consisted of 19 students who studied for three consecutive semesters without taking a break. This group took a placement test in the fall of 2014 (pre-test), a final test at the end of the same semester (Fall 2014, 15-week point), another final test at the end of the winter semester (April 2015, 30-week point), and another final test at the end of the summer semester (August 2015, 45-week point, post-test).

Group # 2. This group consisted of 20 students who, following their second semester of study, were allowed to take a 15-week break and then resume studies for a third semester. This group took a placement test in May of 2014 (pre-test), a final test at the end of the same semester (August 2014, 15-week point), another final test at the end of the fall semester (December 2014, 30-week point), a 15-week break (between January and April of 2015), and another final test at the end of the summer semester (August 2015, 45-week point, post-test).

Setting

The focus of the IEP in which the participants studied for three semesters is to assist international students in developing overall English proficiency with an emphasis on English for academic purposes. It has two major instructional blocks of classes: the foundations level block and the academic level block. The foundations program consists of four levels: Foundations Prep, Foundations A, B, and C with Foundations C being the

highest level in the block. The academic program also consists of four levels: Academic Prep, Academic A, B, and University Prep (the highest level in the academic block). Five major language skills are taught in the language program that correspond with the ACTFL OPI levels: Listening, Speaking, Reading, Writing, and Linguistic Accuracy (Applied Grammar).

Instruments

At the beginning of the course and at the end of each semester, students took a speaking test consisting of 12 tasks. For each level of proficiency from novice to intermediate, advanced, and superior, a total of three different tasks were used to elicit oral communication skills of the participants. These tests and tasks will now be described in detail.

Placement test. Prior to the beginning of each semester, the IEP administers a placement test. From the results of the test, students are placed in various proficiency levels consistent with the results of the test. In the program, each student is tested on a variety of different levels with tasks ranging from novice questions to superior questions. The placement test contains 12 speaking tasks scaled in difficulty from Novice to Superior (see Appendix A).

Level achievement test. Program level achievement tests are given at the end of each 15-week semester. Survey data identifying ESL student demographic factors (native language, age, etc.) are also collected. The program administers a final test at the end of every semester, which covers all levels in both the Foundations and Academic programs. The final exam is administered via a computer. Various tasks such as speaking (open oral and elicited response), writing (integrated writing, 30-minute essay and other tasks

depending on the level of proficiency), listening (listening for the main idea and supporting details), reading (reading passages and answering multiple-choice questions), and grammar (filling-in-the-blanks and error detection) comprise the computer-based test.

For the speaking test, the students are given 12 prompts to which they must respond in a certain period of time. Unknown to the learners, these prompts are divided into three major sections: Novice (with the Novice and Intermediate sublevels), Intermediate (with the Intermediate and Advanced sublevels), and Advanced (with the Advanced and Superior sublevels). The learners move from tasks that would be appropriate for lower-level learners, and continue with ones that increase in complexity. At task seven, the tasks then decrease in complexity gradually returning to tasks appropriate for a novice level. (See Appendix A for an example of the final test format.)

During the speaking portion of the placement test, all new students, who come to study at the language program, answer 12 questions that range from the Novice to Superior levels.

Novice level questions. The first two questions asked (1 and 2) are novice level questions, with the first one being a warm-up question. Students are asked to describe the weather, scenery, clothes they wore on the test day, etc.

Intermediate level questions. Questions 3 and 4 are intermediate level, and they both take more preparation and response time. These intermediate level tasks require the students to show their ability to use the future tense by describing the plans for the future and planning a party for their teacher.

Advanced level questions. The next group of questions (5 and 7) are at the advanced level, with Question 6 being a superior-level question. For these questions, the

students describe a popular holiday in their home countries, or resolve a potential unexpected problem.

Superior level questions. Questions 8 and 9 ask the students to discuss more abstract topics and choose a side of a hypothetical debate or dilemma. Furthermore, students discuss a philosophical quote.

Final questions. Starting from Question 10, the difficulty level of the questions decreases, and the examinees are asked to perform certain tasks such as retelling a story from their life (Advanced level), describing a routine (Intermediate level), and sharing their plans for the rest of the day (Novice level).

Following the scoring of students' responses by trained raters, the overall score received allows the student to either move up to the next proficiency level or remain at the same level for another semester. During this part of the placement test, trained raters used an ACTFL proficiency rubric to assign a specific score (from 0 to 6) to each student's performance.

Finally, depending on the average of their assigned scores given by multiple raters, students were placed into one of the three levels as identified below (see Table 1). Table 1 shows how the 12 questions in the level achievement test were divided into three proficiency levels or groups: Novice, Intermediate, and Advanced. Each level was then further divided into two subcategories that represented the level's floor (left column) and ceiling (right column). This type of division depicts the student's oral language abilities at each proficiency level. For instance, if the examinee's score is an average of 0 to 1.9 points on the speaking portion of the test, he/she is assigned to the Novice level; if the student receives a score between 2.0 and 3.9 points, he/she is assigned to the Intermediate

level; students who score between 4.0 and 5.9 are placed in the Advanced level group (see Table 1).

Table 1

Placement Test Level Assignment and Tasks

Proficiency Question #	Novice (0.0–1.9)		Intermediate (2.0–3.9)		Advanced (4.0–5.9)	
	Floor	Ceiling	Floor	Ceiling	Floor	Ceiling
	Novice	Intermediate	Intermediate	Advanced	Advanced	Superior
	1	3	3	5	5	6
	2	4	4	7	7	8
	12	11	11	10	10	9

Procedure

The data for this study were collected by the assessment coordinator at the IEP over four different intervals: the first data collection period occurred at the beginning of the first semester, and collection points for the other three time periods occurred at the end of each consecutive semester. For Group 1, the first interval was August 2014; the data for the other three intervals were collected in December 2014, April 2015, and August 2015, respectively. For Group 2, the first interval was May 2014; the other three phases were August 2014, December 2014, and August 2015, respectively.

Permission to gather data for research purposes was granted by the assessment coordinator and BYU's Institutional Review Board (IRB) for the purposes of conducting a longitudinal analysis of student performance over the course of three 15-week semesters. The data consisted of audio files that had been previously recorded. All student identification information was removed by the assessment coordinator and replaced with a student code number prior to the researcher receiving the data. Students' individual scores from each of the test periods were compared across all four data points. These four data points correspond to the four time periods at which data was collected (see Table 2).

Table 2

Data Collection Time Periods

Period	Type of Test
Time 1	Initial placement test (Pre-test, Week 1)
Time 2	Final exam (Week 15)
Time 3	Final exam (Week 30)
Time 4	Final exam (Posttest, Week 45)

Once the data were obtained, they were placed into four separate computer folders that corresponded to each of the four data points (e.g., Folder 1 – Time 1, Folder 2 – Time 2, and so forth). Each folder was further divided into three additional folders that corresponded to each of the three levels identified in Table 1 (e.g., Time 1 – Advanced, Intermediate, and Novice; Time 2 - Advanced, Intermediate, and Novice, etc.).

