THE PENNSYLVANIA STATE UNIVERSITY
SCHREYER HONORS COLLEGE

DEPARTMENT OF SPANISH, ITALIAN, AND PORTUGUESE

SPEECH PERCEPTION OF /S/-LENITING SPANISH DIALECTS BY LATE ENGLISH-SPANISH BILINGUALS

CAITLIN E. Y. REICHARD
SPRING 2020

A thesis
submitted in partial fulfillment
of the requirements
for baccalaureate degrees
in Spanish and Biology
with honors in Spanish

Reviewed and approved* by the following:

Matthew T. Carlson
Associate Professor of Spanish and Linguistics
Thesis Supervisor

John Lipski
Edwin Erle Sparks Professor of Spanish and Linguistic
Honors Adviser

* Signatures and electronic approvals are on file in the Schreyer Honors College.

ABSTRACT

This thesis investigated speech perception of Spanish dialectal variation to word-final /s/ by late English-Spanish bilinguals. The experiment tested low proficiency Spanish (L2) learners' sensitivity to the sibilant, aspirated, and deleted variants, which represent a lenition pattern of Spanish /s/. Aspiration is realized by an audible [h] sound, and deletion induces traceable, phonetic changes to neighboring sounds, which shall distinguish plural and singular stimuli. Participants completed a four-picture visual world paradigm with eye-tracking to assess plurality of auditory phrases as an implicit test of /s/ perception. Results showed categorical responses to singular, sibilant, and deleted stimuli: canonical [s] was perceived as "plural", while deletion was perceived as "singular". Accuracy for aspirated stimuli revealed distinct subgroups: a few participants who responded "plural" to most aspirated tokens and the majority who responded "singular" to such tokens. The former showed direct correlation between dialect contact and perception, and there was some evidence that one of these participants changed her behavior in favor of "plural" responses as the task progressed. Response times and eye-tracking data told a similar story as accuracy, in which some participants were sensitive to [h] and none were sensitive to the phonetic cues in a deleting dialect. Overall, this sample indicates low proficiency learners, with no prior dialect contact, are unlikely to adapt to L2 variation upon initial contact with a native speaker.

Keywords: second language acquisition, speech perception, dialectal variation, perceptual learning, bilinguals, Spanish

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

ACKNOWLEDGEMENTS

Endless thanks go out to Dr. Matthew Carlson for letting me explore my intellectual curiosity in Spanish linguistic, since my background had largely been biology-focused. Thank you for your active guidance, positivity, and endless patience as we navigated this project. I also would like to thank Dr. Katrina Connell for her hours of technical assistance with the experimental design. They both were always available, incredibly enthusiastic, and provided thorough advice and feedback. I really value the research process for challenging me to think and problem solve, but I'm especially for their expertise and encouragement in making this such memorable and rewarding learning experience. Additional thanks are extended to Lucía Pérez de Zabalza, Juliana Cruz Martínez, and Gabriela Puscama for taking time to lend their voice for the audio recordings. Finally, I would like to thank the Schreyer Honors College and the College of the Liberal Arts Career Enrichment Network for providing enrichment funding for this study.

Introduction

Comprehending speech is complicated. Listeners are surrounded by a diverse environment of stimuli that they must be able to interpret, from ambient noise to regional speech variation to the idiosyncrasies of individual speech production. Accurate interpretation of these sounds and the phonological components of a language are critical for optimal listening comprehension. Furthermore, certain sounds in a language system may be realized in different or multiple forms depending on the speaker's geographic region, age, gender, or socioeconomic status. Contextual factors surrounding how the speaker may want to be perceived by certain social groups or the formality of the occasion also influence which speech sounds may be produced in a given setting. This is in reference to one's native language (L1); these challenges are compounded by learning additional languages. Second language (L2) learners must not only decipher such sounds in their native language but integrate and categorize the sounds of their second language, as well. This begs the overarching question: how does a second language learner navigate the multiple uses of their second language? The present study investigated this by confining its focus on speech perception of a regional variant of Spanish by English-Spanish bilinguals.

Speech perception refers to a listener's interpretation of acoustic speech signals. Speech perception in a second language is expected to be difficult when listeners learn new sounds that may differ from or not exist in native speech categories. This problem frequently occurs while learning a second language. However, not only may the L2 sound categories differ from a learner's L1 categories, but they also vary across different communities where the L2 is used.

Regional and other such variability may make it harder for listeners to perceive L2 sounds if it requires them to generate new categories, or shift or divide native speech categories. Interestingly, even when L2 sounds resemble native speech sounds, speakers may articulate them in ways that vary from the anticipated realizations. A particularly useful example for studying sound variation is through *coda /s/ lenition* in Spanish. In some varieties of Spanish, the /s/ sound at the end of a syllable (the coda position) or word is weakened, or lenited. During this process, the coda /s/ can be realized across an acoustic spectrum of variants, from the canonical [s] to [h] to [Ø] (Erker, 2010). Although acoustic cues remain that allow native listeners to distinguish when the /s/ phoneme is present versus absent (even when it is supposedly "deleted"), this variation may be difficult to perceive for native English speakers learning Spanish as a second language since English does not map the [h] sound to the /s/ category or delete word-final /s/. Additionally, the word-final /s/ carries morphological importance in marking plurality and subject agreement in verbs. This variability may cause English-Spanish bilinguals to miss critical phonetic cues that consequently impact comprehension. Finally, coda /s/ lenition represents up to 50% of the existing Spanish varieties (Hammond, 2001). This adds to its importance in understanding how learners respond to this dialectal variation. Recent studies have been completed on English-Spanish bilinguals' perception of coda /s/ aspiration in Spanish (Schmidt, 2018; Saz, 2019). Both studies support L2 learners' sensitivity to [h] and focused on the influence of dialect contact factors on listeners' perception. The current investigation built on this study by expanding the range of variation to include three word-final /s/ allophones, [s, h,

∅], and eye-tracking measurements that may show deliberation in perceptual processes beyond listeners' final categorization.

Research supports listeners' ability to adapt to speech variation in novel scenarios, such as with regional accents. Studies by Kleinschmidt and Jaeger (2015) and Norris et al. (2016) have described that listeners refer to previous experiences to generalize, make predictions, and interpret phonetic variation to optimize their listening comprehension. However, consider the average second language learner, who may enter a study abroad program with limited native L2-speaker experiences to draw upon. How do they generalize, predict, and interpret L2 variants uttered by native speakers on initial contact? Kleinschmidt and Jaeger's "generalizing to the similar and adapting to the novel" continues to offer a promising explanation. Instead of relying on previous (limited) L2 experience, listeners may start to refer to similar sounds in their native phonology to interpret L2 variants. In reference to this thesis, the Spanish /s/ can be represented by both [s] and [h], which may be similar enough to English sounds for listeners to generalize them to the Spanish /s/ and, consequently, learn to adapt to this dialectal feature. The Spanish /s/ can also be realized as a deleted [∅] variant. While this variant retains traceable, acoustic cues, this subtle distinction is predicted to be notably challenging for L2 learners to perceive since there is no fricative noise where they would expect. The current investigation's participant pool reflected that many late L2 learners, who started learning their second language at or after adolescence, are taught in classroom settings that often have limited exposure to regional variation.

This thesis evaluated perception and adaptation to regional variation of word-final /s/ in Spanish. It exposed low proficiency Spanish (L2) learners to three variants: a sibilant, aspirated, and deleted form. They were not only tested on perceiving the variant, but also placed in conditions where they might learn to better perceive the variants. This was performed with a four-picture visual world paradigm and eye-tracking to assess plurality of auditory phrases as an implicit test for word-final /s/ perception.

The paper is organized as follows: The first section discusses previous research of L2 learner's acquisition of dialectal variants, predictions suggested by L2 learning models, the role of perceptual learning in L2 acquisition, and the regional variability of coda /s/ in Spanish dialects. The following section describes the materials, methods, and procedures of this investigation. Then, the predictions and experimental results are presented, succeeded by a discussion of their interpretation in context of the initial research questions. Finally, the paper concludes with the study's significance in L2 learning research and how it can be expanded upon in the future.

Background

Second Language Learners' Acquisition of Dialectal Variants

Previous research on the acquisition of dialectal variants in the second language have mostly focused on speech production, but not perception. This is frequently in the context of the influences of studying abroad on production of local, regional variants by L2 learners (Wolfram, Carter, and Moriello 2004; Drummond, 2012; Henriksen, Geeslin, and Willis, 2010). These studies support the idea that second language learners can potentially acquire dialectal variants in their speech production. Although only a limited number of individuals may actually exhibit this in their L2 speech, as observed by Wolfram, Carter, and Moriello (2004) of the unglided /ai/ diphthong of Southern American English in native Spanish speakers residing in North Carolina. Similarly, Fox and McGory (2007) studied native Japanese speakers, learning English as their L2, that had lived in Alabama for at least 2 years. They found that only two of ten English vowels were pronounced like that of the Southern American English variety.

