Lexical Access as a

Predictor of Oral Fluency

by

Aaron S. Carriere

A Thesis Presented in Partial Fulfillment
of the Requirements for the Degree
Master of Arts

Approved April 2017 by the
Graduate Supervisory Committee:

Álvaro Cerrón-Palomino, Chair
Michael Gradoville
Sara Beaudrie

ARIZONA STATE UNIVERSITY

May 2017

ABSTRACT

The present study investigates the role lexical access plays in the oral fluency of intermediate second language (L2) learners. In order to do this, I utilized a picture-naming task (PNT) in the target language to assess lexical access and generated spontaneous L2 speech through two narration tasks to assess oral fluency. The response times from the PNT were correlated with the two fluency measures analyzed from the narration tasks, the frequency of filled pauses and the overall rate of speech. The results revealed that intermediate learners with faster PNT response times used fewer filled pauses in spontaneous L2 speech but did not reveal a significant relationship between intermediate learners' PNT response times and their rate of speech.

DEDICATION

For my mother and father.

ACKNOWLEDGMENTS

Non quia difficilia sunt non audemus, sed quia non audemus difficilia sunt.

TABLE OF CONTENTS

CHAPTER Page

APPENDIX

LIST OF TABLES

LIST OF FIGURES

CHAPTER 1

INTRODUCTION

There are numerous elements involved in the production of speech. The most frequently investigated of the theoretical elements is working memory with a large body of research examining its role in speech production (Daneman, 1991; O'Brien, Segalowitz, Freed, & Collentine, 2007; Mota, 2003; Weissheimer & Mota, 2009, 2011; Wen, 2012), yet studies of the relationship among other elements involved in speech production, such as lexical access, have been limited to simultaneous bilinguals (Bialystok, Craik, & Luk, 2008; Ivanova & Costa, 2007) or as a means to explain speech disfluencies (Hartsuiker & Notebaert, 2010). These studies have not investigated lexical access as a potential predictor of individual differences in fluent speech production as it relates to L2 learners. This thesis studied the relationship between the lexical access ability and the fluency of the spontaneous speech of collegiate L2 learners in order to contribute to SLA pedagogy, explain individual differences in, and add to the overall understanding of, the complex systems involved in speech production.

CHAPTER 2

REVIEW OF LITERATURE

Theoretical Framework

According to Levelt's (2001) production model, lexical access has two major components: (a) a system for *lexical selection* and (b) a system for *form encoding*. These two systems combine to provide the language that is spoken during speech production. *Lexical selection* is the process by which words are selected out of the mental lexicon, which has been estimated to contain between 50,000-100,000 words in a normal literate adult person (Levelt, 2001). It is quite normal for a person to produce two to four words per second in normal speech, and the average speaker errs, on average, no more than once or twice every 1,000 words in L1 speech (Levelt, 2001). Nevertheless, prior to being spoken, the selected word's articulatory shape must be prepared and this is known as *form encoding*. This process is sensitive to context as a word's syllabification is not stored in the mental lexicon but created simultaneously, dependent on the current context (Levelt, 2001). The author uses the example of *horse* vs. *horses* to demonstrate how syllabification is context dependent. The singular *horse* is monosyllabic yet the plural *horses* has multiple syllables and would therefore be encoded for pronunciation differently than *horse*. Another core assumption of this theory is the existence of a mental syllabary, a repository of highly practiced syllabic gestures (Levelt, 2001). Once syllabification takes place, the corresponding syllabic patterns are selected from the syllabary for execution. This string of syllabic gestural routines that correspond to the phonological word are called its articulatory score. The auditory production of this *articulatory score*, or spoken word, is the final step in lexical access and results in the creation of overt speech (Levelt, 2001).

The steps described in lexical access occur in sequence. This sequence, and the speed at which it takes place, gives rise to *reaction time* data by speakers as measured in the laboratory (Levelt, 2001). Reaction time is the time that passes from the moment in which a participant is shown an image to when they utter a response to that image, a name, a categorical term, a description, etc. Reaction time also influences spontaneous speech production. During spontaneous speech a person is selecting and encoding words but the stimulus might come from semantic or syntactic factors instead of being presented visually. As utterances are produced, the speaker has to react to the changing conditions of their speech and continually select and encode new words. In this context, reaction time is occurring in parallel with articulation as speakers are selecting and encoding words simultaneously. The entire utterance does not need to be selected and encoded prior to articulation for a person to begin speaking. This is the very nature of spontaneous speech.

The way that spontaneous speech is produced is what is known as *fluency*. The first step to take is to differentiate fluency from overall language proficiency (Chambers, 1997), so that fluency can be compartmentalized as an individual, measurable skill. Fluent speech seems to have two main characteristics: *speed* and *effortlessness*. While speed, or the amount of speech in a given time, can be measured, terms such as smoothness, ease, or effortlessness are qualitative judgments (Chambers, 1997). Chambers (1997) identifies three quantitative variables that can be used to measure fluency, speech rate, frequency of pauses, and the location of pauses. Wood (2001) agrees with these proposed factors as he feels it is necessary to standardize the metrics used to measure fluency. This study will focus on two of these quantitative aspects of fluency, specifically, rate of speech and frequency of filled pauses as they appear in the spontaneous speech of the target language of second language learners.

Key Concepts

Fluency is a term that can have different meanings in different contexts. In a casual conversation amongst non-academics it can be used as an all-encompassing term to describe a person's competence with a language. When fluency is used by linguists, it refers to specific aspects of speech production, such as rate of speech, length of pauses, and duration of pauses. It is important to define the use of the term *disfluencies* in this context. One component that contributes to the quantity of words uttered in a given amount of time, rate of speech, is pause phenomena, both the amount of pauses and their subsequent length. There are other components that play a role in rate of speech besides pause phenomena and they are repetitions and self-corrections (Hartsuiker & Notebaert, 2010). A pause can be filled when the speaker utters words or sounds such as *um* or *uh* and a repetition can consist of a single word or an entire phrase that a speaker repeats. Self-corrections can occur when a speaker notices an error in pronunciation, conjugation, syntax, etcetera, during an utterance. Repetitions can be viewed as a type of self-correction.

There are multiple measurable aspects of produced speech, which can be analyzed from both quantitative and qualitative perspectives: therefore researchers must identify which feature, or features, they will utilize.

Chambers (1997) and Wood (2001) discuss the different areas of speech production that can be associated with fluency. In their work, they identify *rate of speech* as either the number of words uttered per second or the number of syllables uttered per second. The authors also mention *frequency and duration of pauses* during speech and *location of pauses within utterances* as measurable factors of fluency. These features are all related to the final measurable component of spontaneous speech. This production does not happen on its own; on the contrary, there are systems in place that are responsible for planning and organizing the lexical content of the utterances. One of these systems is called *lexical access*. Levelt's (2001) production model identifies two steps involved in lexical access. The first step is *lexical selection*, where the item is selected from the mental lexicon: the second step is *form encoding*, where the selected item's articulatory shape is prepared. The steps work together to choose appropriate items and modify, pluralize, conjugate, etc. them for syntactic purposes. Based on these definitions, lexical access' role in speech production can be seen as the system 'behind the scenes' that is responsible for supplying the vocal apparatus with content to produce. Fluency, when viewed as a metric of speech production, can then be seen as dependent upon this content. After all, one cannot measure produced speech if nothing is uttered.

It is also necessary to explain the use of the term *intermediate* in this study when it is used to describe the participants as language learners. The proficiency guidelines from the American Council on the Teaching of Foreign Languages (ACTFL) describe three levels of intermediate proficiency, high, mid, and low (Breiner-Sanders, Lowe, Miles, & Swender, 1999). Among other proficiency levels described in these guidelines, ranging from Advanced high to Novice low, these labels can be applied to any language learner who has had their speaking ability in the target language properly evaluated. The participants of this study were not evaluated by the ACTFL proficiency guidelines and therefore they are not being categorized as such. In this study, the use of the term *intermediate* refers to the level of Spanish class they were attending at the time of their participation in this investigation.

Previous Research

The body of research that focuses on lexical access and fluency, speech production, has utilized lexical access to explain speech disfluency and has also investigated the role of lexical access in bilingual speech production. Hartsuiker & Notebaert (2010) investigated how problems with lexical access, in this case meaning-based lexical retrieval and form-based lexical retrieval, can lead to disfluencies in speech. Utilizing their definition of disfluencies and given the high frequency with which disfluencies occur, approximately six times per 100 words, they were surprised that psycholinguistic models of language production did not account for them. Due to this lack of information in the literature, they asked if disfluencies could be localized at processing levels in cognitive models of speech production. In order to be precise, they asked if difficulties at a specific processing level, lexical access, would lead to an increased rate in disfluencies. Acknowledging that there are many potential independent causes for disfluencies, such as intentional disfluencies used for pragmatic purposes, they utilized a picture-naming experiment. In their experiment, the authors manipulated the degree of difficulty of lexical access (Hartsuiker & Notebaert, 2010) to isolate it as the only potential cause for disfluencies under those circumstances. In order to do this, they manipulated name agreement because this has been shown to cause difficulty at the initial stage of lexical access. Name agreement refers to the number of possible responses that speakers can access to refer to an object. A picture of a ship can be described as a boat, schooner, barge, or yacht, and this large number of possible labels equates to the picture having a low name agreement. Meanwhile, a picture of a clown will most likely be referred to as a clown and this element of specificity gives this image a high name agreement (Hartsuiker & Notebaert, 2010). Name agreement has been shown to affect naming latencies as pictures with high name agreement have faster naming latencies and pictures with low name agreement correspond to slower naming latencies.

In their study, the twenty participants, fifteen men and five women, were all native speakers of Dutch with a mean age of 21.5 years. The picture-naming task was a network of eight images, four low name agreement and four high name agreement, connected by lines that the participants described while a computerized dot moved through the network from image to image via the connecting lines. The pace of the dot was determined by a pilot study and the participants were told to use complete sentences while they described the network (Hartsuiker & Notebaert, 2010). The network descriptions were transcribed by two research assistants and coded for disfluencies. The disfluencies were collapsed into three categories, self-corrections, pauses, and repetitions.