Analysis

Once the data was organized into files, a fluency analysis was performed on the speaking data using a Praat script designed to analyze the following measures of oral fluency: speech rate (total # of syllables divided by the total response time in seconds), articulation rate (total # of syllables divided by the sum of speech time and total pause time multiplied by 60), number of syllables (average number of syllables per response), number of pauses (average number of pauses per audio file), and phonation time (average total time in seconds spent speaking). Once these data were obtained, a repeated measures ANOVA was performed to determine whether significant differences were achieved for the different levels of proficiency over the course of the 45 weeks.

Results

The major purpose of this research was to examine how ESL students' fluency changed over a 45-week period by measuring particular features such as speech rate,

articulation rate, number of syllables, number of pauses, and phonation time. Another aim was to determine how changes in students' oral fluency features corresponded with their level of proficiency in the program during the 45-week time period. This section will present the analysis of the data.

Changes in Fluency Over 45 Weeks of Instruction

The first research question involved analyzing students' fluency over the course of 45 weeks of instruction. Data from the analysis of each feature will be presented along with a table and a graph showing the changes for each proficiency group during each of the time periods. Then, there will be a discussion of what interactions occurred in the ANOVA between the factors of level and time.

Speech Rate

As mentioned earlier, speech rate is a total number of syllables divided by the total response time in seconds. Table 3 provides the means and standard deviations for students in all three proficiency levels regarding speech rate. Figure 1 provides a plot of these same scores for speech rate for the three different proficiency levels across the 45-week period.

Table 3

Descriptive Statistics for Speech Rate Across the 45-Week Period

Level	Time	Mean	Std. Dev.
Advanced	Pretest	2.4177	.43068
	15 Weeks	2.4389	.49722
	30 Weeks	2.9425	.73132
	45 Weeks	3.1318	.82844
Intermediate	Pretest	1.8374	.69287
	15 Weeks	2.4009	.56091
	30 Weeks	2.6538	.58970
	45 Weeks	3.1322	.39651
Novice	Pretest	1.5034	.87848
	15 Weeks	1.9761	.58469
	30 Weeks	2.7139	.74929
	45 Weeks	2.8733	.98980

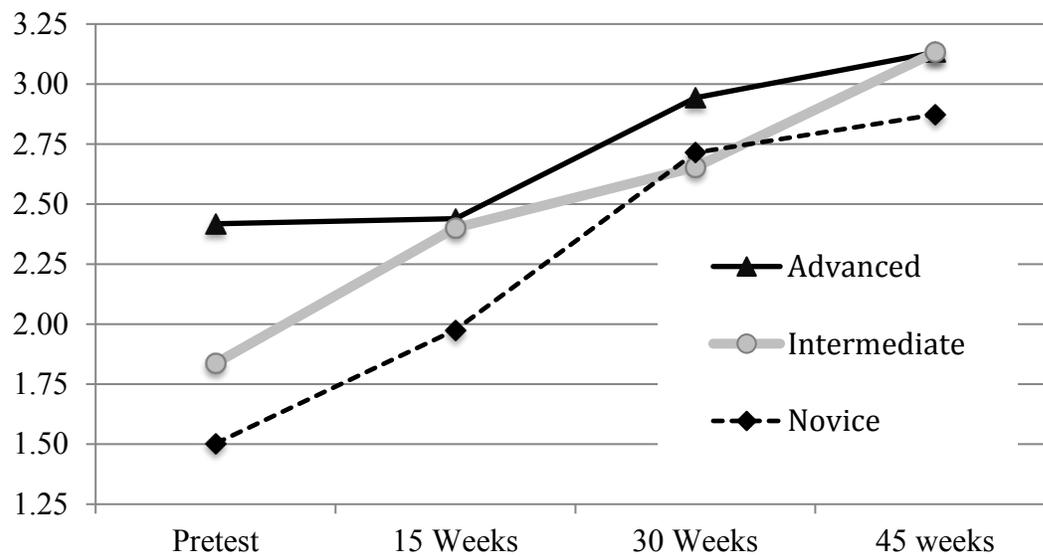


Figure 1. Speech rate changes over time by level of proficiency.

For learners at all three proficiency levels, there were significant differences in their overall speech rate from the pretest to the final exam period at the end of 45 weeks. The Cohen's d showed a large effect for all three groups (Advanced, $d = 1.08$; Intermediate, $d = 2.293$; Novice $d = 1.463$) indicating that each group had made significant progress from the first test to the last test three semesters later (see Figure 1).

Pretest. As expected, the comparison of speech rates across levels at the pretest was statistically significant, $F(2, 465) = 48.060, p < .001$ showing that the groups statistically differed from each other with regards to the feature of speech rate. The effect size¹ between the advanced group and the intermediate group was large ($d = 1.005$) while the difference between the intermediate and novice learners was on the border of moderate ($d = .4223$).

Posttest (45 weeks). After 45 weeks of instruction, the performance of the three proficiency levels for speech rate was much more closely clustered than at the pretest. For example, while all three groups had improved, there was no meaningful difference between the advanced group compared to the intermediate group ($d = 0.0$), and only a small effect size in speech rate when comparing the intermediate and novice level learners ($d = .344$).

Level by Time. The ANOVA results for the factor of time, showed that there were statistically significant differences for all learners in their speech rate from the pretest to the final testing period at the end of 45 weeks of instruction, $F(3, 1859) = 232.604, p < .001, \eta_p^2 = .273$. This finding is further supported by fact that the eta squared showed a large effect size. When checking to see if there was an interaction between proficiency

¹ Cohen's guidelines for interpreting effect sizes (d): .20 = small, .50 = medium, .80 = large.

level and time, the comparison was statistically significant, but the effect size was negligible, $F(6, 1859) = 10.517, p < .001, \eta_p^2 = .033$. These findings indicate that all three proficiency levels improved the speed at which they produce utterances. Interestingly, the intermediate and advanced level learners' oral fluency was almost identical (3.13 syllables per second) after 45 week of formal instruction. This could be explained by the advanced students' lack of motivation or having taken the TOEFL test and having received a score sufficient to be accepted to an American university.

Articulation rate

As mentioned in the previous chapters, articulation rate is a total number of syllables divided by the sum of speech time and total pause time multiplied by 60. Table 4 provides the means and standard deviations for the students of all three proficiency levels regarding changes over 45 weeks. Figure 2 presents articulation rate results for the three different proficiency levels across the 45-week period.

