STARS

Electronic Theses and Dissertations, 2004-2019

2013

The Effect Of Input Modality On Pronunciation Accuracy Of English Language Learners

Marcella A. Farina PhD
University of Central Florida, marcella.farina@ucf.edu

Part of the Education Commons

Find similar works at: https://stars.library.ucf.edu/etd

University of Central Florida Libraries http://library.ucf.edu

University of Central Florida

STARS
Showcase of Text, Archives, Research & Scholarship

THE EFFECT OF INPUT MODALITY ON PRONUNCIATION ACCURACY
IN ENGLISH LANGUAGE LEARNERS

by

MARCELLA A. FARINA
B.A. Rollins College, 1987
M.A. University of Central Florida, 1999

A dissertation submitted in partial fulfillment of the requirements
for the degree of Doctor of Philosophy
in the College of Education
at the University of Central Florida
Orlando, Florida

Summer Term
2013

Major Professor: Joyce W. Nutta

ABSTRACT

The issues relative to foreign accent continue to puzzle second language researchers, educators, and learners today. Although once thought to be at the root, maturational constraints have fallen short of definitively accounting for the myriad levels and rates of phonological attainment (Bialystok & Miller, 1999, p. 128). This study, a *Posttest-only Control Group Design*, examined how the pronunciation accuracy of adult, English language learners, as demonstrated by utterance length, was related to two input stimuli: auditory-only input and auditory-orthographic input. Utterance length and input modality were further examined with the added variables of native language, specifically Arabic and Spanish, and second language proficiency as defined by unofficial TOEFL Listening Comprehension and Reading Comprehension section scores.

Results from independent *t* tests indicated a statistically significant difference in utterance length based on input modality ($t(192) = -3.285.$ $p = .001$), while with the added variable of native language, factorial ANOVA results indicated no statistically significance difference for the population studied. In addition, multiple linear regression analyses examined input modality and second language proficiency as predictors of utterance length accuracy and revealed a statistically significant relationship ($R^2 = .108$, adjusted $R^2 = .089$, $F(3, 144) = 5.805$, $p = .001$), with 11% of the utterance length variance accounted for by these two factors predictors. Lastly, hierarchical regressions applied to two blocks of factors revealed statistical significance: (a) input modality/native language ($R^2 = .069$, adjusted $R^2 = .048$, $F(2, 87) = 3.230$, $p = .044$) and ListenComp ($R^2 = .101$, adjusted $R^2 = .070$, $F(3, 86) = 3.232$, $p = .026$), with ListenComp

increasing the predictive power by 3%; (b) input modality/native language (R^2 = .069, adjusted

R^2 = .048, F(2, 87) = 3.230, p = .044) and ReadComp (R^2 = .112, adjusted R^2 = .081, F(1, 86) =

3.629, p = .016), with ReadComp increasing the predictive power by 4%; and (c) input

modality/native language (R^2 = .069, adjusted R^2 = .048, F(2, 87) = 3.230, p = .044) and

ListenComp/ReadComp (R^2 = .114, adjusted R^2 = .072, F(2, 85) = 2.129, p = .035), with

ListenComp/ReadComp increasing the predictive power by 4%.

  The implications of this research are that by considering issues relative to input modality

and second language proficiency levels especially when teaching new vocabulary to adult second

language learners, the potential for improved pronunciation accuracy is maximized. Furthermore,

the heightened attention to the role of input modality as a cognitive factor on phonological output

in second language teaching and learning may redirect the manner in which target language

phonology is approached.

To honor my mother, Anna Costantino Farina

May the culmination of this endeavor stand for the both of us.

ACKNOWLEDGMENTS

manner so as to significantly strengthen research validity of not only my dissertation research but also, and more importantly, the feedback ultimately benefiting the students of the Center. These individuals have blazed this trail with me, all the while demonstrating sincere enthusiasm and open vision for the "shared creation". Because of the efforts of all of us, the Center currently possesses its first set of audio data of non-native-speaker English and has forged new methodological territory in terms of data collection for pronunciation research. Kudos to all!

I also thank all my dissertation committee members for having the perfect balance between innovation and tradition. They have offered me stellar guidance throughout this research process, held me accountable to rigorous standards, strengthened my perceptions through their pointed inquiries, and allowed me to flourish in my own right as a researcher. I am appreciative to Dr. Florin Mihai, who has grounded in me the importance of research dissemination through presentations and publications. In addition, I thank the inter-disciplinary members, Dr. Barbara Ehren for her sincere enthusiasm and reflective contributions to methodology, and Dr. Jack Ryalls for opening my eyes to the constructs of speech perception and speech production. Furthermore, I acknowledge Dr. Lihua Xu for all the time and patience she extended me during the statistical analysis phase. By answering my endless questions, she has strengthened the power of my results while also often filling critical gaps in my own learning. Finally, I whole-heartedly thank Dr. Joyce Nutta, my Committee Chair, PhD adviser, and friend, for her vision of establishing a PhD in Education/TESOL at the University of Central Florida, an undertaking which has been instrumental in not only adding language research to the University's accolades but also providing a canvas not only for me but for many who aspire to place a dot in the pointillism of SLA.

acknowledge all those who have allowed me to peek inside, gingerly tip-toe about, and at times

even boldly prod. Yes, above all I acknowledge the learners.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

LIST OF ABBREVIATIONS & ACRONYMS

EFL	English as a Foreign Language
ESL	English as a Second Language
IEP	Intensive English Program
L1	First, or Native, Language
L2	Second Language
SLA	Second Language Acquisition
TESOL	Teaching of English to Speakers of Other Languages
TOEFL	Test of English as a Foreign Language

CHAPTER 1:

INTRODUCTION TO THE STUDY

Scientific inquiry often challenges long-standing theoretical constructs. What once may have appeared logical and true may eventually buckle under the pressure of the perpetual desire of humanity toward advancement and discovery. As knowledge grows and science homes in on a theory, it becomes evident from a deeper, more precise understanding that the construct can no longer sustain in its present state. Often traditionalists resist this newly-found knowledge, preferring instead to cling to past notions. However, most often science and researchers prevail and expose so much more of a phenomenon that it is impossible to suppress the eminent change. As with Christopher Columbus' voyage, this heightened understanding inevitably leads others to acknowledge that past theories were incomplete and inaccurate and that novel ideas warrant consideration and further investigation. Similarly, attributed to the increase of near-native second language learners throughout the world and an augmented knowledge base with regard to pronunciation constructs over the last 50 years, one such notion in need of re-booting, in need of a fresh perspective with updated expectations, is that of ultimate phonological attainment for adult second language learners.

The phenomenon of foreign accent is comprised of poorly understood notions of cognition and perception which directly impact the degree of success of phonological attainment in adult second language learners (Lippi-Green, 1997). In fact, in his preface to *Speech Acoustics and Perception*, Boothroyd states when referring to speech in a general sense, "So spontaneous is the acquisition and use of spoken language skills that most people are unaware of the special

nature of this activity, or of its complexity" (1986, p. v). The known constructs of phonological acquisition identify a finite inventory of sounds, of which only a small percentage is combined to create the phonology associated with any given language. In addition, the notion of accent rarely hovers over first language acquisition as it does second language even though it is not uncommon for a native speaker, even a child, to possess a particular regional version of the native language phonology.

When it comes to second language phonology, a widely accepted theory is that children exposed to more than one language during the first language acquisition phase can, in fact, acquire more than one language, but from adolescence on the ease of acquisition decreases (Lippi-Green, 1997). "When a child learns a second language right along with the first, pronunciation for both languages proceeds authentically at the same pace. But once the pronunciation pattern of the first language is established in normal monolinguals, it becomes easy to use these patterns as first approximations to new sounds when learning a second language" (Walsh & Diller, 1979, p. 517). The "Joseph Conrad Phenomenon" is a term coined by Scovel (1981) after the Polish-born writer, who so eloquently learned his third language English, preceded by French his second, that he was able to achieve high acclaim authoring books in English but was never able to refine the accuracy of his pronunciation of English to a comparable level of attainment (Scovel, 1981, 1988). This phenomenon can be explained perhaps by the idea that second language learners rely on pre-existing perceptions of phonetic categories from the native language, and it is this first language phonological anchoring that can inhibit the learner from perceiving, and ultimately producing, target language phonology (Flege, 1991).

Even though age and length of exposure has been viewed as being at the root of the problem and have driven much of the research over the last fifty years, the enigma of foreign accent persists and remains vastly unaccounted for, offering little consolation to those faced with the obstacle. In addition to age and exposure, factors such as motivation and attitude have been considered contributors to the pervasiveness of foreign accent in second language learners as has the type of language instruction received (Celce-Murcia, Brinton, & Goodwin, 1996), with the latter and language literacy being of particular significance. Regardless of the rigor of research in these many areas of language learning, the existing linguistic theories underlying the second language acquisition process somehow still do not adequately account for the phenomenon of foreign accent (Lippi-Green, 1997).

Part of the problem of understanding this unknown area is distinguishing and defining the appropriate jargon to discuss its characteristics and functions. Pennington (1996) claimed that determining an accurate definition depicting the notion of *accent* is nearly impossible. However, Derwing and Munro (2009) operationalized three key terms in the discussion on pronunciation from the standpoint of listener perception: *accentedness*, a speech sound pattern different from the local version (p. 478); *comprehensibility*, the degree of how easy or difficult it is to understand a speech sound pattern (p. 478); and *intelligibility*, the degree of actual comprehension derived from a speech sound pattern (p. 479). They summarize these distinctions with an analogy: accent is to difference, what comprehensibility is to listener effort, and what intelligibility is to understandability (Derwing & Munro, 2009, p. 480).

Unlike for first language acquisition, second language learning, especially for an adult engaged in an instructional setting, is typically presented in both phonologic and orthographic

perspectives. Therefore, it is appropriate to compare similar and, perhaps more importantly, dissimilar features of both stimuli. Several pertinent and noteworthy traits distinguish literacy from phonology, elements that are at the heart second language education.

The first of these is that sound is unique in that it encompasses inherent dissipation, whereas orthography is marked by visual longevity. An example of this is the preferred use of closed captioning for television viewing by many second language learners. Another oppositional feature is that sound reaches the interior and exterior senses of the human being, while sight occurs in separation of the being, taking on an external, isolated sense. As quoted by Ong (2002, p. 71), "Sight isolates, sound incorporates." Furthermore, a literate person usually conjures up a spontaneous, typographic image of a word and, therefore, cannot fully recover the sense of a word in its solely auditory identity. Therefore, the phenomenological nature of sound is a highly somatic component and establishes core sensations and existence completely unique from that of text (Ong, 2002).

Another aspect holds that although orality precedes literacy in the first language acquisition process, with second language acquisition it is usually the opposite. Second language learners often develop the graphemic sense of the target language first, before fully acquiring a phonological understanding. It is conjectured that this reversed sequence of input modalities, specifically in opposition to that of first language acquisition, may contribute to the manner in which a second language is initially processed both acoustically and visually.

Although reading is an integration of phonology and orthography, it is also important to distinguish features of orality and literacy in that literacy employs different degrees of top-down and bottom-up processing, depending on the native language. For example, the native Arabic

speaker learning English will use bottom up processing because there is little overlapping knowledge from which to anchor the reading experience presented by English, while the native Spanish speaker learning English will employ predominantly top-down processing due to the great similarity of English, not only semantically but also orthographically (Pike & Edgar, 2005). These differences in processing stemming from native language reading weigh significantly on the discussion of the orality-literacy interaction for second language learning.

Consequently, in order to address the most basic elements of phonological acquisition of a second language, it is imperative to explore the primary cognitive processes of language, focusing on the initial stimuli of sound and sight and how they interface to activate short-term and long-term memory mechanisms for eventual retrieval and ultimate language attainment. The purest form of input-to-output auditory connection occurs when the learner has no previous knowledge or preconceived notion of the word and encounters the element for the very first time. Hence, pseudo and neo-language used to measure of how input modality affects pronunciation can best replicate the initial experience for second or foreign language learners. By better understanding the factors influencing the input processing of auditory and orthographic inputs for pronunciation, second language educators can optimize on the sequence the exposure of new vocabulary and language in such a manner so that learners can reach acquisition at a semantic level while concurrently increasing the possibility of peak phonological attainment.

During the last two decades of the twentieth century, the concomitant retreat from pronunciation research and pedagogy coincided with the decline in audiolingualism (Derwing & Munro, 2009). Minimal work was conducted by researchers and practitioners with regard to second language pronunciation (Brown, 1988) as compared to areas of semantics, for instance.

5

More recently though, studies have indicated that pronunciation teaching has an obvious and positive effect on intelligibility and comprehensibility (Derwing, Munro, & Wiebe, 1998), finally arriving at the conclusion that "pronunciation is learnable" (Derwing & Munro, 2005). The current general consensus with regard to ultimate phonological attainment is that near-native pronunciation is achieved only with significant effort by the learner and the ultimate goal of sounding as native like as possible rather than just being intelligible (van Els & de Bot, 1987). Long (2005) posits the dilemma of a critical period for phonological attainment by late-age learners as the dichotomous question of whether learners "do not" attain native like pronunciation due to an array of causes or whether they simply "cannot". This, along with more sophistication in the technology of brain activation imaging, has sparked an interest in returning to the notions of a critical period, especially in phonology, to better understand its underpinnings.

In his preface to *Biological Foundations of Language*, Lenneberg (1967) juxtaposed the study of language as a phenomenon pertaining to numerous fields from both the natural and social sciences, with interlocking elements of both nature and nurture. "When SLA research is carried out with a cognitive perspective, the L2 is viewed as a skill, and its acquisition as a linguistic system is assumed to be built up gradually through processes of attention, conscious awareness, and practice" (Pica, 2005, p. 273). Therefore, a thorough knowledge of the related theories from various disciplines is essential in order to attempt to derive associations and dependencies influencing the degree of accentedness in nonnative language speakers. Consequently, the present study represented a merging of the empirical research of second language acquisition and cognitive psychology, establishing a multi-disciplinary theoretical framework for the brain processes relative to both auditory and visual stimuli and their

contribution toward the degree of accentedness in the phonological production of second language learners.

In addition, this study looked to the area of speech science for its fundamental theories of speech perception and speech production as well as methods for measuring degree of accentedness. The notion of foreign accent is as much about speech production as it is about speech perceptive in that the existence of foreign accent is reliant on the fact that the native speaker can recognize it (Scovel, 1969, p. 248). Work by Flege (1984) demonstrated that phonetically untrained listeners were able to recognize French-accented English from native English with very short speech samples, as little as 30 milliseconds of speech. Likewise, identification of native versus nonnative speech was recognizable even when played backwards regardless of the length of the sample or the foreign accent heard (Munro, Derwing, & Burgess, 2003). To further reinforce the pertinence of speech perception to pronunciation, Major (2007) found that listeners were able to distinguish native for nonnative speech even in a language that they did not speak. Hence, an interdisciplinary approach overlapping the relevant yet unique constructs from these fields provides a comprehensive lens with which to address the origin and malleability of foreign accent and further enlighten issues relevant to this phenomenon.

The present study examined pronunciation accuracy, here being measured by output length, and how it is fostered by input stimuli by comparing an auditory-only stimulus to a combined, auditory-orthographic stimulus. The findings were further examined for how they manifest in learners of differing native languages and second language proficiency levels. The latter is of particular relevance due to previous research pointing to language proficiency as an indicator of sound perception (van Els & de Bot, 1987) and also because the relationship

between stimulus length and memory capacity and the relationship between stimulus length and learner proficiency account for the correlation between imitation accuracy and second language proficiency (Bley-Vroman & Chaudron, 1994). Therefore, second language proficiency was also examined as a measure for predicting pronunciation accuracy in terms of utterance length.

The implications of this research have pedagogical bearing on best practices for teaching new vocabulary to adult second language learners so as to maximize the potential for rate and level of phonological attainment. In addition, this research posits a different approach to examining the influences guiding novel language production by investigating the role of input modality as a cognitive factor affecting the initial stages of phonological processing in a second language. Lastly, with this study, it is hoped that language educators will weigh factors of native language and second language proficiency as they address issues of pronunciation in their classrooms and also open their own views about phonological attainment so as to aim for a higher ultimate standard, not only in their own teaching of pronunciation but also for the outcome of their students' pronunciation.

The climate of ultimate phonological attainment has changed. This is due, in part, to the persistent inability to definitively determine why so many fail yet others succeed at mastering the phonology of a language in addition to that of their native language. This shift away from the acceptance of status quo is also due to the growing and changing knowledge base around neuroplasticity and the "plastic paradox" (Doidge, 2007, p. 298). These new findings are making their mark in numerous contexts, from the soldier that returns from war without limbs or with post-traumatic stress disorder and needs to develop new neurological pathways in order to regain faculties, to the child with a learning disability struggling in school who needs to develop

atypical neurological "routes" in order to be a successful learner, to the general population of baby-boomers eager to fend off age-related cognitive decline. These populations are learning to "blaze new neurological trails" in order to live more successfully and content. It stands to reason that in this new understanding of neuroplasticity, there may also be some of the answers for which foreign accent researchers have been looking.

All these developments have brought about glimmers of hope and much curiosity on the phenomenon of foreign accent as well as a common belief that the pronunciation successes that so many enjoy could be, in fact, accessible to all language learners. As with the movement that gradually rescinded the long-standing belief that the world was flat and fueled Columbus for his inaugural voyage, many have come to conclude that what was perceived about phonological attainment may actually have been inaccurate all along. Numerous possibilities as to the degree of ultimate attainment in phonology are on the horizon, a phonological frontier awaiting exploration and discovery in the hope of improving the language learning experience and outcome.

Research Questions

1. Is there a statistically significant difference in pronunciation accuracy as measured by the utterance length of pseudo words produced from a single modality of auditory-only input as compared to those produced from a dual modality of auditory-orthographic input?

2. Is there a statistically significant difference between native language and pronunciation accuracy as measured by the utterance length of pseudo words

produced from a single modality of auditory-only input as compared to those produced from a dual modality of auditory-orthographic input?

3.   Is there a statistically significant relationship between second language proficiency level and pronunciation accuracy as measured by the utterance length of pseudo words produced from a single modality of auditory-only input as compared to those produced from a dual modality of auditory-orthographic input?

4.   Is there a statistically significant relationship between native language and second language proficiency level in terms of pronunciation accuracy as measured by the utterance length of pseudo words produced from a single modality of auditory-only input as compared to those produced from a dual modality of auditory-orthographic input?

Research Hypotheses

1.   In response to Research Question 1, it is hypothesized that there is a statistically significant difference ($p < .05$) in pronunciation accuracy as measured by the utterance length of pseudo words produced from a single modality of auditory-only input as compared to those produced from a dual modality of auditory-orthographic input.

2.   With regard to Research Question 2, it is hypothesized that there is a statistically significant difference ($p < .05$) between native language and pronunciation accuracy as measured by the utterance length of pseudo words produced from a single modality of auditory-only input as compared to those produced from a dual modality of auditory-orthographic input.

3.   It is further hypothesized that there is a statistically significant relationship ($p <$ .05) between second language proficiency level and pronunciation accuracy as measured by the utterance length of pseudo words produced from a single modality of auditory-only input as compared to those produced from a dual modality of auditory-orthographic input.

4.   Lastly, it is hypothesized that there is a statistically significant relationship ($p <$ .05) between both native language and second language proficiency level in terms of pronunciation accuracy as measured by the utterance length of pseudo words produced from a single modality of auditory-only input as compared to those produced from a dual modality of auditory-orthographic input.

Definition of Terms

In understanding the research questions and hypotheses proposed, it is important to first operationalize the terms pronunciation accuracy, utterance length, and pseudo words. The first of these and the focus of this study, *pronunciation accuracy*, was defined as the degree to which the participant utterance matches in length as compared to the model utterance as heard via the audio recording. Pronunciation accuracy was measured through the oscillographic and spectrographic depiction of the digitized audio recordings as displayed by Audacity 2.0.2 software comprising the audio data used in this study.

Another significant term, *utterance length*, or the duration of each utterance, was represented by word boundaries measured in millisecond numerical values, specifically demarcated from the onset of each utterance through the release of the offset of each utterance. These demarcations were identifiable through oscillographic and spectrographic representations

as displayed by Audacity 2.0.2 software and manually documented by the researcher and research assistant for inter-reliability purposes. Utterance length was ascertained by subtracting the numerical value in milliseconds of the utterance onset from the numerical value in milliseconds of the release of the utterance offset, an automated feature included in the free version of the Audacity 2.0.2 software.

The term *pseudo word* used in this study refers to a unit of speech and text which adheres to the phonotactic constraints of Standard American English. A selection of 100 pseudo words was randomly-generated using the ARC Nonword Database (Rastle, Harrington, & Coltheart, 2002) to present mono-morphemic units, or logatomes, units of language which cannot be distributed into smaller parts. In addition, the pseudo words were structured with legal bigrams, adjacent, two-letter spelling patterns, and orthographically existing onsets and bodies, all according to the permissible phoneme spelling patterns of American English. Simply put, all the pseudo words chosen adhered to the phonotactic rules of American English.

The list of 100 pseudo words was further filtered to render only those items with stop onsets and offsets, which were then used to create the pronunciation diagnostic instrument from which the audio data set was ultimately derived (see Appendix B). Hence, the pronunciation diagnostic instrument consisted of only those pseudo words with */p/b/t/d/k/g/* starting and ending sounds, rather than other more sonorous onset and offset sounds, such as continuant or vowel sounds (see Appendix D). The phonetic aspect of a stop, or plosive, onset and offset facilitated the identifying of the beginning and end of each pseudo word during the data analysis phase.

Furthermore, the pseudo words chosen for the pronunciation diagnostic instrument were audio recorded by a native speaker of American English, according to the accompanying

phonetic transcription provided by the ARC Nonword Database and following the pronunciation patterns typical of Standard American English (see Appendix C). Additionally, all stop offsets in the pseudo words were performed and audio recorded with subdued release bursts, rather than unreleased final stops as commonly spoken by American English speakers, so as to signify a clear ending to the utterance while still typifying American English pronunciation.

Also key to this study is the term *input modality*, the receptive language of auditory and orthographic input, that which is heard and read respectively. The grouping variables for the present study were distinguishable by these two input modalities. The audio data set on which this study was based consisted of audio recordings of utterances modeled after either an auditory-only input or an auditory-orthographic input of the pseudo words contained within the pronunciation diagnostic instruments. In both instruments, the same audio recording of the pseudo words was used, and the only difference between the two instruments was that the auditory-orthographic instrument had the added feature of the orthographic representation of the pseudo words. In assigning the study groups, one group of participants was assigned the instrument with the auditory-only input, whereas the other was assigned the instrument with the auditory-orthographic input. In other words, both groups experienced the same audio recording of the pseudo words; however, only the experimental group was exposed to the accompanying stimulus of the on-screen textual depiction of the pseudo words.

Other terms characteristic of second language acquisition research and fundamental to this study are the independent variables of *native language* and *second language proficiency*. Native language is defined as the mother tongue of the participant, information which was self-reported by all incoming students to the Intensive English Program as part of the application and

registration process. This information was provided by IEP administrative personnel to the researcher in the form of a demographic data set. Of the many native languages represented in the data set, the two native languages which emerged abundant enough to enable for specific analyses were Arabic and Spanish. It is important to note that one commonality between American English and these two native languages is that they both employ a phonemic/phonetic writing system, whereas unique to each is the dissimilarity of their respective alphabetic principles: the consonantal-alphabet of Arabic and the alphabetic-transparency of Spanish versus the alphabetic-opaqueness of English (Hedgcock & Ferris, 2009, p. 11). These distinctions in script along with the added consideration of directionality characteristic of the Arabic language enriched the discussion of the implications of study findings and the impact of orthography on the pronunciation learning of these two second language learner populations.

Lastly, the notion of second language proficiency level, although definable through various means, was determined for this study via the scores of unofficial administrations of the Paper-based TOEFL as administered as part of normal placement procedures of the Intensive English Program. These placement tests consisted of retired Paper-based TOEFL test forms, marketed by ETS and previously purchased for the purpose of level placement in the Intensive English Program. Moreover, placement tests were administered and scored according to standard Educational Testing Services protocol for administering and scoring Paper-based TOEFL. The scores obtained consisted of the standard composite TOEFL score, ranging from 220 to 680, as well as separate scores for the three sections of the test: Listening Comprehension, Structure and Written Expression, and Reading Comprehension. Scores from each of these three test sections were represented on a continuum, ranging from 22 to 68 (see Appendix E).

For the data analysis of this study, two of the three TOEFL section scores were examined, the TOEFL Listening Comprehension and the TOEFL Reading Comprehension, due to their analogous relationship to the two input modalities being studied, auditory and orthographic respectively. The TOEFL scores included in the data set were identifiable by student identification numbers assigned through the IEP database and void of examinee names or other identifying information. In conclusion, the operationalization of these key research terms concretized the parameters of the research questions and hypotheses and served to specify the notions being researched by the present study.

Potential Limitations of the Study

As with all scientific inquiry, pronunciation research also includes limitations and delimitations, in particular because of the sociolinguistic and psycholinguistic implications. The mere presence of non-nativeness in English second language learner speech patterns stretches much beyond notions of cognitive psychology to include other areas, such as identity, culture, and behavior. Moreover, adult English language learners differ from young English language learners in that they are more extensively and profoundly grounded in their native language and culture, and perhaps learning style. Therefore, some delicate considerations emerge when discussing foreign accent with respect to adult second language learners.