This may be interpreted as weak support for acquiring L2 variants in production. Yet, production is arguably subject to many sociolinguistic factors as speech – like writing, music, or art – is another mode of expression. For instance, certain speech patterns may be more (or less) appropriate based on formality of occasion or may be stigmatized by association with social or economic status. When evaluating the incorporation of L2 variants into one's speech, the speaker ultimately dictates production to control how they are viewed by others. This makes it less reliable in understanding L2 acquisition. The more relevant investigation may lie upstream of production in *perceptual* processes. Consider the common scenario in which an L2 learner, who only had classroom exposure, studies abroad for the first time in the nation of their second language. Alternatively, consider an individual immigrating to the United States with a limited education in English. In both scenarios, the L2 learner must perceive L2 sounds produced by native speakers for the first time. These sounds may also differ from how the learner was taught the language due to regional variation. Perception of variants becomes harder if they resemble sounds in the listener's native phonology, but are used differently in the L2. Thus, the ability to acquire dialect variants in the L2 becomes important to an aspiring second language user.

Previous studies offer evidence of perception to L2 dialectal variants by learners. Categorization tasks in which a listener evaluates vowel contrasts that reflect different L2 dialects demonstrate perceptual judgements tend to have highest accuracy for the dialect of most exposure or dialects with distinct phonetic cues. Such studies have been performed with dialects of English, German, and French as the L2 (Escudero and Boersma, 2004; Baker and Smith, 2010; Smith and Baker, 2011). Studying L1 variation through foreign-accented speech is also useful in understanding acquisition of variants. Bradlow and Bent (2007) performed a study examining native English speakers' comprehension of Chinese-accented English. Results showed that listeners are not only capable of adapting to foreign-accented speech sounds, but were more successful perceiving the Chinese-accented speech when trained with multiple speakers. This indicates that exposure to a larger variety of speakers, who each have individual, idiosyncratic variations but exhibit the same foreign-accent, potentially increases adaptability to any new speakers that listener may encounter with the same accent.

Recent studies explicitly support second language learners' ability to perceive L2 dialectal variants. Schmidt (2018) conducted a study on native English speakers learning Spanish

(L2) in a traditional foreign language program with particular interest in how dialect contact influences perception. She evaluated their perception of syllable-final /s/ aspiration, a widespread variant in Spanish dialects. She divided her participants into five levels based on Spanish proficiency, which correlated with study abroad experience and contact with /s/ aspirating dialects. Results supported the robust effect of experiential factors on the categorization of aspirated /s/ by L2 learners. Saz (2019) also evaluated L2 perception of aspiration in Western Andalusian dialects of Spanish in accordance with studying abroad and had similar findings that learners can perceive variation in the L2, even with limited study abroad experience.

The aforementioned studies are insightful, yet they provide limited understanding of L2 dialect perception at the point of initial contact with a native speaker, which is an arguably common situation in the active and social world we live in. The next research consideration is how language learning develops when there is limited exposure to regional varieties, as is often the case of late L2 learners taught in classroom settings. How do learners identify these speech sounds initially, if they differ from what they were formally taught? As late bilinguals, who started L2 acquisition in adolescence or adulthood, what role does the native language play in identifying these L2 varieties that may correspond or differ from native phonology? The current study aimed to answer these questions by building upon the final two mentioned studies.

Models of Speech Perception in Second Language Learners

Second language learners are frequently introduced to speech sounds that may be different or non-existent in their native language. Models in speech perception were first developed to explain an inexperienced or naïve listener's perception of non-native speech sounds. These models describe speech sounds in a *phonological space*, a region in which the categorical representation of sounds in a language or multiple languages are organized based on their phonetic properties. These models use this concept as a basis for making predictions about the degree of difficulty a listener may experience when perceiving certain speech sounds. Nevertheless, let's first review frequently referenced symbols and terminology to aid in the discussion of such phonological models. Throughout this document, slashes (/ /) will enclose *phonemes*. This is the categorical representation (also known as "*phonetic category*") of a speech sound from which we, as language users, derive meaning from a sound. Brackets ([ ]) will enclose phonetic transcriptions. This is how the *speech sound* will be realized by the speaker acoustically. Some speech categories have a one-to-one relationship with a specific sound, while other speech categories, like Spanish word-final /s/, can be realized in multiple forms. Each sound can be called an *allophone* of that speech category. Next, this section discusses popular models in L2 speech perception: The Speech Learning Model (SLM), the Perceptual Assimilation Model (PAM), and the Second Language Learning Perception model (L2LP). Then, it makes general predications based on the phonological theories from the models in relation to the present investigation.

Flege (1995) proposed the **Speech Learning Model (SLM)** to describe the organization of speech sounds in the bilingual mind to explain production, as well as offer some predictions about the accuracy of L2 speech perception. Best (1995) provided the **Perceptual Assimilation Model (PAM)** to outline the perceptual patterns by naïve listeners when introduced to non-native speech sounds, which was later modified to cover subsequent L2 learning in the PAM-L2 (Best and Tyler, 2007). Escudero (2005; 2009) expanded upon both the previous models with the

Second Language Learning Perception model (L2LP) by explaining perception of L2 sounds across a spectrum of bilingual listener proficiency levels. The PAM and L2LP are most similar of the three models and will be discussed together before introducing the SLM.

The L2LP is founded on the premise that second language learners "duplicate" their L1 perceptual system of sounds and categories as a basis from which they will interpret their second language (Escudero, 2005). This interferes in learning a new language as the learner applies certain L1 characteristics to the L2. It can also facilitate learning when L2 properties are similar to that of the L1. The L2LP offers three predictions for perception based on the integration of two languages: (1) the "new sounds" scenario, (2) the "similar sounds" scenario, and (3) the "subset" scenario. In the *new sounds* scenario (known as the *single category assimilation type* in the PAM), two L2 sounds are mapped to one L1 sound. This is evident in the /r/-/l/ challenge for Japanese speakers. The English postalveolar approximant [ɹ] or alveolar lateral approximant [l] do not cleanly match the Japanese alveolar tap [ɾ]. Thus, Japanese speakers will map them to the [ɾ] in the absence of closer matches (Bradlow, 1997). In contrast, the *similar sounds* scenario (known as the *two-category assimilation type* in the PAM) describes two L2 sounds that map to two different L1 sounds. For example, Spanish [s] to English /s/ and Spanish [h] to English /h/. Finally, the *subset* scenario describes when there are fewer L2 categories than L1 categories (van Leussen and Escudero, 2015). This relates to PAM's *uncategorizable* or *categorizable-uncategorizable assimilation types*. This is evident in the case of native English speakers perceiving Spanish aspiration. English (L1) has separate /h/ and /s/ categories that need to assimilate with a single /s/ category in Spanish (L2).

The L2LP and PAM predict that the *subset* scenario is less challenging than the *new sounds* scenario, which is the most challenging for L2 learners. Recall that the L2LP postulates

L2 learners "duplicate" their L1 as a reference point for interpreting their L2. Thus, the *subset* scenario merely requires a division of the L2 category, the phonology which is less stable for late bilinguals, into the L1 categories. In contrast, the *new sounds* scenario requires generation of an entirely new L2 category or division of the existing (foundational) L1 category. This task was empirically demonstrated by Escudero et al. (2014) to be more challenging than simply shifting pre-existing L1 sounds. The SLM makes a similar claim, explaining the further an L2 sound is from the closest L1 sound in phonological space, the more likely a new category will be created. It also states as L1 phonology is developed throughout life, the development of new L2 categories becomes less likely. In other words, the L2 category will assimilate to, or merge with, the stabilized L1 category (Flege, 2005), which corresponds with the dominant role of the L1 pointed out by the L2LP.

The current investigation studied word-final /s/ perception by native English speakers learning Spanish as a second language. Participants listened to native Spanish speakers, simulating three dialectal variants for word-final /s/, [s, h, ∅], and performed a plurality test to assess their perception of the variants. The claims made by the L2 learning models mentioned above helped develop the following predictions for dialectal variant perception in the L2. **Figure 1** shows how the three allophones map to certain phonemes to aid in discussing these predictions.

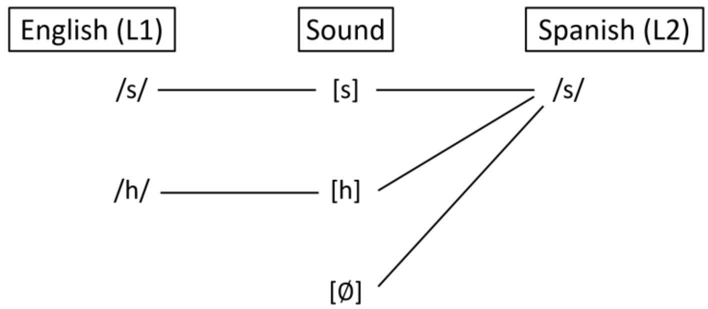

Figure 1. Basic map of this study's phonetic categories

The L2LP, PAM, and SLM would predict the sibilant [s] and aspiration [h] to be less challenging for English-Spanish bilinguals than deletion [Ø]. The former two sounds replicate the "subset scenario" for listeners, where they must assimilate L1 categories /h/ and /s/ to a single, Spanish /s/ category. Both [s] and [h] are allophones representing Spanish /s/ that exist in English phonology, albeit English /s/ is realized by [s] but not [h]. (**Figure 1**) Because these categories already exist in a combined phonological space, they are expected to be easier to re-adjust than in the "New scenario", where a category must be generated by either creating a new one or separating an existing one. Additionally, [s] is predicted to be easier for L2 learners to perceive than [h]. The canonical [s] is present in both English and Spanish and similarly used to mark plurality. They overlap in phonological space, so categorical adjustments are little to none.