Table 4 Descriptive Statistics for Articulation Rate Across the 45-Week Period

Level	Time	Mean	Std. Dev.
Advanced	Pretest	3.5672	.52087
	15 Weeks	3.3550	.70318
	30 Weeks	3.6781	.67252
	45 Weeks	3.9078	.85650
Intermediate	Pretest	3.2038	.64286
	15 Weeks	3.3999	.58101
	30 Weeks	3.6560	.52570
	45 Weeks	3.9738	.35220
Novice	Pretest	2.7678	1.09947
	15 Weeks	3.2844	.54344
	30 Weeks	3.6977	.58971
	45 Weeks	3.5797	1.09794

² Cohen's guidelines for interpreting effect sizes (η_p^2): .01 = small, .06 = medium, .14 = large.

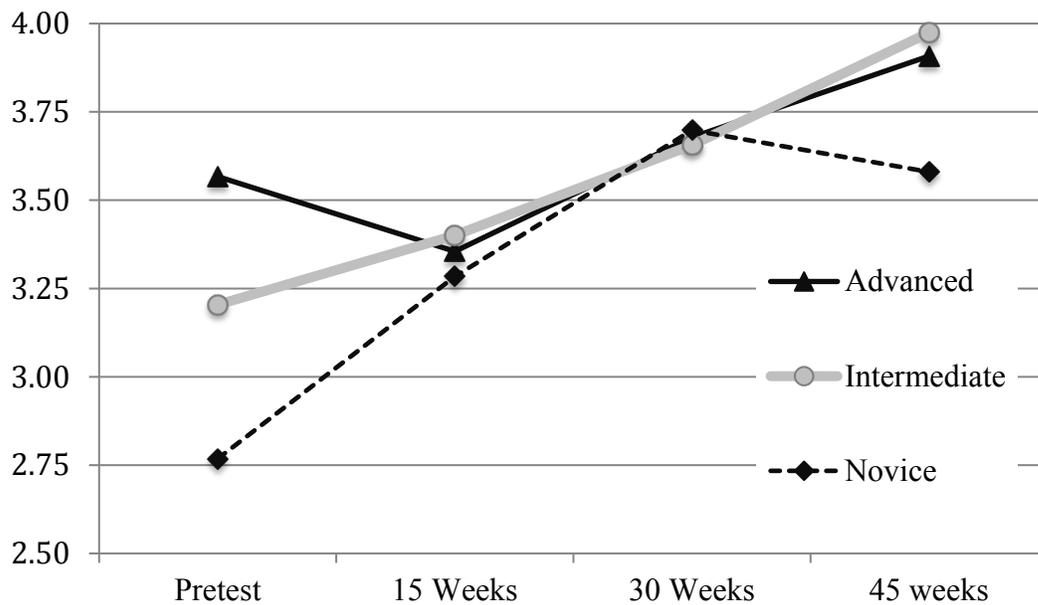


Figure 2. Articulation rate changes over time by level of proficiency.

The data showed a significant difference in articulation rate for all three proficiency levels from the pretest to the final test after 45 weeks of instruction. The Cohen's d showed borderline moderate to large effect sizes for the different levels: Advanced, $d = .480$; Intermediate, $d = 1.486$; Novice, $d = .738$. This data suggests that the students at all three proficiency levels were able to become more efficient in producing speech (number of syllables per second) by saying more within the time period in which they spoke.

Pretest. The comparison across levels at the pretest was statistically significant for articulation rate, $F(2, 465) = 30.788, p < .001$, with a medium effect size between the advanced and the intermediate groups ($d = .620$) and the intermediate and novice group of learners ($d = .484$). This data again showed that students in all three levels of proficiency were quite different from each other for this fluency feature at the beginning of the study.

Posttest (45 weeks). After 45 weeks of instruction, there was no significant difference in performance between the advanced and intermediate proficiency level groups, differing in their articulation rates by only 0.074. There was also a negligible effect size between the advanced group and intermediate groups ($d = .100$), but there was a medium effect size between the intermediate and novice groups of learners ($d = .482$).

Level by Time. The ANOVA results across the factor of time from the pretest to the final exam at the end of week 45 showed significant differences for all three levels of proficiency, $F(3, 1859) = 74.500, p < .001, \eta_p^2 = .107$, with a medium effect size for all three levels of proficiency. In analyzing the interaction between level of proficiency and time for articulation rate, the results were statistically significant, but the effect size was negligible, $F(6, 1859) = 10.095, p < .001, \eta_p^2 = .032$. These findings indicate that all three proficiency levels improved the speed at which they produce utterances. Interestingly, the intermediate level students' articulation rate (3.97 syllables per second) exceeded the advanced level students' (3.90 syllables per second) over the 45-week period.

Number of Syllables

Table 5 provides the means and standard deviations for the fluency feature *number of syllables* as produced by the students at all three proficiency levels over the course of 45 weeks of instruction. Figure 3 presents average number of syllables results for the three different proficiency levels across the 45-week period.

Table 5

Descriptive Statistics for Number of Syllables Across the 45-Week Period

Level	Time	Mean	Std. Dev.
Advanced	Pretest	135.1354	57.81579
	15 Weeks	136.7500	63.74125
	30 Weeks	165.6875	83.46740
	45 Weeks	172.1789	87.32697
Intermediate	Pretest	101.5833	57.77335
	15 Weeks	134.7222	64.14081
	30 Weeks	147.0833	69.92809
	45 Weeks	172.2593	75.07700
Novice	Pretest	81.7436	59.64582
	15 Weeks	110.9038	57.52327
	30 Weeks	150.2244	77.32517
	45 Weeks	158.9167	90.84037

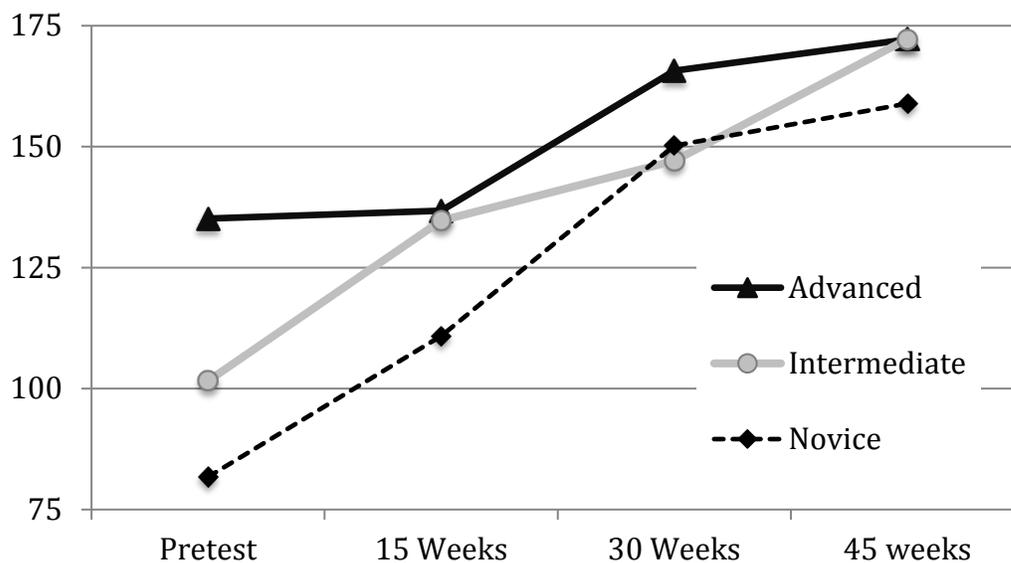


Figure 3. Average number of syllabus per response over time by level of proficiency.