Some schools of thought in second language acquisition perceive the notion of pronunciation as an expression of identity and view modifying one's speech analogous to modifying one's self-image and, therefore, unethical (Porter & Garvin, 1989, p. 8). However, others do not view pronunciation instruction and identity preservation as incompatible. In fact, Timmis (2002) surveyed over 600 students and teachers regarding their views on pronunciation

and identity and found that 67% of students truly aspired to have native like pronunciation. In another study (Derwing, 2003) of 100 adult English as a second language learners in Canada, not only did 97% express views that learning to pronounce English well was important and 95% as desirous of attaining native-speaker pronunciation ability, but there was also overwhelming disagreement with the notion that changing pronunciation would threaten identity; in fact, their wish was to be as fluent as possible in both languages. Jenkins warns, "It is important to emphasize that we should all guard against political correctness, in the sense of telling our learners what their goals should be: in particular that they should not want to sound like native speakers if they clearly wish to do so" (Jenkins, 1998, p. 123).

In an investigation of advanced proficiency second language learners, Piller (2002) found that many claimed to "pass as native speakers" and perceived this as a marker of success, leading her to conclude that a native accent is not what someone is, but rather what someone does. This study addresses a unilateral approach to understanding the phenomenon of foreign accent and does not encompass more individualized factors, such as sense of self and interpersonal interaction, which could actually supersede in importance and relevance to those of purely auditory and orthographic stimuli. In other words, it is intended for the potential understanding gained here of how input modality may impact pronunciation accuracy to be coupled with other relevant and purposeful constructs in order to more fully contribute to the overarching discussion of phonological attainment in adult language learners.

Another concern with regard to delimitations relates to aspects of the study methodology and the notion of generalizability. First of all, the administration of a pronunciation diagnostic assessment was administered to existing language proficiency groups, established Speaking and

Communication Skills courses in the Intensive English Program. The benefit, of course, was to be able to attend to each group of language learners in a language-appropriate manner. However, in stratifying the random assignment of the students within each course section to one of the two comparison groups, the two versions of the instrument were consequently administered simultaneously, meaning that students may have become aware that others within their group had a slightly different instrument, the one with the added orthographic feature. This could have affected internal reliability in terms of *Differential Selection*. In an attempt to minimize this concern for delimitation, students were positioned so that the auditory-only group was assigned to one side of the data collection facility while the auditory-orthographic group to the other. Although this may have not completely removed the threat to internal validity, it is believed that this effort did in fact minimize exposure to the dichotomous nature of the instrument and task.

Furthermore, to be able to generalize the results relating to native language obtained from the present study to other adult English language learners in Intensive English Programs in the U.S., it was important that the sample size per native language was both abundant yet diverse enough to replicate that which is found in other Intensive English Program settings. In an attempt to effectively manage this variable, the researcher considered the most current demographic trends of student enrollment in the Intensive English Program, hence selecting the native languages of Arabic and Spanish as values for the independent variable. It is hoped that researchers in other geographic regions of the U.S. and with access to IEP populations with differing demographic tendencies will attempt to address the issue of pronunciation accuracy as derived from input modality with native languages other than those represented in this study.

Potential Contribution of the Study

The contribution of knowing how auditory and orthographic inputs interact and impact phonological attainment in second language learners is relevant to both teachers and students from numerous standpoints. First of all, a better understanding of this process can impact language teaching with regard to the methods and sequence of exposure implemented by educators so as to optimize initial presentations of new language in a manner that assists the student in attaining a more accurate perception and production of the pronunciation of the target language while also acquiring other language features, such as semantics and syntax. Furthermore, it is hoped that with gained knowledge regarding the optimal sequence and balance of language exposure, learners will embark on the journey of acquiring another language with the potential for successful acquisition in not only lexical and syntactic aspects but also phonological.

Secondly, it is also hoped that this study will contribute to teacher preparedness in pronunciation instruction. Although unfortunate, results from survey research gathered in Canada (Breitkreutz, Derwing, & Rossiter, 2002), Great Britain (Burgess & Spencer, 2000), and Australia (MacDonald, 2002) reported that many second language instructors are actually fearful of teaching pronunciation, partially because they feel unprepared in this area of language teaching. In addition, Murphy (1991) further reported that only about half of the Master of Arts in TESOL programs includes a course on phonology in the program of study. It is believed that with a greater understanding of the phenomenon of pronunciation and teaching strategies to facilitate student improvement and success in this area, language teachers will find confidence

and embrace the notion that although not thoroughly defined and understood, pronunciation instruction has a vital role in the language curriculum and every language lesson.

Thirdly, with this knowledge it is believed that the learner can be empowered with a greater awareness of how the brain processes auditory and text input and how the interaction of the two may affect pronunciation acquisition. A longitudinal study in Canada (Munro & Derwing, 2008) examined the progress of production of 10 vowels in English as performed by 20 Mandarin and 24 Slavic (Russian, Ukrainian, and Croatian) adult English language learners. These individuals were exposed to 20 delayed-repetition tasks of CVC-patterned words, focusing on "vowel target intelligibility". Results of phonetically expert judges indicated that production improved significantly ($p < .05$) during the first year of exposure, in particular during the first six months. In addition, even with low spoken-language proficiency ability and no specific pronunciation instruction, the Mandarin speakers improved in intelligibility from 8% to 37% and the Slavic language speakers from 5% to 27% (Munro & Derwing, 2008). Although these improvement trends did not continue proportionately throughout the study period, the naturalistic inclination toward pronunciation awareness and ability characteristic of students during the early-exposure stage is indicative of the potential for successful pronunciation habits, which is all too often overlooked in most language instruction today.

Lastly, the development of speech perception and production is exemplar-based in a way so that instances of exposure of various forms are stored in memory where they are the basis for long-term representations (Nygaard & Pisoni, 1998). In addition, cognitive representations of language are frequency-tuned on the basis of language input (Ellis, 2006a). Thus, it is believed that an increased awareness of the bearing which input modality can have on output will direct

the learner toward a heightened sensitivity for the nuances of target language phonology and more effective strategies for acquiring new phonology accurately at the initial stage.

The present study proposed a variation on the methodology relating to how pronunciation is not only measured and ascertained but, perhaps more importantly, stimulated and gathered. It is believed that by examining the effect of auditory-only input as compared to the dual modality of auditory-orthographic input and how the two modalities interact with factors of native language and second language proficiency, language educators and learners can more effectively sequence exposure of new language during instruction and learning. It is also the hope of the researcher that more effective exposure will optimize what is retained in long term memory with regard to the phonological structure of the target language so as to, in turn, improve pronunciation and overall communication effectiveness. All in all, with a greater understanding of the phenomenon of foreign accent, second language teaching and learning can induce ameliorated rates and levels of phonological attainment, making the second language experience a more gratifying and successful one for all.

CHAPTER 2:

LITERATURE REVIEW

Ultimate attainment in second language phonology is a phenomenon that has puzzled many researchers from varying areas of expertise. Empirical studies have looked at phonological acquisition through the lenses of psycholinguistics, sociolinguistics, and neurolinguistics, as well as other areas of inquiry, in search of the path or paths that lead to its origin. Since the initial postulations about how it is that some learners do attain native like ability in second language phonology while many others do not, numerous uncertainties have burdened research and researchers. Stemming from the second half of the twentieth century, inquiry has looked to impaired individuals, starting with the early studies by Lenneberg (1975) of aphasiac patients and those with congenital deafness to studies of normally-functioning populations as with the works of Baddeley and Hitch (1974), which address the processes of working memory. The timeline of research places the 1950s as having cited native language interference as the ultimate link, and the thirty years that followed were characterized by a shift toward notions of biological constraints and lack of brain plasticity.

However throughout this entire period, the notion of input, its quality and quantity, has actually been underestimated and received little significance regardless of the fact that with first language acquisition it is considered essential (Flege, 2012). In short, it is apparent that over the past half-century, no one field of inquiry has adequately surmised the origin of this enigma nor understood the dichotomous extremes of levels of attainment, leaving second language

acquisition research unresolved in terms of the phenomena of foreign accent and the factors that lead to its existence, manipulation, and limitations (Ioup, 2005).

Cognitive psychology, the branch of science devoted to the study of the mind and encompassing research of perception, attention, and memory, contributes much to what is known about the language acquisition process. Inquiry conducted in cognitive psychology attempts to isolate one process, such as audition or imagery, from another (Braisby & Gellatly, 2005, p. 7) and can, therefore, contribute a framework on which to create research in target language phonology of, at least from the cognitive perspective, how the brain receives, processes, and produces auditory language.

Memory Systems

The idea of how the brain acquires new phonological patterns can best be understood by considering studies which examine the presence of differing memory systems for verbal stimuli. Baddeley (1966) demonstrated that, by brief exposure to word lists, short-term memory performed better at recalling words that were semantically similar as compared to words that were phonologically similar. Conversely with rehearsal and a longer retention interval, semantic similarity had a negative effect on recall while phonology similarity had no effect. In addition to what is now known about the negative effects of teaching vocabulary in semantic sets (Tinkham, 1993, 1997; Waring, 1997), these results primarily suggest two separate storage systems with different encoding, short-term memory with acoustic properties, and long-term memory with meaning-based encoding.

To add to the topic of short-term and long-term memory, various models have depicted the limited capacity of short-term memory, the relevance of sub-vocal rehearsal for short-term

memory, and the gradual transfer of short-term memory to long-term memory (Murdock, 1967; Atkinson & Shiffrin, 1971). A significant finding to add to the distinctions of verbal capacity between short-term memory and long-term memory is the fast rate of forgetting characteristic of short-term memory as compared to long-term memory (Brown, 1958). This factor, coupled with the fact that echoic signals do not go directly to long-term memory as do iconic signals (Naish, 2005), intensifies the plight of second and foreign language learners, struggling to acquiring vocabulary, syntax, and pronunciation. The approach of varying study populations has also contributed to the knowledge base of the role of short-term and long-term memory in language learning. For example, systematic evaluation of individuals with deficits has also provided valuable input with regard to language and memory (Lenneberg, 1967). In fact, research depicting a brain injury patient sustaining severe deficits in short-term memory but whose long-term memory remained intact re-affirms the notion of a separate neuroanatomical localization of short-term memory (Shallice & Warrington, 1970).

By the mid-1970s, generally accepted theory was grounded in the idea that memory is comprised of these two distinct faculties and that these two memory systems functioned differently, in particular for language. Specific to audition, the temporal features of acoustic properties in short-term memory and the manner in which short-term and long-term memory affects phonological retention provide important constructs relative to pronunciation teaching and learning.

Working memory. Further research was directed at obtaining a more precise understanding of short-term memory specifically, which evolved into what is now called working memory. An empirical study by Baddeley and Hitch (1974) utilized the dual-task

paradigm to explore the effects of simultaneous performance of irrelevant brief memory tasks and cognitive tasks of verbal reasoning. From their observations, it was determined that working memory is more precisely comprised of both information storage and ongoing mental operations. In other words, short-term memory is actually a subsystem of working memory that serves the purpose of storing information for brief durations.

Moreover, working memory includes an additional resource, which is not shared with short-term memory, the ability for mental operations. Baddeley and Hitch (1974) concluded that working memory can be partitioned into two components: one to hold temporary information and the other for cognitive processing. Second language learners toggle between these two functions, storing information and processing information, in order to eventually transfer the target language to long-term memory; this is known as the process of language retention.

In order to understand the role of visual memory and imagery in short-term memory and their interaction with verbal performance, other studies by Baddeley and Hitch (1974) found that production generated from mixed modalities outperformed that of same modalities, thus suggesting two circuits, or loops, for the storing of small amounts of temporary information. The tripartite model by Baddeley (1983) culminated in the understanding that two subsystems hold information, the articulatory loop and the visuo-spatial sketchpad, and feed in and out of a single central executive system, which processes mental operations. Furthermore, information obtained from cases involving individuals with atypical neurological conditions demonstrated that auditory and visuo-spatial abilities can operate independently depending on lesion locations (de Renzi & Nichelli, 1975). This indicated strong evidence for a separate, non-verbal store as well

as the presence of a visual store that work independently and inter-dependently to aid in remembering verbal items.

Articulatory loop. Based on Baddeley's model of working memory function and further research (Baddeley & Hitch, 1974), the storage sector specifically for verbal input was identified as consisting of a passive buffer for storing verbal information according to a phonological code and of an active mode of articulatory rehearsal based on internal language (Baddeley, Lewis, & Vallar, 1984; Baddeley, 1986; Cowan, 1988; Martin, 1993; Trojano & Grossi, 1995; Baddeley, Gathercole, & Papagno, 1998).

The articulatory loop is comprised of a grapheme-phoneme correspondence system which transfers the visually perceived grapheme into phonological information and transfers it to the phonological, or passive, buffer (Vallar & Papagno, 1995). The functions of the phonological loop, from input through the stores and processor to output, are depicted in Baddeley's more recent depiction (2003) of the working memory system. Analogously put, the closing of the phonological loop is to the ears what visualization is for the eyes.

A closer look at the phonological loop and the effect of similar and dissimilar phonology resembles what was found for semantics. Studies of neurologically-typical individuals suggest the function of a *phonological similarity effect* in the passive buffer whereby phonologically similar words are less accurately remembered than those with phonological dissimilarity (Conrad, 1964; Baddeley & Wilson, 1985). However, further examination of phonemic similarity on reasoning and comprehension in short-term memory revealed only a mild disruption. Therefore, researchers concluded a subsystem of working memory, the articulatory rehearsal loop, was capable of holding small memory loads during cognitive functioning and was

responsible for the phonemic similarity effect, while the second component, the central executive, handled operational tasks (Baddeley & Hitch, 1974). The phonological similarity effect can also be described in acoustic versus articulatory terms. Results obtained by Schweppe, Grice, and Rummer (2011) indicated no statistically significant difference between similar and dissimilar phonology as a result of articulation. However, a statistically significant decline in performance for phonologically similar items when performance is acoustically perceived, implying an impact of acoustic features on short-term memory.

Another feature contributing to the processing of the phonological loop is word length. Studies indicating a *word length effect* (Baddeley, Thomson, & Buchanan, 1975; Baddeley, Lewis & Vallar, 1984) describe a tendency toward better recall for words of shorter length, suggesting that the articulatory process rehearses shorter words more efficiently than longer words (Rastle & Coltheart, 1998). An investigation regarding the nature of word length effect being attributable to retroactive interference in short-term memory was conducted by Campoy (2011), who analyzed the accuracy of immediate serial recall of medium-length words followed by either short or long interfering words. With Campoy's work, recall diminished when the target words were followed by long words. Additional results indicated that long words impaired short-term memory content retention to a greater degree. Therefore, word length effect may stem significantly from interference rather than length alone, both considerations being relative to second and foreign language classrooms.

Lastly, the memory span of verbal information is also impacted by the *word frequency effect*, the influence of word familiarity on word retention based on semantic-lexical type coding (Vicari, Carlesimo, Brizzolara, & Pezzini, 1996; Hulme, Roodenrys, Schweickert, Brown,

Martin, & Stuart, 1997), favoring high-frequency versus low-frequency words based on the support offered through long-term memory features. Miller and Roodenrys (2009) also suggested that the word frequency effect is influenced by the concreteness of items such that the more concrete the content of the word, the less the effect of word length on serial recall in short-term memory and the greater the ease with which the phonological items are retrieved from long-term memory. A study of immediate serial recall of nonwords (Nimmo & Roodenrys, 2002) suggested that long-term memory interfaces with short-term memory in cases where nonwords actually occur in polysyllabic environments in English lexicon and that high-frequency single syllable nonwords were more easily retrieved than low-frequency nonwords. These findings substantiate the significance of using mono-syllabic pseudo words in the pronunciation diagnostic instrument of the present study to elicit speech for data analysis purposes. It also, and perhaps more importantly, speaks to the pedagogical implications of first grounding accurately-produced, second language phonology of shorter words in long-term memory so as to create a point-of-reference from which the learner will more accurately be able to later predict the pronunciation of longer or less-frequently used words.

Other contributions to the understand of the articulatory loop include research conducted by Carlesimo, Galloni, Bonanni, and Sabbadini (2006), who examined the articulatory loop of individuals with developed dysarthria and anarthria due to cerebral lesion as compared to individuals with parallel intelligence levels. Results indicated that patients with these conditions performed with approximately the same accuracy on recognizing similar and dissimilar phonological patterns, different from the phonological similarity effect as presented in normal individuals. Furthermore, Carlesimo et al. (2006) reported no statistically significant difference

between groups for the word-length effect but a noteworthy difference between groups for word-frequency effect. All in all, the researchers sustain that not only the articulatory deficit but also intelligence deficiency contribute to the development of the phonological buffer, and that malfunction in aspects of the articulatory loop and cognitive processes impair performance of verbal short-term memory.

   Visuo-spatial sketchpad. Because the present study is juxtaposing the effects of auditory-only input against those generated by auditory-orthographic input, it is also important to understand the role of the visuo-spatial sketchpad in reading and second language learning. Cole and Pickering (2010) investigated Chinese and English language users and the interaction of the two stores in short-term memory, the phonological loop and the visuo-spatial sketchpad. Of particular interest was the logographic writing system characteristic of Chinese versus the alphabetic orthography of English. Results indicated that the phonological and visual similarity effect did not manifest in the same way in the two language groups tested. The Chinese learners in the Cole and Pickering 2010 research showed no difference in performance based on similarity or dissimilarity, suggesting the presence of a dual-coding strategy for memory processes in the subjects studied. In addition, Palladino and Ferrari (2008) examined 12- and 13-year-old Italian students displaying foreign language learning difficulty, or FLLD, a disability that manifests mostly in phonological/orthographic and syntactic, but not semantic, aspects of foreign language learning. Study outcomes indicated that deficits in phonological processing and memory, in contrast to the lack of deficits in visuo-spatial processing and central executive abilities, contributed to the learner's challenges in learning a second language.

To further understand the extent of the separateness of the auditory and visual components, a case study by Lenneberg (1964) depicts the personal account of a 5-year-old child with congenital anarthria, who was totally inarticulate and illiterate, and who after having received one summer of private tutoring at the age of nine was successful in learning to read, exemplifying articulation as an dispensable element in reading. In conclusion, the framework of working memory and in particular of the phonological loop and visuo-sketchpad has bearing on the discussion of foreign accent in that these constructs are at the core of how phonology is received, processed, and produced.

Attention

The role of attention is critical to the language acquisition process because clearly the brain has to attend to audition in order to produce output and potentially arrive at meaning. The auditory system has the capacity to filter a single signal when many are present. This sort of attenuation was first examined by Broadbent (1952, 1954), who produced early studies that utilized dichotic listening tasks, hearing different sounds in each side of a headset, to understand how the brain attends to, or doesn't attend to, diverse simultaneous auditory stimuli. Along with studies by Treisman (1960), it was determined that the brain attends to only one auditory signal at a time, serial processing, because subjects in these studies were unable to recall only from the attended signal and nothing about the unattended signal. However, in studies whereby the signal was shortened to just a few seconds, the echoic memory in subjects allowed them to attend to both signals, parallel processing (Naish, 2005). This finding led toward further exploration of the attenuation process, whereby the brain filters out unattended sounds but is alerted when a permanent sensitive signal occurs, such as hearing one's name spoken amidst a noisy room

(Treisman, 1960). Most relevant to a language learning classroom is the sliding scale feature of two auditory attention constructs: the longer the signal, the less parallel the processing.

Much like hearing extends over time and possesses disentangling properties for auditory attention, visual attention extends over space and also allows for visual filtering. Sperling's (1960) studies of visual memory suggested that subjects stored entire sets of letters for just a short time in iconic memory, much the same way echoic memory works with auditory attenuation. A very common example of how visual stimuli compete for processing attention is the Stroop task, first published in English by Stroop (1935) to demonstrate where incongruent and congruent conditions effect processing. Color-coded color words are presented in coinciding and differing color-schemes, making it difficult for the brain to process. The incongruence experienced by the Stroop tasks may resemble what is experienced by second language learners whose native language is based upon the same alphabetic system of English, the Latin alphabet. The incongruence present in this case, the orthographic element, even if appearing in a second language context correlates to an auditory element associated to the native language and may lead to the same attention competition as experienced by the Stroop task.

What is most interesting is what happens when the two attention processes, auditory and visual, occur simultaneously. Research by Jones (1999) indicated a 30% drop in visual recall performance if visual stimuli were presented along with unrelated auditory stimuli even if the subjects were told to ignore the speech. This suggests that the unattended auditory material is nonetheless processed, which thus impairs visual processing. The meaning of the auditory stimuli appears to have no impact on visual memory (Buchner, Irmen, & Erdfelder, 1996), but the patterning does. A seminal work by Jones, Saint-Aubin, and Tremblay (1999) indicates that a

continually changing sound sequence is more disruptive than one that follows a repetitive

pattern. In other words, concurrently presented auditory and visual stimuli cause a toggling in

attention processing, comprising the retention of the attended stimulus.

These findings are significant to language learning in that language instruction often

delivers the auditory and orthography elements simultaneously, and for a nonnative speaker of a

language, especially perhaps a beginning level learner, one or both of the inputs could, in fact,

appear unrelated as stated with Jones' research (1999). Hence, one or both the inputs could

constitute a distraction for the learner, causing competition for attention to ensue. Given the

temporal constraints of auditory stimuli, one might predict it to fall short of retention more so

than orthographic stimuli, due to the apparent differences in longevity and direct transference to

long-term memory. The relevance of the research regarding disruption in attention processing

caused by a continually changing sound versus a repetitive one (Jones, Saint-Aubin, &

Tremblay, 1999) could manifest for a second language learner as interference in that the

orthography might be seen as repetitive in its familiarity, as with native Spanish speakers,

whereas the audition most likely will seem irregular. Differences between native language and

second language orthography and phonology may cause the toggling of input to sequence the

orthography first and prominent. With a native language such as Arabic, where the orthography

is quite dissimilar not only in symbol but also direction, it might appear logical that an opposite

or no toggling pattern might occur.

Change detection. Further discussion with regard to attention leads to the theory of

change detection, a counterintuitive phenomenon that states that the brain does not detect

changes unless attention is placed on the change. This theory is typically associated with visual

attention, called change blindness, and has been studied extensively (Levin, Momen, Drivdahl, & Simons, 2000; Simon & Levin, 1998; Rensink, O'Regan, & Clark, 1997; Grimes, 1996). However, the auditory parallel of change deafness also exists, and research shows similar findings. The attentiveness to auditory changes was studied by Vitevich (2003) via a shadowing task of subjects repeating lexical items of varying complexity but which changed from a male to a female voice input 40% throughout the instrument. Statistically significant results indicated that attending to lexical components compromised attention to auditory changes, and vice versa.

In addition, under certain conditions event-related brain potential technology (ERPs) reports that subjects had a delayed response to significant semantic changes in auditory input, deferring comprehension 500-600 milliseconds (Nieuwland & Van Berkum, 2005). However, an earlier study of subjects exposed to one channel of a headset presenting speech in English, which was to be shadowed or repeated aloud, and the other channel presenting speech in German reported minimal interference from the unattended channel, the German input side. It was reported that subjects were able to accurately repeat the attended channel in English (Cherry, 1953). With regard to prosody, one can look to findings of change deafness in music as described by Agres and Krumhansl (2008) whereby several music variables and most significantly those of tonality went undetected by both professional musicians and non-musicians, indicating that listeners form a memory representation for tone that supersedes in attention over the actual representation of sound, leaving it unnoticed (Agres & Krumhansl, 2008).

In conclusion, the implications for the theory of change detection, both change deafness and change blindness, have bearing in a language classroom because of issues relative to the manner in which auditory and visual stimuli, whether orthographic or pictorial, are presented and

are potentially sequenced, possibly competing for attention. Applying these constructs to pronunciation research fuels decisions regarding data collection methodology and can assist in minimizing the effects of input distraction during speech production.

Second Language Phonological Attainment

The point of departure for discussions about ultimate phonological attainment in a second language has always been about the same, to delve into familiar constructs from first language acquisition. Like second language acquisition, even first language acquisition theories have evolved over time. Early work in cognitive psychology from the 1960s and 1970s has molded numerous of the most prominent constructs about how a child acquires a first language. Chomsky (1965) theorized that, unlike other living things, a human child will always acquire the ability to understand and produce language if exposed to linguistic data, labeling his theory the "language acquisition device"

Adding to the notions of the sequence of understanding to producing language, Lenneberg (1967) believed that the human infant does not actually imitate language but rather creates novel sound sequences which follow similar rules and, therefore, resemble speech and language. The presence of an appropriate response system, random babbling, is shaped into words, phrases, and complete mature utterances, implying that responding actually precedes understanding (Premack & Schwartz, 1966). Therefore, as part of the language process, it is a necessary phenomenon that the child not be concerned with first attaining mastery of constituent elements but rather the child should focus on the fundamental principle of pattern and structure acquisition (Lenneberg, 1967), exemplifying some degree of innate tolerance for pure phonological processing and acquisition without the necessity of meaning on the part of the first

language learner. In other words, native language listening is not explicitly taught (Rost, 2005, p. 503), and included in that is phonological acquisition. What is more is that the innateness of listening allows the first language learner to also develop cognitively.

Maturational factors influencing acquisition. Even the debate of a critical period for second language learning is anchored in the precepts and findings of first language acquisition research. Since the appearance of Lenneberg's 1967 construct of a critical period for first language learning starting at about age 2 and closing with puberty, much second language acquisition research has battled to establish a similar critical period for second language learners, a "unitary account for nonnative like outcomes" (Birdsong, 1999, p. 9). The decline in the natural acquisition of language coinciding with puberty was thought to signal neuro-chemical changes, causing the end of neural plasticity and the completion of hemispheric lateralization in the brain and assignment of specific language functions to either the right or left hemispheres (Lenneberg, 1967).

However, recent neurophysiological investigations exploring brainwave activity in response to speech stimuli have demonstrated a progressive, developmental change in how speech is processed across participants aged five years to adulthood (Cunningham, Nicol, Zecker, & Kraus, 2000; Sharma, Kraus, McGee, & Nicol, 1997). In addition, differing acquisition rates and levels between children and adults may lie in the cerebral cortex anatomy, neurolinguistic feature detectors, brain maturation, developmental and critical stages, local neuron circuitry, and the distinction between lower and high order language processes. Lower-order language processes are consolidated early in development and include the analyses of speech in Wernicke, the patterning of encoded information and expressive speech in Broca, and

the visual perception leading to reading and writing. Meanwhile higher-order processes develop later and are more adaptive to complex linguistic demands, such as semantics and word-object relationships (Walsh & Diller, 1979). Myelination, the sheathing around brain cells, also seems to play a role in that the neurons of higher-order thinking skills, such as morphology, semantics, and syntax, continue to develop over time whereas those of the Golgi type 1 cells associated with articulatory movement seem to reach maturity sooner (Abuhamdia, 1987).