Deletion [Ø] shall be the hardest as it replicates the "new scenario". Word-final /t/ and /d/ deletion have been studied in English (Bayley, 1996), and colloquial omission of word-final <g> is possible in English gerunds (e.g. "how you doin'?"). However, word-final /s/ deletion is not observed in English, so the English-Spanish bilingual must first produce a new category for the phonetic cues of a deleted variant, then learn to associate those phonetic cues to an L2 category. In addition to such cognitive load, the L2 learning models describe that development of second language categories becomes harder as a late L2 learner, in general. Thus, it's reasonable to expect a deleting dialect would be more challenging for L2 learners to perceive than an aspirating dialect. These predications are further supported by acoustic energy where [s] and [h] are both consonants characterized by aperiodic vibrations, while the deleted variant does not produce vibrations, making it naturally harder to perceive by listeners. Though this is not to say deletion is imperceptible, as it has been studied in Puerto Rican dialects (Hochberg, 1986). Deleting variants are generally accompanied by other phonetic changes, such as in the durations or stress of neighboring sounds that native users are sensitive to (File-Muriel and Brown, 2011). Thus, it was an aim of the current study to observe if English-Spanish bilinguals are sensitive to this fine acoustic cue. Nevertheless, it was predicted that [s] shall be easiest for L2 listeners to perceive, [h] shall pose an intermediate challenge, and [∅] shall be most difficult.

Perceptual Learning

The SLM noted that as L1 phonology becomes stabilized throughout life, L2 categories are less likely to form, and L2 interpretation relies more on pre-existing L1 phonology. Though this is not to be a hinderance, as the model also notes that increased L2 experience helps refine associations between L1 and L2 phonology. Strengthening these associations are likely to be a more gradual than rapid process for the late bilingual, yet there must be individual moments of L2 development. For the average L2 learner, with only classroom instruction in a foreign language, one of these moments may be encountering their first native L2 speaker, who realizes sounds different from what the learner studied. The previous models explain that the listener will create and shift phonological categories to make predictions about the L2 sounds, but does not elaborate as much on how the listener may adapt their phonologies to account for L2 variation. This is important as speech is highly variable due to factors from an individual speaker and between multiple speakers. A speaker aims to convey some target speech sound, yet its realization is subject to physiological features of the vocal apparatus, speech rate, intonation, and overall linguistic experience. Contextual factors such as regional accent, education level, and social identity can also influence speech between individuals.

Adapting to variation may be explained by perceptual learning, which refers to the plasticity of the perceptual system to adjust to speech sounds that are difficult to interpret. Studies have demonstrated such adjustments by exposing participants to a sufficient number of stimuli and assessing the accuracy of their perception over time (Greenspan, Nusbaum, and Pisoni, 1988; Dupoux and Green, 1997). Kriengwatana et al. (2016) found that listeners find it harder to adapt to variation from sociolinguistic sources (accents) than that of anatomical sources (sex differences). Researchers have also evaluated perceptual learning to foreign-accented speech when accents are consistent (Clarke, 2002; 2003) and inconsistent (Witteman, 2014). Participants in both studies showed quick adaptation when accents were consistent in the L1. However, there are limited studies of perceptual learning of dialect variation in the L2. This thesis hoped to explore this area by testing English (L1)-Spanish L2 learners' perceptual ability to three dialects of Spanish.

It's also worth noting linguistic experience is closely associated with perceptual learning. Numerous studies have been published supporting the role of dialect familiarity on perceptual ability in both monolingual and bilingual populations. Some indicate how the standard variety of languages is easier for listeners to perceive (Sumner and Samuel, 2009; Eisenstein and Berkowitz, 1981), while others have reported that L2 learners better perceive dialects that are similar to their L1 (Brown, 1968; Wilcox, 1978). Ultimately, they indicate that L2 learners are more likely to be better at perceiving the dialect to which they have the most exposure (Tauroza and Luk, 1997). Experience also is a critical component in probabilistic inference theories by Kleinschmidt and Jaeger (2015) and Norris et al. (2016). They describe that listeners refer to previous experiences to generalize, make predictions, and interpret phonetic variation to optimize their listening comprehension. They also highlight the role of word frequency, syntax, semantics, and pragmatics in helping us to make these predictions.

What requires further development is understanding how this model may hold on initial contact, when listeners' linguistic experience with native L2 sounds is limited. This becomes a relevant scenario when the listener is in a new dialectal environment, such as moving to a more diverse community or traveling to a new country. Many late L2 classroom learners are taught in settings that often have limited exposure to regional variations and aspire to study abroad for immersion in their developing L2. When faced with a new L2 dialectal variant, these learners may, instead, start relying on similar sounds in their L1 for interpretation. This study aimed to observe perceptual learning in the average second language learner by testing low proficiency Spanish learners' initial response to three word-final /s/ variants exhibited in Spanish dialects: sibilant, aspirated, and deleted. In regards to probabilistic inference theories, the Spanish /s/ can be represented by both [s] and [h], which may be similar enough to English sounds for listeners to generalize them to the Spanish /s/ and, consequently, learn to adapt to this dialectal feature. These conditions create a "novel scenario" in which participants will not only be tested on perceiving the variant, but also placed in conditions where they may learn to better perceive the variants. This occurred through a visual world paradigm that measured eye movements in which low proficiency Spanish listeners heard a native speaker, producing one of three word-final /s/ varieties, and selected the image with correct plurality.

Regional Variability of Coda /s/ in Spanish Dialects

The *coda* position refers to the consonant appearing at the end of a syllable. In Spanish, the coda /s/ can be realized across a spectrum of allophones from full sibilant [s], aspirated [h], or deleted [∅] variants, depending on the variety of Spanish being used. The full sibilant is produced by placing the tip of the tongue in the alveolar region, while the rest of the tongue is in a concave shape. The aspirated and deleted forms are realized from a weakening in the articulatory force, which will be referred to as *lenition*. Example transcriptions are provided below. (Recall that slashes (/ /) enclose phonological transcriptions of the categorical representation of a speech sound. Brackets ([ ]) enclose phonetic transcriptions of how the speech sound will be realized.)

Example 1.1	/gatos/	[ga.tos]	sibilant
Example 1.2	/gatos/	[ga.toh]	aspiration
Example 1.3	/gatos/	[ga.to∅]	deletion

/s/ lenition is a widespread dialectal variant in the Spanish language (Hammond, 2001), posing a relevant challenge to English speakers learning Spanish as a second language and makes it prime for investigation. Latin America can be divided into the low lands and the highlands, linguistically. Low land regions often exhibit leniting dialects, while highland regions – such as Mexico, Central America, and the Andes – and many national capitals tend to exhibit the sibilant dialect (Schwegler et al., 2010). In other regions of the Spanish-speaking world, many Caribbean Spanish varieties exhibit the aspirated variant, while the southern Andalusian region of Spain is known for its deleted variant (Schwegler et al., 2010). It should be noted that while such regions have been generalized to have one of three obvious allophones [s, h, ∅], in actuality, the realization of word-final /s/ is on a continuum (Erker, 2010). For instance, evaluating the acoustic wave reveals notable differences between similar sounds. This is apparent even during deletion, where it is possible to see acoustic traces of [h].

/s/ lenition patterns can be influenced by formality, socioeconomic status, geographic origin, age, and social identity, but also the position of /s/ in the word (Schwegler et al., 2010). If the coda /s/ is positioned within the word, aspiration is more likely. In contrast, if the coda /s/ is positioned at the end of the word, deletion is more likely (Terrell, 1979). It is also influenced by placement beside consonants or vowels. This study will focus on word-final coda /s/ examples that may be easier for low level Spanish learners to recognize since similar final /t/ or final /d/ deletions occur in English (Bayley, 1996). Word-final /s/ also holds morphological significance in Spanish. It denotes the second-person singular verb form and is the plurality marker. Additionally, this study focuses on word-final /s/ because a plurality assessment is a simple way to measure listeners' perception of the various /s/ allophones.

It becomes clear that the phonetic variability from regional dialects may make it difficult for native English speakers learning Spanish as a second language to perceive word-final /s/. Consequently, English-Spanish bilinguals may miss critical verbal acoustic cues that possibly impact their comprehension of person or number marking. The /s/ variant has three allophones, [s, h, ∅], which may be a perceptual challenge to English speakers learning Spanish as an L2 since not all the allophones correspond to the word-final /s/ category in English. Overall, because regional variability of word-final /s/ is overwhelmingly prevalent in the Spanish-speaking world and essential to listening comprehension, it is exactly this perceptual challenge that the present study wished to investigate.