Learners at all three proficiency levels showed a significant difference in the numbers of syllables produced in their responses between the time of the pretest and the end of instruction at 45 weeks. A medium to large effect size occurred for each of the

levels: Advanced, $d = .500$; Intermediate, $d = 1.055$; Novice, $d = 1.004$. The large effect size for the intermediate and the novice-level students showed that their speech became more multisyllabic and more language was being produced per speaking response on average by these learners.

Pretest. For the fluency feature, *number of syllables*, the comparison across proficiency levels at the pretest was statistically significant, $F(2, 465) = 24.830, p < .001$. The effect size between the advanced group and the intermediate group showed a medium effect size ($d = .581$) while the difference between the intermediate and novice learners showed a small effect size ($d = .338$).

Posttest (45 weeks). After 45 weeks of instruction, the results were quite similar for the advanced and intermediate students regarding the number of syllables produced. The difference in effect size between the two groups was negligible ($d < .001$). When comparing the intermediate level speakers to the novice learners, a small effect size was produced ($d = .160$).

Level by Time. Data from the ANOVA showed that that the interaction between Level and Time for the number of syllables produced by the learners was significant, $F(6, 1859) = 3.226, p = .004, \eta_p^2 = .010$. The effect size, however, was small. For the factor of time alone, the differences were statistically significant with a medium effect size, $F(3, 1859) = 364.050, p < .001, \eta_p^2 = .094$.

Number of Pauses

Table 6 provides the means and standard deviations for the students of all three proficiency levels for the number of pauses they produced for each response. Figure 4 shows the results for each group across the 45-week period.

Table 6

Descriptive Statistics of Number of Pauses Across the 45-Week Period

Level	Time	Mean	Std. Dev.
Advanced	Pretest	18.8021	10.07824
	15 Weeks	17.8438	11.60066
	30 Weeks	14.2813	10.95417
	45 Weeks	14.6737	9.70571
Intermediate	Pretest	19.7083	10.20548
	15 Weeks	19.6667	11.56016
	30 Weeks	19.3009	11.04103
	45 Weeks	17.5046	10.69579
Novice	Pretest	15.2821	10.75762
	15 Weeks	22.9231	11.72180
	30 Weeks	18.3846	11.13464
	45 Weeks	15.8397	11.02987

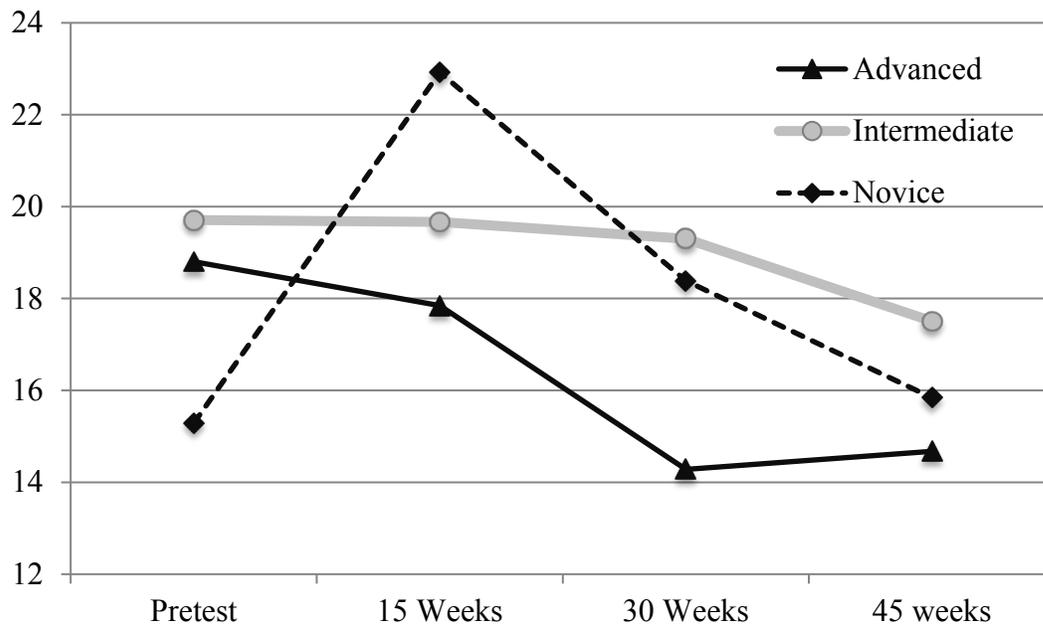


Figure 4. Average number of pauses per response over time by level of proficiency.

Over a period of 45 weeks, a small effect size was found to occur for the intermediate and advanced level learners. For novice learners, the effect was negligible

with the pretest average score and the score at the end of 45 weeks being virtually the same (Advanced, $d = .417$; Intermediate, $d = .211$; Novice, $d = .050$).

Pretest. The comparison across levels at the pretest was statistically significant, $F(2, 465) = 8.576, p < .001$. The effect size between the advanced group and the intermediate group was negligible ($d = .089$), however, while the difference between the intermediate and novice learners was borderline moderate ($d = .422$).

Posttest (45 weeks). After 45 weeks of instruction, only slight differences were seen in the mean scores for each of the proficiency levels. The ANOVA results at the posttest were not significant when comparing the three groups and the effect sizes were small for the feature of pausing (advanced group compared to the intermediate group produced a small effect size ($d = .277$), as did the intermediate compared to the novice learners ($d = .153$)).

Level by Time. When considering the ANOVA data for the interaction between level and time, the results were statistically significant, but the effect sizes were very small, $F(6, 1859) = 5.623, p < .001, \eta_p^2 = .018$. For the factor of time alone, again the results were statistically significant, but the effect size was negligible, $F(3, 1859) = 10.462, p < .001, \eta_p^2 = .017$ implying that no measureable change was made for the proficiency levels when it came to reducing the average number of pauses used in their speaking samples from the pretest to the end of instruction at 45 weeks.