Scovel (1977) took a different approach to the terminus for producing a native like accent by investigating instead the starting point for recognizing a native or nonnative accent, a critical period for acquiring the ability to recognize foreign accents. First with studies of children of differing ages and later with adults, aphasiacs, and foreign language learners (1981), he asked participants to judge speech sample for nativeness or non-nativeness and discovered that foreign language learner recognition rates were directly related to language proficiency and that pre-pubescent children were not able to effectively recognize foreign accent. Scovel's findings further suggested a tendency on the part of these two populations not to compare phonology when perceiving and producing speech as adults do. Perhaps the question of phonological attainment and a critical period is not one of age but rather of whether the language being learned is a first or second language (Scovel, 1969, p. 247). Therefore, it is not surprising that the research in second language acquisition first pointed to age of onset as the primary factor for the persistence of nonnative pronunciation in adult second language speakers.

Also among the early studies was that of Asher and Garcia (1969), who assessed 71 Cuban learners of English for pronunciation accuracy and determined that entering the United States at the age of 6 or earlier had the highest likelihood of attaining a native like accent. Ten

years later, a similar study by Oyama (1976) of 60 Italian learners of English with 6-20 years as age of onset and 5-18 years length of residence indicated that ultimate phonological attainment could be acquired if target language learning began at the age of 10 or earlier, somewhat later than had been reported by Asher and Garcia in 1969. Variations on the critical period window have also extended to encompass the mid-to-late teens (Scovel, 1988; Patkowski, 1980, 1990; Johnson & Newport, 1989) as the point at which ultimate attainment in a second language subsides. The second language learner's inability to lose a foreign accent has been attributed to an age-induced lack of plasticity in the brain's component for phonology (Selinger, 1978), which occurs around puberty. In sum, critical period research seems to repeatedly turn to age of onset as the cause for the persistence of nonnative-ness in spoken language and maturational constraints as contributing to the limitations to ultimate attainment, in particular for phonology. Connecting the critical or sensitive period hypotheses to biologically-grounded explanations for the differences between child and adult second language learning processes, although appealing, has proven to be quite difficult (Baker, Trofimovich, Flege, Mack, & Halter, 2008) perhaps because, as Steinhauer, White, and Drury (2009) stated, "age of onset is not the problem, proficiency is".

However, when it comes to second language learning, especially for adults, the stages of learning are quite different in that cognitive development is typically already beyond a formative stage. Research juxtaposing the learning process of children with that of adults has actually indicated that older learners and adults show advantages in their rate of learning over younger learners (Krashen, Long, & Scarcella, 1982). Findings from a longitudinal study by Snow and Hoefnagel-Höhle (1982) also found that learners aged 8-10 and 12-15 and adults scored higher

on pronunciation assessments that employed imitation and spontaneous tasks, strengthening the argument that older learners can perform as well as younger learners.

Research in neurodevelopment also supports the notion that adults may actually be better equipped for different specific kinds of learning as compared to children. In a study of native Japanese English as a second language children and adults, the older learners initially scored higher than the younger learners for segmental perception and production after 6 months of target language exposure but were surpassed by the younger learners once they reached 18 months of target language exposure. In another study, results of fricative production between the two periods indicated that younger learners improved significantly while older learners did not, suggesting that adults have an advantage at the onset of second language learning but that children improve more quickly over time (Aoyama, Guion, Flege, Yamada, & Akahane-Yamada, 2008). Flege's Speech Learning Model (1995), or *interaction hypothesis*, further asserts that bidirectional influence between the native and second language particularly impacts the realm of target phonology. This tenet further holds that the interaction between languages changes based on the state of development of the native language phonetic system when target language learning begins.

The quest for answers as to how the brain actually acquires language has compelled researchers to also examine cases of dysfunctional language, such as language experienced by feral subjects and impaired individuals. Similar to the way in which damage sustained to different areas of the brain results in different types of aphasia, so do differing rates and types of maturation result in differing acquisition patterns (Walsh & Diller, 1979). Among the most famous feral child case is that of Genie (Curtiss, 1977), confined to a room by her father and

ultimately discovered at an early-pubescent age. She endured severe isolation and abuse resulting in extreme development deficiencies linguistically and otherwise. Despite extensive language learning measures, it was reported that Genie remained significantly limited in language ability.

Because it is believed that cases of feral children give rise to noteworthy psychological issues contributing to learning disabilities, other areas of inquiry of the language learning process have preferred to focus on research of deaf populations. Curtiss (1994) and Grimshaw, Adelstein, Bryden, and MacKinnon (1998) both researched cases of deaf adults who had had their hearing restored after puberty but that demonstrated similar results to Genie; acquiring little language and with severe deficits.

Regardless of the impact made by these critical period proponents, the question still remains unanswered as to why some learners do, in fact, reach native like pronunciation levels in a second language, while others do not. As Bialystok and Miller stated, "If the evidence fails to support the existence of such a biological constraint on language acquisition, then the options for language acquisition are more diverse but are based on a much larger role for general cognitive mechanisms and environmental influences throughout all stages of language acquisition" (1999, p. 128). The issue of a critical period for second language acquisition has evolved over the past 60 years, however still leaving the enigma of phonological attainment predominantly phenomenological.

Non-maturational factors influencing acquisition. To this end, Schumann's (1978) taxonomy of factors influencing second language acquisition homed in on other aspects of language acquisition and details the specific roles of neurology, personality, cognition, instruction, and social factors in phonological attainment. Some have also incorporated notions

that while activated neurocognitive mechanisms for language learning remain active, inactive

mechanism cannot be regained, citing Birdsong's (1999) "use it or lost it" approach and the

"exercise hypothesis" of Johnson and Newport (1989). Flege's Speech Learning Model (1995)

reported that child/adult differences in phonological acquisition stem from the fact that adults

rely on pre-existing phonetic categories and are less sensitive to the phonology unique to the

target language. Flege further suggests that the adult perceives the target phonology but has

difficulty producing it (1987). In a 1995 paper, Flege, Munro and MacKay divided interference

into two components. The first of these was *habit formation*, where first language sounds are

plugged in for second language sounds. The other factor they identified as *incorrect perception*,

where the learners fail to perceive accurately the phonetic details of a second language.

A variety of sociolinguistic models have also addressed the variability in learner

language. For example, Labov (1970), based on his systematic observation of New Yorkers,

devised the Labovian paradigm which points to the features of differing speech styles and

variation rules employed by the learner as a result of the setting or topic. Similarly, others have

conducted inquiry into the roles of identity (Baker, 1992; Young & Gardner, 1990). Salient

research by Guiora, Beit-Hallahmi, Brannon, Dull, and Scovel (1972) delved into the notion of

empathy and language permeability. Guiora et al. (1972) posited that language ego is the barrier

to native like pronunciation and that pronunciation appears to be the most critical aspect to self-

representation. During the same period, Stevick coined the term *lathophobic aphasia* to define

the "unwillingness to speak for fear of making a mistake" (Stevick, 1976, p. 78). All these

studies speak to the sociological implications of learning a new phonology.

The element of learner motivation and exposure to the target language has also been deemed significant in terms of the possibility of ultimate phonological attainment (Moyer, 1999). Research by Coates (1986) suggested that motivation perceived as the need for achievement was strongly correlated with phonological attainment. James Flege (1995) contributed to the discussion of non-maturational causes of foreign accent by identifying contributing factors, such as the *motivation* to produce the exact sounds required, which decreases if articulatory errors do not impede communication, *individual differences*, which include language history, language habits, and resistance to sounding native like, and finally, the *phonetic input* received from exposure to native and nonnative accents The latter resonates the notion that input, both type and amount, is an essential factor in second language pronunciation.

Second language instruction is also considered a component in the notion of phonological attainment (Iandoli, 1990; Moyer, 1999; Pennington & Richards, 1986; Bongaerts, Van Summeren, Planken, & Schils, 1997). Research of adult learners of French exhibited native like performance on grammaticality judgments regardless of late age of onset (Birdsong, 1992) even though grammaticality judgment tasks have been highly criticized as inadequate means of measuring true language ability (Sorace, 1996). These studies have attempted to broaden the perspective beyond age of onset and length of exposure to the possibility of other contributing factors as the reasons why "…many fall far short of a native speaker ideal, while others succeed beyond all expectations" (Moyer, 2004, p 1).

The subjectivity crippling the critical period hypothesis for second language learning and ultimate phonological attainment is diminishing today due to recent developments in neuro-imaging technology, making the brain "more amenable to direct psychological investigation"

40

(Schumann, 2001). Two common neuroimaging techniques are the positron emission tomography (PET), which measures brain activity via the increase in blood flow, and functional magnetic resonance imaging (fMRI), which measures brain activity via the decrease in local deoxyhemoglobin concentration (hemoglobin not combined with oxygen because it has been released to the tissues). The growing dearth of studies using PET and fMRI analyses is providing a unique glimpse into brain activation patterns during second language learning and contributing to the debate over the issue of biological constraints for ultimate phonological attainment. To illustrate, Bloch, Kaiser, Kuenzli, Zappatore, Haller, Franceschini, & Nitsch (2009) applied fMRI analyses to study the cortical representations of multilinguals and implemented color-coding to represent brain activity patterns of participants during free speech tasks performed in the first, second, and third languages, respectively.

Several such studies have also focused on age as a factor of language development. An fMRI comparative study of younger and older second language learners found differences in terms of the overlapping between brain areas when children process native and second languages versus that of adults (Kim, Relkin, Lee, & Hirsch, 1997). Results from yet another fMRI study of younger second language learners showed a larger composition of grey matter in the language processing areas of the brain as compared to that of older learners (Mechelli, Crinion, Noppeney, O'Doherty, Ashburner, Frackowiak, & Price, 2004). In addition, fMRI results of early and late age English-Mandarin Chinese bilinguals administered word completion tasks in each of the two languages showed that common cortical areas were activated and similar pattern of brain activations regardless of the language and age of onset (Chee, Tan & Thiel, 1999), rejecting results found of auditory sentence comprehension tasks by English-French bilinguals. These

results imply differing hemispheric activations based on first or second language (Dehaene, Dupoux, Mehler, Cohen, Paulesu, Perani, Van de Moortele, Lehericy, & Le Bihan, 1997).

In some instances, research has been conducted to compare the results of PET to those of fMRI to determine consistency and net study findings for even more informed observations and conclusions. As the number of stimuli increased, so did the activations associated with visual processing and response generation as revealed by both analysis methods (Mechelli, Friston, & Price, 2000). Additionally, the results of PET- and fMRI-based research clearly indicates the role of second language proficiency as accessing native like neurocognitive processes (Perani, Paulesu, Galles, Dupoux, Dehaene, Bettinardi, Cappa, Fazio, & Mehler, 1998; Abutalebi, Cappa, & Perani, 2001; Wartenburger, Heekeren, Adutalebi, Cappa, Villringer, & Perani, 2003). In fact, whereas it was once thought that multiple languages were represented in different cerebral entities within the same individual (Penfield & Roberts, 1959), with advanced neurofunctional imaging, such as PET and fMRI, it is now possible to obtain depictions of bilinguals (Kovelman, Baler, & Petitto, 2008; Frenck-Mestre, Anton, Roth, Vaid, & Viallet, 2005), trilinguals (Vingerhoets, Borsel, Tesink, van den Noort, Deblaere, Seurinck, Vandemaele, & Achten, 2003), and even quadrilinguals (Briellmann, Saling, Connell, Waites, Abbott, & Jackson, 2004). All appear to indicate that a unified, neural language area is in use.

Beyond positron emission tomography and functional magnetic resonance imaging are also studies applying magnetoencephalography (MEG) analysis to second language research. The advantage over other cross-sectional methods is that MEG offers superior temporal resolutions (Schmidt & Roberts, 2009). Research applying event-related brain potentials (ERPs) displaying real-time electrophysiological brain dynamics of the cognitive processes in

millisecond time resolutions are also serving as a valuable method for measuring the impact of the second language process and factors, such as age of onset. In addition to contrasting ERP patterns to native speakers, this technology is also used to compare results to other second language learners of different native language groups and young learners of a second language as well as to execute longitudinal studies of the second language acquisition process for the purpose of observing the processes relative to miniature or artificial languages (Steinhauer, White, & Drury, 2009).

Recently, a longitudinal study employing ERPs was conducted to delve into the previously established notion that P600, the neurocognitive processes of "grammaticalization", would remain unchanged when faced with a grammatical structure unique to the second language. After studying 16 Korean and 16 Chinese adult English language learners for a grammatical structure of English not present in Korean or Mandarin Chinese, results revealed that although there was no P600 response for either group prior to treatment, after 9 weeks of intensive English language instruction, significant P600 levels in both groups were observed, suggesting that P600 neuro-cognitive processes are modulated by performance and learning (White, Genesee, & Steinhauer, 2012). This reinforces the notion of the role instruction plays in the language learning process.

Work by Steinhauer, White, and Drury (2009) reported studies of late-age second language learners revealing native like ERP patterns for morphosyntax. Similarly, ERP-based studies of late age French and Chinese learners of English (Steinhauer, White, Cornell, Genesee, & White, 2006) and university-age English speakers learning Spanish (Bowden, Sanz, Steinhauer, & Ullman, 2007) further indicate that proficiency, rather than age of onset,

influences the cognitive patterns of second language processing and that the target language appears to have no bearing on performance.

The most valuable contribution made by these variations in research has been the determination that non-maturational constraints are most likely at the core of differences in attainment (Ioup, 2005). This body of research has greatly contributed to the different perspectives of second language acquisition research in terms of adult learner variability in phonological attainment (Ellis, 2008). As the many facets that contribute to the process of phonological attainment in second language acquisition are being exposed, the probability of establishing a critical period theory diminishes. This trend toward continued inquiry into the workings of phonological attainment and an unspoken reluctance to succumb to the notion of a terminus in what an adult can phonologically achieve leads one to believe that researchers will continue to identify and define the numerous factors uniquely inherent in second language phonological acquisition in adult learners.

Pronunciation Research Methodology

Empirical research of second language pronunciation is concerned with which aspects of speech to observe, how to elicit those aspects, and how to objectively measure the aspects. In measuring degree of accentedness, researchers in the past have relied heavily on the comparison of the phonological performances of native speakers to that of nonnative speakers in order to better understand the phenomenon (Flege, Munro, & MacKay, 1995; Olson & Samuels, 1973; Oyama, 1976). In addition, a great deal of effort has been placed on contrastively-structured study designs (Lado, 1957), emphasizing the similarity or dissimilarity of two languages and how these features have a bearing on pronunciation attainment in a target language.

Often the elicitation instruments and tasks used to generate spoken language for the purpose of foreign accent observation are based on orthographically depicted language. There is general concern of how orthography may affect the data being gathered and analyzed in that English as a second language learners can be "…particularly hampered by the irregular orthographic representation…" of the target language (Brett, 2004, p. 103). Another critical component of pronunciation research is the method in which the data is analyzed, juggling the objectivity of digital analysis of sound with the naturalistic benefits of rater judgment. All these methodological factors come together to posit a challenging feat for researchers keen on maintaining objectivity and rigor while preserving relevance.

What to look at. To measure the degree of variation in pronunciation, over the years researchers have turned to the exploration of segmental aspects (Flege & Port, 1981). One such example is the measurement of formant frequency variations of vowel segments, which goes beyond identifying the variations of vibration in the vocal folds associated with fundamental frequency to further reflecting resonance in the vocal cavity, both in front and behind the tongue, characteristic of vowel production (Ryalls, 1996). Other examples of vowel-based segmental study designs are the Simon and D'Hulster (2012) investigation of the perception and production of /æ/ and /ɛ/ and the effect of language experience as investigated with three groups of native Dutch speakers of English as a Foreign Language in Belgium. In addition, a study of perceptual performance involved measuring the effect of two-months of computer-based, learner-focused training of three English vowels. It was administered to sixteen native Mandarin and Cantonese speakers and, three months after the study period, examined for retention (Wang & Munro, 2004).

Consonant-based research has also helped shape an understanding of the workings of foreign accent. Saito (2011) studied the impact of explicit instruction on the pronunciation of specific consonant segments in the target language in an experiment with 20 Japanese English language learners randomly assigned to control or experimental groups, the latter of which received four hours of direct instruction on */f/v/w/l/θ/ð/æ/ô/*. When later evaluated by native English listeners-raters for accentedness and comprehensibility, results indicated that explicit instruction had a significant effect on comprehensibility in sentence-reading tasks. Another experiment based on consonant analysis was conducted on the segments /s/ and /θ/ to determine degree of perception and production from Japanese adults (n = 16) and children (n = 16) learning English (Aoyama, Guion, Flege, Yamada, & Akahane-Yamada, 2008). Lastly, in a study of 51 Brazilian Portuguese speakers learning English as a foreign language, results indicated the three factors contributing to an increased probability of perception for post-vocalic plosive offsets were advanced language proficiency level, alveolar */t/d/* and bilabial */p/b/* segments, and offsets preceded by a lax vowel (Cardoso, 2011).

Relevant to segmentally-based research, functional load is an area that has been investigated for its relevance in pronunciation, demonstrating that certain segments are more impacting on intelligibility than others (Catford, 1987; Brown, 1988). Such examples include the functional load with */s/* and */ʃ/* as in the minimal pairs *so* and *show*, versus the erroneous replacement of */d/* for */ð/* as with *day* and *they*, with the former more greatly impacting intelligibility (Derwing & Munro, 2005). However, the segmental approach to teaching and researching pronunciation has been equally criticized (De Bot, 1983; James, 1976; Neufeld & Schneiderman, 1980), preferring instead suprasegmental analyses.

Research by Derwing and Munro (1997) and later supported by Derwing and Rossiter (2003) determined that native-speaker judgments of foreign accent perceived greater improvement in nonnative comprehensibility when the focus was on grammar and prosodic proficiency correction rather than with phonemic. Furthermore, Anderson-Hsieh, Johnson and Koehler (1992) concluded that suprasegmental variations, such as speaking rates, intonation, and pausing tendencies, are more fundamental to comprehensibility and effective rater assessment than segmental phonetic features, indicating a greater tolerance for inaccuracy in vowel and consonant production when certain prosodic features are replicated appropriately. Moreover, an investigation of sentence stress errors revealed a negative effect on intelligibility Hahn's (2004), while a recent study by Zielinski (2008) demonstrated the relevance of segments and syllable stress in strong syllables.

The value of the focus on prosody can be seen in an investigation of two groups of English as a second language learners, one receiving segmental instruction and the other prosodic instruction, whereby results indicated that the group receiving prosodic instruction demonstrated improved comprehensibility ratings on free speech tasks as compared the group that had received segmental instruction (Derwing, Munro, & Wiebe, 1998). Derwing and Munro (2005) further state the priority in pronunciation effectiveness should be an emphasis on the macroscopic aspects, the suprasegmental features. The relevance of examining prosodic features when identifying degree of foreign accent is additionally established by Kormos and Denes (2004) in a study which revealed that speech rate, including utterance length and phonation time ratio, are the most effective predict of nonnative speaker fluency. In fact, the body of research

examining naming latency of nonwords points specifically to string length as the primary

determiner (Frederiksen & Kroll, 1976).

How to look at it. Another important factor to consider in conducting pronunciation

research is to examine the tasks and instruments used to elicit speech. For instance, in a

pronunciation study by Moyer (1999) of 24 highly proficient and highly motivated learners of

German from various native languages, four speech elicitation techniques were used,

representing four different speech modes: word-list reading, sentence reading, paragraph reading,

and free speech production. Results revealed that the word-list task produced the highest

incidence of native like pronunciation as judged by a panel of four native German listeners.

Another study design employed oral retelling tasks of four prepared texts and a written

composition of a silent film (Hyltenstam, 1992) to measure the performance Swedish second

language learners who had an age of onset of either prior to age 6 or between 7 and puberty.

Other methodological approaches include read aloud tasks of 162-word standard passages in

English (Brennan & Brennan, 1981). It is important to note that the elicitation approach used in

many studies of pronunciation has employed tasks of a visual nature, whether pictorial or

orthographic, and it is the impact of this sort of elicitation that is being questioned through the

present inquiry.

Along with elicitation task types, it is necessary to look at the language itself being used

in the study. Often pronunciation research relies on pseudo word instruments, an intriguing

phenomenon when it comes to language acquisition. Learners seem to intuitively apply existing

word recognition processes to unknown lexical constructions regardless of the fact that they are

void of meaning. In an attempt to address the long-standing debate of word recognition of

pseudo words, researchers have examined various aspects of word recognition. One such study

published by Coltheart (1978), depicting a "dual-route theory" whereby text is converted to

speech via either a lexical route, a system by which orthographic input is read aloud by retrieving

the word's pronunciation, or a non-lexical route, whereby a graphemic representation is

converted into phonemes. These notions are also substantiated by other simulation studies based

on the Dual-route Cascaded (DRC) model (Coltheart, Curtis, Atkins, & Haller, 1993; Coltheart

& Rastle, 1994). It is believed that nonwords primarily employ the non-lexical route as a

fundamental process for naming, bypassing a stage of cognitive functioning (Rastle & Coltheart,

1998) and offering perhaps a cleaner look at how input modality effects phonological output.

Taken from another perspective, work by Bongaerts, Planken, and Schils (1995)

researched the degree of pronunciation monitoring by native and nonnative speakers of English

by exposing participants to four tasks: 3 minutes of spontaneous speech on a given topic, read

aloud of 84-word English texts, read aloud of 5- to 10-word sentences, and read aloud of 25

mono-/poly-syllabic English words. Their assumption was that pronunciation monitoring would

be minimal for spontaneous speech because of the semantic and syntactic demands of the task.

Study results indicated that there are, in fact, cases where nonnative speakers can acquire native

phonology (Bongaerts, Planken, & Schils, 1995), challenging Scovel's claim of biological

constraints on pronunciation preventing nonnative speakers from attaining native-enough

pronunciation to "pass themselves off as native speakers" of that language (Scovel, 1988, p.

185).

To understand the depth of word length and phoneme effects on word recognition and

how these dynamics present themselves in speakers of languages other than and in addition to

49

English, Rey, Jacobs, Schmidt-Weigand, and Ziegler (1998) investigated learners of English and French. Their study revealed statistically significant results in terms of word length and phoneme effect on word recognition, concluding that visual word recognition is strongly governed by functional orthographic units. Another study examined the Stroop effect with Catalan and Spanish to determine if the cross-language identity effect and the phono-translation effect persist and if, in fact, they are equally reliable and generalizable across speech production models beyond English applications. The results indicated that where the cross-language identity effect presented in this study generated the same naming latencies and error rates, the phono-translation effect did not manifest in the same manner, weakening the notion of lexical competition across languages (Costa, Albareda, & Santesteban, 2008).

The body of research examining naming latency of pseudo words points to string length as the primary determiner of latency (Frederiksen & Kroll, 1976). For example, Weekes (1997) reported that the number of letters and the number of graphemes had a statistically significant effect on nonword naming latency, with word neighborhood size having a facilitating impact. In addition, Rastle and Coltheart (1998) found that of the length effect for nonwords and the number of graphemes plays a role in predicting nonword naming latency, thus adding another variable to the DRC model. In conclusion, the issue of instrument design particularly relevant to this study is the concern that the speech derived from orthographic stimuli might vary from that derived from auditory-only input in that the visual and decoding stimuli of orthography triggers other cognitive functions and may alter the speech produced and output accuracy.

How to analysis it. To measure and analyze the patterns by sound and speech, several approaches are implemented in pronunciation research. These methods range from scientific

equipment to speech analysis software to rater judgment. Inherent in each of these methods is a delicate balance between objectivity and subjectivity.

To start with, equipment can be utilized to examine the multifarious aspects of sound and speech, such as time, frequency and amplitude, by generating electrical signals to represent sound patterns. One such instrument is the *sound level meter*, used for measuring the amplitude of sound patterns. The sound level meter converts sound patterns to electrical signals, magnifies the signals, and then measures the magnified signals in terms of decibels. Another approach is with an *oscilloscope*, an instrument designed to convert vibration patterns into electrical signals called *oscillograms*, which depict time and air movement patterns in respective horizontal and vertical visual displays. Yet another instrument used when examining and analyzing sound and speech is the *sound spectrograph*. This instrument created by Potter, Kopp and Green in 1946 analyzes rapidly changing sound patterns in speech and creates graphic depictions, called *spectrographs*, indicating time, frequency and amplitude (Boothroyd, 1986, p. 9-14). The latter of these three is most commonly used in foreign accent measurement and analysis.

Another approach to measuring degree of accentedness has been the idea of an accent index, which dates back to Labov's 1963 study of the frequency and distribution of phonetic variants in the speech changes of community life on the island of Martha's Vineyard, Massachusetts. Labov (1963) was able to correlate these linguistic tabulations to ethnic, social, and attitudinal variables. A later study employing an accent index was conducted on the bilingual Puerto Ricans in a barrio of New York, producing a reliable listing of the speech patterns characteristic of the population being studied, which in turn related to background and aspects of speech style (Fishman, Cooper, & Ma, 1971). In a correlational experiment of perceived degree

of accent and status judgment regarding the speech of nine Mexican American readers, Brennan and Brennan (1981) implemented their Accentedness Index, a 9-point Likert-scale ranged from High Accent to Low Accent and addressing 18 pronunciation variables, which was applied by forty-three Mexican American and 37 Anglo high-school student raters. Results indicate a significant correlation between evaluated degree of accent and status judgments (Brennan & Brennan, 1981).