The findings that the images with low name agreement produced more disfluencies than the images with high name agreement are compatible with the claim that the pauses reflect an act of choice between lexical items with similar semantic features (Hartsuiker & Notebaert, 2010). Crucially, the authors were able to isolate the lexical selection phase of speech production through the use of images with low name agreement to demonstrate that difficulties in the initial stage of lexical access can result in patterns of disfluency (Hartsuiker & Notebaert, 2010). These results can be applied to future studies that attempt to use lexical access as a predictor of oral fluency. While demonstrating lexical access' role in disfluencies, this study shows the impact of lexical selection in spontaneous speech production. In doing so, these results open the door for more research into the role that lexical access may have in the other elements involved in speech production, namely oral fluency as measured by rate of speech, and frequency of pauses.

An examination of the role of lexical access in the speech production of bilinguals investigated if bilinguals suffered a linguistic disadvantage when compared to monolinguals. Ivanova & Costa (2007) studied this disadvantage by comparing the lexical access of monolinguals to that of bilinguals in both their dominant and non-dominant languages. In doing so, the authors demonstrated that bilinguals show a naming disadvantage in their dominant language in comparison to monolingual speakers.

Their investigation was based on recent evidence of the cognitive advantages associated with speaking two languages, which researchers generally agree stem from the need to maintain two separate representational systems and use each one appropriately (Ivanova & Costa, 2007). They took this conclusion one step further and asked if the maintenance of these two systems would cause a processing disadvantage for bilinguals during speech production. Specifically, they asked if bilingualism affects lexical access, the ease with which representations are retrieved from the lexicon, during speech. Previous research into this bilingual disadvantage has shown that bilinguals were slower and less accurate in a picture-naming task than monolinguals but that this disadvantage disappeared after several repetitions of the same stimuli. There have been several explanations offered to describe why this disadvantage existed and then disappeared. Viewing it as a factor of word frequency would show that bilinguals use their dominant language less than monolinguals and this may affect availability of lexical items. Cross-language interference proposes that multiple representations from both linguistic systems get activated and thus increase naming latency as lexical selection has more competing items to deal with (Ivanova & Costa, 2007). However, the participants in these studies must be taken into account when discussing/considering the conclusions of the study. Previous studies of bilingual disadvantage have used switch-dominance bilinguals, who are usually characterized by the fact that their dominant language is not the language that they first acquired. Ivanova & Costa (2007) argue that broad conclusions about bilinguals having a disadvantage cannot be generalized until the appropriate group of bilinguals is tested. Therefore, they studied whether lexical access in the first and dominant language of bilingual speakers is less efficient than that of monolingual speakers, as this would allow them to explore the existence of a genuine bilingual disadvantage in the lexical access of bilinguals during speech production (Ivanova & Costa, 2007).

For their study, Ivanova & Costa (2007) tested one hundred and eleven participants separated into three groups of thirty-seven participants each: monolingual Spanish speakers, Spanish dominant bilingual speakers, and Catalan dominant bilingual speakers. The students were all university students at two different universities in Spain, which allowed the researchers to control for age and education level. The picture-naming task consisted of fifty total images taken from the International Picture-naming Project database, twenty five images with high-frequency names, and twenty five images with low-frequency names. The pictures were shown in blocks of five images with the restriction that no images that started with the same phoneme could be presented consecutively. The participants were instructed to name the images as quickly and as accurately as possible. Four categories of errors surfaced during the data elicitation phase, stuttering, utterance repairs, missing responses, or wrong responses (Ivanova & Costa, 2007).

The monolinguals had faster naming latencies than the bilinguals using their first and dominant language in both the low-frequency and high-frequency words. However, the monolingual advantage was greater for the low-frequency words than for the high-frequency words. When the responses of the bilingual participants, in their dominant and non-dominant languages, were compared, the results show that the bilinguals naming latencies were faster in their dominant language than the naming latencies in their second and non-dominant language (Ivanova & Costa, 2007). These latencies did not disappear with repetition, which had been observed in previous research with switched dominance bilinguals. The authors conclude that the bilingual disadvantage demonstrated in their study can be attributed to a frequency effect in disguise. Ivanova & Costa (2007) state that this effect is due to the fact that bilinguals need more time to retrieve lexical items from their dominant language lexicon because they use these words less often than monolingual speakers. This study adds to the research on lexical access and how it affects the speech production of monolingual and bilingual speakers. By demonstrating that monolingual speakers are able to retrieve lexical representations and articulate them faster than bilingual speakers in their dominant language, this study supports the claim that bilingualism has an effect on lexical access. This study makes this claim by comparing response times among speakers and by comparing response times in the multiple languages of the same speaker. Nevertheless, the authors fail to hypothesize how these findings can be applied to future research on other factors related to lexical access such as, fluent speech production of sustained speech, and not just picture-naming tasks. This study demonstrates how a person's lexical access can be isolated with a picture-naming task, yet it does not attempt to investigate the other areas of speech production that are influenced by lexical access.

Bialystok, Craik, & Luk (2008) examined monolingual and bilingual speakers to further investigate how bilingualism affects lexical access. The authors report the results of two studies that explore speech production differences between monolingual and bilingual speakers and use these results to draw conclusions on the effects of vocabulary size and executive control on lexical access. In the first study, monolinguals performed better than bilinguals on tests of naming and letter fluency, where they were asked to list as many words as possible that begin with the letters F,A, and S, but not on category fluency, where the students were asked to list as many words as possible that belonged to a certain category, animals in this case. In the second study, the letter fluency task had a more restrictive structure to increase executive processing involvement and the bilinguals with matched vocabulary scores outperformed monolinguals under these conditions (Bialystok et al., 2008).

Prior research has shown differences in aspects of linguistic processing between monolingual and bilingual speakers (Bialystok et al., 2008), yet none of these studies has attempted to investigate vocabulary size and executive control as potential predictors of these differences. The differences in naming latencies between monolinguals and bilinguals have been documented but, again, none of these studies has investigated the speakers' vocabulary size as a potential intervening variable contributing to the latency differences. Also, neuroimaging data has substantiated some of the behavioral differences in lexical access between monolinguals and bilinguals. Bilinguals were shown to engage two frontal areas while performing a picture-naming task, yet these areas were not engaged in the monolinguals performing the same task. This difference arises because a simple naming task for a bilingual involves conflict between the two language systems. This conflict carries costs in both time and accuracy, and the frontal cortex responsible for executive functioning is recruited to solve this conflict (Bialystok et al., 2008). Their study investigated both of these factors with the assumption that, when compared to monolinguals, the bilinguals would experience a deficit in the automaticity of word retrieval due to overall vocabulary size between the multiple language systems while experiencing an advantage in their level of executive control (Bialystok et al., 2008).

In their first study, Bialystok et al. (2008) included forty eight participants, half of which were monolingual speakers and the other half bilingual. All participants were students at the same university who received their education in English: the bilingual students spoke another language at home. The additional languages of the bilingual speakers were Polish, French, Spanish, Hebrew, Portuguese, Mandarin, Tamil, Cantonese, Croatian, Korean, Persian, Punjabi, and Ukrainian. The participants were all administered a spatial span task. The span task was administered forward, where the participants repeated a sequence in the order it was demonstrated, to establish a measure their memory span, and backward, where the participants repeated the sequence in the reverse order that it was demonstrated, to establish a measure of their working memory. This span task was used to equate the participants on a non-verbal cognitive measure.

The participants were administered two different picture-naming tasks and a letter and category fluency task. The results of this study show that even though there was no difference in the forward, memory span, task of the participants, bilinguals obtained higher scores in the backward, working memory, task. The monolinguals outperformed the bilinguals on the lexical retrieval tasks, but, due to the scores on the memory span task, these results cannot be attributed to superior memory, or superior intelligence. The authors, therefore, state that the differences between bilinguals and monolinguals in verbal fluency may be attributed to vocabulary size as the overall vocabulary size of the bilinguals leads to less automaticity in lexical retrieval (Bialystok et al., 2008).

The second part of their study was designed to pursue the results of the first using a more detailed assessment of language proficiency. For this additional study, fifty bilingual and sixteen monolingual students from a university were recruited as participants. The bilingual participants were classified into subgroups based on their English proficiency. They were divided based on their scores on the Peabody picture vocabulary test (Dunn & Dunn, 1997) and subsequently placed into high-proficiency (HP) and low-proficiency (LP) groups. The non-English languages for the HP bilinguals included Cantonese, Hindi, French, Hebrew, Korean, Gujarati, Urdu, Italian, Russian, and Tagalog, while the non-English languages for the LP bilinguals included Cantonese, Hebrew, Arabic, Russian, Sinhala, Indonesian, Punjabi, Portuguese, French, Mandarin, Ukrainian, Vietnamese, Toisan, Korean, Spanish, Polish, Gujarati, and Tamil. The participants were exposed to the aforementioned Peabody picture vocabulary test, an Expressive vocabulary task, a Spatial span subtest, similar to the spatial span task used in their first study, and a letter and category fluency task. The letter and category fluency task was similar to the letter fluency task in the first study with one change. The researchers instructed the participants to generate as many words as possible after being shown a single letter, F, A, and S. However, they were told to exclude names of people, places, numbers, and verbs with different endings/inflections (Bialystok et al., 2008). The results of this study showed no difference between the groups in terms of memory on the span task, yet there was a significant difference in the performance on the letter fluency task. Unlike in the first study, the HP bilinguals outperformed the monolinguals and LP bilinguals in letter fluency. The letter fluency task was modified to demand higher levels of executive control than the task employed in the first study. This allowed the HP bilinguals to perform better than the monolinguals and LP bilinguals.