Phonation Time

Table 7 provides the means and standard deviations for the students of all three proficiency levels showing the changes in values over 45 weeks of instruction. Figure 5

presents a graph demonstrating the changes in average mean scores for each of the three levels of proficiency.

Table 7

Descriptive Statistics of Phonation Time Across the 45-Week Period

Level	Time	Mean	Std. Dev.
Advanced	Pretest	38.3921	17.03478
	15 Weeks	41.5921	18.66793
	30 Weeks	44.8438	21.22195
	45 Weeks	43.7063	18.91994
Intermediate	Pretest	31.6115	16.56582
	15 Weeks	39.9858	18.02394
	30 Weeks	40.0692	17.41609
	45 Weeks	43.2112	18.19856
Novice	Pretest	32.8989	21.21246
	15 Weeks	33.9953	16.82987
	30 Weeks	40.6385	19.30146
	45 Weeks	44.5100	19.12198

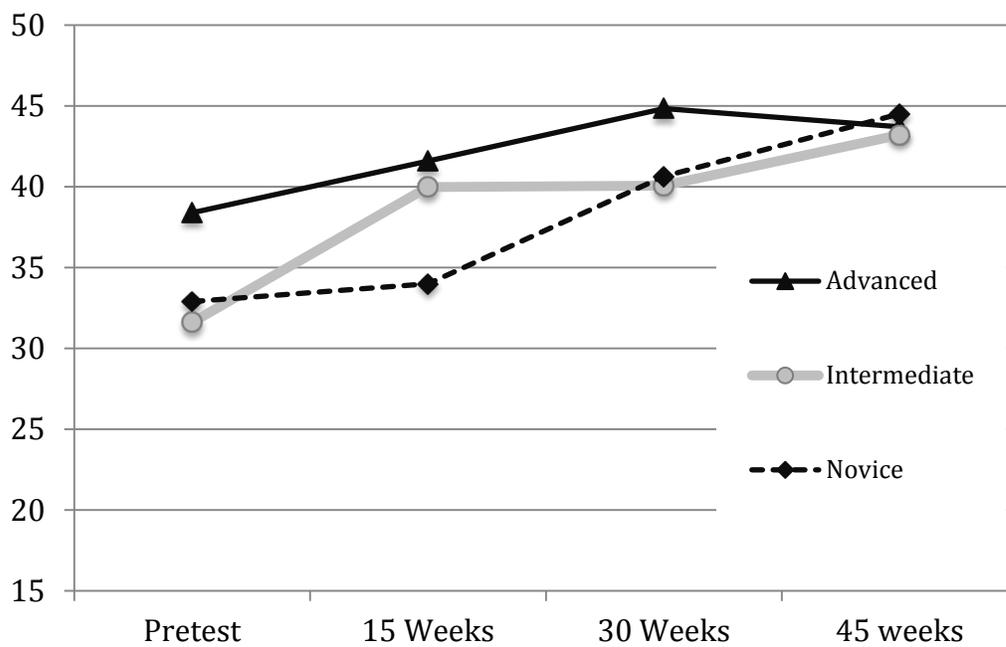


Figure 5. Phonation time changes over time by level of proficiency.

Over a period of 45 weeks of instruction, all three levels of proficiency demonstrated a small to medium effect size (Advanced, $d = .295$; Intermediate, $d = .666$; Novice, $d = .575$) when considering changes in the fluency feature of phonation time.

Pretest. The ANOVA results showed that at the pretest, there differences between the three levels of proficiency were again statistically significant, $F(2, 465) = 4.844, p = .010$. When comparing the advanced group to the intermediate group, there was an effect size of borderline moderate ($d = .404$). The effect size comparing the difference between the intermediate and novice learners, however, was too small to be meaningful ($d = .067$).

Posttest (45 weeks). At the end of 45 weeks of instruction, the three proficiency levels were much more closely clustered than at the pretest for this feature. The ANOVA results showed that differences between the three groups was not significant and the effect sizes comparing each of the three levels of proficiency were very small (advanced group compared to the intermediate group - ($d = .026$) and the intermediate compared to the novice learners - ($d = .069$)).

Level by Time. In looking at the ANOVA results for the interaction between level and time, the results were statistically significant, but the effect size was negligible, $F(6, 1859) = 2.469, p = .022, \eta_p^2 = .008$.

Discussion

The purpose of this study was to explore the extent to which five different fluency features changed for adult ESL learners enrolled in an English for academic purposes IEP. The features investigated in this research included speech rate (syllables/second), articulation rate (syllables per second), number of syllables, number of pauses, and phonation time (seconds per syllable).

As can be seen from the results of the repeated measures ANOVA, all three levels of proficiency did make significant improvement over the 45-week instructional period for the fluency features of speech rate, articulation rate, and number of syllables produced per response. Each of these features will be discussed in greater detail below.

Gains Based on Fluency Features

Speech rate. Based on these particular results, one can infer that unlike the rest of the fluency variables, learners at all three proficiency levels improved their speech rate, in particular the intermediate and novice level learners. Figure 1 depicts the progress of speech rate for all three levels from the placement test to the final Level Achievement Test. It is critical to note that even though all three levels, as expected, had very different speech rates at the pretest ($d = 1.005$), the advanced and intermediate speakers' scores were identical or nearly identical at the 15-week period as well as the 45-week period ($d = 0.0$). While continued improvement in this variable as identified at week 45, it does appear that between weeks 30 and 45, learners at the advanced and novice levels are beginning to plateau in their rate of speech. After having spent nearly 12 months in an English-speaking environment, learners may be developing a rate of speech with which they feel comfortable communicating in their L2.

Articulation rate. The data shows that all three groups were significantly different from each other at the pre-test, with the novice speakers having the lowest mean score (2.77 syllables), intermediate learners' mean score was 3.20 syllables, and advanced learners had the highest mean score (3.57 syllables). It is interesting to note that at the 15-week point, the mean articulation rate for all three groups converges with the mean rate dropping for the advanced learners over the first 15 weeks, while the

intermediate and novice level learners are still improving. During the second semester (weeks 15 to 30), all three groups appear to have parallel growth, with steady improvement in the number of syllables produced and along with a few pauses. In terms of articulation rate, the intermediate learners make consistent progress over the 45-week period ending with producing a larger average of syllables per number of pauses multiplied by 60 (3.97 syllables) versus the advanced group (3.91 syllables).

Number of syllables. Similar to the speech rate, all three proficiency levels were significantly different from each other in terms of number of syllables per response produced at the pre-test. Figure 3 indicates more considerable progress of number of syllables for the novice and intermediate learners, with the novice speakers exceeding the intermediate group after 30 weeks. The number of pauses for the advanced students appeared to start plateauing between week 30 and week 45. As a result, the novice and intermediate groups were progressing steadily and consistently while the advanced students did not show much progress after 30 weeks of instruction for this variable.