Research Questions

Interpreting speech stimuli is complicated due to variation from a range of physiological, social, and contextual factors that include geographic region. Learning a second language poses an additional challenge as L2 speech sounds may not exist or be perceived as valid tokens of native speech sounds. Similarly, an L2 learner with limited exposure to variation in the L2 may ignore or misinterpret valid tokens of the L2 itself. This scenario is common to the native English speaker who learned Spanish from nonnative Spanish speakers and is exposed to native Spanish speakers for the first time. This thesis aimed to investigate the following, using word-final /s/ lenition as a tool for study.

1. How do low proficiency Spanish learners respond to different realizations of word-final /s/ lenition as spoken by three different native Spanish speakers?
2. Can low proficiency Spanish learners exhibit perceptual learning of this dialect variant?
3. What insights does this provide to dialect perception during initial contact scenarios?

Methods

Participants

Fourteen late bilinguals, who were native American English speakers studying Spanish as a second language at the Pennsylvania State University, participated in this study (mean age 19.4 years, SD = 1.1; 10 females, 4 males). Participants routinely started learning Spanish in middle school, except for one participant that started learning Spanish in high school. They described their Spanish classes before entering university as being taught by predominantly nonnative Spanish instructors. They completed an average of 2.54 Spanish classes at the university level (range = 0-4 classes) and represented low L2 proficiency levels. The participant pool included 1 Spanish major, 9 Spanish minors, and 4 with other degree interests. They were not proficient in any additional languages. The study excluded students who had taken linguistics classes or had extensive exposure to native Spanish speakers. This excluded heritage speakers and those who studied abroad for six weeks (or longer) in Spanish-speaking countries. Data from 3 participants were dropped for being trilingual, having completed a Spanish linguistics course, and extensive experience with native Spanish-speakers. Participants were recruited from Spanish classes and campus organizations via in-person visits and flyers, starting at 200-level courses which represent the introductory Spanish courses. Individuals from the 300-level were included, as needed, to try to achieve sufficient sample size, and these individuals strictly met the previously described profile. They received $12 for their participation in a 75-minute session.

Language proficiency was assessed using a lexical decision task, Lextale-Esp, of 90 Spanish words and non-words. The assessment itself was empirically tested and reported to be an accurate measure of Spanish proficiency (Izura, Cuetos, and Brysbaert, 2014). Participants were asked to indicate whether they thought each item was a real word or non-word in Spanish by hitting “sí” or “no” on the button box. This task was implemented via E-Prime 3.0 software (Psychology Software Tools, Pittsburgh, PA) on a Windows computer. **Table 1** shows the LexTale scores for each participant. Grey boxes correspond to data from the 3 participants that were dropped for not meeting profile criteria. Scores fit a normal distribution centered between 0 and 10. According to Izura et al. (2014), a normalized score of 0 corresponds to guessing. It is unlikely such a large, negative value would be possible, and it is suspected that Participant 7 may have switched the buttons on the response box, though this is hard to say for sure. Participant 13 was rescheduled as Participant 15 due to eye-tracking calibration problems. This study accepted her first LexTale score since she would have had prior exposure to the same list the second time she completed the task. No participants were excluded by LexTale performance.

Table 1. LexTale scores

Participant	**Normalized**	*Correct reject*	*False alarm*	*Hit*	*Miss*
1	**19**	28	2	23	37
2	**6**	18	12	30	30
3	**11**	23	7	25	35
4	**-3**	16	14	25	35
5	**11**	26	4	19	41
6	**11**	20	10	31	29
7	**-26**	1	29	32	28
8	**4**	24	6	16	44
9	**5**	24	6	17	43
10	**6**	17	13	32	28
11	**5**	19	11	27	33

12	**8**	20	10	28	32
13	**9**	22	8	25	35
14	**1**	20	10	21	39
15	**4**	22	8	20	40

Materials

Auditory stimuli

Word selection. The sentence structure "Pincha en su(s) (target noun) (adjective)" (*Click on his/her…*) was intentionally selected. Specifically, possessive pronouns *su* and *sus* avoided vowel cues towards plurality that would occur with masculine articles *el* and *los*. Adjectives were included to offer a phonetic cue towards plurality for the lenited forms as research indicates coda /s/ deletion is often accompanied by changes to the durations of neighboring sounds. For example, the first consonant of the following word or syllable is lengthened, in some dialects (File-Muriel and Brown, 2011).

Target nouns were generated from memory and easily imageable objects. The EsPal Spanish Lexical Database was used to order the list from most to least frequent. A subjective cut off was made for low proficiency Spanish learners, where words lower than that relative frequency were deemed unlikely to be known by our population. The top 180 most frequent images were used as the target nouns.

Target nouns ended in vowels, which results in a simple -s morpheme instead of -es morpheme when marking plurality that is consistent across the entire word items. They did not contain stress on the final syllable to avoid interference with phonetic cues that might appear as a result of lenition. The majority of these nouns were two to three syllables long to control for word length. The phonetic restraints for target nouns are as follows: Word initial onsets and final syllable onsets began with obstruents to provide an identifiable point for acoustic measurements. There were no onset /s/ or internal coda /s/ in the target nouns. However, word internal onset /s/ were occasionally used as a final effort to generate word items, since they are unlikely to lenite (File-Muriel and Brown, 2011; Terrell, 1979). The carrier phrase *Pincha en…* did not include any /s/.

180 competitor nouns were selected after the target nouns from the noun list previously described. Competitor nouns and adjectives followed similar phonological constraints as the target nouns, but they were less restricted since competitor images will not be heard and adjectives merely serve as a phonetic cue. There was one adjective that included an internal coda /s/. If participants became sensitive to the dialectal feature, we anticipated that the accuracy for this particular item would be higher.

Recording and processing. Verbal stimuli were recorded by three different female speakers whose native language is Spanish. They were faculty or graduate students at the university, and they reported these variants as natural speech patterns for them. The speakers included a woman from northern Spain to produce the canonical [s] variant. A woman from Puerto Rico to produce the deleted variant, [Ø]. The third was a was Argentina to produce the aspirated variant, [h]. The deleted variant should create phonetic changes on neighboring sounds, as observed by File-Muriel and Brown (2011), which provide an acoustic cue for participants. The token of interest from each talker was verified by the appearance of (or lack of) acoustic waves corresponding to the spoken /s/ (**Figure 2**, red box corresponds to /s/ position). The complete sentence with the appropriate tokens were recorded for all 180 targets in singular and plural forms. All singular stimuli were recorded first, before recording the group of plural stimuli. Recordings were made in duplicates, and the best version was selected for the experiment. If stimuli were unfit for the experiment, the original speaker was asked to return for another recording session.

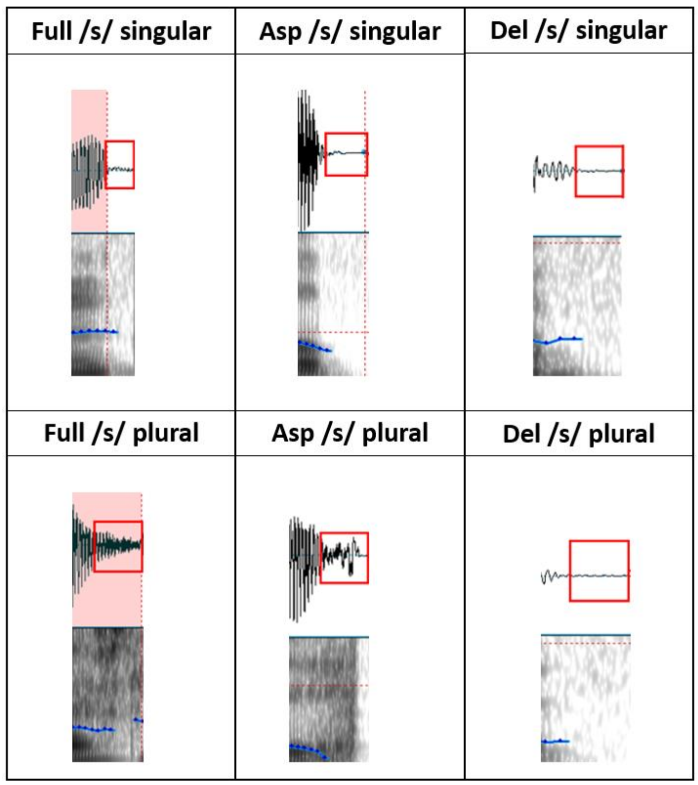

Figure 2. Acoustic features of variant tokens

All audio processing occurred in Praat on a Windows computer. First, all recordings were reduced to a single channel. The onset and offset boundaries of the sentences were marked to extract the desired recording and remove silences at the beginning and end of the recordings. These trimmed recordings were normalized to an intensity of 65 dB. The experiment used the best singular and plural version of each sentence. Finally, the onset of the “su(s)” was marked to provide an anchor for analysis of the eye-gaze at the beginning of the first word where participants have evidence for plurality.