However, the discussion of measuring accentedness often encompasses the related construct of *comprehensibility*, generally regarded as the listener's ability to understand the meaning of the utterance (Jenkins, 2002; Smith & Nelson, 1985). Therefore, it is not surprising that the most common approach applied in research attempting to measure degree of accentedness has been rater assessment. However, in spite of precise rater training efforts and close examination of intra- and inter-rater reliability, listener ratings of speech are vulnerable to inconsistencies between rater judgments and rater severity primarily due to the feature of individuality (Bonk & Ockey, 2003). One case in point is a study by Bachman, Lynch and Mason (1995) whereby noteworthy differences in rater severity between individual raters were evident but which appeared normalized when rater pool data was generalized, thus masking important inconsistencies in rating. In addition, concerns emerge with regard to studies of "expert" native speaker listener ratings that rely on the application of Likert-based criteria (Piske, MacKay, & Flege, 2001). A single 7- or 9-point Likert measurement tool employed to ascertain degree of accentedness, as in works by Derwing and Munro (1997) and Munro and Derwing (1995), proposes methodological weaknesses in terms of measurement reliability (Kang, Rubin & Pickering, 2010).

Other factors weakening the effectiveness of native speaker assessment of nonnative speaker phonological attainment are rater familiarity of nonnative accents (Rajadurai, 2007), rater attitudes about specific native language groups (Lippi-Green, 1997), rater expectation based on negative stereo-typical perceptions (Lindemann, 2003), rater exposure to languages and number of languages spoken (Baetens-Beardsmore, 1979), and rater tolerance of one accent over another (Bent & Bradlow, 2003). Scovel (1988) also questioned findings of degree of nativeness as reported by Neufeld (1977) based on a rater related design flaw, the prior rater bias that the subjects were native speakers. Examination of the reliability of rater assessment also has shown that rater consistency diminishes over time due to the dissipation of rater preparedness (Lumley & McNamara, 1995) and how rater consistency correlates to inexperience (Weigle, 1998), demonstrating the volatility inherent in individualized rater judgment of degree of foreign accent (Bonk & Ockey, 2003).

As an alternative to the subjectivity of judgments obtained through native speaker rating and to better align the time and cost benefits typically associated by listener/rater evaluation, researchers increasingly look to computerized analysis methods for measuring features of speech and determining degree of foreign accent (Pickering, 2004; Schuetze-Coburn, Shapley, & Weber, 1991; Wennerstrom, 1998). One such approach is through the use of programs such as Praat (the Dutch word for *talk*), which consists of audio analysis software utilized by Brett (2004) to develop real-time feedback to Italian adult English language learners by calculating and plotting the vowel formants on a graph and normalizing as needed for the differences in vocal tract dimensions of participants.

With the advent of speech-recognition technology, research is currently focusing on exploring the calibration of this technology to the rater judgment method by studying the correlation between the two. VILTS (Voice Interactive Training System), as well as the French version called ECHOS, is another technological tool playing a role in pronunciation research. This tool employs speech-recognition technology with that of pronunciation scoring to examine for correlations between technological and human pronunciation scoring methods (Neumeyer, Franco, Digalakis, & Weintraub, 2000). For example, work done by Cucchiarini, Stirk and Boves (1997) examined the telephone speech of 20 native and 60 nonnative speakers of Dutch as compared to pronunciation ratings assigned by experienced phoneticians and resulted in a high correlation of $r = 0.79$, demonstrating the potential of this technological approach to measuring degree of accentedness.

In addition to devising tools for objective foreign accent assessment, research is also being done on foreign accent conversion, a transformation of nonnative speech into native speech by removing only the accented features of the speech while preserving the speaker's voice identity. Being able to re-create the speaker's voice but with native like speech parameters means that the speaker will be able to hear what they would sound like without the accent, ultimately closing the ever-important phonological loop (Felps & Gutierrez-Osuna, 2010). The prospect that with time a speech recognizer can be devised to compute measurements aimed at evaluating pronunciation is promising.

CHAPTER 3:

METHODOLOGY

The methodology underlying this study was guided by the six interrelated principles of inquiry of the scientific process as indicated by the National Research Council (Shavelson & Towne, 2002), (a) posing questions that are relevant and can be empirically investigated; (b) being grounded in current theoretical constructs; (c) employing methods of direct investigation; (d) presenting explicit, sound, and connected logic; (e) providing for generalization and replication, and (f) submitting to professional scrutiny and criticism. This study proposed to have structured an inquiry of pertinent, valid research aimed at filling conceptual gaps with regard to foreign accent amidst current SLA theory which could be answered through direct, observable yet objective investigation with a potential for generalization and which could serve as a base for similar research in other settings.

Research Questions

1. Is there a statistically significant difference in pronunciation accuracy as measured by the utterance length of pseudo words produced from a single modality of auditory-only input as compared to those produced from a dual modality of auditory-orthographic input?

2. Is there a statistically significant difference between native language and pronunciation accuracy as measured by the utterance length of pseudo words produced from a single modality of auditory-only input as compared to those produced from a dual modality of auditory-orthographic input?

3. Is there a statistically significant relationship between second language proficiency level and pronunciation accuracy as measured by the utterance length of pseudo words produced from a single modality of auditory-only input as compared to those produced from a dual modality of auditory-orthographic input?

4. Is there a statistically significant relationship between native language and second language proficiency level in terms of pronunciation accuracy as measured by the utterance length of pseudo words produced from a single modality of auditory-only input as compared to those produced from a dual modality of auditory-orthographic input?

Research Design

The research design was a *Posttest-only Control Group Design* as described by Campbell and Stanley (1963) and administered in the form of a diagnostic assessment to adult nonnative speakers of English enrolled in an Intensive English Program at a public, "top tier" research university in the United States. Participants in existing class groups were randomly assigned to either an auditory-only or an auditory-orthographic comparison group. Both groups were administered a diagnostic language activity whereby participants listened to an audio recording of spoken pseudo words and were asked to repeat the pseudo words they heard. The only variation between these two groups was that participants in one group were exposed to the auditory-only input of the pseudo words, comprising a single-modality stimulus, while those in the other group were exposed to the same auditory input but with the added feature of the textual representation of the pseudo words, creating a dual-modality stimulus. The language activity, conducted under experimental conditions, had a total duration of 1 minute, 46 seconds. In both

cases, participant utterances were audio recorded for later evaluation. The audio files were labeled with the participant's student identification number and saved to the University's secure server in one of two secured folders: one for the auditory-only audio files and the other for the auditory-orthographic audio files. Throughout the data collection and data analysis processes, access was restricted to primary IEP administrative personnel, the researcher, and the research assistant.

Upon completion of the data collection phase, the two secure folders containing the audio files were combined into one data set, void of notations indicating from which instrument each speech sample was derived; the audio files contained within the combined secure folder were then analyzed for utterance length by the researcher, creating the audio data set of values used within the study.

Subsequently, 20% of the audio files from each of the two secure folders were randomly selected and combined into another secure folder and purged of notations indicating from which instrument each sample was derived. These audio files were later analyzed by the research assistant for inter-reliability purposes. In addition, correlational statistical analysis was applied to establish reliability and validity of the analysis process.

The demographic data set provided to the researcher and on which this study was based consisted of information with regard to participant name, student identification number, date of birth, gender, native language, and IEP placement scores consisting of unofficial combined and section TOEFL scores. As part of normal procedure, the Intensive English Program regularly maintains this data from all students enrolled in the Intensive English Program.

Lastly, the IRB approval (Appendix A) obtained for this study allowed for use of both data sets by the researcher, the audio data set of utterance length analysis and the demographic data set. Once obtained, the researcher proceeded with matching each individual audio data results with the appropriate demographic details so that descriptive and inferential statistical analyses could ensue.

Encompassed within this study were several variables; one dependent and four independent variables. Applying the criteria of measurement scales defined in 1951 by psychologist Stanley Smith Stevens (Lomax, 2007), the dependent variable, utterance length, was the feature being analyzed in order to determine pronunciation accuracy. This variable is a continuous variable, rendering an interval scale of measurement. Utterance length was determined by applying the subtraction method to waveform counter values, demarcating the beginning of each utterance and the end of each utterance. The difference between the two values yielded the measurement of utterance length. The free version of the Audacity 2.0.2 software used to analyze the spectrographs of the audio files contains a standard feature that automated the subtraction process in order to calculate for utterance length. This method was applied to each pseudo word of each audio file and utilized by both the researcher and the research assistant following the same research analysis protocol.

With regard to the independent variables, the grouping variable consisted of the two versions of the pronunciation diagnostic instrument. One group was administered the instrument with the single modality of auditory-only input, while the other group was exposed to the instrument with the dual modality of auditory-orthographic input. The utterances generated from these two groups were digitally recorded and saved for later use.

The other independent variables, native language and second language proficiency, described demographic information regarding each participant and were included in the demographic data set obtained from the IEP database. The native language variable consisted of a discrete variable of nominal scale with multiple values, representing the various native languages as self-reported by the participants. The variable second language proficiency level consisted of a continuous variable of an interval scale with multiple values. In the present study, second language proficiency was defined by unofficial scores from TOEFL Listening Comprehension and TOEFL Reading Comprehension sections as depicted in the demographic data set provided by IEP administrative personnel.

The data sets gathered in this study were examined using parametric analyses. Initially, the audio data set was analyzed by applying one-sample t tests to examine for differences in utterance length between each of the two values of the input modality variable in comparison to the test value, consisting of the utterance length as modeled in the language activity instrument. After the preliminary examination, statistical analyses were applied to attempt to answer each of the four research questions.

In order to address Research Question 1, a series of independent t tests were executed to analyze for differences in pseudo word utterance length based on the two values of the input modality variable, auditory-only and auditory-orthographic. Research Question 2 was examined via factorial ANOVA analyses applied to evaluate for differences in utterance length based on the independent variables of input modality, defined by auditory-only input and auditory-orthographic input, and the native language variable. With the latter, two languages were identified for inclusion, Arabic and Spanish. Although the data set reflected the 14 native

language values corresponding to the entire group of participant, Arabic and Spanish were specifically examined because of not only their prevalence within the data set but also their similarity in phonemic/phonetic writing systems.

Because Research Question 3 addressed relationship, multiple regressions were applied to determine the prediction power of input modality and the two second language proficiency variables of listening comprehension and reading comprehension, the unofficial TOEFL scores reported as IEP placement data. Although the composite TOEFL score was originally considered for inclusion in the statistical analysis phase, it was ultimately omitted from the analyses because of issues relating to collinearity due to the duplicating nature of the section scores within the composite TOEFL score.

Lastly, hierarchical regressions were applied in order to respond to Research Question 4. The impact of three factors in predicting utterance length was examined: input modality, native language, and second language proficiency. In applying the hierarchical regression analysis, two blocks were constructed with the variables. The first block consisted of the input modality variable, defined as auditory-only input and auditory-orthographic input, and the native language variable, defined as Arabic and Spanish. The second block added the two second language proficiency variables of unofficial TOEFL Listening Comprehension and Reading Comprehension section scores. Again, the second language proficiency factor was limited to the listening comprehension and reading comprehension variables only and did not include the composite TOEFL score in order to avoid a violation of fixed-effect model assumption of data independence.

Internal Validity

The issue of internal validity in this study was controlled and accounted for primarily by the *Posttest-only Control Group Design*, which regulated the eight classes of extraneous variables as described by Campbell and Stanley (1963): (a) History, (b) Maturation, (c) Testing, (d) Instrumentation, (e) Statistical Regression, (f) Selection, (g) Experimental mortality, and (h) Selection-maturation interaction. The effects of *History*, *Maturation*, *Testing*, and *Experimental mortality* were controlled for by the mere fact that there was only one measurement in the study, the post-test, and therefore also no passage of time. The effect of the *Instrument* variable was statistically controlled for by tests of inter-rater reliability to ascertain the calibration of the interpretation and application of date analysis protocol. The other three effects, *Statistical regression*, *Selection*, and *Selection-maturation interaction*, were controlled for in that participants were randomly assigned to either the auditory-only or auditory-orthographic group from existing Intensive English Program class groups.

Other concerns relative to internal validity were also controlled and accounted for. For example, the effects of *Differential selection*, *Experimental treatment diffusion*, *Compensatory rivalry by the control group (John Henry effect)*, and *Resentful demoralization of the control group* were all controlled for by the fact that participants were not specifically informed of the dichotomous structure of this study or the presence of two versions of the instrument enabling the two input modalities, auditory-only or auditory-orthographic. Nor were participants informed of the random assignment to either of the two comparison groups. Furthermore, the concern of treatment replication was not of concern because the reported number of independent replications of the study was the same as the reported number of participants. Lastly, the effect of

Compensatory equalization of treatment was accounted for because the study employed pseudo

word exposure, not likely to be perceived as desirable goods and/or services by either of the

comparison groups.

External Validity

The research design proposed for this study was also scrutinized for the four factors

jeopardizing external validity (Campbell & Stanley, 1963): (a) *Interaction effect of testing*, (b)

Interaction effect of selection, (c) *Reactive effect of experimental arrangements*, and (d)

Multiple-treatment interference effect. The lack of a pre-test inherent to the *Posttest-only Control*

Group Design again controlled for the *Interaction effect to testing* and *Multiple-treatment*

interference effect in that participants were administered the instrument only once, while the

Reactive effect of experimental arrangements was accounted for by the fact that listening to and

repeating pseudo words was an unlikely activity by any language learner, much less any human,

and that the words used in the instrument were randomly-generated by ARC Nonword Database

(Rastle, Harrington, & Coltheart, 2002). Lastly, the *Interaction effect to selection* was controlled

for by simply limiting the generalizability of experiment findings to the population of ESL

learners studying in Intensive English Programs in the United States, rather than attempting to

generalize to the other populations of adult English language learners, such as those in refugee

language programs or basic adult education English language programs.

Other considerations with regard to external validity included accounting for the

Hawthorne effect, which with the current study was lessened because the participants were

involved as part of typical diagnostic assessment typifying any new term of study in the IEP and

not administered specifically for a research focus or hypotheses. Moreover, pseudo word

production was limited to the participant's own natural abilities and was not subject to repetition or measured for improved performance over time. On the other hand, the *Novelty and disruption effect* may have been to some degree a threat to external validity in that listening to and repeating pseudo words is not a typical occurrence, and the novelty of repeating pseudo words may have been at an optimal level during the data collection phase because of the uniqueness of the items.

Population

There was no recruitment in this study because the data collection phase was administered as part of the typical diagnostic evaluation of all students enrolled in the Intensive English Program at the University. These assessments are generally administered to students by instructors in their respective course sections within the first few days of the term. However, IEP administrative personnel believed that the incorporation of a Program-wide language activity administered during the first few days of the term would provide useful information with regard to the pronunciation accuracy of not only individual students but also the overall pronunciation ability within each course group. It was also believed that the additional baseline information would prove useful to instructors in guiding pronunciation instruction throughout the course curriculum and, therefore, further supplement normal diagnostic steps taken within each course section. Lastly, the administrators of the Intensive English Program were eager to experience large-scale diagnostic assessment for speaking, so the program-wide pronunciation activity was well suited to meet this objective.

Therefore, the data used in the present study represented the entire population of students enrolled in the Intensive English Program during the second term of spring semester of 2013. Furthermore, because no sampling procedures were implemented, guidelines established by

Krejcie and Morgan (1970) with regard to sampling, which maintain a 95% confidence level and thus enable generalizability to adult English language learners in similar IEP populations, were not applied.

The population on which this study is based consisted of adult English language learners enrolled in a University Intensive English Program in the U.S. during the spring semester of 2013. Based on enrollment, the data gathering was predicted to include a rather large population size ($N = 262$), with all students enrolled in at least one of 16 existing IEP course groups. The gender distribution of the population at the time of the administration of the language activity was 58% male and 42% female, with an average student age of 20. Due to cases of absence ($N = 57$) and technical difficulty ($N = 11$), post-administration data indicated that the total number of audio files retrieved from participants and analyzed was less than expected but still remained substantial ($N = 194$), with a fairly equitable distribution between input modalities ($N_{AuditoryOnly} = 99$, $N_{AuditoryOrthographic} = 95$). Table 1 depicts the projected and actual population values across second language proficiency levels as determined via course enrollment. The second language proficiency level of the population in these courses ranged from beginning to advanced, as defined by the IEP placement criteria (see Appendix E). These are listed in Table 1, in generally progressive order, starting with beginning language proficiency level and ending with advanced.

Table 1

Predicted & Actual Input Modality Groups

| | Projected | | Actual | |
	Control N	Experimental N	Control N	Experimental N
Speaking 1B1	5	5	4	5
Speaking 1A2	7	8	7	7
Speaking 2B1	6	7	6	7
Speaking 2B2	9	8	8	7
Speaking 2A1	6	7	6	6
Speaking 2A2	9	9	6	6
CommSkills 3B1	9	9	7	9
CommSkills 3B2	8	8	6	0
CommSkills 3B3	9	9	8	8
CommSkills 3A1	12	12	11	10
CommSkills 3A2	11	11	8	6
CommSkills 4U1	6	5	2	3
CommSkills 4U2	9	8	2	3
CommSkills 4V1	10	10	7	7
CommSkills 4W1	9	8	7	5
CommSkills 4Y1	6	6	4	6

Note. CommSkills = Communication Skills courses.

In addition, fourteen native languages were reflected within the population of the data used in this study and are depicted in Table 2.

Table 2

Number of Speakers per Native Language

L1	*N*
Arabic	93
Spanish	35
Japanese	28
Chinese	10
French	5
Portuguese	5
Turkish	5
Korean	3
Thai	3
Guajarati/Hindi	2
Russian	2
German	1
Tagalog	1
Vietnamese	1

Of the total number of audio samples retrieved, 52% were from Arabic L1 speakers, and 17% were Spanish L1 speakers. Ultimately, the stratified random assignment to the two input

modality values rendered a fairly proportionate distribution of the two native languages at the

focus of this research: Arabic ($N_{AuditoryOnly} = 48$, $N_{AuditoryOrthographic} = 45$) and Spanish ($N_{AuditoryOnly}$

$= 18$, $N_{AuditoryOrthographic} = 17$).

Instrumentation

The pronunciation diagnostic instrument used in this study was created by the researcher

and was comprised of a pseudo-word language activity. Presented to both the auditory-only and

auditory-orthographic comparison groups, the tasks consisted of a PowerPoint presentation in

slideshow mode that included an accompanying sound file containing the pronunciation of the

pseudo words. The launching of the instrument was set on a timer and executed via Scheduled

Task in the operating system of the University's server so that, with each round of data

collection, the instrument would launch as simultaneously as possible at all workstations. The

audio recording device used to record participant utterances was also launched with Scheduled

Tasks to begin slightly before the instrument but also as simultaneously as possible at all

workstations. Both timers closed automatically at the end of the language activity.

During the instrument design phase, the researcher randomly generated one hundred

pseudo words (see Appendix B) from the ARC Nonword Database (Rastle, Harrington, &

Coltheart, 2002), all with mono-morphemic, legal bigrams, and orthographically existing onsets

and bodies. In other words, all the pseudo words randomly generated consisted of only one

syllable and contained letter sequences that adhered to the rules of standard American English.

From the original randomly-generated listing of 100, only those with stop onsets and offsets

were selected for inclusion in the pronunciation diagnostic instrument, yielding a list of pseudo

words both beginning and ending with the consonant sounds */p/b/t/d/k/g/*. Based on Nation and

Webb's guidelines for vocabulary research (2011, p. 210), stating the preference for smaller test sizes and larger groups rather than larger test sizes and small groups and considering that the population size already established by the existing IEP enrollment was large ($N = 262$), it was determined that 15 pseudo words would suffice. These pseudo words chosen were then audio recorded as spoken by the researcher, a native speaker of American English, according to the assigned phonetic transcription provided by the ARC Nonword Database (see Appendix C).

On the other hand, the visual aspect of the pronunciation diagnostic instrument presented against a light, beige background and began with a title page followed by brief, written and spoken, instructions explaining the language activity. The instructions were expressed with basic vocabulary and syntax, appropriate for beginning-level English language learners. The instructions lasted 27 seconds and were immediately followed by the 15 pseudo words.

In this next section of the language activity, each pseudo word was first heard by the participant via the headset. Participants were then asked to repeat the pseudo words as they heard them. The study instrument allowed approximately two seconds after the pseudo word modeling in order for the participant to repeat the pseudo word. Meanwhile, the Audacity 2.0.2 software audio recorded the participant utterances. The first three pseudo words in the pronunciation diagnostic instrument served as warm-up and, therefore, were not transcribed. The remaining 12 pseudo words that comprised the remainder of the instrument were later analyzed for utterance length, which constituted the audio data set subsequently used in statistical analyses.

The distinction in instrument versions between the two comparison groups lay in the fact that the experimental group instrument had the added feature of on-screen text for the pseudo words heard. The on-screen text appeared in a sans-serif font against the light beige background

with no other visual markers, and it appeared and disappeared in a manner that mirrored the presentation of the auditory modeling of each pseudo word. It was intended that the fade-in-fade-out feature of the orthographic representation on the screen resemble the manner in which the auditory models of the pseudo word utterances were produced.

From the audio files created via the pronunciation activity, the dependent variable of utterance length was determined, which generated a numeral measure of the distance in milliseconds from the onset through the offset of each participant utterance. Of primary concern was whether the sheer presence of the orthographic representation of the pseudo words would affect the pronunciation accuracy of the utterance length as compared to the model utterance of the instrument and between study groups. Seen from the other perspective was the issue of whether the auditory-only input would generate a more accurate utterance on the part of the subject as compared to the model utterance or the auditory-orthographic group. In language learning and teaching, this could have numerous implications, especially in the way new vocabulary is presented and sequenced with regard to auditory and orthographic representations so as to maximize pronunciation accuracy.

Data Collection Procedures

In spring semester of 2013, the Intensive English Program administered a pronunciation diagnostic assessment to all IEP students, adult non-native speakers of American English. This diagnostic assessment, or language activity, consisted of gathering audio recorded speech samples of all IEP students enrolled in Speaking and/or Communication Skills courses as part of a program-wide pronunciation diagnostic assessment administered within the first week of the term. The data collection was scheduled so as to be completed within two consecutive days.

In preparation of this language activity, all IEP instructors were given general information at the faculty meeting prior to the start of the term. Once the term had begun, Speaking and Communication Skills teachers were provided with a schedule of when their class would participate and basic information about the logistics. No specific information regarding the nature of the instrument or input modalities was discussed so as to avoid risks to internal validity. Instructors were, however, informed that feedback from the language activity would be provided via email to each participant and that instructors would also receive summarized feedback regarding the overall performance of their group of students. Two days before the administration of the language activity, instructors were reminded of the upcoming activity and pertinent details.

In order to facilitate the administration of the study, the researcher scheduled each class for either the first or the last 20-minute section of the class period. The researcher, therefore, administered the language activity to two groups per 50-minute class period, with a brief time in between to prepare the workstations for the next group. The researcher was present throughout the entire administration of each language activity to each participant and assisted if needed; technical personnel were also present in case technical support was needed.

The language activity was scheduled to take place in a multimedia lab located on the main campus of the University. It was equipped with 32 workstations, each with a radial distance between workstations of about five feet. The software specifications for the stations in the multimedia lab facility consisted of a Windows 7 Pro 64-bit Operating System. Microsoft Office 2010 was used to execute the instrument, and Audacity 2.0.2 audio recorded the participant speech. The operating system feature of Scheduled Tasks was used to launch both the audio

recording device and the instrument PowerPoint. The audio files of the instrument and participant utterances were recorded at 48 kHz, and a low-pass filter was applied at 24 kHz with an amplitude resolution of 16 bits. Moreover, each station was equipped with Eforcity 370648 Handsfree Stereo Headset Mic Headphones. All the hardware and software used in the execution of the study were individually tested prior to the actual administration to assure for accuracy and quality. Lastly, all audio files were stored in the University's server, a secure location with restricted access for the researcher, the research assistant, and the IEP administrative personnel only.

The workstations were previously scheduled to first launch the Audacity 2.0.2 audio recording device; then after a 15-second delay, the study instrument launched as displayed by the PowerPoint slideshow (see Appendix D). The audio recording device captured only microphone audio of the participant utterances and not the headset audio of the instrument. Each computer workstation was assigned to run either the control, auditory-only, or experimental, auditory-orthographic, version of the instrument. The control instrument was uploaded to run from the stations on the left side of the lab facility, while the experimental instrument ran from the stations on the right side of the facility.

On the day of the language activity, all instructions for the administration of the pronunciation diagnostic were read by the researcher from a script (see Appendix F). The researcher went to each classroom and, reading from the script, explained to students that the group was going to adjourn to the multimedia lab for a language activity. Participants were informed that they were going to be audio recorded as part of a language activity for all IEP students and that feedback would later be provided to them as to their performance. They were

asked to leave all personal effects, including cellphones, in the classroom for the brief duration

of the activity. The teacher remained in the classroom with the personal items and to retain any

students that happen to arrive late; no accommodations were made for late comers and absent

students.

Before entering the multimedia lab, the researcher read from a script (Appendix F) and

explained that they would each be identified by student identification number and assigned to a

computer station. Students were also asked not to touch the keyboard or mouse because the

headset and microphone were the only necessary devices. As students entered the multimedia lab

and were identified by student identification number, they were randomly assigned to one of the

two input modalities, either the stations with the auditory-only instrument or the ones with

auditory-orthographic instruments. The computer stations had previously been set up with the

pronunciation diagnostic instrument so that most of the stations on the left side had been loaded

with the auditory-only instrument, while the stations on the right side had the auditory-

orthographic instrument. Arranging the two instruments among the stations in this manner, on

opposite sides and separated by a ten-foot aisle, enabled the research to attenuate risks to internal

validity.