The overall conclusions drawn from these two studies address two variables affecting the bilingual disadvantage in speech production. The results of the first study suggest that the bilingual disadvantage described in previous research would be more accurately interpreted as reflections of vocabulary size than weaker control of lexical processing (Bialystok et al., 2008). In their second study, the HP bilinguals performance can be attributed to the increased demand placed on executive control, which, based on neuroimaging evidence, put the bilinguals at an advantage. The bilingual speakers were able to perform better than monolinguals on letter fluency because of the boost to frontal executive processes stemming from their bilingualism (Bialystok et al., 2008). The conclusions of this study continue to develop how researchers define the contributions of lexical access in the speech production of both bilingual and monolingual speakers. Bialystok et al. (2008) limited their speech production elicitation to a picture-naming task in order to isolate the contribution of lexical access in bilinguals and monolinguals. Now that this contribution has been illustrated, future research can begin to investigate the role of lexical access in the fluency of sustained speech.

Justification of the Current Study

The body of research investigating the predictors of fluent speech has focused primarily on the role of working memory capacity. As noted above, working memory capacity is related to the amount of information that the individual can store and effectively access in cognition. Studies have shown that there is a relationship between working memory capacity and fluency as the quantity of words that a speaker can access from working memory contributes to an increase in rate of delivery (Weissheimer & Mota, 2009). However, lexical access, the speed at which words are retrieved and prepared for speech, does not appear in the literature of studies on fluency. Due to the fact that rate of delivery is a primary component of fluency (Chambers, 1997; Wood, 2001), the relationship between lexical access and rate of delivery in spontaneous speech could further develop the understanding of the cognitive processes involved in L2 speech production.

Research Question

Based on the results of the research previously discussed that addressed both lexical access and fluency independently and in combination with other factors, the following research question will be addressed in this study: Is there a relationship between lexical access and fluency in the speech of intermediate Spanish L2 learners?

CHAPTER 3

METHODOLOGY

Participants

This study included four university students (N = 4) as participants. The participants were two males and two females between the ages of 18 and 30 with a mean age of 22 (SD = 5.47) and an average of 4.5 years (SD = 1.19) of academic L2, Spanish, instruction. The two male participants began studying Spanish in middle school while the two female participants did not begin until high school. The participants were coded using a three digit system that accounted for their order by sex and the specific course in which they were enrolled (e.g. the first female volunteer, enrolled in 5th semester Spanish, course number SPA313, would be coded as participant 1F3, the second female participant, also enrolled in the course SPA313, would be 2F3, and so on). At the time of the investigation, all participants were enrolled in their fifth semester of university level Spanish, a course designed to focus primarily on developing the students' conversational skills and writing abilities. All participants were born and raised in the United States as monolingual English speakers with English as their primary language at home and at school. One participant (1F3) reported that for the past four years she has spoken Spanish with her stepmother while at home on a daily basis. This interaction with her step mother began after she had begun taking Spanish classes in high school and not prior. None of the participants have spent time in a Spanish speaking country for educational purposes and only one of the participants (1F3) has spent time in Mexico, a combined total of three weeks, during multiple family vacations.

Instruments

Language contact profile (Appendix A). The participants completed the Language Contact Profile (Freed et al., 2004) in order to assess their history and current contact with the target language, Spanish.

Picture naming task (Appendix B). To measure lexical access abilities each participant performed a picture-naming task in Spanish to assess their lexical access ability with their L2. During the picture-naming task the participants were shown a series of images and instructed to respond by naming the image as quickly as possible in the target language (e.g., an image of an apple would elicit "*manzana")*. The latency period for each response was measured in seconds using a stopwatch and these latencies were used to establish each participant's average response time (RT). This instrument was chosen so that the investigators would be able to establish a quantifiable metric for the lexical access ability of the participants. The participant's RT served as a way to rank the participants from strongest, shortest average RT, to weakest, longest average RT, in terms of lexical access ability. The results from this instrument, which will be discussed in detail in the Results section, demonstrate the time required for each participant to go through the *lexical selection* and *form encoding* steps described by Levelt's (2001) production model.

Narration task 1 (Appendix C). To measure fluency, students retold a recently viewed short silent film. The film chosen contained a simple plot with people waiting for a train and included a plot twist where a thief enters to rob the principal female character. The video clip for the first narration task lasted 4 minutes, 59 seconds. Once the video ended, the participants were asked to narrate what took place in the film. The potential variance in the amount each participant was able to remember was mitigated by evaluating their narrations solely on word quantity and not evaluating for accuracy. Another potential variance, lexical density, was not measured. Lexical density is measured by dividing the total numbers weighted lexical items uttered, nouns, verbs, adjectives, adverbs, from the total number of all lexical items uttered (Mota, 2003). Mota (2003) analyzed the contents of the spontaneous L2 speech of her participants in order to evaluate their lexical density and correlate that with their working memory capacity and found that those participants who had less lexical density spoke at a higher rate than those who spoke with more lexical density. For this reason, this investigation prioritized the quantity and rate of the spontaneous L2 speech generated by each participant and not the contents of the speech. These narrations were audio recorded.

Narration task 2 (Appendix D). Upon completion of the first narration task, the participants conducted a second narration task where they were asked by the researcher to tell about a favorite memory or personal anecdote. The investigators of this study used the data from the two narration tasks to establish the participants' fluency, as measured by their rate of speech, the number of words spoken per second, and the frequency of filled pauses used during spontaneous speech.

When evaluating the speech production of second language learners, previous research has given fluency assessing tasks to bilingual participants in both their first and second languages (Bialystok et al., 2008). This investigation did not collect data on the first language (L1) production of the participants. The investigators chose not to incorporate the potential for individual differences in L1 speech production as a factor in L2 oral fluency and focus solely on L2 lexical access. However, the lack of available research on the role of L1 speech production as a predictor of L2 oral fluency could lead to future investigations.

Procedures

The Language Contact Profile (Freed et al., 2004), picture-naming task, and the two narration tasks were administered in a computer lab similar to the classrooms in which the participants attend their Spanish classes. The computer lab was chosen to mitigate any environmental variables such as weather, climate, and background noise levels while offering convenience to the participants as it is located in the same building as their Spanish classes. The data were collected during the second month of the semester. Students were recruited and asked to participate voluntarily. Those who volunteered met with the investigators in the computer lab at a time convenient to them during the school week. All participants chose to meet between the hours of 11:00 am and 2:00 pm. During the recruitment, the participants were told about the general scope of the investigation and given brief details about the instruments. They were given the complete explanation of what they would be doing prior to data collection. The participants volunteered and were not compensated for their participation.

Measurements

Data elicited in two narrations and one picture-naming task were included in the final analysis. Data from the picture-naming task, the latency period between the appearance of the images and responses of the participants, was collected and used to give each participant a quantifiable value for their lexical access ability in their L2. The data from the two narration tasks were transcribed and coded for quantity of words uttered against the total duration of speech and for the quantity of filled pauses to give each participant quantifiable values for their individual fluency. Once transcription took place, data were coded to determine the numbers of words uttered during the two narration tasks. Each word that the participants produced was counted (Chambers, 1997; Wood, 2001). If stuttering occurred the word was not counted unless the participant was able to produce it completely, such as the time participant 1M3 said *inmedia-inmediatamente* 'immedia-immediately' or when participant 2F3 said *cor-cor-corre* 'ru-ru-runs'. Words were counted if the participant repeated the same word successively, provided it was in a completed form each time, as in the example *tarjeta uh tarjeta* 'card uh card' from participant 2F3.

The data from the picture-naming task was collected as a measurement of the amount of time that passed from when the participant was exposed to the visual stimulus, the image, to when they uttered the response. This passage of time was coded as the response time for each image and used to establish the participant's lexical access ability. The participants were shown a series of forty images and instructed to respond in the target language, Spanish, as soon as they identified the image displayed. All images corresponded to vocabulary items on which the students had received instruction during their time studying Spanish at their current university. Responses in English were discarded as were any responses where the participants simply failed to produce a response for an image, three total occurrences out of 200 possible responses. Upon completion of all forty images, the response time data were compiled to find the average naming latency for each participant.

Data Analysis

Upon completion of transcribing and coding the narrations and compiling the naming latency data, the statistical analysis was completed to respond to the research question. The a priori statistical testing was done to find multiple values from the naming latency data; mean, mean standard error (MSE), standard deviation (SD) and variance. The mean was found to determine the central tendency of the data and to give the average RT per participant. The mean standard error is calculated to determine how accurately the mean of the RTs in question would represent a larger set of RTs if the participants were given a picture-naming task with more images and there were more values to average. The standard deviation shows how the RTs are spread around the mean while the variance was found to determine how widely any of the RTs vary from the mean. The data from these instruments were analyzed using Pearson's correlation coefficient to measure the strength of the association between lexical access and oral fluency, both measured on interval scales. Pearson's correlation coefficient exists as a numeric value r ranging on a scale from negative one, a perfectly inverse linear relationship, to one, a perfectly positive linear relationship. An r value of zero means there is no relationship between the variables (Mackey & Gass, 2005). The results, in response to the research question, are presented in the following section.

CHAPTER 4

RESULTS

Individual Instruments

The results of the Picture-Naming task are displayed on Table 1 and show the average response time for each participant.