Visual stimuli

Four pictures were used in the visual world instead of two pictures. Having competing images requires participants to interpret the images' lexical representations in addition to plurality. This helps conceal the plurality assessment to better measure if participants learn to perceive word-final /s/. **Figure 3** provides a sample of the visual world display during a trial. Black-and-white images were obtained from open source catalogs that could be modified such as Google images and the International Picture Naming Project. All images were stored in JPEG format. Experiment Builder modified all images to a uniform 350 x 350-pixel image in the visual world display. Images were selected with similar artistic design and in correspondence to target and competitor noun descriptions. This was to limit any effects that could distract participant gaze that were not intended. It should be noted that some of the adjectives used in this experiment described colors although images were black and white. This was irrelevant to selecting the correct image since competing images were distinct enough for low level Spanish learners.

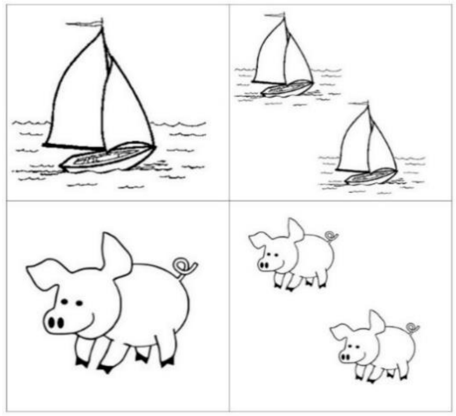

Figure 3. Sample visual world display

Eye-tracking apparatus

An EyeLink 1000 Plus (SR Research Ltd, Ottawa, Ontario, Canada) was used to measure eye movements. Samples were collected a rate of 1,000 Hz. Fixations that landed in a 450 x 450-pixel area of interest (AOI) around the 350 x 350-pixel image were considered valid looks to that image. Eye-movement data was from the right eye only. The EyeLink device placed a camera under the computer monitor, and an infrared light was used to record eye movements. Eye movements were interpreted as pupil size changes in relation to corneal reflection with ocular motion during each trial. It collected data from only the right eye with centroid pupil tracking. A headrest stabilized the forehead and chin to prevent large head movements that may cause inaccurate measurements. The device started recording data as the audio stimuli began, but sample analyses began at the onset of the word “su(s)” since this is the first cue to number. The exact onset times were measured with Praat at the appearance of the sibilant acoustic wave. The right eye movements, duration of gaze, and accuracy of mouse clicks were measured. The experiment was designed and data was collected through Experiment Builder 2.2.1 software (SR Research Experiment Builder 2.2.1, Mississauga, Ontario, Canada) on a Windows computer.

A 9-point calibration procedure occurred before starting the practice round and before each block. The researcher judged quality of calibration via the calibration plot, and if it did not meet recommended values, the calibration was repeated.

Procedure

Participants were tested in a sound attenuated booth in a single session. All tasks were completed on a laboratory computer after receiving verbal, informed consent. Trials began with a display of four pictures on the screen to prepare participants to their orientation for the upcoming trial. A centered fixation (+) was shown to direct eye-gaze to an unbiased image position before displaying the same four-picture arrangement and playing audio stimuli. The visual world remained on-screen until a mouse selection was made or reaching a time limit of 5000 ms. The time limit set to be longer than the average anticipated response times, but participants were encouraged to answer intuitively as soon as they were sure of the utterance. Audio was played through external speakers. After listening to the sentence, participants' task was to click on the picture they thought best correlated to the noun and plurality of the utterance. Responses were made with a mouse click an image. Eye movements were recorded for each trial.

Participants heard speakers of each dialect (sibilant, aspiration, and deletion) in three, separate blocks. Block order alternated with each participant, except the first block was always the sibilant speaker. Within each block, singular and plural tokens were split in half. Participants were neither explicitly informed of this nor about the plurality assessment. This was intentional to measure organic adaption to the dialectal variants. Items that were assigned singular or plural was also counterbalanced. Items were rotated across the three talkers (sibilant, aspirated, or deleted /s/) such that a third of participants heard a third of the items with sibilant /s/. A different third of participants heard the same third of items with aspirated /s/, and the final third of participants heard the same items with deleted /s/. Finally, target images were counterbalanced across the four quadrants of the display. No filler trials were included. No feedback was provided. **Figure 4** summarizes the experiment progression, trial blocking, and proportion of stimuli.

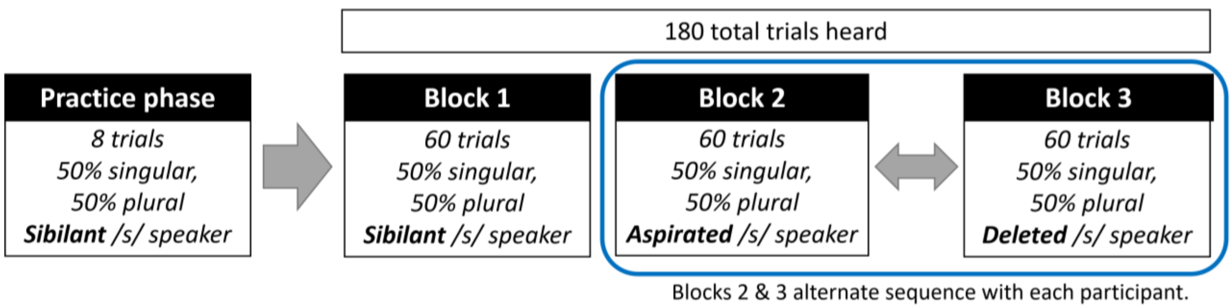

Figure 4. Schematic illustrating the flow of the experiment and proportions of stimuli.

The final two blocks were counterbalanced across participants. They either heard the aspirated speaker in the second block and the deleted speaker in the final block, or vice versa.

The experiment started with 8 practice trials to familiarize participants with the task. Practice trials were only by the sibilant speaker. Participants received on-screen and verbal instructions by the researcher in English. The practice nouns were reused in the experiment, but these tokens not necessarily uttered by the same speaker for each participant. No feedback was provided after responses. The visual world experiment alone lasted about 30 minutes.

Afterwards, participants completed a word familiarity task and LHQ. The word familiarity task presented participants with a digital word list of all 180 target nouns from the experiment. They viewed words on the computer screen, one at a time, with a slider appearing below each word. They rated their familiarity with the nouns on a scale of 1 to 10 (1 = “I’ve never seen this word before”; 10 = “I know this word well and can use it easily”). This task was created with E-Prime 3.0 software (Psychology Software Tools, Pittsburgh, PA) on a Windows computer. The language history questionnaire (LHQ) was completed as a form through Microsoft Word. It collected information on both their English and Spanish languages concerning factors such as age of acquisition, frequency of use, context of use, and contact with other varieties of Spanish. A copy of the questionnaire is included in **Appendix B** of this document.

Statistical analysis

Behavioral data was measured through mouse clicks. The dependent variable was mouse response, and the independent variable was the Spanish dialectal variant. Accuracy per condition was calculated as the number of trials the participant clicked the target divided by the total number of trials (for that condition). No statistical tests were performed for this measure since responses were essentially categorical across nearly all conditions.

Response time (RT) analyses excluded non-dominant responses in the full /s/, singular, and deletion conditions since this composed less than 1% of total responses per corresponding condition. RTs fit a skewed distribution that trailed off near 5000 ms, but no trials exceeded this limit and needed to be removed from analyses. RTs were log-transformed to approximate a normal distribution. RTs were analyzed using (1 predictor) four-category, linear mixed effects regression with the lme4 (Bates, Maechler, Bolker, and Walker, 2015) and lmerTest (Kuznetsova, Brockhoff, Christensen, 2017) packages in R. Residuals were roughly symmetrical and normally distributed. Random intercepts by participant and noun were included, but random slopes of condition could not be included due to lack of convergence. Significance was assessed using the Satterthwaite method.

Eye measurements were recorded through Experiment Builder. Fixations that landed in a 450 x 450-pixel area of interest (AOI) around the 350 x 350-pixel image were considered valid looks to that image. Data were analyzed in bins of 10 milliseconds, and fixations more than 2500 ms after onset of the critical portion of the sentence were discarded for visualization purposes. Calculation of the proportion of samples in each interest area was across all on-screen samples. This included samples outside of the AOIs to avoid inflating results. Data selection excluded samples during saccades and bins that contained non-fixation samples because the participant is unlikely to be analyzing visual information during those events. Statistical analysis was conducted on the proportion of fixations to the target image over the time window of 600 to 1800 ms after "su(s)" onset, as this is the first cue to plurality in utterances. The dependent variable was the proportion of fixations to target during this time window. (A separate analysis of looks to competitor was also conducted, and yielded the same results). Eye-tracking data were analyzed using a generalized linear mixed model with the logit link function (using the Laplace approximation) with glmerMod package in R. Significance was assessed via likelihood Wald Z.

Predictions

Low proficiency Spanish (L2) learners were anticipated to be most accurate when hearing the sibilant [s] form, of intermediate accuracy when hearing the aspiration [h] form, and least accurate when hearing the deletion [∅] form. Both dependent variables are expected to agree with these predictions, although eye movements should provide finer details on their decision-making process.