Study participants were admitted to the multimedia lab ten minutes prior to the activity

launch time so as to allow time for headphone/microphone set up. Once all participants were

seated, simple instructions were read from the script (Appendix F) to describe the task; questions

were answered as needed although few were presented. In addition, students were informed that

the language activity had been scheduled to start automatically as was the audio recording device

documenting their voice. This was handled via the Scheduled Tasks feature in the operating

system. No reference was made as to the two versions of the instrument or the added orthographic feature of the pseudo words on certain workstations. However, participants were informed that the language activity comprised difference instruments and encouraged to focus on their task only and not be distracted by other participants. Once the instructions had been read and the timers were approaching the launch time, each participant was asked to put on the headset equipped with audio microphone and wait for the start of the language activity. No other persons were admitted to the study area other than the researcher and technical assistants.

The instrument for the activity was administered via a PowerPoint SlideShow which started with brief audio and on-screen text instructions. Students were only able to adjust the volume of the audio of the headset. The instructions were followed by 15 pseudo words presented in either an auditory-only input modality or an auditory-orthographic input modality depending on station assignment. Students were instructed to listen and repeat the pseudo words they heard. Their production of the pseudo words was audio recorded using Audacity 2.0.2 software and automatically saved in the University's server in the appropriate secured folder.

The instrument began with a very brief introduction, written with the beginning language learner in mind. Both the auditory-only and auditory-orthographic versions of the instrument had the exact same auditory and orthographic depictions of the instructions. The introduction section lasted 27 seconds and was immediately followed by the pseudo word section of the instrument. The first three pseudo words were intended for practice and were not analyzed; only the remaining 12 pseudo words were scored. The only difference between the two instruments was the textual depiction accompanying each of the 15 pseudo words found on the instrument assigned to the experimental group only. After each pseudo word was heard, there was a two-

second delay before the next pseudo word so that the participant could repeat the pseudo word that had been heard. This alternating pattern of listening to the pseudo word and then speaking the pseudo word continued for all 15 items on the instrument and was the same for both versions of the instrument. The entire pseudo word section lasted 1 minute, 15 seconds. The slideshow ended with a brief salutation indicating that the language activity was finished. The total time for the execution of the instrument was 1 minute, 46 seconds.

Upon completion of the language activity, the researcher read from the script (Appendix F) to inform participants how the results would be disseminated and to return to their classrooms. Once the participants had left, the researcher assisted by IEP technical personnel proceeded to label each audio file with the participant's student identification number. These files were then saved to the appropriate subfolders, auditory-only or auditory-orthographic, located in a shared folder on the University's secure server.

Two weeks after the data gathering phase, the researcher disseminated the results individually to each participant with regard to pronunciation accuracy and collectively to each instructor with regard to class performance; these communications were delivered via electronic mail. The entire execution of the pronunciation diagnostic assessment, consisting of the activity organization, collection of participant audio files via the study instrument, later analysis of audio files gathered, and dissemination of results to both participants and teachers, was handled exclusively by the researcher as part of administrative duties for the Intensive English Program.

The second data set utilized in this study consisted of demographic data furnished to the researcher by the IEP administrative personnel and contained the student name and identification number, date of birth, native language, and placement TOEFL scores. Both the audio data set

and demographic data set were made available to the research via the University's secure server. Permission to use both data sets, that of utterance length values and that of demographic information, was granted by IEP administrative personnel as was the IRB approval (Appendix A) to obtain and utilize the data sets for the present study.

Data Analysis Procedures

The preparation of data for statistical analysis consisted of three stages: the researcher analysis of utterance length values from the audio files ($N = 194$), the research assistant analysis of utterance length values from a random sample of the audio files (n = 38) for inter-rater reliability purposes, and the union of the audio data set with the demographic data set. Although numerous features of pronunciation could have been analyzed in determining pronunciation accuracy, the prosodic aspect of utterance length was elected because it depicts one of the broadest features of speech and is considered a valid good general measure of learner pronunciation ability (Kormos & Denes, 2004). The first three pseudo words of the language activity served only as warm-up items and were not analyzed, leaving the remaining twelve pseudo words for data analysis.

In the first phase, the researcher determined the utterance length of each pseudo word of the model, the audio recording associated to the instrument. Utterance length was determined by applying the subtraction method to the numerical values for the onset and offset of each pseudo word. The free version of the Audacity 2.0.2 software used to analyze the spectrographs of the audio files contained a standard feature that automated the subtraction process in order to calculate for utterance length.

The researcher then analyzed all the participant audio files, void of notations indicating from which instrument each speech sample was derived. All pseudo word utterance length values were defined to the millisecond, and only complete and audible sound tracks were considered. In short, the program-wide audio data set obtained from the audio files ($N = 194$) consisted of 2,328 utterance length values derived from 4,657 tokens (194 cases x 12 pseudo word repetitions x 2 tokens each [1 onset and 1 offset]). Once the researcher had analyzed all the audio files and had created the audio data set of all utterance length values, the inter-rater reliability phase proceeded.

In this second data preparation phase, the researcher started by randomly selecting 20% of the participant audio files from each of the two folders of input modality ($n_{AuditoryOnly} = 20$, $n_{AuditoryOrthographic} = 18$). These audio files were set aside for later reliability analysis by the research assistant. From the remaining files, the researcher next created two sets of audio files for calibration purposes. This was done by randomly selecting four audio files from each of the two input modalities to create one calibration set, and then repeating the process to create a second calibration set. This process generated three sets of audio files: two calibration sets and one inter-rater reliability set

Once the researcher had instructed the research assistant as to the protocol for analyzing utterance length, the research assistant practiced the analysis protocol with the two calibration sets of audio files. Correlation statistics run on both calibration sets revealed statistically significant correlations between the utterance length analyses of the researcher and research assistant. With the completion of the calibration phase, the research assistant proceeded to analyze the 38 randomly-chosen audio files, which comprised the inter-rater reliability set. A

Pearson product-moment correlation coefficient analysis was applied to the utterance length values determined by the two raters, and results indicated a positive correlation, r = .854, n = 38 $p < .001$, demonstrating the reliability and validity of the utterance length values of the complete audio data set.

The third preliminary phase of data analysis treated the demographic data of the participants. Intensive English Program administrative personnel created a report using the IEP database of student information, consisting of participant name, student identification number, age, gender, and native language. The demographic data set was supplemented with second language proficiency data comprised of unofficial TOEFL scores obtained also through the IEP database as part of placement procedures for incoming students to the Intensive English Program. The placement tests administered and scored according to ETS protocol rendered both composite TOEFL scores and individual TOEFL section scores.

The second language proficiency level variable for this study initially intended to look at composite TOEFL scores as well as TOEFL Listening Comprehension and TOEFL Reading Comprehension scores. However, issues of collinearity and violations of fixed-effect model of assumptions related to dependence of data prevented the use of the composite TOEFL score. Therefore, the variable of second language proficiency was ultimately defined by the TOEFL Listening Comprehension and Reading Comprehension scores only. To sum up, the audio data set, consisting of utterance length values in milliseconds for all twelve pseudo words of the audio files, and the demographic data set, containing participant details and second language proficiency indicators, were combined to create a comprehensive set of data for use in the statistical analyses of this study; only complete cases were included.

Once the complete data set had been entered into *IBM SPSS Statistics 20*, the statistical analyses proceeded. In the preliminary data analysis for differences, one-sample *t* tests were applied to compare differences in utterance length values between the test value and those of the input modality variable, auditory-only and auditory-orthographic. Results were classified into one of two categories: AuditoryOnly/Model or AuditoryOrthographic/Model. Effect size *d* for the Control/Model and the Experimental/Model comparisons was manually computed using (Green & Salkind, 2008, p. 165)

$$d = \frac{t}{\sqrt{N}}$$

Subsequently, each of the four research questions was addressed in terms of statistical analyses. The first two examined differences in utterance length, the third research question looked at the relationship between utterance length and second language proficiency, and the last assessed the relationship between utterance length and native language to second language proficiency.

Data analysis 1. In order to address Research Question 1, independent *t* tests investigated differences in utterance length between the two input modalities of auditory-only and auditory-orthographic. This statistical procedure was conducted for both combined utterance length and individual utterance length values for each pseudo word.

Data analysis 2. The next area addressed native language. Results of a pilot study (Farina, 2012) of input modality and native language, as seen in Figures 1 and 2, demonstrated statistically significant differences in utterance length performance, $F(1, 15) = 9.817$, $p = .007$, with an eta-squared value of .396 further indicating that 40% of the variance in pseudo word utterance length was attributable to the two variables of input modality and native language. This pilot study served to guide the statistical approach applied in the current study.
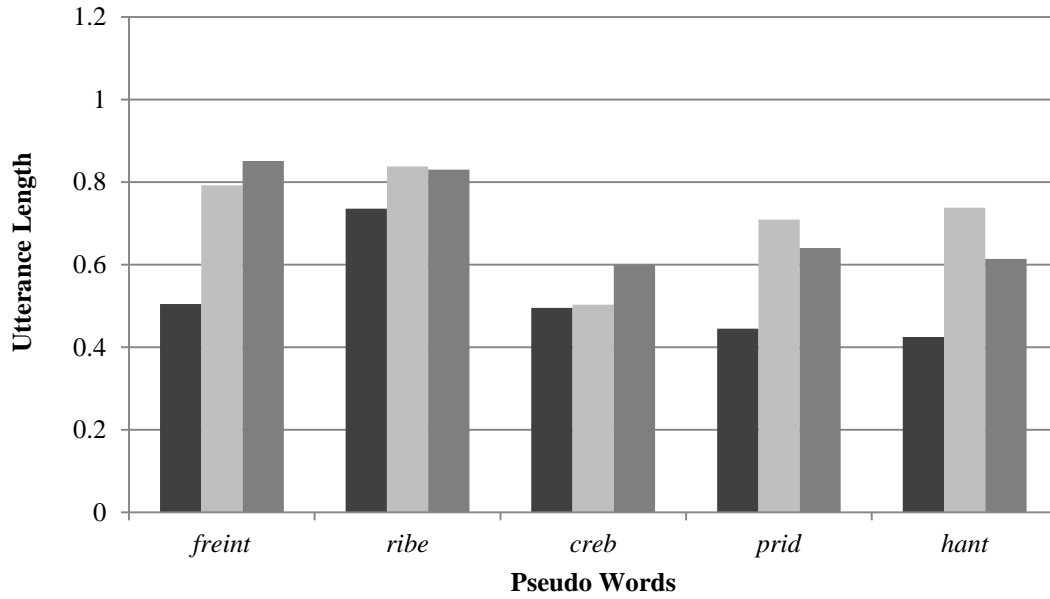
Figure 1: L1/Arabic pseudo word utterance length per input modality.
Model (dark bar) AuditoryOnly (light gray bar), AuditoryOrthographic (medium gray bar)
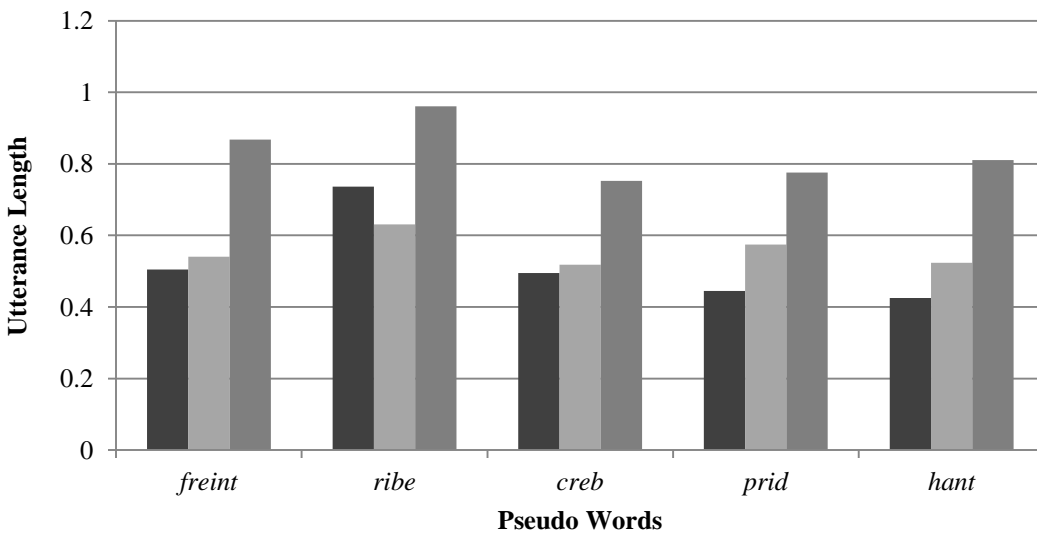
Figure 2: L1/Spanish pseudo word utterance length per input modality.
Model (dark bar) AuditoryOnly (light gray bar), AuditoryOrthographic (medium gray bar)

With the current study, the demographic data set was first examined to determine native language frequencies within the population represented in the audio data, and it was determined that Arabic ($N = 93$) and Spanish ($N = 35$) native language speakers were most represented in the data. Other predominant native language values were also be considered, such as Japanese ($N = 28$), but were not included in the analyses due to their dissimilarity in writing scripts as compared to American English.

Once the native languages were definitively identified, they were re-coded into *IBM SPSS Statistics 20*. With the native language factors determined and in response to Research Question 2, two-way analysis of variance tests were applied to explore differences in utterance length between input modality and native language, as defined in this study as Arabic and Spanish. These two-way ANOVAs were executed on the total utterance length values as well as individual utterance length values for each of the pseudo words.

Data analysis 3. As for the issue of relationship presented by Research Question 3, multiple regression analyses were used to examine utterance length from the standpoint of the predictive power of input modality and second language proficiency, as measured by TOEFL Listening Comprehension and TOEFL Reading Comprehension scores. Statistical procedures were first applied with TOEFL Listening Comprehension values, then TOEFL Reading Comprehension values, and finally for values of the two combined.

Data analysis 4. With input modality and native language potentially accounting for a bulk of predictive power, the researcher specified *a priori* a sequence for the sets of predictor variables. The first block was determined to be comprised of input modality and native language, while the second block of predictors contained the two second language proficiency factors of

listening comprehension and reading comprehension. Hierarchical regressions were employed to ascertain the predictive power of these two blocks and respond to Research Question 4. The concern with using composite TOEFL came about due to concerns of collinearity, in that the composite TOEFL variable as well as two of its three component variables, Listening Comprehension and Reading Comprehension, were to be used as predictors (Lomax, 2007, p. 403). Special regression analyses indicated severe collinearity with this arrangement of values due to the fact that all these variables were so highly correlated. It was, therefore, determined that TOEFL Listening Comprehension and Reading Comprehension values would be used only because they would better serve the purpose of examining for relationship between the dependent and independent variables. Statistical procedures were first applied with TOEFL Listening Comprehension values, then TOEFL Reading Comprehension values, and finally for values of the two combined.

CHAPTER 4:

RESULTS

The statistical analyses applied to the data set of this study consisted of five procedures and were based on the values of the single dependent variable of utterance length and multiple independent variables. For the dependent variable, the statistical procedures were conducted on not only the combined utterance length values of all twelve pseudo words (TotalUL) but also the individual utterance length values of each of the twelve pseudo words (e.g., twaspUL, crebUL) in order to expose additional patterns that may emerge at the pseudo-word level. The independent variables are derived from the four research questions of this study: input modality, defined by the two values of the single modality of an auditory-only stimulus (AuditoryOnly) and the dual modality of an auditory stimulus accompanied by an orthographic stimulus (AuditoryOrthographic); native language, defined by the two values of Arabic and Spanish; and second language proficiency, defined by the two variables TOEFL Listening Comprehension score (ListenComp) and TOEFL Reading Comprehension score (ReadComp).

Preliminarily, one-sample *t* tests were used to examine differences between utterance length and input modality as compared to test values, the modeled utterance length values derived from the instrument audio files. Subsequently, independent *t* tests were used to investigate Research Question 1, whether a statistically significant difference in utterance length existed between the two values of input modality, AuditoryOnly and AuditoryOrthographic. To respond to Research Question 2, a two-way analysis of variance test was implemented to determine whether a statistically significant difference in utterance length existed between input

modality and the two values of native language, Arabic and Spanish. As for relationship, multiple linear regression analyses were conducted in response to Research Question 3 to investigate whether utterance length could be predicted by input modality and the two variables of second language proficiency, ListenComp and ReadComp. Lastly, a hierarchical regression analysis was used to collectively delve further into how utterance length might be predicted by the interaction of the input modality, native language, and second language proficiency variables as proposed by Research Question 4.

Preliminary Results

The one-sample t tests were conducted initially on the TotalUL mean values for each of the two input modality values, AuditoryOnly ($N = 100$) and AuditoryOrthographic ($N = 94$), to determine if a statistically significant difference existed as compared to that of the test value. Results indicated that the AuditoryOnly/Model mean of 7.811 ($SD = 1.058$) was significantly different from the test value of 8.702, $t(99) = -8.415$, $p < .001$, with a 95% confidence interval ranging from -1.101 to -.681. The effect size d reported -.841, a large effect size, based on the conventional interpretation of .2, .5, and .8, as small, medium, and large effect sizes, respectively. Similarly, the AuditoryOrthographic/Model mean of 8.343 ($SD = 1.195$) was significantly different from the test value of 8.702, $t(93) = -2.912$, $p = .004$, with a 95% confidence interval ranging from -.604 to -.114. The effect size d showed -.300, a small effect size. This baseline difference points to a consistent and noteworthy deviation in utterance length as compared to the modeled utterance.

For a closer look, one-sample t tests were conducted on the utterance length mean values for each of the twelve pseudo words as compared to their relative test value. Table 3 indicates the

descriptive analyses of the individual UL data, depicting comparisons between the test values

and the two values of the input modality variable, AuditoryOnly and AuditoryOrthographic.

Table 3

Descriptive Statistics of Individual Pseudo-word Utterance Length in Milliseconds

Individual Pseudo-word UL	Test Value (*N* = 1) 8.702	Input Modality	
		AuditoryOnly (*N* = 100) 7.811 (1.059)	Auditory/Orthographic (*N* = 94) 8.343 (1.195)
twaspUL	.810	.703 (.143)	.782 (.143)
crekUL	.697	.635 (.112)	.683 (.116)
dapeUL	.631	.567 (.135)	.605 (.132)
clabUL	.746	.727 (.148)	.739 (.160)
tregUL	.757	.711 (.134)	.718 (.157)
pridUL	.680	.644 (.156)	.654 (.150)
clartUL	.656	.650 (.135)	.694 (.123)
trusqueUL	.768	.693 (.115)	.713 (.106)
crebUL	.659	.615 (.143)	.672 (.140)
qwipeUL	.713	.639 (.120)	.671 (.131)
dwimpUL	.913	.593 (.162)	.702 (.140)
tealtUL	.672	.636 (.142)	.710 (.107)

Note. UL = utterance length. *M* (*SD*) = population mean (standard deviation).

The one-sample *t* test results for individual pseudo-word utterance lengths, as illustrated in Table 4, indicated a statistically significant difference for the AuditoryOnly/Model comparison for ten of the twelve pseudo words, while results for the AuditoryOrthographic/Model comparison revealed a statistically significant difference for six out of twelve pseudo words, with two additional pseudo words approaching significance. By examining the twelve pseudo words individually, a pattern emerged whereby the significance of the AuditoryOnly/Model comparison was consistently more as compared to that of the AuditoryOrthographic/Model. Only with the pseudo word *clab* was this not the case. In fact, *clab* did not report significance that was contrary to this pattern but rather reported no statistically significant difference at all, regardless of input modality.

Table 4

One-sample *t* Test Results for Individual Pseudo-word Utterance Length in Milliseconds

	t	*df*	*p*	Mean Difference	95% CI of the Difference Lower	Upper	*d*
twaspUL							
AudioOnly	-7.506	99	.000***	-.107	.675	.731	-.751
AudioOrtho	-1.919	93	.058	-.028	.753	.811	-.198
crekUL							
AudioOnly	-5.553	99	.000***	-.062	.613	.657	-.555
AudioOrtho	-1.166	93	.247	-.014	.659	.707	-.120
dapeUL							
AudioOnly	-4.784	99	.000***	-.064	.540	.593	-.478
AudioOrtho	-1.893	93	.061	-.026	.578	.632	-.195
clabUL							
AudioOnly	-1.311	99	.193	-.019	.697	.756	-.131
AudioOrtho	-.409	93	.683	-.007	.707	.772	-.042
tregUL							
AudioOnly	-3.436	99	.001	-.046	.685	.738	-.344
AudioOrtho	-2.404	93	.018	-.039	.686	.750	-.248
pridUL							
AudioOnly	-2.335	99	.022	-.036	.613	.675	-.234
AudioOrtho	-1.655	93	.101	-.026	.624	.685	-.171
clartUL							
AudioOnly	-.468	99	.641	-.006	.623	.677	-.047
AudioOrtho	3.011	93	.003	.038	.669	.720	.311
trusqueUL							
AudioOnly	-6.583	99	.000***	-.075	.670	.715	-.658
AudioOrtho	-5.045	93	.000***	-.055	.691	.734	-.588
crebUL							
AudioOnly	-3.089	99	.003	-.044	.587	.643	-.309
AudioOrtho	.888	93	.377	.013	.643	.701	.092
qwipeUL							
AudioOnly	-6.187	99	.000***	-.074	.615	.663	-.619
AudioOrtho	-3.141	93	.002	-.043	.644	.698	-.324
dwimpUL							
AudioOnly	-19.785	99	.000***	-.320	.561	.626	-1.979
AudioOrtho	-14.579	93	.000***	-.211	.673	.731	-1.504
tealtUL							
AudioOnly	-2.526	99	.013	-.036	.608	.664	-.253
AudioOrtho	3.448	93	.001	.038	.688	.732	.356

Note. UL = utterance length. CI = confident interval.
*** *p* < .001, two-tailed.

In conclusion, one-sample *t* tests applied to the means of both the TotalUL values and the

individual pseudo-word UL values consistently revealed statistically significant differences for

AuditoryOnly/Model comparisons and AuditoryOrthographic/Model comparisons, suggesting a

baseline difference in UL exists for both groups as compared to the model.

Statistical Results

  Research hypothesis 1. In response to Research Question 1, it was hypothesized that

there is a statistically significant difference ($p < .05$) in pronunciation accuracy as measured by

the utterance length of pseudo words produced from a single modality of auditory-only input as

compared to those produced from a dual modality of auditory-orthographic input.

  Result 1. In order to address Research Question 1, a series of independent *t* tests were

conducted to examine for a difference between the means of both TotalUL values and individual

pseudo-word UL values as compared to the values comprising the input modality variable,

AuditoryOnly ($N = 100$, *M/SD* = 7.811/1.059) and AuditoryOrthographic ($N = 94$, *M/SD* =

8.343/1.195). The results of the independent *t* tests applied to the remaining data showed a

statistically significant difference in TotalUL between the two comparison groups, $t(192) = -$

3.285. $p = .001$, with a 95% confidence interval for the difference in means ranging from -.851 to

-.212. The effect size η^2 of .050 indicated that 5% of the variance in utterance length was

accounted for by input modality, a moderate effect size based on the conventional interpretation

of .01, .06, and .14 representing small, medium, and large effect sizes, respectively.

  Figure 3 indicates the distribution of utterance length for the two input modalities. As

depicted in the boxplot, the AuditoryOnly group included two outliers, case 20 and 73. However,

independent *t* tests reflecting the elimination of these outliners presented minimal variation,

$t(190) = -3.447.$ $p = .001$, with a 95% confidence interval for the difference in means ranging

from -.870 to -.237. With the outliers removed, the effect size η^2 of .06 indicated that 6% of the

variance to utterance length was attributable to input modality, a stronger but still moderate

effect size.
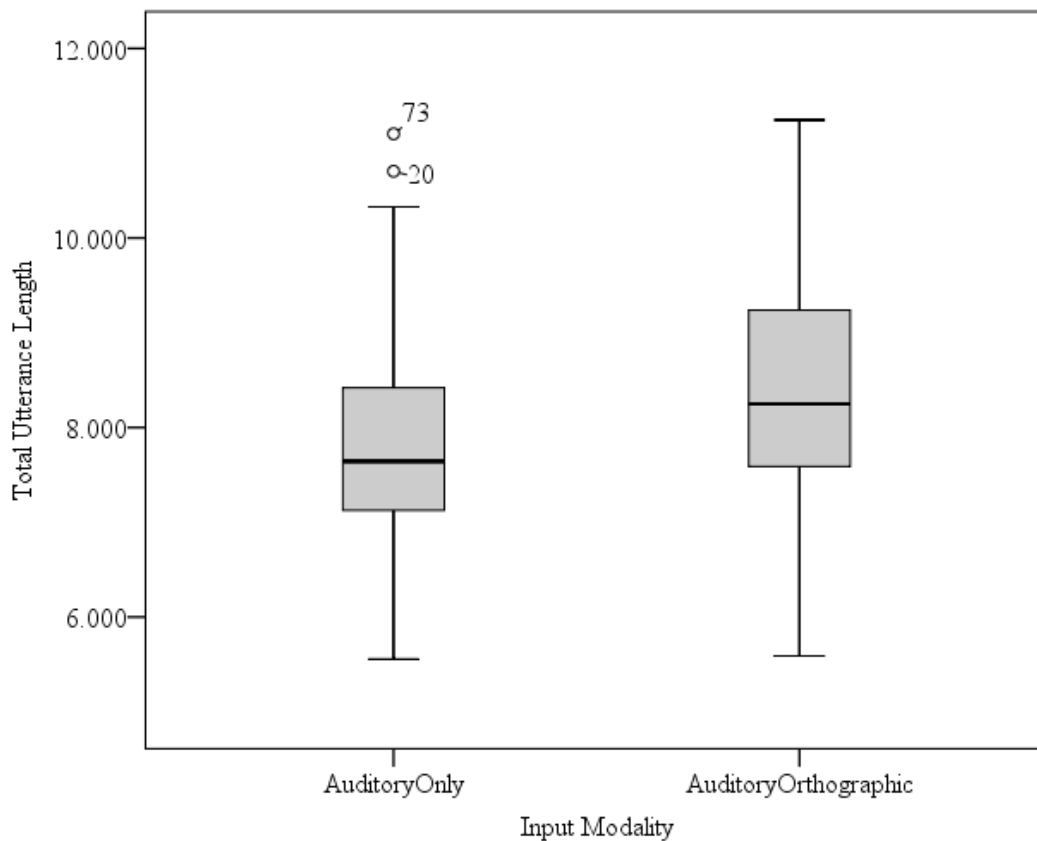
Figure 3: Utterance length for input modalities with outliers.