Table 1

Picture-Naming Task Results per Participant

	Participants			
	1F3	2F3	1M3	2M3
Mean	1.677	1.953	2.557	3.038
MSE	0.152	0.262	0.367	0.5
SD	0.937	1.675	2.350	3.161
Variance	0.878	2.805	5.52	9.99
Min	0.69	0.66	0.79	0.72
Max	4.850	9.65	12.97	16.21
Range	4.160	8.99	12.18	15.49
Median	1.28	1.41	1.97	1.72
Quartile 1	1.098	1.06	1.32	1.22
Quartile 3	1.965	2.205	3.015	3.687

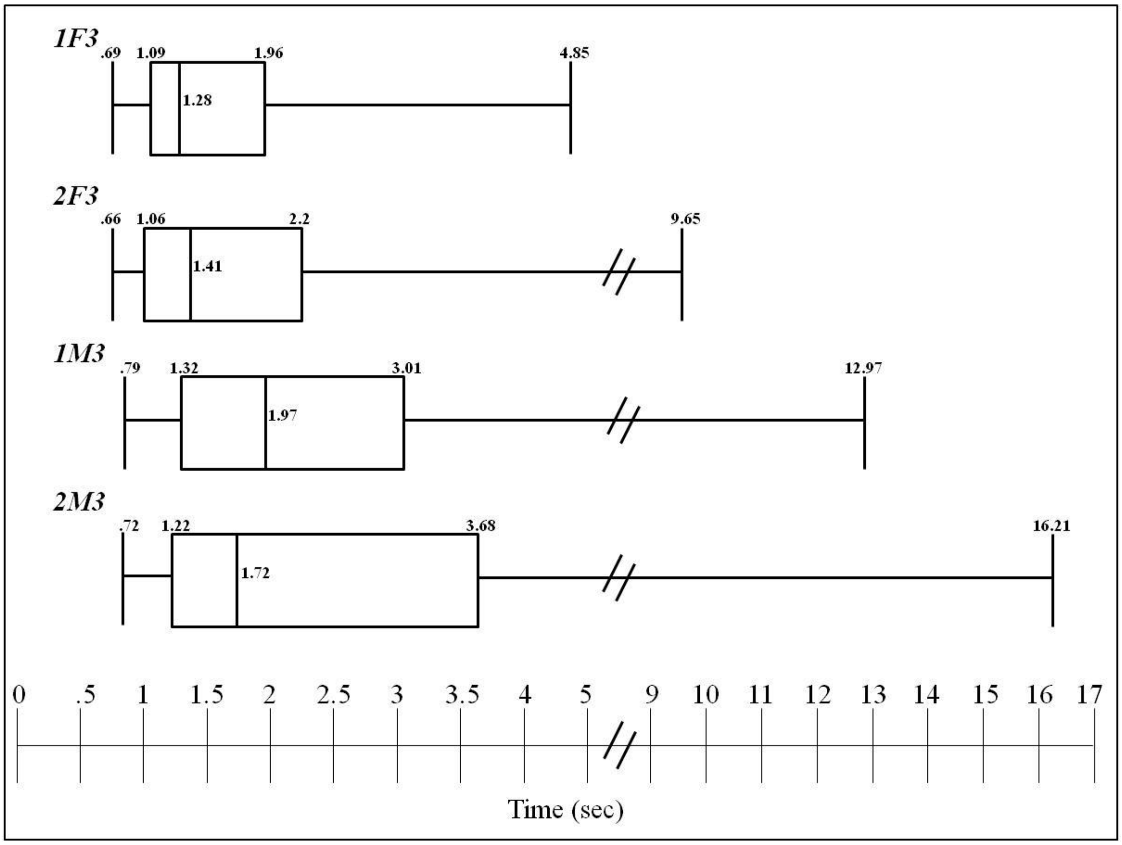
Participant 1F3 recorded the fastest average response time, lowest mean value, of 1.677 seconds (SD = 0.937). Participant 2M3 had the slowest average response time, largest mean value, of 3.038 seconds (SD = 3.161). Participants 2F3 and 1M3 scored response times between these two values with response times of 1.953 (SD = 1.675) and 2.63 seconds (SD = 2.35) respectively.

Table 1 shows the overall fastest and slowest response times from all four participants represented under the Min, fastest response time for each participant, and Max, slowest response time for each participant. The response times ranged from the fastest with a latency of 0.66 seconds from participant 2F3 to the slowest by participant 2M3 with a latency of 16.21 seconds. Participant 1F3 had the most consistent response times as demonstrated by her variance of 0.878 while participant 2M3 demonstrated the largest variance at a value of 9.99. All participants were able to achieve their minimum latency in under a second with each of the participants' fastest response times occurring within a range of 0.13 seconds, .66 for 2F3 to .79 seconds for 1M3. Their slowest response times displayed a larger variation and the difference between participant 1F3, maximum latency of 4.85 seconds, and participant 2M3, maximum latency of 16.21 seconds, was a range of 11.36 seconds. As the average of all response times, the mean response time for each participant was used as the value for their naming latency score. For this task, the participants can be ordered from fastest to slowest, 1F3, 2F3, 1M3, and 2M3.

The boxplots in Chart 1 provide a visual representation of the RT data, presented in Table 1, of each participant. The boxplot for each participant contains, from left to right, their fastest response time, the bracket extending to the left of the box, their first quartile, the left edge of the box, their median response time, the line bisecting the box, their third quartile, the right edge of the box, and their slowest response time, the bracket extending to the right of the box. The first and third quartiles do not correspond to a specific response time. The first quartile represents the midway point between the fastest and median response times. Equally, the third quartile represents the midway point between the median and slowest response times.

Figure 1

Boxplot of RT Data



The results of the two narration tasks were analyzed separately. The results of the movie narration task are displayed in Table 2.

Table 2

Speech Rate and Pause Data from Movie Narration

	Participants			
	1F3	2F3	1M3	2M3
Total Words	160	175	226	174
Total Duration (seconds)	244	174	271	190
Words Per Second	0.656	1.006	0.834	0.916
Total Pauses	30	42	64	48
Words Per Filled Pause	5.333	4.167	3.531	3.625

The results of the story narration task are presented in Table 3.

Table 3

Speech Rate and Pause Data from Story Narration

	Participants			
	1F3	2F3	1M3	2M3
Total Words	105	116	149	165
Total Duration (seconds)	178	119	149	159
Words Per Second	0.590	0.975	1	1.038
Total Pauses	36	30	39	41
Words Per Filled Pause	2.917	3.867	3.821	4.024

Tables 2 and 3 present the total number of words spoken and filled pauses uttered during each narration and the duration of each narration as those values varied for each participant.

For example, comparing the Total Words and Total Duration values from Tables 2 and 3 we see that each participant's movie narration lasted longer and included more words than their story narration yet their rate of speech, Words per Second, and use of pauses did not correspond to the two narration tasks as uniformly. Each participant spoke for a minimum of 175 (2.92 minutes) seconds to fully summarize the movie and no participant spoke longer than 271 (4.52 minutes) seconds in their retell narration. Participant 1F3 produced the highest rate of speech with 1.006 words per second followed closely by participant 2M3 at a rate of 0.916 words per second. Participant 1M3 achieved a rate of speech with 0.834 words per second while participant 1F3 spoke with the slowest rate at 0.656 words per second.

In Figure 2, the line of best fit is graphed shown a scatter plot of the first fluency measure, words per second, from the movie narration task and the participants' results from the PNT. The slope of the line of best fit corresponds to the correlation coefficient between these two variables and is given as value r.

Figure 2

Speech Rate vs. Naming Latency (Movie)

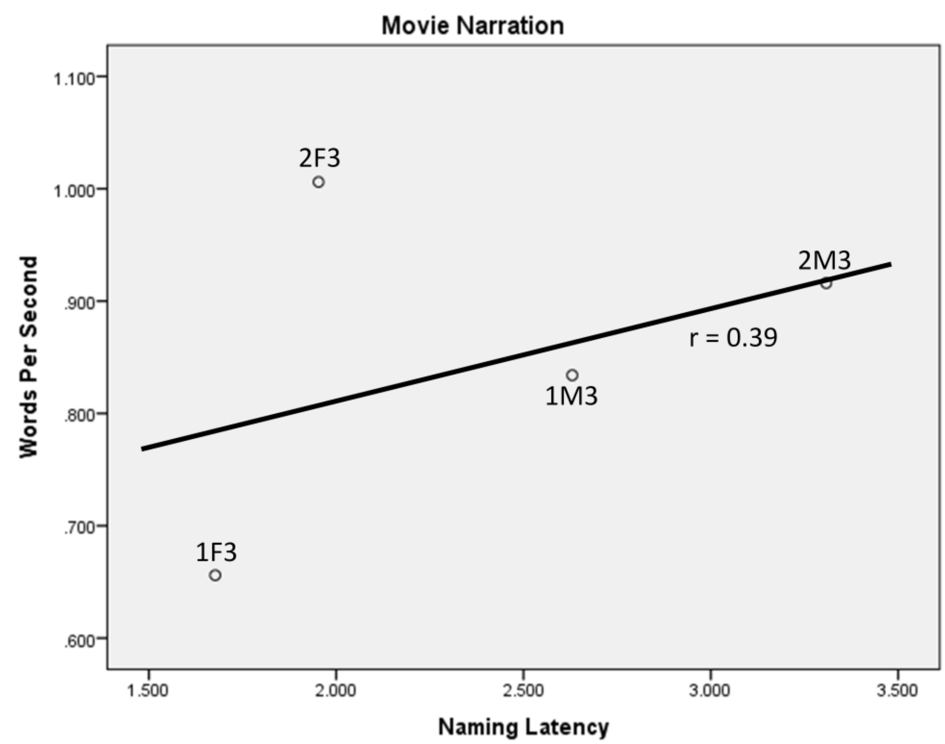
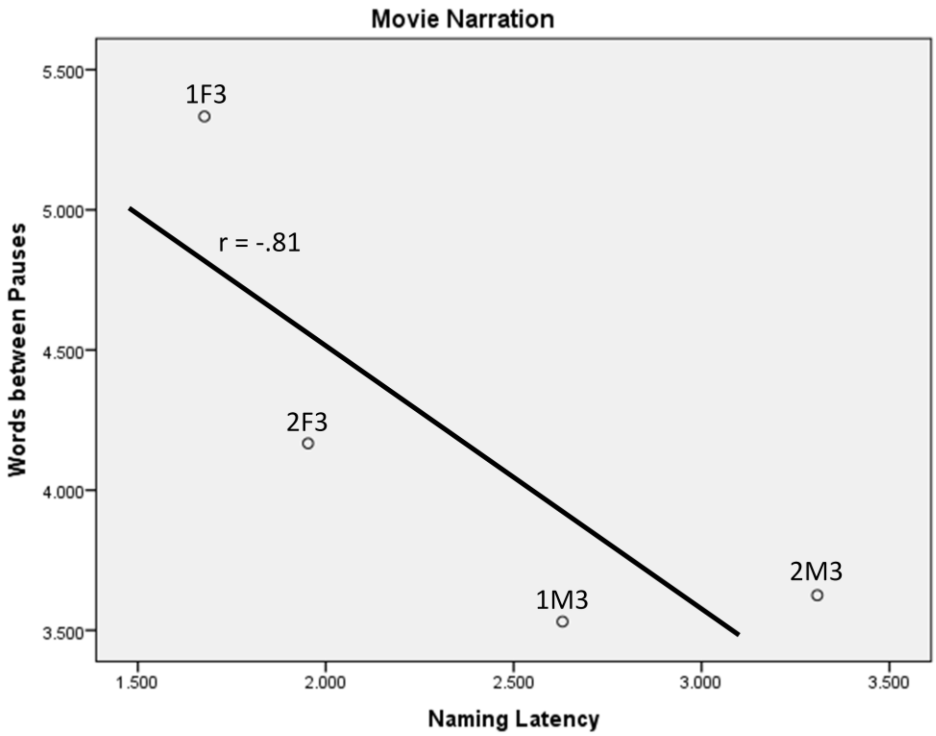


Figure 3 follows the same format as Figure 2 but contains the data from the second fluency measure, words per filled pause, and contains the line of best fit as that data is presented in a scatter plot with the PNT results.