The sibilant [s] is both a phoneme in Spanish and English that represents /s/. Word-final /s/ is also similarly used as a plural morpheme in both languages, making it a highly familiar cue for participants to interpret. As such, the data for the [s] variant should represent baseline performance in word-final /s/ perception. Full /s/ stimuli should have high accuracy and quick response times relative to other conditions. A proportion of fixations (to target) graph for the full /s/ condition should show the curve starting at 25% since there are four pictures. Once the listener hears the first plural cue in “su” or “sus”, the curve should start to move towards 50% of fixations as the listener narrows in on either the singular or plural images. After hearing the noun, there should be a steep increase in looks to target. This slope should be greatly positive compared to curves for other conditions to reflect a quick decision on the image.

[h] is also a sound in both Spanish and English, although it represents /s/ in Spanish and /h/ in English. Participants should perceive this as a valid speech sound, but the L1-L2 competition and the fact that /h/ is relatively unexpected in this part of English words, could initially make it difficult for listeners to recognize the sound as Spanish /s/. However, the participant will have already completed a practice block and the first block with the “standard”, sibilant speaker, so it is predicted that listeners will anticipate plural stimuli and learn to map [h]

to /s/. Accuracy to aspirated stimuli should be high but less than that to full /s/ and singular stimuli. Response times to aspirated stimuli should be slower compared to the other conditions. The deliberation in [h] perception should be reflected in a more gradual positive slope on the proportion of fixations curve than that for full /s/. It will similarly start at 25% and most noticeably move towards 50% fixation to target after the listener hears the noun. The fixation curve should reflect strong commitment to the target in the end, yet it is expected to take longer due to L1-L2 competition.

The deletion [∅] condition was predicted to be most challenging for listeners of low Spanish proficiency. While English word-final /t/ and /d/ deletions are known to exist (Bayley, 1996), word-final /s/ deletions do not occur in the English language. With the lack of experience and limited acoustic cue (recall that the preceding vowel may exhibit acoustic cues that differ from the singular form of the word), listeners were not expected to be accurate beyond chance for this variant. Response times were predicted to be about the same speed as that of the aspirated condition. The fixation curve will start at 25%, but will not exceed 50% of fixations to target once the noun is uttered. It was expected that listeners will not notice any acoustic cues in neighboring syllables or the adjectival phrase given their limited language experience. This curve is expected to lie right of that for full /s/, reflecting a longer response time, similar to the aspirated condition. Another possible outcome could be that listeners become suspicious of the absence of plural stimuli and learn to detect the more subtle cues such as a period of silence where the /s/ should be, or changes in the duration or other features of the neighboring segments. Accuracy would still be low, though higher than if they though all stimuli were singular. Response times would be slower than those of singular or full /s/ stimuli, reflecting deliberation. The fixation curve would either have a more gradual slope and be more ambivalent, indicating moderate competition between target and competitor, or fixations would settle on the competitor, if participants take these tokens to be singular.

This experiment uses a pair of images corresponding to the target noun and a pair of images corresponding to an unspoken, distractor noun. It is possible that participants may click distractors if they do not know the target noun, although the study's vocabulary should be representative of low proficiency Spanish learners. Yet, they should still be able to determine if they though stimuli were singular or plural. Their plurality responses should be similar across target or distractor images for a certain speaker. In the end, data should reflect that limited language experience makes it challenging for L2 learners to correctly perceive L2 variants, but they may be able to learn to perceive variants that are similar to speech categories in their L1.

Results

5.1% of data were excluded from analysis due to incorrect mouse clicks to distractor objects. (The remaining 67.8% of clicks were to the target and 27.1% of clicks to the competitor.) An additional 21.41% were excluded from three participants for being trilingual, having previously taken a Spanish linguistics course, or extensive exposure to native Spanish-speakers. No trials were excluded based on the word familiarity assessment since there was no clear way of deciding how familiar participants should be with a word for it to be considered valid for this task. Exploring distractor data, in which participants selected the wrong image, revealed they answered “plural” only upon hearing [s] and answered “singular” upon hearing [h] or [∅]. This pattern was consistent when participants selected the correct image, as well. Ultimately, understanding the word has little influence on their responses, so it seemed reasonable to keep all the data.

Accuracy

Accuracy results were nearly categorical and undeniably distinct across all conditions, so no statistical analyses were performed. **Table 2** shows the total number of responses to target or competitor per condition and their corresponding percentages. Responses in the singular and full /s/ conditions were essentially at ceiling. This was expected since the [s] plural morpheme is similarly used in English. Responses in the deletion condition were essentially categorical with less than 1% answering “plural” for these stimuli. This minute percentage was similar to that for non-dominant responses in the singular and full /s/ conditions, so such responses were interpreted as clicks made accidentally or when participants were not paying attention. Participants seem sure: when [s] is present, stimuli are plural. When [s] is absent, whether deleted or truly singular, stimuli are interpreted as singular. Deleted variants are frequently accompanied by acoustic changes to neighboring sounds (File-Muriel and Brown, 2011) that were expected to differentiate singular and plural stimuli. We considered the possibility of listeners developing sensitivity to such an acoustic cue. This would've been evident by an accuracy that was higher than if they interpreted stimuli as singular, but lower than that for the aspirated condition. Results do not support this. Participants only responded "plural" to aspirated stimuli 21% of the time. This is inconsistent with the initial predictions as participants were expected to quickly learn to perceive [h] as a valid word-final /s/, and accuracy results were expected to be very high, albeit less than for the singular and full /s/ conditions.

Table 2. Accuracy of mouse clicks

	Condition			
Item Clicked	**Aspirated**	**Deleted**	**Full /s/**	**Singular**
Competitor	302 (78.44%)	397 (99.002%)	4 (1.01%)	5 (0.42%)
Target	83 (21.56%)	4 (0.997%)	394 (98.99%)	1,193 (99.58%)

Data for the aspiration condition were explored to determine if all participants performed similarly or if certain participants were more sensitive to aspirated stimuli. The majority (8/11) of participants nearly categorically answered "singular" for aspirated stimuli, despite the presence of [h]. (There was only one "plural" response out of all trials for this group.) In contrast, the remaining 3 participants answered "plural" 71.9% of the time. This warranted that the rest of experimental data should be analyzed as two separate subsets to reflect the distinct participant groups. Trial order effects for the latter 3 were examined (clicks to distractors were removed) to determine if they understood aspiration immediately or if they learned to perceive it with short exposure, and results were mixed. Participant 9 responded “singular” to aspirated stimuli in the first 7 trials but only answered as such twice in the remaining 21 trials. This strongly supports this participant started interpreting [h] as singular, quickly realized this was incorrect, and changed her response. Participant 3 responded “singular” for the first 3 trials, a group of 3 trials halfway through, and the very last trial of the block. These responses were more evenly distributed, but also seem to support a gradual realization that [h] represents a valid word-final /s/. Participant 2’s “singular” responses to aspirated stimuli are more randomly distributed to the first 2 trials, a gap of 15 trials with only 1 “singular” response, and 4 somewhat sporadically in the remaining 12 trials. This provides weakest support for adaptation, though she correctly answered “plural” for aspirated stimuli 75.86% of the time. Statistical analyses were not performed for trial order data due to limited sample size.

Their LexTale and LHQ data were reviewed for factors that might explain their performance. Participant 9 reported one of her Spanish courses at the university being taught by a native Spanish-speaker from Chile, a coda /s/ aspirating dialect. She also reported spending 8 hours per week listening to Spanish, which was significantly greater than the average 3.82 hours per week. Her LexTale score was 5, which falls near center of the normal distribution. Participant 3 completed 4 Spanish classes at the university by native speakers from Mexico, Spain, and Venezuela. Aspiration can occur in some dialects of the latter two countries. She directly uses the term “aspiration” in her LHQ, which hints at explicit understanding of this linguistic feature, and her LexTale score was 11, the highest score of participants whose data was kept. Participant 2 rated contact with Caribbean Spanish, an aspirating dialect, as “very often”. She also reported learning Spanish before elementary school, describing that her mother would read books and play songs for her in Spanish as a child (it is not clear whether her mother was a native Spanish speaker). She reported “never” using Spanish at home, however, and her LexTale score of 6 fell in the middle of the normal distribution. Overall, all these participants had notable contact with aspirating dialects that may explain their higher accuracy scores for such stimuli. Interestingly, Participant 3, with the highest LexTale score, only showed a weak learning effect, whereas Participant 9, with an average LexTale score, showed an obvious learning effect. Nonetheless, her high proficiency suggests it is unlikely she would need to revise her responses throughout the experiment.