Furthermore, the independent *t* test analyses of individual pseudo-word UL values per comparison group revealed a statistically significant difference between input modality for seven out of the twelve pseudo words, with significance stronger than .05 for most, as illustrated in Table 5.

Table 5

Independent *t* tests for Individual Pseudo-word Utterance Length in Milliseconds

| | *t(df)* | *p* | *MD* | 95% CI of the Difference | | η^2 |
				Lower	Upper	
twaspUL	-3.839(192)	.000**	-.079	-.119	-.038	.071
crekUL	-2.929(192)	.004*	-.048	-.080	-.016	.043
dapeUL	-2.020(192)	.045	-.039	-.077	-.001	.021
clartUL	-2.395(192)	.018	-.045	-.081	-.008	.029
crebUL	-2.801(192)	.006*	-.057	-.097	-.017	.039
dwimpUL	-4.984(192)	.000**	-.109	-.152	-.066	.115
tealtUL	-4.074(192)	.000**	-.074	-.110	-.038	.080

Note. UL = utterance length; *MD* = mean difference. CI = confident interval.
* *p* < .01, two-tailed. ** *p* < .001, two-tailed.

To sum up, the results obtained from the independent *t*-test analyses suggest that input modality does have a bearing on the pronunciation accuracy of pseudo words as measured by utterance length and, therefore, do support Research Hypothesis 1.

Research hypothesis 2. In response to Research Question 2, it was hypothesized that there is a statistically significant difference (*p* < .05) between native language and pronunciation

accuracy as measured by the utterance length of pseudo words produced from a single modality of auditory-only input as compared to those produced from a dual modality of auditory-orthographic input.

Result 2. The second research question asked whether utterance length differences were attributable to not only input modality but also native language. The native language variable of Arabic ($N = 93$) and Spanish ($N = 35$) was juxtaposed with the input modality variable, AuditoryOnly ($n = 66$) and AuditoryOrthographic ($n = 62$) groups, yielding a fairly proportionate distribution as seen in Figure 4: Arabic ($N_{AuditoryOnly} = 48$, $N_{AuditoryOrthographic} = 45$) and Spanish ($N_{AuditoryOnly} = 18$, $N_{AuditoryOrthographic} = 17$). The means and standard deviations for TotalUL as a function of the two variables, input modality and native language, as well as the distributions are presented in Table 6.
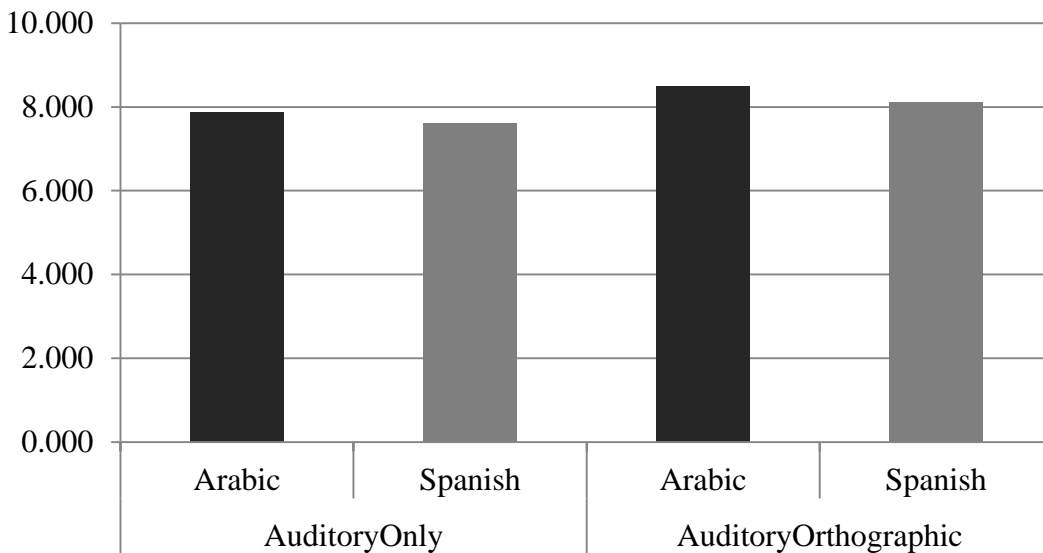
Figure 4: Input modality distribution for native languages.
L1/Arabic (dark bar) and L1/Spanish (light gray bar)

Table 6

Means & Standard Deviations for Total Utterance Length in Milliseconds

Input Modality	L1	*M*	*SD*	*N*
AuditoryOnly	Arabic	7.860	1.100	48
	Spanish	7.619	.891	45
AuditoryOrthographic	Arabic	8.498	1.316	18
	Spanish	8.105	1.171	17

Factorial, or two-way, ANOVAs were applied to both combined utterance length and individual utterance length values. As for TotalUL, results indicated no significant interaction between input modality and native language, $F(1, 124) = .108$, $p = .743$, and no significance main effect for native language, $F(1, 124) = 1.878$, $p = .173$. The very minimal effect size in both cases, $\eta^2 = .1$ and 1.5 respectively, reinforced findings. However, a significant main effect for input modality was found, $F(1, 124) = 5.914$, $p = .016$. The effect size η^2 of .046, a moderate effect size, indicated that almost 5% of the variance in utterance length was attributable to input modality.

Similarly, results as seen in Table 7 reflecting outcomes from the ANOVA tests for each individual UL value further revealed a significant main effect for input modality but no significant main effect for native language nor a significant interaction between the two factors.

Table 7

Two-way ANOVA Results for Individual Pseudo-word Utterance Length in Milliseconds

Pseudo-word UL	F	df	p	η^2
twaspUL	1.231	1, 124	.269	.105
crekUL	.014	1, 124	.907	.111
dapeUL	.173	1, 124	.678	.018
clabUL	.273	1, 124	.603	.007
tregUL	.041	1, 124	.838	.014
pridUL	2.504	1, 124	.116	.041
clartUL	.102	1, 124	.750	.078
trusqueUL	.196	1, 124	.658	.023
crebUL	.127	1, 124	.722	.057
qwipeUL	.108	1, 124	.743	.024
dwimpUL	.001	1, 124	.982	.143
tealtUL	1.147	1, 124	.286	.137

Note. UL = utterance length

In summary, all the results strongly indicate that differences in the population means of utterance length are not attributable to the interaction effect between input modality and native language, defined as Arabic and Spanish as in this study. In addition, significant main effects were found for input modality, but none were found for native language. Therefore, Research Hypothesis 2 is not supported because the data analysis indicates no statistically significant

difference between input modality and native language when it comes to output accuracy as measured by pseudo-word utterance length.

Research hypothesis 3. In response to Research Question 3, it was further hypothesized that there is a statistically significant relationship ($p < .05$) between second language proficiency level and pronunciation accuracy as measured by the utterance length of pseudo words produced from a single modality of auditory-only input as compared to those produced from a dual modality of auditory-orthographic input.

Result 3. The next phase of analysis focused on predictors of utterance length. Research Question 3 addressed the relationship between second language proficiency and input modality in predicting output accuracy as measured by utterance length. The second language proficiency factor consisted of two independent variables, ListenComp and ReadComp, both interval scale data comprised of two-digit numeric values ranging from 21-68. A fixed-effect model was applied for the multiple regression analyses since the data was gathered under experimental conditions. Furthermore, the ListenComp and ReadComp variables were classified as two unordered sets of predictors. Only cases with values for both variables were included. Table 8 indicates the descriptive statistics regarding the dependent and predictor variables.

Table 8

Means & Standard Deviations for Dependent and Predictor Variables in Milliseconds

Variables	*M*	*SD*	*N*
Total Utterance Length	8.040	1.172	148
Input Modality	1.490	.502	148
TOEFL Listening Comprehension score	44.650	8.085	148
TOEFL Reading Comprehension score	40.600	8.294	148

Two multiple regressions were conducted to determine the dependent variable of utterance length, labeled as TotalUL. One analysis included input modality and ListenComp as predictors, while the second analysis included input modality and ReadComp as predictors. Figure 5 presents a graphic depiction of the distribution of the two second language proficiency variables and the two native language values of Arabic and Spanish.
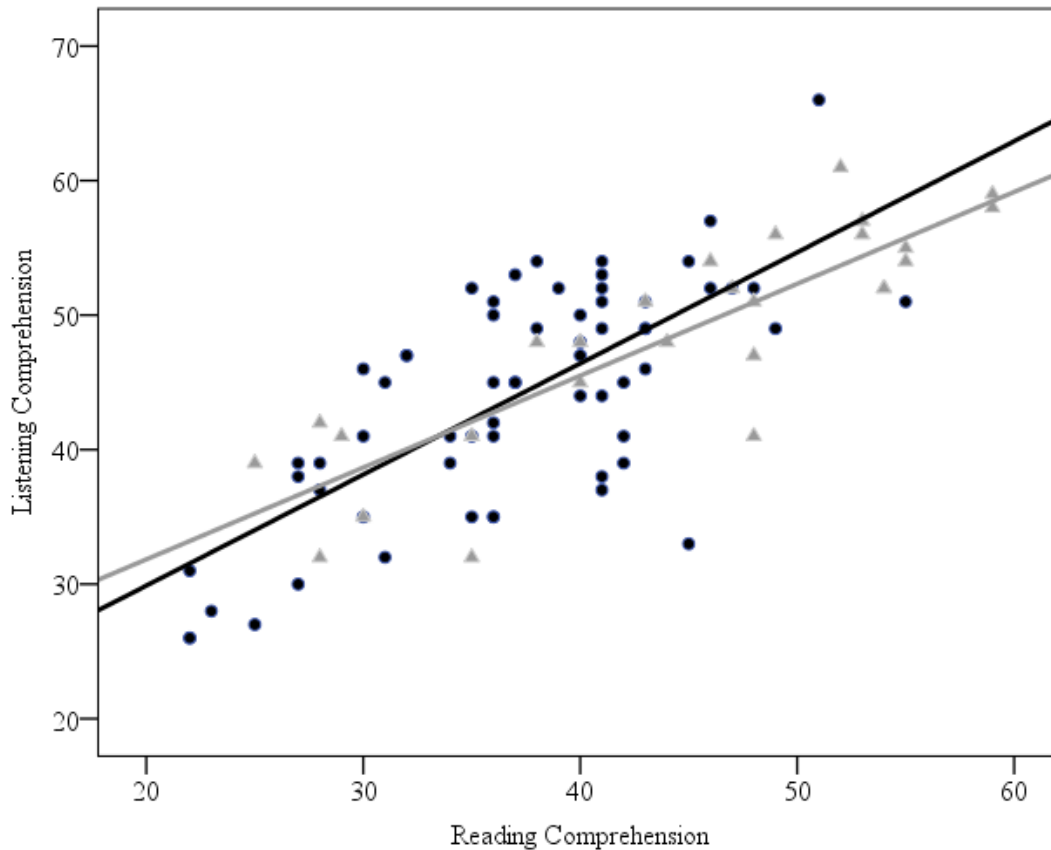
Figure 5: Second language proficiency distribution for native language.
L1/Arabic (light gray) and L1/Spanish (black)

The regression equation with ListenComp was significant, $R^2 = .087$, adjusted $R^2 = .075$, $F(2, 145) = 6.947$, $p = .001$. It was found that 9% of the utterance length variance was accounted for by the predictors of input modality and ListenComp. Likewise, the linear combination with ReadComp was significant, $R^2 = .108$, adjusted $R^2 = .096$, $F(2, 145) = 8.768$, $p < .001$, with 11% of the variance in utterance length attributable to the predictors of input modality and ReadComp.

Next, the multiple regression analysis was conducted with input modality and the combined ListenComp/ReadComp variable. The linear combination of input modality and the two second language proficiency measures was significantly related to utterance length, $R^2 = .108$, adjusted $R^2 = .089$, $F(3, 144) = 5.805$, $p = .001$. It was determined that 11% of the utterance length variance is accounted for by the predictors of input modality and second language proficiency as measured by the combined ListenComp/ReadComp variable.

Based on these results, both input modality and second language proficiency, here defined as listening comprehension and reading comprehension, are predictors of utterance length, with listening comprehension ability offering little additional predictive power over and above that contributed by reading comprehension ability. Hence, Research Hypothesis 3 is supported in that input modality and second language proficiency come together to predict utterance length.

Research hypothesis 4. In response to Research Question 4, it was hypothesized that there is a statistically significant relationship ($p < .05$) between both native language and second language proficiency level in terms of pronunciation accuracy as measured by the utterance length of pseudo words produced from a single modality of auditory-only input as compared to those produced from a dual modality of auditory-orthographic input.

Result 4. The final question, Research Question 4, examined the predictive power of the independent variables, native language, second language proficiency, and input modality, in determining pronunciation accuracy as measured by utterance length. A set of hierarchical multiple regression analyses were conducted to evaluate how well input modality, native language, and second language proficiency predict utterance length. According to the

96

researcher's *a priori* specifications in relation to the pilot study (Farina, 2012), the sequence of predictors was established as the two variables of input modality and native language, specifically Arabic and Spanish, and were listed as the first block of predictors. The second block of predictors consisted of the second language proficiency factors, the ListenComp and ReadComp variables. The analyses were conducted first with ListenComp, then with ReadComp, and finally with the combination, ListenComp/ReadComp.

The first hierarchical regression was based on the first block of input modality and the two native language values, Arabic and Spanish, and the second block of ListenComp scores. The descriptive statistics are seen in Table 9.

Table 9

Means & Standard Deviations with Listening Comprehension Predictor, Utterance Length in Milliseconds

Variables	*M*	*SD*	*N*
Total Utterance Length	8.078	1.255	90
Input Modality	1.480	.502	90
L1 Arabic/Spanish	1.300	.461	90
TOEFL Listening Comprehension score	45.320	8.560	90

Results showed statistical significance for the first block with input modality and the native languages of Arabic and Spanish, $R^2 = .069$, adjusted $R^2 = .048$, $F(2, 87) = 3.230$, $p = .044$. When the ListenComp variable was added, the multiple regression equation showed an increase

in statistical significance, $R^2 = .101$, adjusted $R^2 = .070$, $F(3, 86) = 3.232$, $p = .026$, indicating

that the added predictor of ListenComp increased the predictive power of utterance length over

and above the first block of variables from 7% to 10%, a slight increase in predictive power of

3%.

The second hierarchical regression was conducted with the same first block, input

modality and the native language variable of Arabic and Spanish, but with a different second

block, ReadComp. Table 10 displays the descriptive statistics for this combination of predictors.

Table 10

Means & Standard Deviations with Reading Comprehension Predictor, Utterance Length in

Milliseconds

Variables	*M*	*SD*	*N*
Total Utterance Length	8.078	1.255	90
Input Modality	1.480	.502	90
L1 Arabic/Spanish	1.310	.461	90
TOEFL Reading Comprehension score	39.230	8.899	90

The regression equation showed statistical significance for the first block with input modality

and native language, $R^2 = .069$, adjusted $R^2 = .048$, $F(2, 87) = 3.230$, $p = .044$. When the

ReadComp variable was added, the multiple regression equation showed an increase in statistical

significance, $R^2 = .112$, adjusted $R^2 = .081$, $F(1, 86) = 3.629$, $p = .016$, indicating that the added

predictor of ReadComp increased the predictive power of utterance length over and above the

other variables from 7% to 11%, again a slight increase of 4%. Although both statistically significant, the ReadComp variable indicated a greater significance and also a slightly higher predictive power in demonstrating utterance length over and above that of the ListenComp.

A final hierarchical regression was conducted again with the first block, input modality and native languages defined as Arabic and Spanish. However, the second block of predictors consisted of the combined variable of ListenComp/ReadComp in order to ascertain whether predictive power is increased by holistically adding both variables of second language proficiency. Descriptive statistics are depicted in Table 11.

Table 11

Means & Standard Deviations with ListenComp/ReadComp Predictor, Utterance Length in Milliseconds

Variables	M	SD	N
Total Utterance Length	8.078	1.255	90
Input Modality	1.480	.502	90
L1 Arabic/Spanish	1.300	.461	90
TOEFL Listening Comprehension score	45.320	8.558	90
TOEFL Reading Comprehension score	39.230	8.899	90

The results revealed a statistically significance relationship between utterance length and the predictors of input modality and native language, $R^2 = .069$, adjusted $R^2 = .048$, $F(2, 87) = 3.230$, $p = .044$. The addition of the second block, the second language proficiency factor comprised of the two variables of ListenComp and ReadComp, generated statistical significance,

$R^2 = .114$, adjusted $R^2 = .072$, $F(2, 85) = 2.129$, $p = .035$. Results also indicated that the addition

of ListenComp and ReadComp together increased the predictive power of utterance length from

7% to 11%. All in all, the results suggest that pronunciation accuracy measured by utterance

length can, to some degree, be predicted by considering factors of input modality, native

language and second language proficiency, and therefore, Research Hypothesis 4 is tenable.

CHAPTER 5:

CONCLUSION

The present study attempted to examine the phenomenon of pronunciation accuracy by focusing not on the accuracy of the output, as many prior studies have done, but rather on the input from which the output is derived. According to recent discussions by Flege (2012), a shift toward a focus on the origin of sound is in order. Therefore, the broad perception of output analysis undertaken in this study was an intentional attempt to keep the search light aimed on input.

Furthermore, the sound feature of utterance length was chosen because it was deemed the most macro of speech elements and, simply put, restricted the analysis to either the presence or the absence of language in its most general sense, a kind of "sound versus silence" approach to analyzing output. Again the intention of the researcher was not to be distracted by intricate output analyses but rather to stay pointed in the direction of input stimuli and how they manifest for output.

Along with the novelty of input modality and measuring utterance length as a method for assessing pronunciation accuracy, the researcher also sought to trim away the myriad notions of language, from meaning to spelling to grammar, and utilize study tasks that would capture as exclusively as possible the notion of sound imitation. All too often pronunciation research is based on read-aloud or naming tasks, tasks that represent global aspects of language and activate many cognitive processes far beyond what is needed to process sound. The presence of anything that has meaning guarantees cognitive functioning, making it extremely difficult to parse out

what is merely phonological from what is morphological, syntactic, or semantic. To this end, the instrument utilized in the current study was based on pseudo words, another deliberate attempt to deviate from the methodology previously used pronunciation research. By using unknown words, it was hoped that the activation of other cognitive mechanisms would remain at bay, leaving exposed more of what is auditory in nature. Therefore, pronunciation research that is based on nonsensical language can provide researchers with perhaps a cleaner view of the object being studied. In this case, it was utterance length, but other notions of pronunciation should also be examined under a pseudo language study framework.

Research Implications

The four research questions addressed within this study were addressed via the application of five statistical procedures. One-sample *t* tests were used as precursors in order to address the study questions in that they compared mean differences between the output derived from the two modalities as compared to the test value, the model utterances of the instrument. This was considered of value primarily because the audio heard by participants was comprised of unknown language. It was the researcher's inclination that participant utterances would necessitate direct reliance upon the modeled utterance and that the greater the reliance upon the audio track heard, the more accurate the utterance.

The input-output connection. Research Question 1 asked whether there was a statistically significant difference ($p < .05$) in pronunciation accuracy as measured by the utterance length of pseudo words produced from a single modality of auditory-only input as compared to those produced from a dual modality of auditory-orthographic input. It was further hypothesized that there would be a statistically significant difference ($p < .05$) between auditory-

only versus auditory-orthographic input modalities in terms of pronunciation accuracy as measured by the utterance length of pseudo words produced. Results from the one-sample t tests did, in fact, show that input modality affected utterance length differences in that both modalities reported statistical significance ($p < .05$), with auditory-only showing a large effect size η^2 of .841. Even the results for individual pseudo words replicated this finding for auditory-only over that of auditory-orthographic. The independent t tests reinforced these findings, whereby utterance length differences were attributable to input modality with statistical significance of $p = .001$ and moderate effect size of .050.

In simple terms, these indicators could suggest that as new vocabulary is presented in a language learning setting, the role of input is critical. One might infer that the toggling strategy underlying the presentation of auditory and orthographic stimuli is also an important consideration, further supporting findings of Jones (1999) where attention to one stimulus was compromised when dual stimuli were presented. In light of the temporal nature of auditory stimuli, the optimal sequence in order to maximize the phonological perception on which production will be based may point to an auditory-first approach, followed by orthographic stimuli.

In addition to the sequence, the role of auditory input versus that of orthographic input could also be defined in terms of the number of exposures of each modality. This idea mirrors what is already know about vocabulary exposure and acquisition with regard to the relevance of exposure to attainment. Further research is needed in the area of input modality and its impact on output accuracy in order to determine the optimal sequence and exposure of target language phonology so that more precise pronunciation can be acquired by non-native learners.

The role of native language. Research Question 2 inquired as to the role of native language and if there was a statistically significant difference ($p < .05$) between native language and pronunciation accuracy as measured by the utterance length of pseudo words produced from a single modality of auditory-only input as compared to those produced from a dual modality of auditory-orthographic input. The researcher hypothesized that native language would play a statistically significant role in the differences that would manifest with the two input modalities studied.

Interestingly, previous results from a pilot study (Farina, 2012) indicated statistically significant difference in utterance length based on Arabic and Spanish L1 ($p = .007$), with 40% of the variance in utterance length attributable to the two variables. However, the current study seemed to be in conflict with these results, reporting no statistically significant difference in utterance length based on input modality and the two L1s of Arabic and Spanish. Even with the removal of two outliers and a look at individual pseudo word data were the results not significant. One factor that may be affecting the oppositional findings on these two studies could be that the pilot study included a recruitment protocol and, therefore, was comprised of a sample group, whereas the current study was based on data of an entire study population. This difference in study groups could have also created variations in second language proficiency levels, a feature for which the pilot study did not account. Results from the present study do, in fact, reinforce ERP-based findings by Bowden, Sanz, Steinhauer, and Ullman (2007) of adult English speakers learning Spanish, where native language did not impact performance.

Too much of what is already known about foreign accent relates to the role of the L1, so it stands to reason that this is more to this aspect that meets the eye. Therefore, further research is

needed into not only second language learners but more so perhaps the general effect of input modality on all speakers. Knowledge about native speaker performance and the effect of input modality by all learners would strengthen the baseline understanding from which SLA research could be derived.

In addition, before determining that native language does not interface with input modality to impact pronunciation accuracy, research is also needed with speakers from other native language groups. For example, it would be beneficial to examine native languages of syllabic and logographic scripts to determine if patterns emerge as to the role of L1. It would also be wise to analyze audio data for phonological features other than utterance length before assuming that the native language hypothesis is not tenable. Investigation in these areas will contribute to base knowledge about how input modality affects learners in terms of phonological acquisition and assist in the identifying the role of native language as well.

The role of L2 proficiency. Research Question 3 delved into the possibility of a statistically significant relationship between second language proficiency level and pronunciation accuracy as measured by the utterance length of pseudo words produced from a single modality of auditory-only input as compared to those produced from a dual modality of auditory-orthographic input. It was hypothesized that there would be a statistically significant relationship ($p < .05$) between second language proficiency level and pronunciation accuracy as measured by the utterance length of pseudo words produced from a single modality of auditory-only input as compared to those produced from a dual modality of auditory-orthographic input.

Of the proficiency markers to consider, this study addressed the role of second language proficiency by looking at listening and reading comprehension values. These factors were

considered individually and in combination. The value of combining these two scores is that listening and reading comprehension scores represent ability in the two receptive skills of language learning, the input side of second language acquisition. In addition, the combination of these two areas of comprehension replicates the dual modality of the auditory-orthographic version of the instrument. The multiple regression analysis results indicated that these second language proficiency scores, both individually and collectively, serve as predictors of utterance length when combined with input modality. This result builds on research findings by van Els and de Bot (1987), whereby language proficiency was an indicator of sound perception. Although reading comprehension as used in the current study seemed to be a slightly better predictor over listening comprehension, the results demonstrated by the three multiple regression analyses were not considered strong predictors of pronunciation accuracy as measured by utterance length and mark a clear need for deeper investigation.

Research Question 4 further the exploration of the role of L2 proficiency by coupling the factors of input modality with that of native language to determine if a statistically significant relationship exists in terms of pronunciation accuracy as measured by the utterance length of pseudo words characterized in this study. It was hypothesized that there would be a statistically significant relationship ($p < .05$) between native language and second language proficiency level in terms of pronunciation accuracy as measured by the utterance length of pseudo words produced from a single modality of auditory-only input as compared to those produced from a dual modality of auditory-orthographic input. Hierarchical regressions were applied to examine for the predictive powers of these factors.

The hierarchical regressions consisted of one block with the input modality and native language variables and the other block with the addition of the second language proficiency variables of listening and reading comprehension, examined both individually and collectively. Findings resembled those found of multiple regression analyses and showed a statistically significant relationship between variables but with minimal predictive power. Therefore, of value would be further analysis into the predictive nature of second language proficiency. Such studies could be conducted with a pre/post design, for example, so as to examine for predictive powers over time and increased proficiency. In addition, different measures of second language proficiency could be employed, including minimal pair assessments that focus completely on auditory perception and not meaning as is with the TOEFL. Clearly there is more to the discussion of the role of second language proficiency and native language when it comes to input modality, but how those variables come together in the form of pronunciation accuracy is yet to be determined.