Figure 3

Pause Data vs. Naming Latency (Movie)

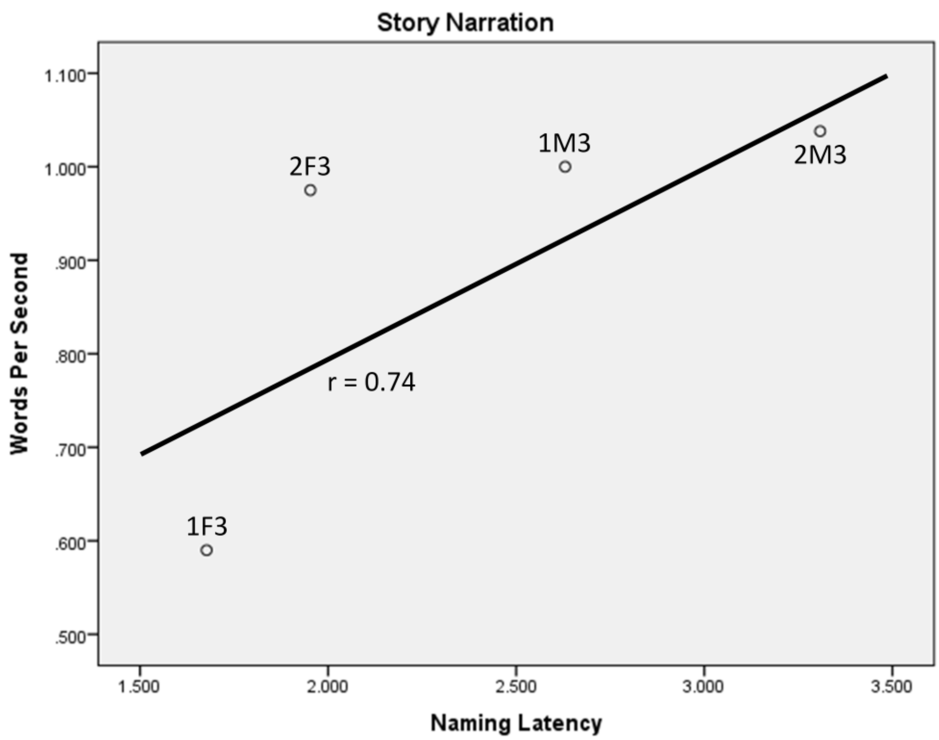


As the rate of speech is not the only contributing factor to a speaker's fluency, the participant's narrations were also analyzed for their use of pauses. The pauses were counted and factored into the total number of words spoken to produce a ratio of the amount of words each participant spoke between each pause during spontaneous speech. During the movie narration, participant 1F3 uttered the least amount of words and produced the slowest rate of speech, relative to the other participants, yet she uttered so few pauses, 30 in total, that she achieved the highest ratio of 5.333 words per filled pause. Participant 1M3 uttered 64 pauses during his movie narration and even though he produced the most total words of all narrations his ratio of 3.531 words per filled pause ranked as the lowest. Participants 2F3 and 2M3 produced 42 and 48 pauses respectively and this contributed to participant 2F3 scoring a ratio of 4.167 words per filled pause while participant 2M3 uttered 3.625 words per filled pause.

The results from the story narration task are presented in the following two figures. The scatter plot and line of best fit of the first fluency measure, words per second, from the story narration task are displayed in Figure 4.

Figure 4

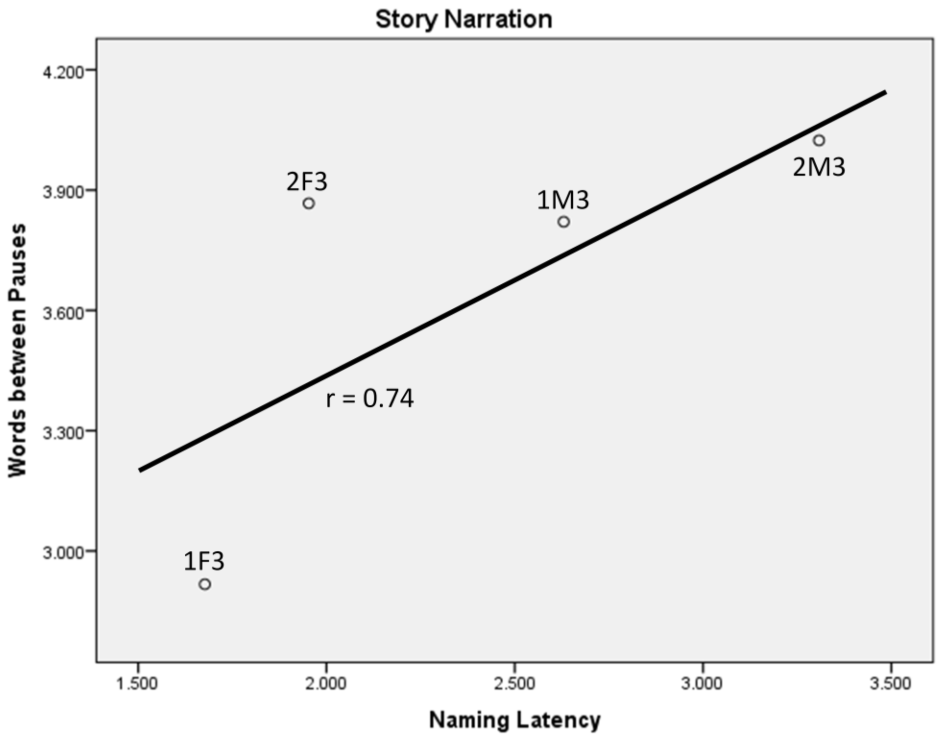
Speech Rate vs. Naming Latency (Story)



The scatter plot and line of best fit of the second fluency measure, words per filled pause, from the story narration task are displayed in Figure 5.

Figure 5

Pause Data vs. Naming Latency (Story)



Again, participant 1F3 produced the fewest total words and scored the slowest rate of speech at 0.590 words per second. Similar to participant 1F3, participant 2F3 produced fewer total words in her story narration and saw her rate of speech register at 0.975 words per second. The two male participants also produced fewer total words than their movie narrations yet both spoke with a higher rate of speech than their movie narrations with participant 1M3 speaking at 1.000 word per second and participant 2M3 increasing to 1.038 words per second. Participant 2M3 produced the most pauses during the story narration, 41, yet still achieved a ratio of 4.024 words per filled pause, the highest of all participants. As with the words per second, the other male participant, 1M3, also increased his ratio of words per filled pause from the movie narration to a value of

3.821. The two female participants saw their ratios of words per filled pause register at a lower value during the story narration to 3.867 words per filled pause for participant 2F3 and 2.917 words per filled pause for participant 1F3.

Compiled Data

The data from the two narration tasks were compiled and are presented in Table 4 as a cumulative data set for each participant.

Table 4

Compilation of Speech Rate and Pause Data

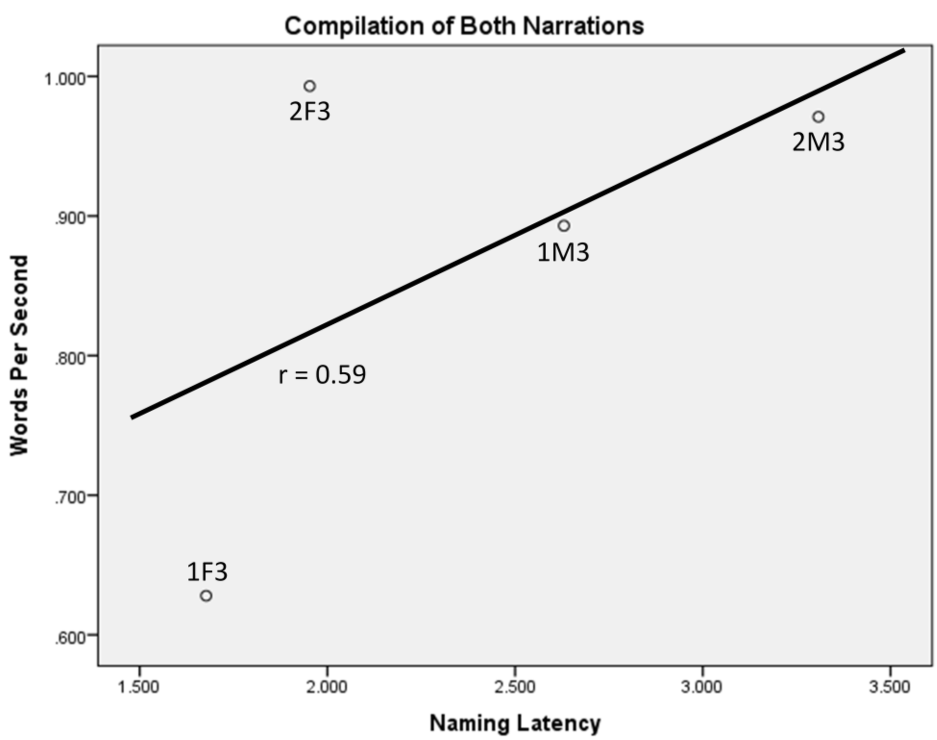
Participants	Total Words	Total Duration (sec)	Total Filled Pauses	Words Per Second	Words Per Filled Pause
1F3	265	422	66	0.628	4.015
2F3	291	293	72	0.993	4.042
1M3	375	420	103	0.893	3.641
2M3	339	349	89	0.971	3.809

Note. sec = seconds

Upon compilation of all data, participant 2F3 demonstrated the highest overall rate of speech at 0.993 words per second and the highest ratio of 4.042 of words per filled pause. Participant 1F3, as she had on both of the narration tasks individually, recorded the lowest rate of speech with a combined count of 0.628 words per second. She had the highest ratio of words per filled pause on the movie narration and then the lowest ratio on the story narration which gave her an overall ratio of 4.015 words per filled pause equating to the second highest in the overall data set. Participant 1M3 used the most pauses during his two narrations and finished with a ratio of 3.641 words per filled pause. His cumulative rate of speech placed him at the third fastest overall with 0.893 words per second. The final participant, 2M3, scored the second highest overall rate of speech with 0.971 words per second and the third highest score for his ratio of 3.809 words per filled pause. These data are represented graphically in the following 2 figures. Figure 6 shows the scatter plot and line of best fit for the compiled rate of speech data.