Response time

Accuracy results showed a clear distinction between participants that are sensitive to aspiration versus those who are not. Thus, the remainder of statistical analyses treated them as separate groups. Since non-dominant responses in the full /s/ (FS), singular (SI), and deletion (DL) conditions composed essentially less than 1% of total responses per corresponding condition, they were removed from analyses. RTs fit a skewed distribution that trailed off near 5000 ms, but no trials exceeded this limit and needed to be removed from analyses. RTs were log-transformed to approximate a normal distribution. They were analyzed using (1 predictor, four-category) linear mixed effects regression with the lme4 (Bates, Maechler, Bolker, and Walker, 2015) and lmerTest (Kuznetsova, Brockhoff, Christensen, 2017) packages in R. Significance was assessed using the Satterthwaite method.

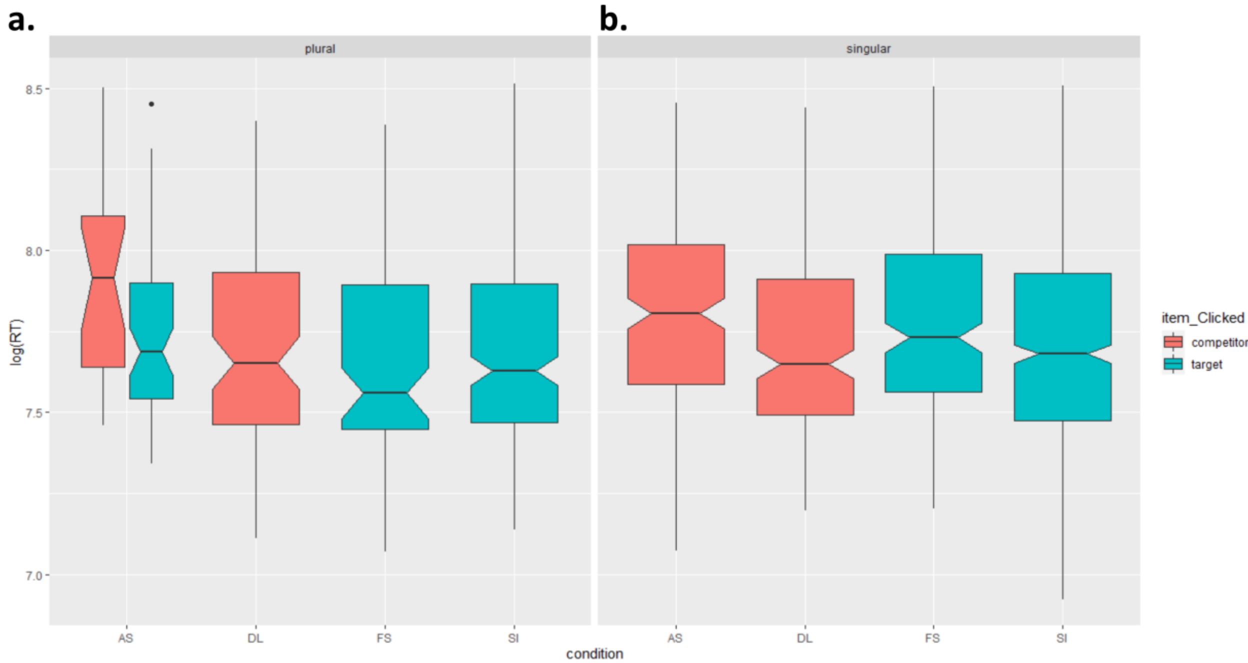

Figure 5. (a) Response times for subset sensitive to Aspiration (b) Response times for subset insensitive to Aspiration

Figure 5a shows the three participants who responded "plural" to aspirated stimuli. When they made clicks to the competitor, there was no difference in RTs between FS and SI conditions, nor DL and SI conditions. It was expected that RTs in the SI and FS condition would be relatively faster than the other conditions since this [s] and non-[s] usage is similar to English plurality markers. Predictions were inconsistent for the DL condition, as slower responses were anticipated to reflect contemplation. Instead, they appear confident that DL stimuli are singular and are quick to respond to stimuli they think represents canonical forms. RTs for the AS condition were significantly slower than those for the SI condition ($\beta = 0.24$, SE = 0.06, $t(421) = 4.26$, $p < 0.001$). This was also found in the majority of participants (**Figure 5b**), who were analyzed later. This could reflect that they were sensitive to [h], contemplated it, but ultimately decided it didn't represent /s/. Alternatively, it could be because it took longer to hear the noun in the utterance since “su[h]” takes longer to hear than “su”. Separate evaluation of clicks to the target showed that they responded “plural” to aspirated stimuli significantly slower than responses to FS stimuli (β = 0.09, SE = 0.04, t(456) = 2.14, p = 0.03), indicating they thought about [h] more than the canonical [s]. (Recall [s] has similar usage in their native language, and therefore, serves as a plurality baseline). There were no differences between the other conditions for this analysis. Finally, within the AS condition, their RTs were significantly faster when answering “plural” versus “singular” (β = -0.18, SE = 0.06, t(75.06) = -2.83, p = 0.006). This is strong evidence that the longer RT to answer “singular” were not merely due to the time it takes to hear [h], but because of deliberating the category of the sound. Additionally, AS responses suggests they were confident (reflected in quicker RTs) with certain aspirated stimuli versus others. However, reviewing AS trials where they selected the target, did not show nouns that were more frequently selected, nor those with more obvious aspirations.

Figure 5b shows the majority of participants who responded “singular” to aspirated stimuli. There was no difference in RTs between the DL and SI conditions, which was unexpected as we hoped participants would become curious about the lack of plurals and start to deliberate on stimuli. Instead they seemed content to conclude that two of the talkers never produced plural target items. RTs for the FS condition were significantly slower than those for the SI condition (β = 0.05, SE = 0.02, t(1314) = 2.27, p = 0.02), possibly due to the time it takes to hear [s]. RTs for the AS condition were significantly slower than those for the SI condition (β = 0.10, SE = 0.02, t(1314) = 4.20, $p < 0.001$). This finding may be attributed to possible [h] sensitivity, since (1) RTS were nearly twice that of FS and (2) the three participants that were more clearly sensitive to [h] also responded slower when answering “singular” in the AS condition. These observations make it seem unlikely that the slower RTs in the AS condition were due solely to the extra time occupied by the [h].

Eye-tracking

Eye measurements started with audio stimuli and ceased after mouse response, and were anchored to the onset of the possessive “su(s)”, which was the first word indicating plurality. For visualization, the eye movement record from this point up to a duration of 2500 ms is shown, but statistical analysis was performed on the window of 600 to 1800 ms, which was identified as spanning the point where looks to target and competitor begin to diverge strongly, and ending where the gaze pattern has stabilized. The dependent variable was the proportion of fixations to target during this time window. (A separate analysis of looks to competitor was also conducted, and yielded the same results). Eye-tracking data were analyzed using a generalized linear mixed model (using the Laplace approximation) with glmerMod (Bates, Maechler, Bolker, and Walker, 2015) in R. Significance was assessed via likelihood Wald Z (Laplace approximation).

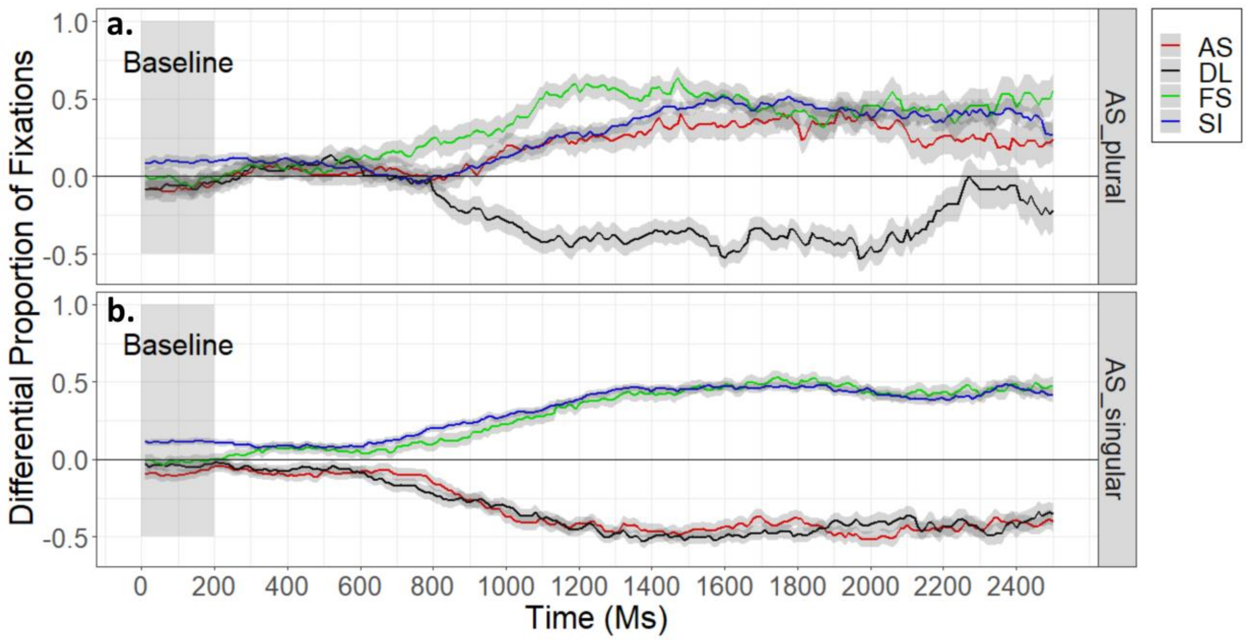

Figure 6. (a) Differential proportion of fixations for subset sensitive to Aspiration (b) Differential proportion of fixations for subset insensitive to Aspiration