Limitations

Of the limitations to be discussed, the first to acknowledge is the over-arching role of cognition. The relevance of overlapping cognitive perspectives with SLA research (Pica, 2005) can assist in establishing baseline notions for phonology acquisition for both first and second language. The present study looked at second language learners, and although much of the results are significant and appear to clearly point to input modality as the cause, no such definitive conclusion can be drawn from the data obtained. For the uncovered stone here is whether these results would manifest differently in a native speaker. To truly understand how much of the difference in utterance length is attributable to input modality, it is necessary to first

account for how much of the behavior observed in second language learners is mere human cognition.

An obvious way to better understand this factor is by studying native speaker behaviors, perhaps even those of monolingual speakers. This could easily be accomplished employing the research methodology of the present study. The inclusion of native speakers of American English would make it possible to determine how much of the SLL outcomes are attributable to any speakers, an important ratio to understand before attempting to push forward on input modality as a critical component to pronunciation accuracy. Via the subtraction method, it is believed that native speaker behaviors relative to input modality could account for how much of the behavior stems from cognition in general versus that relative to learning a second language.

Future Research

In line with this recommendation, it would be useful to also study the effect of input modality with regard to foreign language learning. In practical terms, this could mean studying differences due to input modalities with populations in EFL settings. However, this could also develop into studying foreign languages taught in English-speaking countries. Of course, this would entail the creation of native-speaker auditory input and an appropriate instrument in the target language. Clearly it might be challenging to obtain pseudo words in less common languages, and there may not be an easy solution, such as the ARC database used in this study. However, the importance of understanding how auditory and orthographic inputs interact to contribute to accent is perhaps even more critical in a foreign language classroom where the setting beyond the classroom may not promote acquisition. The understanding how foreign language learners behave with exposure to the auditory and orthographic modalities can help

teachers and learners better deduce the optimal sequence of exposure to language stimuli in order to maximize of phonological acquisition in the target language. Additionally by administering the study protocol with other languages, research could also better address the patterns in utterance length specific to first language, an area that this study was unable to inform but that the previous pilot study showed significance (Farina, 2012).

Furthermore, it is important to also view languages with other scripts. The Arabic and Spanish focus portrayed in this study proposes some interesting components, such as directionality, which warrants deeper investigation. However, equally valuable would be analysis of languages with syllabic writing systems, such as Korean and Japanese, or perhaps logographically-based languages, such as Chinese. The study of other writing scripts adds another layer to the understanding of how auditory language can, at times, be visually perceived and how this perception can impact the auditory input from being absorbed by the learner.

Another cognitively-related limitation of the present study stems from visual stimuli. It is not possible to fully account for the impact orthographic input has on learner output without thoroughly extracting the portion of difference that any visual stimulus would conjure up. Therefore, the present results cannot truly determine with certainly whether the differences in the auditory-orthographic group are due to reading per se, rather than visual stimulation in a broader sense. More research is needed as to how the brain is affected by visual stimuli and even non-language script images, such as that comprised perhaps of simple lines, circles and half-circles arranged in a way to resemble text but that is not actually readable. By trimming out the behaviors of other visual, non-reading stimuli, one could determine just how much of the interference is visual versus something more cognitively engaging, such as reading.

Although the present study does address some key areas of pronunciation, it does not account for how participant performance may have changed throughout the short duration of the pronunciation diagnostic instrument itself. Because the very nature of the language task was based on pseudo language and therefore could have been perceived as contrived, participants could have somewhat "checked out" during the task due to the lack of context or meaning. In addition, by looking at the progressive performance of participants over the course of the task, it would be possible to also determine if learners actually improve in utterance length imitation as they adapt to listening and repeating pseudo language, or in fact if they actually decline in utterance length accuracy because they apply less attention given the language is void of meaning and, therefore, irrelevant.

Ironically, even though the pseudo words are completely unknown and, therefore, the modeled utterance is the only way from which the participant could imitate the sound sequence, many students in the audio recordings analyzed still deviated considerably. This could imply that the natural inclination by leaners is to not listen carefully to the first time language is spoken, and that when the model utterance is not clearly depicted in the mind, pre-conceived patterns of what was probably heard quickly supersede. This coincides with Flege's notion (1991) that the first language phonological anchoring can prevent the learner from perceiving, and in the end even producing, the target language phonology.

For this, it is important to better understand the tolerance the brain has for language without meaning and determine what causes fatigue and impacts phonological acquisition and pronunciation accuracy levels. Knowing more about participant performance throughout pseudo word tasks could inform SLA with regard to the learning behaviors relative to pronunciation,

particularly when no meaning is present as with beginner learners or those learning new vocabulary. Regardless of the outcome, knowing if patterns exist and if those patterns relate to the native language of the learner as well as the second language proficiency level of the learner could bring valuable information to the discussion of pronunciation curriculum and instruction.

Conclusion

The evolutionary path of research seeking to understand the phenomenon of L2 phonological acquisition moved from a time when age was believed to be the determining factor to the present, where it is believed that other factors, such as sociolinguistics, psycho-linguistics, and instruction, significantly impact second language learning and, in a particular way, the level and rate of phonological attainment. Current trends in technology are further pushing along the understanding of how the brain processes a second language and how new sounds are perceived and ultimately processed. Moreover, advancements in neuro-plasticity are pointing to far-reaching brain plasticity with realistic implications relative to the malleability of foreign accent (Doidge, 2007). The additional perspective of the role of input modality in second language learning builds upon existing constructs in cognitive psychology and second language acquisition while also representing a somewhat un-turned stone in pronunciation research. With greater knowledge in the area of input modality in language learning, it is hoped that SLA researchers can better identify the origin of foreign accent and engage in developing new pedagogical approaches that improve pronunciation teaching and learning, raising the bar of ultimate phonological attainment in second language learning.

APPENDIX A:

IRB APPROVAL LETTER

University of Central Florida Institutional Review Board
Office of Research & Commercialization
12201 Research Parkway, Suite 501
Orlando, Florida 32826-3246
Telephone: 407-823-2901 or 407-882-2276
www.research.ucf.edu/compliance/irb.html

Approval of Exempt Human Research

From: **UCF Institutional Review Board #1**
          FWA00000351, IRB00001138

To: **Marcella A. Farina**

Date: **March 18, 2013**

Dear Researcher:

On 3/18/2013, the IRB approved the following activity as human participant research that is exempt from regulation:

Type of Review:	Exempt Determination
Project Title:	THE EFFECT OF INPUT MODALITY ON PRONUNCIATION ACCURACYIN ENGLISH LANGUAGE LEARNERS
Investigator:	Marcella A Farina
IRB Number:	SBE-13-09226
Funding Agency:	
Grant Title:	
Research ID:	N/A

This determination applies only to the activities described in the IRB submission and does not apply should any changes be made. If changes are made and there are questions about whether these changes affect the exempt status of the human research, please contact the IRB. When you have completed your research, please submit a Study Closure request in iRIS so that IRB records will be accurate.

In the conduct of this research, you are responsible to follow the requirements of the Investigator Manual.

On behalf of Sophia Dziegielewski, Ph.D., L.C.S.W., UCF IRB Chair, this letter is signed by:

Signature applied by Joanne Muratori on 03/18/2013 12:23:07 PM EST

IRB Coordinator

APPENDIX B:

LIST OF 100 PSEUDO WORDS

	Word	Pronunciation	Number of letters
1	ruge	ru_	4
2	frirk	fr3k	5
3	charque	J#k	7
4	cliek	kl2k	5
5	phrerp	fr3p	6
6	pluse	plus	5
7	shro	Sr5	4
8	vourth	v9T	6
9	gwusque	gwVsk	7
10	phrice	fr2s	6
11	yelt	jElt	4
12	ike	2k	3
13	loice	l4s	5
14	stroove	struv	7
15	whouse	w6s	6
16	trisc	trIsk	5
17	woove	wuv	5
18	nurch	n3J	5
19	flink	flINk	5
20	sprarn	spr#n	6
21	vom	vQm	3
22	clipe	kl2p	5
23	thweeve	Twiv	7
24	bolf	bQlf	4
25	pive	p2v	4
26	twasp	tw{sp	5
27	smeige	sm1_	6
28	minge	mIn_	5
29	jalc	_{lk	4
30	yash	j{S	4
31	rhoin	r4n	5
32	kna	n#	3
33	druice	drus	6
34	gwame	gw1m	5
35	zooch	zuJ	5
36	creech	kriJ	6
7	skerv	sk3v	5
38	lount	l6nt	5
39	slarc	sl#k	5
40	rherg	r3g	5
41	wulch	wVlJ	5

42	treathe	triD	7
43	knint	nInt	5
44	brurch	br3J	6
45	crek	krEk	5
46	strol	strQl	5
47	phreant	frint	7
48	dape	d1p	4
49	gwurve	gw3v	6
50	smope	sm5p	5
51	thwauve	Tw9v	7
52	gwulge	gwVl_	6
53	thrurg	Tr3g	6
54	zedge	zE_	5
55	cleuth	kluT	6
56	clab	kl{b	4
57	prourse	pr9s	7
58	thelch	TElJ	6
59	flooth	fluT	6
60	rolk	rQlk	4
61	treg	trEg	5
62	prid	prId	4
63	clart	kl#t	5
64	scroise	skr4s	7
65	thweak	Twik	6
66	ghelse	gEls	6
67	psourth	s9T	7
68	trusque	trVsk	7
69	threp	TrEp	5
70	knoff	nQf	5
71	strarse	str#s	7
72	ploul	pl6l	5
73	nusque	nVsk	6
74	frooth	fruT	6
75	streik	str1k	6
76	cred	krEd	4
77	irp	3p	3
78	thrarge	Tr#_	7
79	tharp	T#p	5
80	veige	v1_	5
81	gwech	gwEJ	5
82	strub	strVb	5
83	qwipe	kw2p	5

84	thwuth	TwVT	6
85	harse	h#s	5
86	psud	sVd	4
87	husp	hVsp	4
88	smourt	sm9t	6
89	phloach	fl5J	7
90	phrulb	frVlb	6
91	dwimp	dwImp	5
92	nean	nin	4
93	sprouse	spr6s	7
94	nulp	nVlp	4
95	gawn	g9n	4
96	tealt	tilt	5
97	stount	st6nt	6
98	pseath	siT	6
99	thirque	T3k	7
100	trife	tr2f	5

ARC Nonword Database, search results for January 24, 2013, 19:43; shaded words contain stop onsets & offsets

APPENDIX C:

PHONEME SYMBOLS

The following list represents the phoneme symbols on which the pronunciation of the pseudo words in the audio recording of the diagnostic assessment was based. The symbols are an adaptation of the DISC phonetic character set used in the CELEX Lexical Database. This phonetic character set represents each distinct phoneme by a single symbol taken from the ASCII character set.

1	b**ay**		Q	p**o**t		l	**l**ad	
2	b**uy**		S	**sh**eep		m	**m**ad	
3	b**ur**n		T	**th**in		n	**n**at	
4	b**oy**		U	p**u**t		p	**p**at	
5	n**o**		V	p**u**tt		r	**r**at	
6	br**ow**		Z	mea**s**ure		s	**s**ap	
7	p**ee**r		b	**b**ad		t	**t**ack	
8	p**air**		d	**d**ad		u	b**oo**n	
9	p**oor**		f	**f**at		v	**v**at	
D	**th**en		g	**g**ame		w	**w**hy	
E	p**e**t		h	**h**ad		z	**z**ap	
I	p**i**t		i	b**ea**n		#	b**ar**n	
J	**ch**eap		j	**y**ank		{	p**a**t	
N	ba**ng**		k	**c**ad		_	**j**eep	

119

APPENDIX D:

PRONUNCIATION DIAGNOSTIC INSTRUMENT

Language Activity

In this activity, you will hear 15 words.

The words are not real words in English, but they sound like English words.

Just **listen carefully** and **repeat** what you hear.

Let's begin...

cliek

trisc

clipe

twasp

crek	dape
clab	treg
prid	clart

trusque

creb

qwipe

dwimp

tealt

The activity is finished.

Thanks!

APPENDIX E:

IEP PLACEMENT SCORE DISTRIBUTION

Center for Multilingual Multicultural Studies
Intensive English Program
Placement Test Score Distribution

LEVEL	START	END	SPREAD	Reading	Writing	Essay
1B	216	320	100+			
1A	320	370	50			
2B	370	400	30			
2A	400	430	30			
3B	430	450	20	ReadSk	autobio	
3A	450	480	30	Module	AcaWtg.	
4B	480	510	30	ReadSk	BasicEssay	
4A	510	575+	65+	module	module	

APPENDIX F:

LANGUAGE ACTIVITY SCRIPT

Language Activity Script

Check in

"We are going to the lab for a brief pronunciation activity and then you will return to this room. Please leave your personal things here. Your teacher will stay in the room the whole time, waiting for you to return. Please leave your cell phone here as well. We are going to be audio-recording your speech and do not want cell phone interruptions. Now, I'd like for you to please line up and stay in order as I call your name."

Assignment

"I will now assign each of you to a station. The number is on the bottom left side of the terminal. You will only need the headset is for this assignment, so please do not touch the other equipment. I will give you further instructions in a few minutes."

Activity

If you happen to have your cellphone with you, please shut it off now for the next few minutes...no vibrate/silent please...just off!

You are going to do a 2-minute pronunciation activity. You will hear invented words. They sound like real words in American English, but they're not real words; they are invented. All you need to do is listen carefully and repeat what you hear. Your voice is going to be audio-recorded. In about a week, you will receive an email with information about your pronunciation.

Everyone will be doing different tasks, so please don't get distracted by what others are doing, so make sure to focus on your activity. Plus there is going to be lots of talking because everyone is going to doing this activity together. Again please be careful to stay focused; ignore everyone else. Remember all you need to do is just listen and repeat.

After

"You may not return to your classroom. Results will be emailed to you in about one week. If you have any questions about this activity, please see me personally. Thanks."

LIST OF REFERENCES

Abuhamdia, Z. (1987). Neurobiological foundations for foreign language. *International Review of Applied Linguistics, 25*, 203-211.

Abutalebi, J., Cappa, S. F., & Perani, D. (2001). The bilingual brain as revealed by functional neuroimaging. *Bilingualism, 4*, 179–90.

Agres, K. R., & Krumhansl, C. L. (2008). Musical change deafness: The inability to detect change in a non-speech auditory domain. In *Proceedings of the 30th Annual Conference of the Cognitive Science Society* (pp. 975-980).

Anderson-Hsieh, J., Johnson, R., & Koehler, K. (1992). The difference between native speaker judgments of nonnative pronunciation and deviance in segmentals, prosody and syllable structure. *Language Learning*, *42*, 529–555.

Aoyama, K., Guion, S. G., Flege, J. E., Yamada, T., & Akahane-Yamada, R. (2008). The first years in an L2-speaking environment: A comparison of Japanese children and adults learning American English. *IRAL*, *46*(1), 61.

Asher, J. J., & Garcia, R. (1969). The optimal age to learn a foreign language. *Modern Language Journal, 53,* 334-41.

Atkinson, R. C., & Shiffrin, R. M. (1971). The control of short-term memory. *Scientific American, 225*(2), 82-90. doi:10.1038/scientificamerican0871-82

Bachman, L., Lynch, B., & Mason, M. (1995). Investigating variability in tasks and rater judgments in a performance test of foreign language speaking. *Language Testing, 12*, 238–58.

Baddeley, A. D. (1966). Short-term memory for word sequences as a function of acoustic, semantic and formal similarity. *Quarterly Journal of Experimental Psychology, 18*(4), 362-365. doi:10.1080/14640746608400055

Baddeley, A. D. (1983). Working memory. *Philosophical Transactions of the Royal Society London, B302*, 311.

Baddeley, A. D. (1986). *Working memory.* Oxford, English: Oxford University Press.

Baddeley, A. D. (2003). Working memory: Looking back and looking forward. *Nature Reviews Neuroscience*, *4*(10), 829-839. doi:10.1038/nrn1201

Baddeley, A. D., & Hitch, G. J. (1974). Working memory. In G. Bower (Ed.), *The psychology of learning and motivation: Advances in research and theory Vol. 8* (pp. 47–89). New York: Academic Press.

Baddeley, A. D., & Wilson, B. (1985). Phonological coding and short-term memory in patients without speech. *Journal of Memory and Language, 24*(4), 490-502. doi:10.1016/0749-596X(85)90041-5

Baddeley, A. D., Gathercole, S., & Papagno, C. (1998). The phonological loop as a language learning device. *Psychological Review, 105*(1), 158-173. doi:10.1037/0033-295X.105.1.158

Baddeley, A. D., Lewis, V., & Vallar, G. (1984). Exploring the articulatory loop. *Quarterly Journal of Experimental Psychology: A Human Experimental Psychology, 36A*(2), 233-252. doi:10.1080/14640748408402157

Baddeley, A. D., Thomson, N., & Buchanan, M. (1975). Word length and the structure of short-term memory. *Journal of Verbal Learning and Verbal Behavior, 14*(6), 575-589. doi:10.1016/S0022-5371(75)80045-4

Baetens-Beardsmore, H. (1979). The recognition and tolerance level of bilingual speech. *Working Papers on Bilingualism Toronto*, (19), 115-127.

Baker, C. (1992). *Attitudes and language*. New York: Language Multilingual Matters Ltd.

Baker, W., Trofimovich, P., Flege, J. E., Mack, M., & Halter, R. (2008). Child-adult differences in second-language phonological learning: The role of cross-language similarity. *Language and Speech*, *51*(4), 317-342.

Bent, T., & Bradlow, A. (2003). The interlanguage speech intelligibility benefit. *Journal of the Acoustical Society of America*, *114*, 1600–1610.

Bialystok, E., & Miller, B. (1999). The problem of age in second-language acquisition: Influences from language, structure, and task. *Bilingualism: Language and Cognition*, *2*(02), 127-145.

Birdsong, D. (1992). Ultimate attainment in second language acquisition. *Language*, *68*, 706-755.

Birdsong, D. (1999). Introduction: Whys and why nots of the critical period hypothesis. In D. Birdsong (Ed.), *Second language acquisition and the critical period hypothesis* (pp.1-22). Mahwah, NJ: Lawrence Erlbaum.

Bley-Vroman, R., & Chaudron, C. (1994). Elicited imitation as a measure of second- language competence. In E. Tarone, S. Gass, & A. Cohen. (Eds.), *Research methodology in second-language acquisition* (pp. 245-261). Hillsdale, NJ: Lawrence Erlbaum Associates.

Bloch, C., Kaiser, A., Kuenzli, E., Zappatore, D., Haller, S., Franceschini, R., & Nitsch, C. (2009). The age of second language acquisition determines the variability in activation elicited by narration in three languages in Broca's and Wernicke's area. *Neuropsychologia*, *47*(3), 625-633.

Bongaerts, T., Planken, B., & Schils, E. (1995). Can late starters attain a native accent in a foreign language? A test of the critical period hypothesis. *The age factor in second language acquisition*, 30-50.

Bongaerts, T., Van Summeren, C., Planken, B., & Schils, E. (1997). Age and ultimate attainment in the pronunciation of a foreign language. *Studies in Second Language Acquisition, 19*(04), 447-465.

Bonk, W. J., & Ockey, G. J. (2003). A many-facet Rasch analysis of the second language group oral discussion task. *Language Testing, 20*(1).

Boothroyd, A. (1986). *Speech acoustics and perception*. Austin TX: Pro-ed.

Bowden, H. W., Sanz, C. S., Steinhauer, K., & Ullman, M. T. (2007). An ERP study of proficiency in second language. *Journal of Cognitive Neuroscience*, supplement, 170.

Braisby, N., & Gellatly, A. (2005). Foundations of cognitive psychology. *Cognitive Psychology*, 1-32.

Breitkreutz, J., Derwing, T. M., & Rossiter, M. J. (2002). Pronunciation teaching practices in Canada. *TESL Canada Journal, 19*, 51-61.

Brennan, E. M., & Brennan, J. S. (1981). Accent scaling and language attitudes: Reactions to Mexican American English speech. *Language and Speech*, *24*(3), 207-221.

Brett, D. (2004). Computer generated feedback on vowel production by learners of English as a second language. *ReCALL, 16*(1), 103-113.

Briellmann, R. S., Saling, M. M., Connell, A. B., Waites, A. B., Abbott, D. F., & Jackson, G. D. (2004). A highfield functional MRI study of quadrilingual subjects. *Brain and Language*, *89*, 531–542.

Broadbent, D. E. (1952). Listening to one of two synchronous messages. *Journal of Experimental Psychology, 44, 51-55.*

Broadbent, D. E. (1954). The role of auditory localization in attention and memory span. *Journal of Experimental Psychology, 47*(3), 191-196. doi: 10.1037/h0054182

Brown, A. (1988). Functional load and the teaching of pronunciation. *TESOL Quarterly*, *22*(4), 593-606.

Brown, J. (1958). Some tests of the decay theory of immediate memory. *Quarterly Journal of Experimental Psychology, 10*, 12-21. doi:10.1080/17470215808416249

Buchner, A., Irmen, L., & Erdfelder, E., (1996). On the irrelevance of semantic information for the "Irrelevant Speech" effect. *Quarterly Journal of Experimental Psychology, 49*(3), 765-779.

Burgess, J., & Spencer, S. (2000). Phonology and pronunciation in integrated language teaching and teacher education. *System*, *28*(2), 191-215.

Campbell, D. T., & Stanley, J. C. (1963). *Experimental and quasi-experimental designs for research.* Boston, MA: Houghton Mifflin.

Campoy, G. (2011). Retroactive interference in short-term memory and the word-length effect. *Acta Psychologica, 138*(1), 135-142. doi:10.1016/j.actpsy.2011.05.016

Cardoso, W. (2011). The development of coda perception in second language phonology: A variationist perspective. *Second Language Research*, *27*(4), 433-465.

Carlesimo, G. A., Galloni, F., Bonanni, R., & Sabbadini, M. (2006). Verbal short-term memory in individuals with congenital articulatory disorders: New empirical data and review of the literature. *Journal of Intellectual Disability Research, 50*(2), 81-91. doi:10.1111/j.1365-2788.2005.00725.x

Catford, J. C. (1987). Phonetics and the teaching of pronunciation: A systemic description of English phonology. *Current perspectives on pronunciation: Practices anchored in theory*, 87-100.

Celce-Murcia, M., Brinton, D., & Goodwin, J. M. 1. (1996). *Teaching pronunciation: A reference for teachers of English to speakers of other languages*. New York: Cambridge University Press.

Chee, M. W., Tan, E. W., & Thiel, T. (1999). Mandarin and English single word processing studied with functional magnetic resonance imaging. *Journal of Neuroscience*, *19*(8), 3050-3056.

Cherry, E. C. (1953) Some experiments on the recognition of speech, with one and two ears. *Journal of the Acoustical Society of America, 25,* 975–979.

Chomsky, Noam (1965). *Aspects of the theory of syntax*. Cambridge, MA: MIT Press.

Coates, J. (1986). *Pronunciation and personality: A study of personal factors in German students' pronunciation of English.* Bochum, Germany: AKS-Verlag.

Cole, R. L., & Pickering, S. J. (2010). Phonological and visual similarity effects in Chinese and English language users: Implications for the use of cognitive resources in short-term

memory. *Bilingualism: Language and Cognition, 13*(4), 499-512.

doi:10.1017/S1366728909990587

Coltheart, M. (1978). Lexical access in simple reading tasks. In G. Underwood (Ed.), *Strategies*

of information processing (pp.112-174). New York: Academic Press. Retrieved from

http://ci.nii.ac.jp/naid/10009546232/en/

Coltheart, M., & Rastle, K. (1994). Serial processing in reading aloud: Evidence for dual-route

models of reading. *Journal of Experimental Psychology: Human Perception and*

Performance, 20(6), 1197-1211. doi:10.1037/0096-1523.20.6.1197

Coltheart, M., Curtis, B., Atkins, P., & Haller, M. (1993). Models of reading aloud: Dual-route

and parallel-distributed-processing approaches. *Psychological Review, 100*(4), 589-608.

doi:10.1037/0033-295X.100.4.589

Conrad, R. (1964). Acoustic confusions in immediate memory. *British Journal of Psychology,*

55(1), 75-84. doi:10.1111/j.2044-8295.1964.tb00899.x

Costa, A., Albareda, B., & Santesteban, M. (2008). Assessing the presence of lexical competition

across languages: Evidence from the Stroop task. *Bilingualism: Language and Cognition,*

11(1), 121-131. doi:10.1017/S1366728907003252

Cowan, N. (1988). Evolving conceptions of memory storage, selective attention, and their

mutual constraints within the human information-processing system. *Psychological*

Bulletin, 104(2), 163-191. doi:10.1037/0033-2909.104.2.163

Cucchiarini, C., Strik, H., & Boves, L. (1997, December). Automatic evaluation of Dutch

pronunciation by using speech recognition technology. In *Proceedings of the 1997 IEEE*

Workshop on Automatic Speech Recognition and Understanding (pp. 622-629).

Cunningham, J., Nicol, T., Zecker, S., & Kraus, N. (2000). Speech-evoked neurophysiologic

    responses in children with learning problems: development and behavioral correlates of

    perception. *Ear and Hearing*, *21*(6), 554-568.

Curtiss, S. (1977). Genie: A psycholinguistic study of a modern day "wild child". New York:

    Academic Press.

Curtiss, S. (1994). Language as a cognitive system: Its independent and selective vulnerability.

    In C. Otera (Ed.), *Noam Chomsky: Critical assessment* (pp. 211-255). London:

    Routledge.

De Bot, K. (1983). Visual feedback of intonation I: Effectiveness and induced practice behavior.

    Language and Speech, *26*(4), 331-350.

de Renzi, E., & Nichelli, P. (1975). Verbal and non-verbal short-term memory impairment

    following hemispheric damage. *Cortex: A Journal Devoted to the Study of the Nervous*

    System and Behavior, 11(4), 341-354. Retrieved from

    http://ezproxy.lib.ucf.edu/login?URL=http://search.ebscohost.com/login.aspx?direct=true

    &db=psyh&AN=1977-01476-001&site=ehost-live

Dehaene, S., Dupoux, E., Mehler, J., Cohen, L., Paulesu, E., Perani, D., Van de Moortele, P.