Figure 6

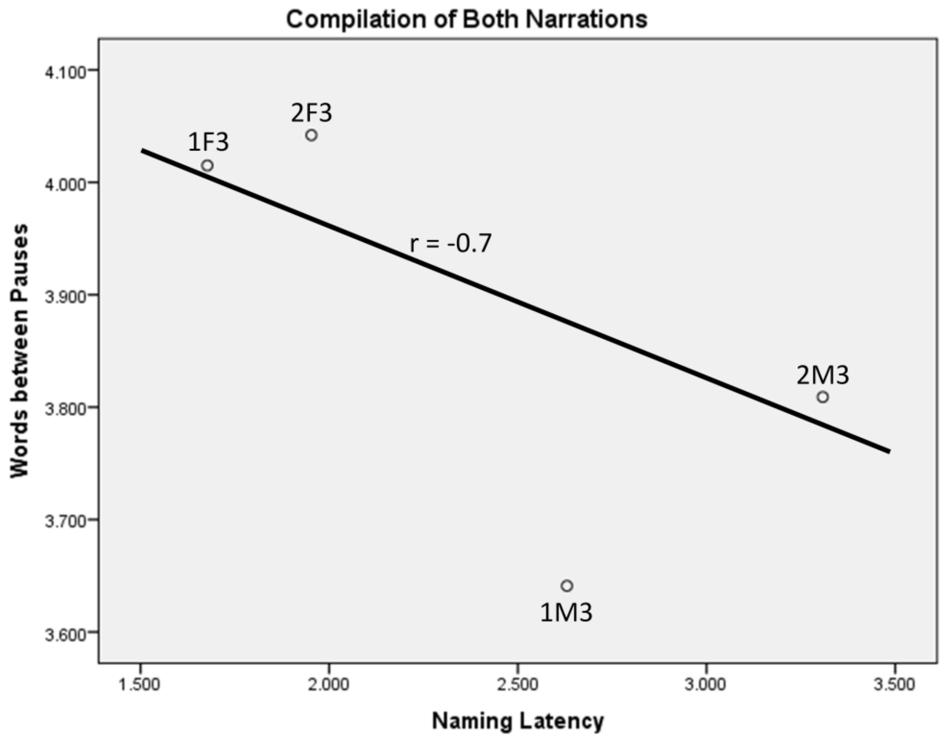
Speech Rate vs. Naming Latency (Compiled)



The scatter plot and line of best fit for the words per filled pause from the compiled narrations is presented in Figure 7.

Figure 7

Pause Data vs. Naming Latency (Compiled)



The analysis of the correlation between the naming latencies and the two fluency measures, rate of speech and frequency of pauses, was conducted using SPSS. The results are presented in Table 5. The data from the narration tasks were analyzed individually and again as a compiled data set against the naming latencies produced by each participant.

Table 5

Correlation Data between Narration and Picture-Naming Tasks

	Narrations					
	Movie		Story		Compiled	
	r	*Sig.*	*r*	*Sig.*	*r*	*Sig.*
Words Per Second	0.389	0.611	0.741	0.259	0.591	0.409
Words Per Filled Pause	-0.817	0.183	0.740	0.260	-0.701	0.299

Note. *r* = correlation coefficient; *Sig.* = significance

When the participants' rates of speech from the movie narration were correlated against their naming latencies they showed a Pearson's correlation coefficient of *r* = 0.389 (*Sig.* 0.611), a value close to zero revealing a weak correlation, while the words per filled pause from the movie narration produced a value of *r* = -0.817 (*Sig.* 0.183), a value close to negative one demonstrating a strong inverse correlation. The same analysis run against the data from the story narrations provided similar values, words per second *r* = 0.741 (*Sig.* 0.259) and words per filled pause *r* = 0.74 (*Sig.* 0.26), both of these values trending toward a strong positive correlation. Lastly, the compiled data from both narrations were analyzed and the words per second correlated with the naming latencies with a value of *r* = 0.591 (*Sig.* 0.409), showing a weak relationship, while the words per filled pause produced a correlation of *r* = -0.701 (*Sig.* 0.299), trending toward a strong inverse correlation. The strongest correlation and significance values appeared when only the pause data from the movie narration were analyzed, *r* = -0.817 (*Sig.*

0.183). That same task, the movie narration, also produced the weakest correlation and significance value at $r = 0.389$ (*Sig.* 0.611).

CHAPTER 5

DISCUSSION

The purpose of this paper was to answer the question about the role of lexical access in the oral fluency of intermediate L2 learners. Presented in Table 5, this investigation found that lexical access, as measured by a Picture-Naming Task, trended toward an inverse correlation ($r = -.701$) with one measure of fluency, words per filled pause, for the participants. For the other measure of fluency, words per second, the lexical access of the participants did not show an inverse correlation like words per filled pause but, instead, trended in the opposite direction toward no correlation ($r = .591$). The Significance tests of each of these correlations show that neither are statistically significant. This suggests that the inverse correlation demonstrated between lexical access and words per filled pause may have occurred due to chance and that the lack of a linear relationship between lexical access and words per second also may have occurred due to chance. The Significance values in Table 5 indicate that the results of this investigation cannot reasonably be used to make predictions about the larger population. The sample size of this investigation, n = 4, was the primary factor in limiting the strength of the Significance values in Table 5.

Due to the fact that these results cannot be generalized, this investigation has not sufficiently answered the research question either in the affirmative or in the negative. Therefore, it is still possible to investigate this research question if changes are made to the methodology of the investigation. There are two ways that can be done. The first would be to control for working memory in the narration tasks. The two narration tasks, the immediate retelling of a movie and the narration of a memory, place different burdens on the participants' working memory. The first narration task, the movie retell, requires that the participant hold the recently received content of the movie in their working memory while they simultaneously use their lexical access to select and encode the lexical items for their spontaneous narration. Meanwhile, the second narration task, the narration of a memory, doesn't require that the participants hold newly received information in their working memory. Instead, the second narration task requires that they first use their long term memory to choose the memory or anecdote to be narrated before their lexical access selects and encodes the appropriate lexical items during the generation of spontaneous speech. By using only one of these narration tasks to assess oral fluency, future investigations of this research question would have fewer intervening variables. Second, changes to the picture-naming task may allow for a more complete assessment of the lexical access of the participants. The present investigation utilized a PNT comprised of images of specific items, thus resulting in the elicitation of nouns, and no other lexical item from the participants. A PNT that requires the participants to generate more lexically diverse responses (e.g., verbs, adverbs, adjectives, prepositions, pronouns), similar to the PNT used by Hartsuiker & Notebaert (2010), may offer a more complete view of their lexical access as it relates to the generation of spontaneous speech. Also, while not a change to the methodology, a future investigation of this research question would benefit from a larger sample size. This would increase the probability of finding a statistically significant p value regardless of the correlation coefficient. In other words, it would provide an answer, whatever it may be, to the research question that could reasonably be applied to the larger population.

Future Research

Another way to address the question of lexical access' role in oral fluency is to study a different population of speakers. One population worth investigating may be monolingual Spanish speakers to function as a control group. In their research, Ivanova & Costa (2007) demonstrated that monolingual participants performed better on a PNT than did the bilingual participants in their dominant language. As previously mentioned, Levelt (2001) stated a monolingual person will generate two to four words per second in normal speech. This rate of speech ranges from two to six times the rate of speech generated by the participants of this study. Given this information, it would be reasonable to assume that monolingual speakers would also perform better, generate spontaneous speech at a higher rate with more words per filled pause, than bilingual speakers with the narration instruments utilized in this study. However, the assumption that monolingual speakers would achieve better results relative to the participants of the current study on each of the instruments doesn't imply those results would demonstrate a statistically significant correlation coefficient. Applying this study's research question and methodology to a population of monolingual Spanish speakers would provide valuable information on the role of lexical access as a predictor of oral fluency and would potentially highlight any variation between populations.

CHAPTER 6

CONCLUSIONS

In the present study, the results showed indications that a relationship between lexical access and oral fluency can exist when confined to a small population. The lack of statistical significance in the correlation coefficients means that these results cannot be applied to the larger population and therefore this research question remains to be answered. A likely solution for the lack of statistical significance can be found in refinements to the methodology and an increased number of participants. Answering the question of lexical access' role in oral fluency will contribute to the body of research on individual differences in SLA as they apply to the speech production of intermediate L2 learners.

REFERENCES

Bialystok, E., Craik, F. I., & Luk, G. (2008). Lexical access in bilinguals: Effects of vocabulary size and executive control. *Journal of Neurolinguistics, 21*(6), 522-38.

Breiner-Sanders, K. E., Lowe, P., Miles, J., & Swender, E. (2000). ACTFL proficiency guidelines—Speaking: Revised 1999. *Foreign Language Annals*, *33*(1), 13-18.