Number of pauses. In order to answer the second research question related the fluency scores corresponding to the students' proficiency level at placement and Level Achievement Test, we can look at Figure 4 and identify that the number of pauses for the intermediate learners (19.71 pauses) was higher than the one for the advanced group (18.86 pauses). It is interesting to note that the novice learners made fewer pauses than the other two groups (15.28). However, this pattern changes over time. Unlike the advanced and intermediate levels, who tend to reduce their pauses very consistently, the novice speakers' score increases to 22.92 pauses after 15 weeks. At 45 weeks, all three

levels were significantly different from each other, but the effect sizes were so small that they make this result meaningless.

Phonation time. The results in Figure 5 suggest that all three levels were able to make progress in terms of phonation time. At the pre-test, the novice and intermediate learners were very close to each other ($d = .067$). Similarly, at the 45-week mark the results were almost identical ($d = .069$); however over the 45-week period both the novice and intermediate speakers were able to make gains in their phonation time (novice: from 32.9 sec to 44.51 sec; intermediate: from 31.61 sec to 43.21 sec), with the novice learners exceeding both the advanced and intermediate groups.

Gains Based on Proficiency Level

Novice. Pretest – 15 weeks. The number of syllables for this particular level kept increasing throughout the entire 45-week study. During the first semester, the novice level learners were able to improve their number of syllables from 81.7 to 110.9 per response; they also increased their speech rate from 1.5 to 1.9 syllables/second as well as their phonation time (from 32.9 to 34.0 syllables/second). Moreover, their articulation rate progressed from 2.8 to 3.3 syllables/second, which is very impressive since the other two groups took more time to improve their articulation rate. On the downside, these beginner students made by far more pauses than their counterparts from advanced and intermediate levels (15.3 vs. 22.9 pauses/response).

15 weeks - 30 weeks. During the second semester, novice students were able to improve the number of syllables from 81.7 to 110.9 syllables per response. Their speech and articulation rates also increased: from 2.0 to 2.7 and from 3.3 to 3.7 syllables/second

respectively. Furthermore, this group of students reduced the number of pauses from 22.9 to 18.4 pauses/response.

30 weeks – 45 weeks. During the last 15 weeks of instruction, novice students improve in almost all aspects of oral fluency. The number of syllables increased from 150.2 to 158.9 syllables per response, which is much higher than the intermediate level in the previous semester. These novice learners were also able to significantly reduce the number of pauses from 18.4 to 15.8 pauses/response. Interestingly, they reduced their pause frequency below that of the intermediate students whose final score was 17.5 pauses/response. Just as during the previous two semesters, this group of students improved their speech rate from 2.7 to 2.9 syllables per second, and phonation time from 40.6 to 44.5 seconds per response. The only two fluency variables that did not show much improvement were articulation rate (3.7 vs. 3.6 syllables per second)

Intermediate. Pretest – 15 weeks. As one can infer from Figures 1-5, the intermediate students were very consistent in their progress and fluency gains. They increased the number of syllables produced in the speaking tasks from 101.6 to 134.7 (per response); speech rate stretched from 1.84 to 2.40 syllables/second; articulation rate went from 3.2 to 3.4 syllables/second; and their phonation time reached 40.0 seconds as opposed to 31.6 on the initial placement test. The number of pauses, on the contrary, remained on the same level – 19.71 vs. 19.67 (per response).

15 weeks - 30 weeks. The second semester or the period between 15 and 30 weeks was also very successful for the intermediate students regarding number of syllables (147.1 vs. 134.7 during the previous 15-week period), speech rate (2.65 vs. 2.40 syllables/second during the first semester), and articulation rate (3.66 syllables/second as

opposed to 3.40). On the other hand, some fluency variables remained on the same level: phonation time (from 39.99 to 40.07 seconds); while others even declined – number of pauses went from 19.7 to 19.3

30 weeks – 45 weeks. During the last 15-week period intermediate students showed rather steady and consistent results. The number of syllables increased from 147.1 to 172.2 per response; speech rate, similarly to the second semester, also rose from 2.65 to 3.13, which is remarkable since the advanced students reached the exact same result (3.13). Unlike the previous 15-week period, where students' phonation time plateaued, this time it increased to 44.2 seconds, which is slightly less significant than the advanced students whose phonation time was 43.7 seconds per response.

Advanced. Pretest – 15 weeks. As the findings suggest, ESL learners in the advanced level have not improved a great deal in terms of speech rate (2.42–2.44 syllables/second), and number of syllables produced (135.8–136.75 per response) in the speaking tasks during the first 15 weeks of instruction. Additionally, their articulation rate went down from 3.57 to 3.36 syllables/sec during the first semester. There was no significant change with regards to the other variables; however, the advanced students did make gradual progress during this time period.

15 weeks - 30 weeks. During the 15 to 30 week period, both the speech rate (from 2.44 to 2.94 syllables/second) and number of syllables produced (from 136.75 to 165.7 per response) began to improve for the advanced speakers. One can also observe, that all five variables tracked in this particular research showed progress.

30 weeks – 45 weeks. During the last 15 weeks of instruction, the results from this group somewhat declined compared to the previous 15 weeks. In terms of phonation

time, the number changed from 44.8 to 43.7 seconds. In addition to the decline in phonation time, the number of pauses (14.3 vs. 14.7) also improved, but to a very small degree. During the last 15 weeks, this group of advanced students did not produce results consistent with those achieved during the second semester of instruction.

Implications

The following section will provide an overview this study's implications. The data indicates that regardless of the proficiency level at which the students entered the IEP, learners made progress in their English language fluency over the course of three semesters. Unlike the study by Llanes and Muños (2009), who found that “participants with lower proficiency levels showed comparatively greater gains ... producing more accurate and fluent speech” (p. 361), learners at each of the proficiency levels in this study made significant gains in their speech rate, articulation rate, and number of syllables produced. While the data showed significant for all of the proficiency groups in the number of pauses produced, the effect size was negligible. Similar results were also found for the feature of phonation time.

The feature of speech rate appeared to be feature in which all learners in the study made the most progress. This finding is in keeping with results from other studies on fluency (Lambert & Kormos, 2014; Ginther, Dimova, & Yang, 2010) where learners also made significant gains in speech rate. Even with this progress though, when compared to the target speech rate of native speakers of English (Derwing and Munro, 2001), the advanced level learners average rate of 3.908 syllables per second was nearly a second slower than the rate produced by native speakers of English (4.7 syllables per second). While all learners did make considerable progress in their average speech rate, further

practice will be needed for all of the ESL learners if they hope to reach the speech rate level identified being used by native speakers of English.