    Lehericy, S., & Le Bihan, D. (1997). Anatomical variability in the cortical representation

    of first and second language. *NeuroReport*, *8*(17), 3809-3815.

Derwing, T. M. (2003). What do ESL students say about their accents? *Canadian Modern*

    Language Review/La Revue canadienne des langues vivantes, *59*(4), 547-567.

Derwing, T. M., & Munro, M. J. (1997). Accent, intelligibility, and comprehensibility: Evidence

    from four L1s. *Studies in Second Language Acquisition*, *19*, 1–16.

Derwing, T. M., & Munro, M. J. (2005). Second language accent and pronunciation teaching: A research-based approach. *TESOL Quarterly*, *39*(3), 379-397.

Derwing, T. M., & Munro, M. J. (2009). Putting accent in its place: Rethinking obstacles to communication. *Language Teaching*, *42*(04), 476-490.

Derwing, T. M., & Rossiter, M. J. (2003). The effects of pronunciation instruction on the accuracy, fluency, and complexity of L2 accented speech. *Applied Language Learning*, *13*(1), 1-17.

Derwing, T. M., Munro, M. J., & Wiebe, G. (1998). Evidence in favor of a broad framework for pronunciation instruction. *Language Learning*, *48*(3), 393-410.

Doidge, N. (2007). *The brain that changes itself: Stories of personal triumph from the frontiers of brain science*. Penguin Books.

Ellis, N. C. (2006a). Cognitive perspectives on SLA: The Associative-Cognitive CREED. *AILA Review,* 19,100-121.

Ellis, R. (2008). *The study of second language acquisition* (2nd ed.). Oxford, England: Oxford University Press.

Farina, M. (2012). *The Effect of Input Modality on Output Length Accuracy*. Unpublished manuscript.

Felps, D., & Gutierrez-Osuna, R. (2010). Developing objective measures of foreign-accent conversion. *Audio, Speech, and Language Processing, IEEE Transactions on*, *18*(5), 1030-1040.

Fishman, J. A., Cooper, R. L., and Ma, R. (1971). *Bilingualism in the Barrio,* Indiana University
Research Center for the Language Sciences, Bloomington, Indiana (Language Science
Monograph Series, No. 7).

Flege, J. E. (1984). The detection of French accent by American listeners. *Journal of the
Acoustical Society of America*, *76*, 692.

Flege, J. E. (1991). Age of learning affects the authenticity of voice-onset time (VOT) in stop
consonants produced in a second language. *Journal of the Acoustical Society of America*,
89, 395.

Flege, J. E. (1995). Second language speech learning: Theory, findings, and problems. In W.
Strange (Ed.), *Speech perception and linguistic experience: Issues in cross-language
research* (pp. 233–277). Timonium, MD: York Press.

Flege, J. E. (2012). The role of input in second language (L2) speech learning. *In Proceedings of
the 6th International Conference on Native and Nonnative Accents of English*, Łódź
(Poland), 6-8 December 2012.

Flege, J. E., & Port, R. (1981). Cross-language phonetic interference: Arabic to English.
Language and Speech, *24*(2), 125-146.

Flege, J. E., Munro, M. J., & MacKay, I. R. A. (1995). Factors affecting strength of perceived
foreign accent in a second language. *Journal of the Acoustical Society of America, 97*(5),
3125-3134. doi:10.1121/1.413041

Frederiksen, J. R., & Kroll, J. F. (1976). Spelling and sound: Approaches to the internal lexicon.
Journal of Experimental Psychology: Human Perception and Performance, 2(3), 361-
379. doi:10.1037/0096-1523.2.3.361

Frenck-Mestre, C., Anton, J. L., Roth, M., Vaid, J., & Viallet, F. (2005). Articulation in early and
late bilinguals' two languages: Evidence from functional magnetic resonance imaging.
NeuroReport, *16*, 761–765.

Green, J. B., & Salkind, N. J. (2008). *Using SPSS for Windows and Macintosh: Analyzing and
understanding data* (*5th ed.*). Upper Saddle River, NJ: Prentice Hall.

Grimes, J. (1996). On the failure to detect changes in scenes across saccades. In: Akins K, (Ed.),
Perception: Vancouver studies in cognitive science, Vol. 2 (pp. 89–110). Oxford,
England: Oxford University Press.

Grimshaw, G., Adelstein, A., Bryden, P., & MacKinnon, G. (1998). First language acquisition in
adolescence: Evidence for a critical period for verbal language development. *Brain and
Language, 63*, 237-255.

Guiora, A. Z., Beit-Hallahmi, B., Brannon, R. C. L., Dull, C. Y., & Scovel, T. (1972). The
effects of experimentally induced changes in ego states on pronunciation ability in a
second language: An exploratory study. *Comprehensive Psychiatry, 13*(5), 421-428.

Hahn, L. D. (2004). Primary stress and intelligibility: Research to motivate the teaching of
suprasegmentals. *TESOL Quarterly*, *38*(2), 201-223.

Hedgcock, J. S., & Ferris, D. R. (2009). *Teaching readers of English*. New York: Routledge.

Hulme, C., Roodenrys, S., Schweickert, R., Brown, G. D. A., Martin, S., & Stuart, G. (1997).
Word-frequency effects on short-term memory tasks: Evidence for a redintegration
process in immediate serial recall. *Journal of Experimental Psychology: Learning,
Memory, and Cognition, 23*(5), 1217.

Hyltenstam, K. (1992). Nonnative features of near-native speakers: On the ultimate attainment of

    childhood L2 learners. *Advances in Psychology*, *83*, 351-368.

Iandoli, L. J. (1990). How can the language teacher help students to approximate native

    pronunciation, intonation, and body language? *Language Quarterly, 28*(1-2), 22-31.

Ioup, G. (2005). Age in second language development. In E. Hinkel (Ed.), *Handbook of research*

    in second language teaching and learning (pp. 419-435). New York: Lawrence Erlbaum.

James, E. F. (1976). The acquisition of prosodic features of speech using a speech visualizer.

    International Review of Applied Linguistics, 14(3), 227-243.

Jenkins, J. (2002). A sociolinguistically based, empirically researched pronunciation syllabus for

    English as an international language. *Applied Linguistics*, *23*, 83–103.

Jenkins, J. 1998. 'Which pronunciation norms and models for English as an international

    language?' *ELT Journal, 52*(2), 119–26.

Johnson, J., & Newport, E. (1989). Critical period effects in second language learning: The

    influence of maturational state on the acquisition of English as a second language.

    Cognitive Psychology, 21, 60-99.

Jones, D. M. (1999). The cognitive psychology of auditory distraction. *British Journal of*

    Psychology, 90, 167-187.

Jones, D. M., Saint-Aubin, J., & Tremblay, S. (1999). Modulation of the irrelevant sound effect

    by organizational factors: Further evidence from streaming by location. *Quarterly*

    Journal of Psychology, 52A, 545-554.

Kang, Rubin & Pickering. (2010). Suprasegmental measures of accentedness and judgments of language learner proficiency in oral English. *Modern Language Journal, 94*(4). doi:10.1111/j.1540-4781.2010.01091.x 0026-7902/10/554–566

Kim, K. H., Relkin, N. R., Lee, K. M., & Hirsch, J. (1997). Distinct cortical areas associated with native and second languages. *Nature*, *388*(6638), 171-174.

Kormos, J., & Denes, M. (2004). Exploring measures and perceptions of fluency in the speech of second language learners. *System*, *32*, 145–164.

Kovelman, I., Baler, S. A., & Petitto, L.A. (2008). Bilingual and monolingual brains compared: A functional magnetic resonance imaging investigation of syntactic processing and a possible neural signature of bilingualism. *Journal of Cognitive Neuroscience*, *20*, 153–160.

Krashen, S., Long, M., & Scarcella, R. (1982). Age, rate, and eventual attainment in second language acquisition. In S. Krashen, R. Scarcella, and M. Long (Eds.). Child-adult differences in second language acquisition (pp. 161-172). Rowley, MA: Newbury House.

Krejcie, R. V., & Morgan, D. W. (1970). Determining sample size for research activities. *Educational Psychology Measure, 30*(3), 607-610.

Labov, W. (1963). The social motivation of a sound change. *Word*, *19*(309), 1966.

Labov, W. (1970). *The study of language in its social context*. Springer, 180-202.

Lado, R. (1957). *Linguistics across cultures: Applied linguistics for language teachers.* Ann Arbor: University of Michigan Press.

Lenneberg, E. H. (1964). Speech as a motor skill with special reference to nonaphasic disorders. *Monographs of the Society for Research in Child Development*, *29*(1), 115-127.

Lenneberg, E. H. (1967). *Biological foundations of language*. New York: Wiley.

Lenneberg, E. H. (1975). *Foundations of language development: A multidisciplinary approach* (Vol. 2). New York: Academic Press.

Levin D. T., Momen, N., Drivdahl, S. B., & Simons, D. J. (2000). Change blindness: The metacognitive error of overestimating change-detection ability. *Visual Cognition, 7* 397-412.

Lindemann, S. (2003). Koreans, Chinese, or Indians? Attitudes and ideologies about nonnative English speakers in the United States. *Journal of Sociolinguistics*, *7*, 348–364.

Lippi-Green, R. (1997). *English with an accent: Language, ideology, and discrimination in the United States*, (2nd ed.). Oxon, England: Routledge.

Lomax, R. G. (2007). *An introduction to statistical concepts for education and behavioral sciences (2nd ed.)*. Mahwah, NJ: Lawrence Erlbaum.

Long, M. (2005). Problems with supposed counter-evidence to the critical period hypothesis. *International Review of Applied Linguistics in Language Teaching 43*(4). 287–317.

Lumley, T., & McNamara, T. F. (1995). Rater characteristics and rater bias: Implications for training. *Language Testing*, *12*(1), 54-71.

MacDonald, S. (2002). Pronunciation – views and practices of reluctant teachers. *Prospect 17*(3), 3–18.

Major, R. C. (2007). Identifying a foreign accent in an unfamiliar language. *Studies in Second Language Acquisition, 29*(4), 539-556.

Martin, R. C. (1993). Short-term memory and sentence processing: Evidence from neuropsychology. *Memory and Cognition, 21*(2), 176-183. doi:10.3758/BF03202730

Mechelli, A., Crinion, J. T., Noppeney, U., O'Doherty, J., Ashburner, J., Frackowiak, R. S., & Price, C. J. (2004). Neurolinguistics: structural plasticity in the bilingual brain. *Nature*, *431*(7010), 757-757.

Mechelli, A., Friston, K. J., & Price, C. J. (2000). The effects of presentation rate during word and pseudoword reading: a comparison of PET and fMRI. *Journal of Cognitive Neuroscience, 12*(Supplement 2), 145-156.

Miller, L. M., & Roodenrys, S. S. (2009). The interaction of word frequency and concreteness in immediate serial recall. *Memory and Cognition, 37*(6), 850-865. doi:10.3758/MC.37.6.850

Moyer, A. (1999). Ultimate attainment in L2 phonology: The critical factors of age, motivation, and instruction. *Studies in Second Language Acquisition*, *21*(1), 81-108. Retrieved from http://ezproxy.lib.ucf.edu/login?URL=http://search.ebscohost.com/login.aspx?direct=true &db=eric&AN=EJ655702&site=ehost-live

Moyer, A. (2004). *Age, Accent, and Experience in Second Language Acquisition: An Integrated Approach to Critical Period Inquiry*. Clevedon, UK: Multilingual Matters Ltd.

Munro, M. J., & Derwing, T. M. (1995). Foreign accent, comprehensibility, and intelligibility in the speech of second language learners. *Language Learning*, *45,* 73–97.

Munro, M. J., & Derwing, T. M. (2008). Segmental acquisition in adult ESL learners: A longitudinal study of vowel production. *Language Learning*, *58,* 479-502.

Munro, M. J., Derwing, T. M., & Burgess, C. S. (2003, August). The detection of foreign accent in backwards speech. In *Proceedings of the 15th International Congress of Phonetic Sciences* (pp. 535-538).

Murdock, B. B. J. (1967). Recent developments in short-term memory. *British Journal of Psychology, 58*(3-4), 421-433. doi:10.1111/j.2044-8295.1967.tb01099.x

Murphy, J. M. (1991). Oral communication in TESOL integrating speaking, listening, and pronunciation. *TESOL Quarterly*, *25*(1), 51-75.

Naish, P. (2005). Attention. In Braisby, N., & Angus, G. (Eds.), *Cognitive psychology* (pp. 37-70). Oxford, England: Oxford University Press.

Nation, I. S. P., & Webb, S. A. (2011). *Researching and analyzing vocabulary*. Heinle, Cengage Learning.

Neufeld, G. G. (1977). Language learning ability in adults: A study on the acquisition of prosody and articulatory features. *Working Papers on Bilingualism, 12,* 47-60.

Neufeld, G., & Schneiderman, E. (1980). Prosodic and articulatory features in adult language learning. *Research in Second Language Acquisition. Rowley, MA: Newbury House*, 105-109.

Neumeyer, L., Franco, H., Digalakis, V., & Weintraub, M. (2000). Automatic scoring of pronunciation quality. *Speech Communication*, *30*(2), 83-93.

Nieuwland, M. S., & Van Berkum, J. J. (2005). Testing the limits of the semantic illusion phenomenon: ERPs reveal temporary semantic change deafness in discourse comprehension. *Cognitive Brain Research, 24*(3), 691-701.

Nimmo, L. M., & Roodenrys, S. (2002). Syllable frequency effects on phonological short-term memory tasks. *Applied Psycholinguistics, 23*(4), 643-659. doi:10.1017/S0142716402004071

Nygaard, L. C., & Pisoni, D. B. (1998). Talker-specific learning in speech perception. *Perception & Psychophysics, 60,* 355-376.

Olson, L. L., & Samuels, S. J. (1973). The relationship between age and accuracy of foreign language pronunciation. *Journal of Educational Research,* 263-268.

Ong, W. J. (2002). *Orality and literacy: The technologizing of the word*. New York: Methuen.

Oyama, S. (1976). A sensitive period for the acquisition of a nonnative phonological system. *Journal of Psycholinguistic Research, 5*(3), 261-283.

Palladino, P., & Ferrari, M. (2008). Phonological sensitivity and memory in children with a foreign language learning difficulty. *Memory, 16*(6), 604-625. doi:10.1080/09658210802083072

Patkowski, M. (1980). The sensitive period to the acquisition of syntax in a second language. *Language Learning, 30,* 449-472.

Patkowski, M. (1990). Age and accent in a second language: A reply to James Emil Flege. *Applied Linguistics, 11,* 73-89.

Pennington, M. C. (1996). *Phonology in English language teaching*. New York: Longman.

Pennington, M. C., & Richards, J. C. (1986). Pronunciation revisited. *TESOL Quarterly, 20*(2), 207-25. Retrieved from http://ezproxy.lib.ucf.edu/login?URL=http://search.ebscohost.com/login.aspx?direct=true &db=eric&AN=EJ338937&site=ehost-live

Perani, D., Paulesu, E., Galles, N. S., Dupoux, E., Dehaene, S., Bettinardi, V., Cappa, S. F., Fazio, F., & Mehler, J. (1998). The bilingual brain: proficiency and age of acquisition of the second language. *Brain*, *121*(10), 1841-1852.

Pica, T. (2005). Second language acquisition research and applied linguistics. In E. Hinkel (Ed.), *Handbook of research in second language teaching and learning* (pp. 263-280). New York: Lawrence Erlbaum.

Pickering, L. (2004). The structure and function of intonational paragraphs in native and nonnative speaker instructional discourse. *English for Specific Purposes*, *23*, 19–43.

Pike, G., & Edgar, G. (2005). Perception. In Braisby, N., & Angus, G. (Eds.), *Cognitive psychology* (pp.71-112). Oxford, England: Oxford University Press.

Piller, I. (2002). Passing for a native speaker: Identity and success in second language learning. *Journal of Sociolinguistics*, *6*(2), 179-208.

Piske, T., MacKay, I. R., & Flege, J. E. (2001). Factors affecting degree of foreign accent in an L2: A review. *Journal of phonetics, 29*(2), 191-215.

Porter, D., & Garvin, S. (1989). Attitudes to pronunciation in EFL. *Speak Out*, *5*, 8-15.

Potter, R. K., Kopp, G. A., & Green, H. C. (1946). Technical aspects of visible speech. *Journal of the Acoustical Society of America*, *17*, 1-89.

Premack, D., & Schwartz, A. (1966). Preparations for discussing behaviorism with chimpanzee. In F. S. Smith & G. A. Miller (Eds.), *The genesis of language* (pp. 295-347). Cambridge, MA: MIT Press.

Rajadurai, J. (2007). Intelligibility studies: A consideration of empirical and ideological issues. *World Englishes*, *26*, 87–98.

Rastle, K., & Coltheart, M. (1998). Whammies and double whammies: The effect of length on nonword reading. *Psychonomic Bulletin and Review, 5*(2), 277-282. doi:10.3758/BF03212951

Rastle, K., Harrington, J., & Coltheart, M. (2002). 358,534 nonwords: The ARC nonword database. *Quarterly Journal of Experimental Psychology A: Human Experimental Psychology, 55A*(4), 1339-1362. doi:10.1080/02724980244000099

Rensink, R. A., O'Regan, J. K., & Clark, J. J. (1997). To see or not to see: The need for attention to perceive changes in scenes. *Psychological Science, 8,* 368–373.

Rey, A., Jacobs, A. M., Schmidt-Weigand, F., & Ziegler, J. C. (1998). A phoneme effect in visual word recognition. *Cognition, 68*(3), B71-B80. doi:10.1016/S0010-0277(98)00051-1

Rost, M. (2005). L2 Listening. In E. Hinkel (Ed.), *Handbook of research in second language teaching and learning* (pp. 503-527). New York: Lawrence Erlbaum.

Ryalls, J. H. (1996). *A basic introduction to speech perception*. San Diego, CA: Singular Publishing.

Saito, K. (2011). Examining the role of explicit phonetic instruction in nativelike and comprehensible pronunciation development: an instructed SLA approach to L2 phonology. *Language Awareness*, *20*(1), 45-59.

Schmidt, G. L., & Roberts, T. P. (2009). Second language research using magnetoencephalography: A review. *Second Language Research*, *25*(1), 135-166.

Schuetze-Coburn, S., Shapley, M., & Weber, E. G. (1991). Units of intonation in discourse: A comparison of acoustic and auditory analyses. *Language and Speech*, *34*, 207–234.

Schumann, J. H. (1978). *The pidginization process: A model for second language acquisition*. Rowley, MA: Newbury House.

Schumann, J. H. (2001). Appraisal psychology, neurobiology, and language. *Annual Review of Applied Linguistics*, *21*, 23-42.

Schweppe, J., Grice, M., & Rummer, R. (2011). What models of verbal working memory can learn from phonological theory: Decomposing the phonological similarity effect. *Journal of Memory and Language, 64*(3), 256-269. doi:10.1016/j.jml.2010.11.006

Scovel, T. (1969). Foreign accents, language acquisition, and cerebral dominance. *Language learning*, *19*(3-4), 245-253.

Scovel, T. (1977). The Ontogeny of the Ability to Recognize Foreign Accents. In *Proceedings of the Second Language Research Forum* (pp. 38-47).

Scovel, T. (1981). The recognition of foreign accents in English and its implications for psycholinguistic theories of language acquisition. In *Proceedings of the 5th Congress of AILA* (pp. 389-401).

Scovel, T. (1988). *A time to speak: A psycholinguistic inquiry into the critical period for human speech.* Rowley, MA: Newbury House.

Selinger, H. W. (1978). Implications of a multiple critical period hypothesis for second language learning. In W. Ritchie (Ed.), *Second language research: Issues and implications* (pp. 11-19). New York: Academic Press.

Shallice, T., & Warrington, E. K. (1970). Independent functioning of verbal memory stores: A neuropsychological study. *Quarterly Journal of Experimental Psychology, 22*(2), 261-273. doi:10.1080/00335557043000203

Sharma, A., Kraus, N., J McGee, T., & Nicol, T. G. (1997). Developmental changes in P1 and N1 central auditory responses elicited by consonant-vowel syllables.

Electroencephalography and Clinical Neurophysiology/Evoked Potentials Section, *104*(6), 540-545.

Shavelson, R. J., & Towne, L. (2002). *Scientific research in education*. Washington, DC: National Academies Press.

Simon, E., & D'Hulster, T. (2012). The effect of experience on the acquisition of a nonnative vowel contrast. *Language Sciences 34*(3), 269-283.

Simons, D. J., & Levin, D. T. (1998). Failure to detect changes to people during a real-world interaction. *Psychonomic Bulletin and Review, 5,* 644–649.

Smith, L., & Nelson, C. (1985). International intelligibility of English: Directions and resources. *World Englishes*, *4*, 333–342.

Snow, C. E., & Hoefnagel-Höhle, M. (1982). School-age second language learners' access to simplified linguistic input. *Language Learning*, *32*(2), 411-430.

Sorace, A. (1996). The use of acceptability judgments in second language acquisition research. In W. C. Ritchie & T. K. Bhatia (Eds.), *Handbook of second language acquisition* (pp. 375-409). New York: Academic Press.

Sperling, G. (1960). The information available in brief visual presentations. *Psychological Monographs, 74* (Whole Number 498), 29.

Steinhauer, K., White, E. J., & Drury, J. E. (2009). Temporal dynamics of late second language acquisition: Evidence from event-related brain potentials. *Second Language Research 25*(1), 13-41.

Steinhauer, K., White, E. J., Cornell, S., Genesee, F., & White, L. (2006). The neural dynamics of second language acquisition: evidence from Event-Related Potentials. *Journal of Cognitive Neuroscience*, supplement, 99.

Stevick, E. (1976) *Memory, meaning, and method.* Rowley, MA: Newbury House.

Stroop, J. R. (1935). Studies of interference in serial verbal reactions. *Journal of Experimental Psychology, 12*, 643-662.

Timmis, I. (2002). Native-speaker norms and International English: a classroom view. *ELT journal*, *56*(3), 240-249.

Tinkham, T. (1993). The effects of semantic clustering on the learning of second language vocabulary. *System, 21*(3), 371-380.

Tinkham, T. (1997). The effects of semantic and thematic clustering on the learning of second language vocabulary. *Second Language Research, 13*(2), 138-163.

Treisman, A. (1960). Contextual clues in selective listening. *Quarterly Journal of Experimental Psychology, 12,* 242-8.

Trojano, L., & Grossi, D. (1995). Phonological and lexical coding in verbal short-term memory and learning. *Brain and Language, 51*(2), 336-354. doi:10.1006/brln.1995.1064

Vallar, G., & Papagno, C. (1995). Neuropsychological impairments of short-term memory. In A. D. Baddeley, B. A. Wilson & F. N. Watts (Eds.), *Handbook of memory disorders* (pp. 135-165). Oxford, England: John Wiley & Sons. Retrieved from http://ezproxy.lib.ucf.edu/login?URL=http://search.ebscohost.com/login.aspx?direct=true&db=psyh&AN=1995-97743-006&site=ehost-live

van Els, T., & de Bot, K. (1987). The role of intonation in foreign accent. *Modern Language Journal, 71*(2), 147-155.

Vicari, S., Carlesimo, G., Brizzolara, D., & Pezzini, G. (1996). Short-term memory in children with Williams syndrome: A reduced contribution of lexical-semantic knowledge to word span. *Neuropsychologia, 34*(9), 919-925. doi:10.1016/0028-3932(96)00007-3

Vingerhoets, G., Borsel, J. V., Tesink, C., van den Noort, M., Deblaere, K., Seurinck, R., Vandemaele, P., & Achten, E. (2003). Multilingualism: An fMRI study. *Neuroimage*, *20*, 2181–2196.

Vitevich, M. S. (2003). Change deafness: The inability to detect changes between two voices. *Journal of Experimental Psychology: Human Perception and Performance, 29*(2), 333-342. doi: 10.1037/0096-1523.29.2.333

Walsh, T. M., & Diller, K. C. (1979). Neurolinguistic considerations on the optimum age for second language learning. In *Proceedings of the Annual Meeting of the Berkeley Linguistics Society* (Vol. 5).

Wang, X., & Munro, M. J. (2004). Computer-based training for learning English vowel contrasts. *System*, *32*(4), 539-552.

Waring, R. (1997). The negative effects of learning words in semantic sets. *System, 25*(2), 261-274.

Wartenburger, I., Heekeren, H. R., Abutalebi, J., Cappa, S. F., Villringer, A., & Perani, D. (2003). Early setting of grammatical processing in the bilingual brain. *Neuron, 37*, 159-170.

Weekes, B. S. (1997). Differential effects of number of letters on word and nonword naming latency. *Quarterly Journal of Experimental Psychology A: Human Experimental Psychology, 50A*(2), 439-456. doi:10.1080/027249897392170

Weigle, S. C. (1998). Using FACETS to model rater training effects. *Language Testing*, *15*(2), 263-287.

Wennerstrom, A. (1998). Intonation as cohesion in academic discourse: A study of Chinese speakers of English. Studies in Second Language Acquisition, 42, 1–13.

White, E. J., Genesee, F., & Steinhauer, K. (2012). Brain responses before and after intensive second language learning: Proficiency-based changes and first language background effects in adult learners. *PloS one, 7*(12), e52318.

Young, M., & Gardner, R. C. (1990). Modes of acculturation and second language proficiency. *Canadian Journal of Behavioural Science/Revue Canadienne Des Sciences Du Comportement, 22*(1), 59-71.

Zielinski, B. W. (2008). The listener: No longer the silent partner in reduced intelligibility. *System*, *36*(1), 69-84.