Chambers, F. (1997). What do we mean by fluency? *System, 25*(4), 535-544.

Daneman, M. (1991). Working memory as a predictor of verbal fluency. *Journal of Psycholinguistic Research*, *20*(6), 445-464.

De Bot, K., Lowie, W., & Verspoor, M. (2007). A dynamic systems theory approach to second language acquisition. *Bilingualism: Language and cognition*, *10* (1), 7-21.

Dunn, L. M., & Dunn, L. M. (1997). *Peabody picture vocabulary test--III*. Circle Pines, MN: American Guidance Service.

Freed, B. F., Dewey, D. P., Segalowitz, N., & Halter, R. (2004). The language contact profile. *Studies in Second Language Acquisition*, *26*(02), 349-356.

Hartsuiker, R. J., & Notebaert, L. (2010). Lexical access problems lead to disfluencies in speech. *Experimental Psychology, 57*(3), 169.

Ivanova, I. & Costa, A. (2007). Does bilingualism hamper lexical access in speech production? *Acta Psychologica*, *127*(2), 277-288.

Levelt, W. J. (2001). Spoken word production: A theory of lexical access. *Proceedings of the National Academy of Sciences*, *98*(23), 13464-13471.

Mackey, A., Gass, S. M. (2005). *Second language research: methodology and design*. New York, Routledge.

Mota, M. B. (2003) Working memory capacity and fluency, accuracy, complexity, and lexical density in L2 speech production. *Fragmentos, 24*, 69-104.

O'Brien, I., Segalowitz, N., Collentine, J., & Freed, B. (2006). Phonological memory and lexical, narrative, and grammatical skills in second language oral production by adult learners. *Applied Psycholinguistics*, 27 (03), 377-402.

O'Brien, I., Segalowitz, N., Freed, B., & Collentine, J. (2007). “Phonological memory predicts second language oral fluency gains in adults.” *Studies in Second Language Acquisition*, *29* (2), 557-582.

Weissheimer, Janaina., & Mota, M.B. (2009) Individual Differences in Working Memory Capacity and the Development of L2 Speech Production. *Issues in Applied Linguistics, 17*(2), 93-112.

Weissheimer, J., & Mota, M. B. (2011). Working memory capacity and the development of L2 speech production: a study of individual differences. In G. Granaena, (Eds.), *Selected Proceedings of the 2010 Second Language Research Forum,* (pp.169-181).

Wen, Z. (2012) Working memory and second language learning. *International Journal of Applied Linguistics, 22*(1), 1-22.

Wood, D. (2001). In search of fluency: What is it and how can we teach it? *The Canadian Modern Language Review, 57*(4), 573-589.

APPENDIX A

LANGUAGE CONTACT PROFILE

PRETEST VERSION OF THE LANGUAGE CONTACT PROFILE

PROJECT: ACQUISITION OF SPANISH AS A SECOND LANGUAGE

The responses that you give in this questionnaire will be kept confidential. This cover sheet is to allow the researcher to associate your responses with your name if needed. However, only the people entering your responses into the computer will see this name. An identification number will be used in place of your name when referring to your responses in publications. Every effort will be made to keep your responses confidential.

Thank you for your cooperation. The information that you provide will help us to better understand the backgrounds of students who are studying Spanish in various contexts. Your honest and detailed responses will be greatly appreciated.

Name: ________________

Part 1: Background Information

1. Gender: Male /Female
2. Age: ___
3. Country of birth:
4. ____ What is your native language? 1) English 2) Spanish 3) Other ________________________________
5. ____ What language(s) do you speak at home? 1) English 2) Spanish 3) Other ____________________________

5a. If more than one, with whom do you speak each of these languages? __________________

6. In what language(s) did you receive the majority of your *precollege* education?

6a. If more than one, please give the approximate number of years for each language. _______
1) English 2) Other _____________________

7. Have you ever been to a Spanish-speaking region *for the purpose of studying Spanish*? Circle one: Yes / No

7a. If yes, when? ______________________7b. Where? ___________________

7c. For how long? 1 semester or less 2 semesters more than 2 semesters

8. Other than the experience mentioned in Question 7, have you ever lived in a situation where you were exposed to a language other than your native language (e.g., by living in a multilingual community; visiting a community for purposes of study abroad or work; exposure through family members, etc.)? Circle one: Yes / No

If Yes, please give details below. If more than three, list others on back of this page.

	Experience 1	Experience 2	Experience 3
Country/region			
Language			
Purpose			
From when to when			

9. In the boxes below, rate your language ability in each of the languages that you know. Use the following ratings: 0) Poor, 1) Good, 2) Very good, 3) Native/nativelike.
How many years (if any) have you *studied* this language in a *formal school* setting?

Language	Listening	Speaking	Reading	Writing	Number of years of study
English					
Spanish					
Other					

10. Have you studied Spanish in school in the past at each of the levels listed below? If yes, for how long?
a) Elementary school: _No _Yes: _less than 1 year _1–2 years _more than 2 years
b) Junior high (middle) school: _No _Yes: _less than 1 year _1–2 years _more than 2 years
c) Senior high school: _No _Yes: _less than 1 year _1–2 years _more than 2 years
d) University/college: _No _Yes: _less than 1 year _1–2 years _more than 2 years
e) Other (Please specify) :
_No _Yes: _less than 1 year _1–2 years _more than 2 years
11. What year are you in school? (circle one):
Freshman Sophomore Junior Senior Graduate student Other
12. What is your major?

Part 2: All of the Questions That Follow Refer to Your Use of Spanish, Not Your Native Language, Unless the Question Says Otherwise

13. On average, how often did you *communicate* with native or fluent speakers of Spanish *in Spanish* in the year prior to the start of this semester?
0) never 1) a few times a year 2) monthly 3) weekly 4) daily
14. Use this scale provided to rate the following statements.
0) never 1) a few times a year 2) monthly 3) weekly 4) daily
Prior to this semester, I tried to speak Spanish to:
a. my instructor outside of class
b. friends who are native or fluent speakers of Spanish
c. classmates
d. strangers whom I thought could speak Spanish
e. a host family, if living in a Spanish-speaking area
f. service personnel (e.g., bank clerk, cashier)
15. For each of the items below, choose the response that corresponds to the amount of time you estimate you spent on average doing each activity *in Spanish* prior to this semester.
a. watching Spanish language television
0) never 1) a few times a year 2) monthly 3) weekly 4) daily
b. reading Spanish language newspapers
0) never 1) a few times a year 2) monthly 3) weekly 4) daily
c. reading novels in Spanish
0) never 1) a few times a year 2) monthly 3) weekly 4) daily
d. listening to songs in Spanish
0) never 1) a few times a year 2) monthly 3) weekly 4) daily

e. reading Spanish language magazines
0) never 1) a few times a year 2) monthly 3) weekly 4) daily

f. watching movies or videos in Spanish
0) never 1) a few times a year 2) monthly 3) weekly 4) daily

16. List any other activities that you commonly did using Spanish prior to this semester.

17. Please list all the Spanish courses you are taking this semester. This includes Spanish language courses as well as content area courses taught in the Spanish language.

Course name	Course number	Brief description

APPENDIX B

PICTURE-NAMING TASK

Picture Naming Task

Instructions:

I am going to ask you to name some pictures. **Your job is to name the picture as soon as you see it.** Please use only one word. We'll practice several pictures before we begin.



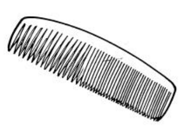




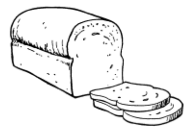
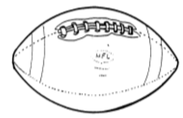