For teachers and administrators, the data shows that while learners continue to make improvement in the fluency features of speech rate, articulation rate, and number of syllables produced over the course of 45 weeks of instruction, the rate of progress does shift within individual semesters depending on the proficiency level of the learners. In general, the effect sizes show that learners tend to make less progress during the third semester of study. In this particular IEP program, possible reasons for this could be that students have started to become weary of studying the second language. At this point in their studies, they would have lived in the United States for nearly nine months. For the advanced level learners, another possibility is that there are students who have passed the TOEFL exam with a sufficiently high score to enable them to be admitted to an American university. For student who accomplish this goal, there may be less motivation to continue to push themselves to perform well in the intensive program. Regardless of the reason for why a plateau effect appears to be happening, teachers and administrators would be wise to make sure that for students enrolled in a third semester of study, they have the opportunity to participate in alternative courses or program activities which encourage students to continue improving their fluency development.

Limitations and Directions for Future Research

One of the limitations of this study was the limited number of participants who had studied at the EIP for 45 weeks. In this particular intensive English program, an average of 245 students study each semester. While a large percentage of these students study for two semesters, fewer than a fourth of the students typically study for a period of

three semesters. The findings in this study are based on the scores of 39 students, a small but representative sample. This sample size may affect the reliability and accuracy of the results. One way of continuing to verify the results of this study would be to continue to track additional students who have studied over three semesters and add their scores to this existing data pool. Additional data collected from future semesters would allow researchers and administrators to see if the trends identified in this study remain consistent.

Another limitation is related to the use of Praat in measuring fluency. While Praat has the capability to generate numerical values for fluency features, it showed some limitations in processing data for this particular research study. Certain file extensions such as *mp3* were not recognized by Praat, and therefore were not generated properly without warning the user. As a result, the number of files *uploaded* on Praat was different from number of files *analyzed* by Praat. Expanding the types of files Praat that can process would greatly benefit data processing.

Finally, obtaining accurate measurements through the use of the Praat script can be problematic due to noise in speech samples. If the noise in the speech sample is sufficiently high, the Praat script cannot segment the speech sample properly. In this particular study, some of the audio files that included long pauses were confused with syllables; consequently, certain audio files had to be deleted from the data due to inaccuracies in processing specific features such as number of syllables.

While this study carefully analyzed several important fluency features consistent with those analyzed by other studies investigating oral fluency, additional features could be analyzed in future research. One of these features is mean length of run, which is a

mean number of syllables occurring between unfilled pauses of no less than 0.2 seconds” (Cucchiari, 2000, p. 994). This particular feature would be helpful because it would provide data regarding the learners’ ability to produce more complex utterances (Ginther et al., 2010; De Jong et al, 2013) during speaking tasks.

In addition to measuring mean length of run, data from native speakers completing the same 12 speaking tasks of the placement and LAT tests could be gathered and compared to the results from the non-native English speakers (NNES) involved in this study. Comparing the NNES’ speech rate and other fluency features with those of native speakers, for example, would provide an opportunity to see how closely the NNES participants’ fluency features have reached a target like comparison.

It is important to recognize that fluency is only one measure of language proficiency. Further study into other dimensions of language proficiency could also be done, namely, investigating longitudinally the dimensions of accuracy and complexity. This data would provide a more holistic view of ESL learners’ language development over an extended period of time.

Conclusion

The principle objective of this longitudinal study was to investigate fluency gains in oral speech by ESL learners enrolled in an intensive English language program. Five fluency variables – speech rate, articulation rate, number of syllables, number of pauses, and phonation – were measured in order to monitor the changes over the course of three semesters and identify which features changed the most and when in the course of instruction did these gains occur.

While improvement in oral fluency is a gradual process in which instruction appears to have an important affect, the results imply that the first 30 weeks of instruction seem to be the most crucial time period for learners studying for an extended period of time. Novice and intermediate level learners also appear to make the most dramatic progress when compared to advanced level learners. This finding is not surprising given that they are still in the process of becoming functionally communicative in the target language, actively learning large amounts of vocabulary and increasing the number of contexts in which they can appropriately function with the language.

Longitudinal research of second language learning provides important findings that inform learners, teachers, and administrators. It is hoped that further studies such as this one will be done to help to validate and clarify the changes in oral fluency and proficiency that second language learners do make through prolonged periods of instruction.

References

- Baker-Smemoe, W., Dewey, D. P., Bown, J. & Martinsen, R. A. (2014). Does measuring L2 utterance fluency equal measuring overall L2 proficiency? Evidence from five languages. *Foreign Language Annals*, 47, 707-728.
- Baró, À. L. & Serrano, R. S. (2011). Length of stay and study abroad: Language gains in two versus three months abroad. *Revista Española de Lingüística Aplicada*, 24, 95-110.
- Blake, C. (2009). Potential of text-based internet chats for improving oral fluency in a second language. *The Modern Language Journal*, 93, 227. Retrieved from <http://search.proquest.com/docview/197294219?accountid=4488>.
- Boersma, P. & van Heuven, V. (2001). Speak and unSpeak with PRAAT. *Glott International*, 5(9-10), 341-347.
- Chambers, F. (1997). What do we mean by fluency? *System*, 25, 535-544. Retrieved from <http://search.proquest.com/docview/85652604?accountid=4488>.
- Christensen, C. V. (2012). *Fluency features and elicited imitation as oral proficiency measurement*. Retrieved from <http://scholarsarchive.byu.edu/etd/3114/>
- Crystal, D. (1987). *The Cambridge Encyclopedia of Language*. Cambridge, UK: Cambridge University Press.
- Cucchiaroni, C., Strik, H. & Boves, L. (2000). Quantitative assessment of second language learners' fluency by means of automatic speech recognition technology. *The Journal of the Acoustical Society of America*, 107, 989-999.

- De Jong, N. H., Steinel, M. P., Florijn, A., Schoonen, R. & Hulstijn, J. H. (2013). Linguistic skills and speaking fluency in a second language. *Applied Psycholinguistics*, 34, 893-916.
- De Wet, F., Van der Walt, C. & Niesler, T. R. (2009). Automatic assessment of oral language proficiency and listening comprehension. *Speech Communication*, 51, 864-874.
- Delaney, T. (2012). Quality and quantity of oral participation and English proficiency gains. *Language Teaching Research*, 16, 467-482.
- Derwing, T. M. & Munro, M. J. (2001). What speaking rates do non-native listeners prefer? *Applied Linguistics*, 22, 324-337.
- Derwing, T. M., Munro, M. J., & Thomson, R. I. (2008). A longitudinal study of ESL learners' fluency and comprehensibility development. *Applied Linguistics*, 29, 359-380.
- Fillmore, C. J. (1979). On fluency. In D. Kempler, and W.S.Y. Wang (Eds.), *Individual differences in language ability and language behavior*. New York, NY: Academic Press.
- Ginther, A., Dimova, S. & Yang, R. (2010). Conceptual and empirical relationships between temporal measures of fluency and oral English proficiency with implications for automated scoring. *Language Testing*, 27, 379-399.
- Götz, S. (2013). *Fluency in native and nonnative English speech*. Amsterdam/Philadelphia: John Benjamins Publishing Company.
- Ingham, R. J., Bothe, A. K., Wang, Y., Purkhiser, K. & New, A. (2012). Phonation

- interval modification and speech performance quality during fluency-inducing conditions by adults who stutter. *Journal of Communication Disorders*, 45, 198-211. Retrieved from <http://search.proquest.com/docview/1221441116?accountid=4488>.
- Kormos, J. & Dénes, M. (2004). Exploring measures and perceptions of fluency in the speech of second language learners. *System*, 32, 145-164.
- Lambert, C. & Kormos, J. (2014). Complexity, Accuracy, and Fluency in Task-based L2 Research: Toward More Developmentally Based Measures of Second Language Acquisition. *Applied Linguistics*, 35, 607-614.
- Llanes, A. & Muñoz, C. (2009). A short stay abroad: Does it make a difference? *System*, 37, 353-365. Retrieved from <http://ejournals.ebsco.com.erl.lib.byu.edu/Direct.asp?AccessToken=9515XI18XI X9KRI9KR5MPZ9ZEIQ48J9JI&Show=Object>
- McCarthy, M. (2009). Rethinking spoken fluency. *Estudios De Linguística Inglesa Aplicada (ELIA)*, 9, 11-29. Retrieved from <http://search.proquest.com/docview/744447353?accountid=4488>.
- Miller, J. & Schwanenflugel, P. J. (2008). A longitudinal study of the development of reading prosody as a dimension of oral reading fluency in early elementary school children. *Reading Research Quarterly*, 43, 336-354.
- Nakano, M., Kondo, Y. & Tsutsui, E. (2008). Fundamental research on automatic speech evaluation of L2 English. *Age*, 23(4.42), 18-38.

- O'Brien, I., Segalowitz, N., Freed, B. & Collentine, J. (2007). Phonological memory predicts second language oral fluency gains in adults. *Studies in Second Language Acquisition, 29*, 557-581.
- Pinget, A. F., Bosker, H. R., Quené, H. & de Jong, N. H. (2014). Native speakers' perceptions of fluency and accent in L2 speech. *Language Testing, 31*, 349-365.
- Polat, B. & Kim, Y. (2014). Dynamics of complexity and accuracy: A longitudinal case study of advanced untutored development. *Applied linguistics, 35*, 184-207.
- Rossiter, M. J., Derwing, T. M., Manimtim, L. G. & Thomson, R. I. (2010). Oral fluency: The neglected component in the communicative language classroom. *Canadian Modern Language Review, 66*, 583-606. Retrieved from <http://search.proquest.com/docview/881456311?accountid=4488>.
- Sato, M. (2014). Exploring the construct of interactional oral fluency: Second language acquisition and language testing approaches. *System: An International Journal of Educational Technology and Applied Linguistics, 45*, 79-91. Doi: <http://dx.doi.org/10.1016/j.system.2014.05.004>.
- Skehan, P. (2009). Modeling second language performance: Integrating complexity, accuracy, fluency, and lexis. *Applied Linguistics, 30*, 510-532. Retrieved from <http://search.proquest.com/docview/61815690?accountid=4488>
- Swender, E., Conrad, D., & Vicars, R. (2012). *ACTFL proficiency guidelines 2012*. Alexandria, VA: American Council for the Teaching of Foreign Languages. Retrieved from http://www.actfl.org/sites/default/files/pdfs/public/ACTFLProficiencyGuidelines2012_FINAL.pdf

- Ushigusa, S. (2009). *The relationships between oral fluency, multiword units, and proficiency scores* (Order No. AAI3344157). Available from Linguistics and Language Behavior Abstracts (LLBA). (85699039; 200921334). Retrieved from <http://search.proquest.com/docview/85699039?accountid=4488>.
- Xi, X., Higgins, D., Zechner, K. & Williamson, D. M. (2008). Automated scoring of spontaneous speech using SpeechRaterSM v1. 0. *ETS Research Report Series, 2008(2)*, i-102.
- Zechner, K., Higgins, D., Xi, X. & Williamson, D. M. (2009). Automatic scoring of non-native spontaneous speech in tests of spoken English. *Speech Communication, 51*, 883-895.

APPENDIX: Sample Placement/LAT Prompts at ACTFL Proficiency Levels

Item Level ³		Prompt	Time (sec)	
			Prep	Speak
1	N	Describe the weather and scenery as you came to take this test today.	15	30
2	N	Describe what you are wearing today. List the clothes and identify their color, material, and other characteristics. Also include your reason for choosing to wear them.	15	30
3	I	Do your best to describe where you will be and what you will be doing one year in the future. How will your life be different? How will it be the same? What events will happen between now and one year from now?	15	45
4	I	You and your classmates want to plan a party for one of your teachers who is moving after this semester. What are several questions you should ask your teacher in order to plan a party that she would like?	15	30
5	A	Describe a holiday in your country that the U.S. does not celebrate. What is the reason for the holiday? How do people celebrate? What are things that a person would see, do or eat if they visited your country during that holiday?	30	60
6	S	Two friends are having a debate. One friend believes that playing video games is a waste of time, and parents should prohibit their use. The other friend believes that children can acquire valuable skills from video games, and parents should facilitate their use. Choose one side of this argument to support and explain your reasons for having your opinion.	30	90
7	A	You are working with a group of classmates to complete a presentation. Your responsibility was to create a media presentation with information and pictures that other group members researched. On the day of the presentation, you lose the USB drive containing the presentation and all the information the group had collected. Explain to your group members what happened and describe a series of actions that the group should do to reach the best result.	30	60
8	S	In many countries, people are moving from rural areas into urban areas. Discuss the short term and long term consequences of this type of population movement.	30	90

³ The letters denote the level of the task as identified by the ACFTL proficiency scale: N – Novice, I – Intermediate, A – Advanced, S – Superior.

9	S	Edwin Land, an American inventor, said, “An essential aspect of creativity is not being afraid to fail.” Discuss the principle behind this expression. In that way is it true or accurate? Who should learn from it and how should their actions change?	30	90
10	A	Retell a story from your life when you or someone you know won a prize or award. Include a detailed description of the events before, during and after this experience. How or why was this experience memorable to you?	30	60
11	I	A friend from your hometown asks about what you do on the weekend now that you live in the U.S. Describe your routine on a typical Saturday from the morning to the evening. What do you do? Where do you go? Who are you with? How is it different than weekends in your hometown?	15	45
12	N	What are your plans for the rest of the day? What will you do to relax and enjoy the time following your test?	15	30