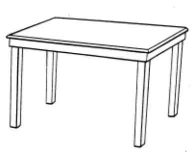



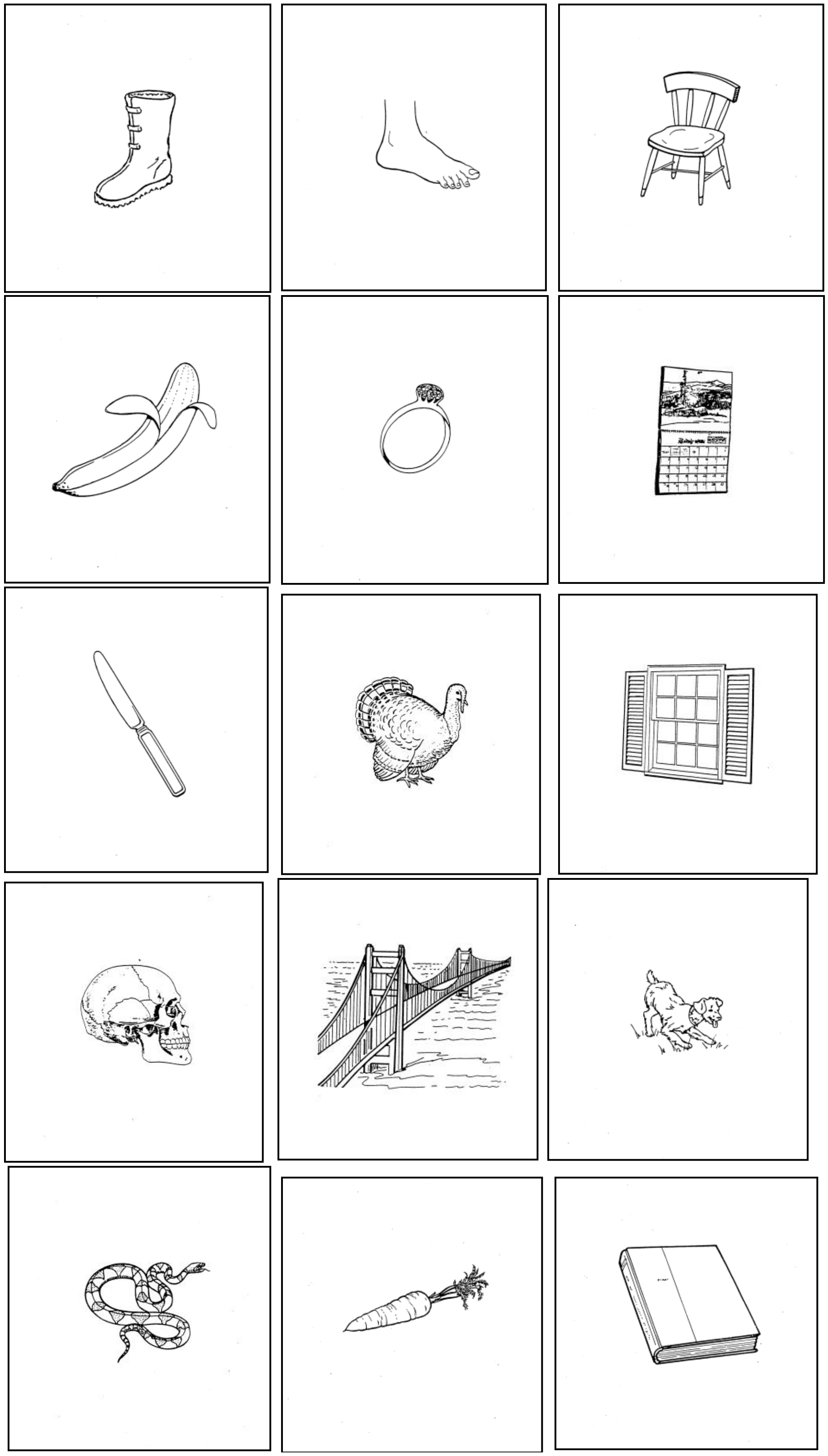


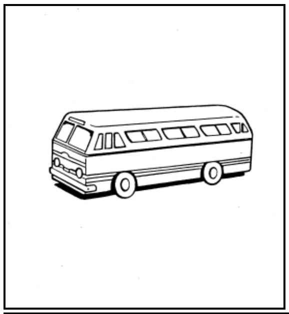
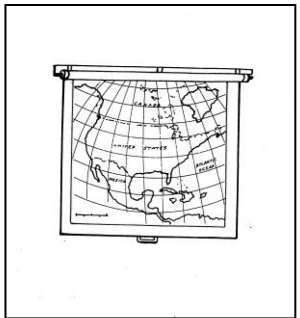
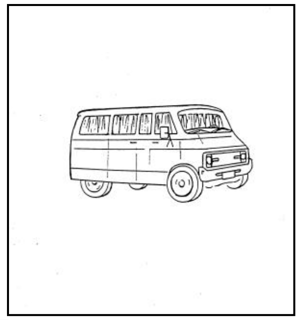
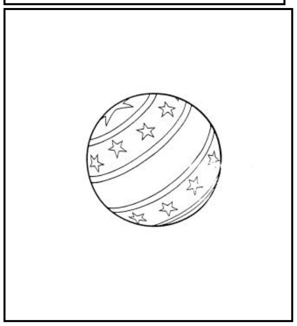


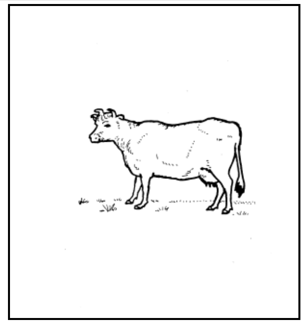










APPENDIX C

NARRATION TASK 1

Script:

-Researcher: After watching the following silent film you will be asked to retell what took place.

-(https://www.youtube.com/watch?v=P5MLKUnnT_A)

-Researcher: ¿Qué pasó?

APPENDIX D

NARRATION TASK 2

Script:

Researcher: What is your favorite memory? If you can think of a funny story, please tell me about it.

APPENDIX E

IRB PROTOCOL

Instructions and Notes:

- Depending on the nature of what you are doing, some sections may not be applicable to your research. If so, mark as "NA".
- When you write a protocol, keep an electronic copy. You will need a copy if it is necessary to make changes.

1 Protocol Title

Include the full protocol title: Lexical Access as a Predictor of Oral Fluency

2 Background and Objectives

Provide the scientific or scholarly background for, rationale for, and significance of the research based on the existing literature and how will it add to existing knowledge.

- Describe the purpose of the study.
- Describe any relevant preliminary data.

The purpose of the study is to investigate the relationship between 2 linguistic constructs: Lexical Access and Oral Fluency. Lexical access is defined as the process by which the mind retrieves, prepares, and sends words to the vocal apparatus for pronunciation. Oral Fluency is a measure of the rate of speech (amount of words pronounced in a given amount of time) by a speaker during spontaneous speech.

No preliminary data

3 Inclusion and Exclusion Criteria

Describe the criteria that define who will be included or excluded in your final study sample. If you are conducting data analysis only describe what is included in the dataset you propose to use.

Indicate specifically whether you will target or exclude each of the following special populations:

- Minors (individuals who are under the age of 18)
- Adults who are unable to consent
- Pregnant women
- Prisoners
- Native Americans
- Undocumented individuals

Will target: Students currently enrolled in SPA313, 5th semester Spanish, who are native monolingual speakers of English that are learning Spanish in a classroom setting. This subset may include, but is not limited to, adults who are able to consent, pregnant women, native Americans, or any other individual who is currently enrolled in SPA313.

Will exclude: Minors, prisoners, undocumented individuals

4 Number of Participant

Indicate the total number of participants to be recruited and enrolled*:* 20

5 Recruitment Methods

- Describe when, where, and how potential participants will be identified and recruited.
- Describe materials that will be used to recruit participants. (Attach copies of these documents with the application.)

Students will be recruited when investigator visits classrooms to request participation in study. The recruitment materials are attached (short consent form)

6 Procedures Involved

Describe all research procedures being performed and when they are performed. Describe procedures including:

- Surveys or questionnaires that will be administered. (Attach all surveys, interview questions, scripts, data collection forms, and instructions for participants.)
- What data will be collected including long-term follow-up?
- Lab procedure and tests and related instructions to participants
- The period of time for the collection of data.
- Describe the amount and timing of any compensation or credit to participants.
- If the research involves conducting data analysis only, describe the data that that will be analyzed.

LCP (Language contact profile)
Data will be collected via a computer and digital recorder, no long-term follow up.
Data will be collected when students are available, over weekends or after class time during the week in the computer lab of the G. Homer Durham Languages Building.
Students will not be financially compensated for participation, I estimate data collection will take 1 hour per student
The data will be collected via a picture naming task and a narration task. The picture naming task will take place at a computer where the students will be shown visual stimuli (single images) and asked to respond to the image by naming it as soon as they can. There will be approximately 30-50 images. There will be 2 narration tasks. The students will watch a short silent film (https://www.youtube.com/watch?v=P5MLKUnnT_A) and be asked to narrate what they saw and the second narration task will be a response to a question about a memory (What is your favorite memory? If you can think of a funny story, please tell me about it).
The images that will be used for the picture naming task are attached to the IRB submittal document under the Supporting Documents section. Not all of these images will be used, but they provide the sample from which the images will be selected.

7 Risks to Participants

List the reasonably foreseeable risks, discomforts, or inconveniences related to participation in the research. Consider physical, psychological, social, legal, and economic risks.

There are no potential risks to participants

8 Potential Benefits to Participants

Realistically describe the potential benefits that individual participants may experience from taking part in the research. Indicate if there is no direct benefit. Do not include benefits to society or others.

There are no potential benefits to participants

9 Prior Approvals

Describe any approvals – other than the IRB - that will be obtained prior to commencing the research. (e.g., school, external site, or funding agency approval.)

No other approvals are required

10 Privacy and Confidentiality

Describe the steps that will be taken to protect subjects' privacy interests. "Privacy interest" refers to a person's desire to place limits on with whom they interact or to whom they provide personal information.

Describe the following measures to ensure the confidentiality of data:

- Where and how data will be stored?
- How long the data will be stored?
- Who will have access to the data?
- Describe the steps that will be taken to secure the data (e.g., training, authorization of access, password protection, encryption, physical controls, certificates of confidentiality, and separation of identifiers and data) during storage, use, and transmission.

Data will be coded to ensure participant names are never disclosed or used in the publication of the investigation. The participants will be coded as follows; by sex and their number in each sex group so that the first male participant will be "M1" and the third female participant will be "F3." Only the principal investigator and assistant investigator will have access to the data. The data will be password protected on the personal computer of the assistant investigator, no other person will have access to this machine. The data will be stored for the duration of the study.

11 Consent Process

Indicate the process you will use to obtain consent. Include a description of:

- Where will the consent process take place
- How will consent be obtained

Non-English Speaking Participants

- Indicate what language(s) other than English are understood by prospective participants or representatives.
- If participants who do not speak English will be enrolled, describe the process to ensure that the oral and/or written information provided to those participants will be in that language. Indicate the language that will be used by those obtaining consent.

Waiver or Alteration of Consent Process (written consent will not be obtained, required information will not be disclosed, or the research involves deception)

- Review the "CHECKLIST: Waiver or Alteration of Consent Process (HRP-410)" to ensure you have provided sufficient information for the IRB to make these determinations.

Participants who are minors (individuals who are under 18)

- Describe the criteria that will be used to determine whether a prospective participant has not attained the legal age for consent to treatments or procedures involved in the research under the applicable law of the jurisdiction in which the research will be conducted.

Consent will be obtained when participants volunteer. After recruiting the participants (see attached recruitment document) the participants will be informed that they can withdraw from the study at any time without the potential for any penalties or negative action. Non-English speaking individuals will be excluded

12 Process to Document Consent in Writing

If your research presents no more than minimal risk of harm to participants and involves no procedures for which written documentation of consent is normally required outside of the research context, the IRB will consider a waiver of the requirement to obtain written documentation of consent.

(If you will document consent in writing, attach a consent document. If you will obtain consent, but not document consent in writing, attach the short form consent template or describe the procedure for obtaining and documenting consent orally.)
I will be using the short form consent template which has been modified for this study and is attached.

13 Training

Provide the date(s) the members of the research team have completed the CITI training for human participants. This training must be taken within the last 3 years. Additional information can be found at: http://researchintegrity.asu.edu/training/humans

The principal investigator completed the CITI training in September of 2011