Figure 6a and **6b** displays the differential of fixations to the Target minus Competitor by condition for the subsets who responded "plural" or "singular" to AS stimuli, respectfully. In both figures, none of the curves start at a zero differential, which would have reflected impartiality towards all images. (In a non-differential data assessment, chance proportion of fixations would be 25% since there were four pictures.) Recall that data was anchored to the "su(s)" onset, not the very start of the trial. Thus, the initial nonzero differential could be from participants looking around as soon as the pictures were displayed at trial onset. There is also an obvious singular bias for the target in both subsets. This can be explained as singular nouns are more frequent in languages and become the "default" plurality. (Consider nouns such as bed and dog versus pants and shoes.) **Figure 6a** refers to the three participants who responded "plural" to AS stimuli. Fixations for SI, AS, and DL conditions all diverge by 800 ms, while fixation for FS appear to diverge earlier, near 450 ms. Since AS and DL stimuli were the acoustically weaker variants, it makes sense they might draw participants' attention later. Conversely, fixations to FS stimuli diverged sooner because it is easier to confirm a high-energy sound [s] than a lower energy sound. The slope for FS and SI differential curves appears alike, reflecting equal confidence in their commitment to these targets. These similarities are, again, due to the ease of [s] and non-[s] distinction for plurality in participants' native language. There was no difference in fixations between the FS and SI conditions after committing to the target. The sibilant [s] was as easy for participants to recognize as a plural /s/ as it was for them to recognize singular stimuli.

There were substantially fewer fixations to competitor in the DL versus SI condition (β = -2.19, SE = 0.09, z = -24.67, $p < 0.001$). It also takes less time for the participants to commit to the competitor in the DL condition than the target in the SI condition, as the DL curve approaches -0.5 at 1200 ms versus when the SI curve breaks 0.5 near 1550 ms. The magnitude of fixations is identical for both the SI and DL, which is reasonable if participants interpreted DL stimuli as singular. The differential curve for the DL condition is somewhat choppy, which might reflect some interest in target images. However, this is insignificant by the magnitude at which the competitor is favored. Overall, results do not support the prediction that participants become interested in the target for DL stimuli since they actually appear unfazed. As will later be shown with the rest of the participants, when [∅] was present, it was interpreted as singular stimuli.

While points of divergence are comparable and there is a clear commitment to target, there were appreciably fewer fixations to target in the AS versus SI condition, overall (β = -2.06, SE = 0.09, z = -23.25, $p < 0.001$). This is likely due to interest in the competitor, if the [h] seemed somewhat of an ambiguous sound to listeners. Reviewing only AS trials where participants responded “singular” did not show certain nouns that were more frequently selected. Thus, participants were unlikely to be responding to specific items that may have been produced with exceptionally weak aspirations, which reaffirms participants’ uncertainty in perceiving [h]. When this subset responded to AS stimuli as “plural”, there continued to be no difference in fixations to target in the AS versus SI condition. When this subset responded “singular”, less than 50% of fixations were to the target in the SI condition. There were significantly more fixations to FS targets than that of SI ($\beta = 0.54$, SE = 0.20, z = 2.71, p = 0.007). There were significantly fewer fixations to AS targets than that of SI ($\beta = 2.87$, SE = 0.23, z = -12.719, $p < 0.001$) and to DL targets than that of SI ($\beta = -2.33$, SE = 0.20, z = 11.71, $p < 0.001$). This means participants were more frequently looking at the response they selected. Cross-evaluation of RT with the AS condition as a function of response supported that their gaze corresponding to mouse clicks ($\beta = 2.86$, SE = 0.60, z = 4.76, $p < 0.001$).

Figure 6b displays the differential of fixations for the subset who responded “singular” to AS stimuli. Clear categorical effects were seen as there was no difference in fixations to target for FS and SI conditions and substantial difference in fixations to competitor for AS and SI conditions and DL and SI conditions. The general smoothness of the curves reflects little deliberation in participants’ decisions. FS and SI stimuli were easy for participants to perceive in a plurality assessment due to similar phonotactics in English. In contrast, the lack of a clear sibilant [s] led listeners to unquestionably favor the singular competitor. These results directly correspond with this subset’s accuracy results.

In summary, Accuracy, response times, and eye-tracking consistently support that low proficiency Spanish leaners respond to SI, FS, and DL stimuli in a categorical fashion. The canonical sibilant is interpreted as “plural”. However, its absence is interpreted as “singular”.

This was evident as SI and FS accuracy results were near ceiling, while DL accuracy results were near 1%. Participants responded “plural” to AS stimuli 21% of the time, but closer evaluation revealed this was driven by a small subset who seemed sensitive to AS stimuli while the majority was insensitive to AS stimuli. RTs were consistent with accuracy, but slower for AS stimuli, suggesting greater deliberation or uncertainty. Eye-tracking data told the same story as accuracy and RTs. Participants insensitive to aspiration showed essentially identical fixations to target in the SI and FS conditions, as well as identical fixations to competitor for in the AS and DL conditions. Participants sensitive to aspiration didn’t show any suspicion towards DL stimuli. However, in the AS condition, while they greatly favored the target, there was notable interest in the competitor. This is the only reflection of deliberation of leniting dialects by eye-tracking data. Overall, participants were confident DL stimuli were singular. A few realized AS stimuli were plural, but if they did not make this realization, AS stimuli were unquestionably considered singular. The following section hopes to explain these responses.

Discussion

This thesis aimed to assess (1) how low proficiency Spanish learners respond to different realizations of word-final /s/ lenition as spoken by three different native Spanish speakers, (2) if such learners can exhibit perceptual learning to these dialectal variants, and (3) what insights these results provide about L2 perception during initial contact with native speakers. Listeners were predicted to be most accurate when hearing the sibilant form, of intermediate accuracy when hearing the aspirated form, and least accurate when hearing the deleted form. Dependent variables told a similar story, but there were inconsistencies with predictions for leniting conditions. This portion of the thesis will address these initial questions, limitations of the study, and future improvements.

Low proficiency Spanish learners respond to SI, FS, and DL stimuli in a categorical fashion. While a few participants responded “plural” to AS stimuli, those who did not make this realization, unquestionably responded “singular” to AS stimuli. The results of this thesis support claims by speech perception models and effects of L1 constraints. The L2LP built upon the SLM by proposing an L2 learner “duplicates” their L1 phonology to use for reference when learning an L2. The L2LP, SLM, and PAM also collectively make predictions about perceptual difficulty based on the location of speech sounds in phonological space. The ceiling level accuracy for the FS and SI conditions support the idea that perceptual challenges may be reduced as the English /s/ and Spanish /s/ categories overlap in phonological space, and no other phonological categories need to be created or removed. L1 constraints also explain the frequent similarities between SI and FS results throughout the experiment. Because word-final /s/ in Spanish (L2) is used as a similar plural morpheme and [s] phoneme in English (L1), it is a highly practiced perceptive and productive feature for native English-speakers. Furthermore, it may explain why DL stimuli are imperceptible for participants, though lenition often induces phonetic changes to neighboring sounds (File-Muriel and Brown, 2011), and deletion are likely to show traces of aspiration in their acoustic wave due to the spectral nature of lenition patterns (Erker, 2010). While similar word-final /t/ and /d/ deletions occur in English (Bayley, 1996), as well as colloquial omission of <g> in English gerunds (e.g. "how you doin'?"), coda /s/ does not delete in English. This is likely why English-Spanish bilinguals were unable to consider [Ø] a valid allophone to /s/.

The challenge reflected in AS results also support claims by perceptual models. AS perception would require participants to reduce the L1 categories /h/ and /s/ to correspond with a single, Spanish /s/ category. This describes the "Subset scenario" of the L2LP (Escudero, 2005; van Leussen and Escudero, 2015) and "categorizable-uncategorizable assimilation type" of the PAM (Best, 1995). This is naturally harder than if the categories had similarly aligned in phonological space, but the Spanish /s/ category has a continuum of allophones that extend throughout the space in a way that the English /s/ category does not consider. An additional L1 constraints is that it's uncommon for English words to end with [h], so native English speakers would not be practiced in hearing this cue.

When comparing all participants, AS stimuli were marked "plural" 21% of the time. This is in line with the 27% accuracy reported by Saz (2019), who studied Western Andalusian aspiration by Spanish learners that had not studied abroad, and 21% accuracy reported by Schmidt (2018), who evaluated the effects of a variety of linguistics experiences on aspiration perception. Saz reported that of those who were sensitive to aspiration, accuracy results ranged from 0% to 57%, which differed from this study's 71% accuracy by those sensitive to aspiration.

No deeper evaluation of this range was provided in her results. Schmidt also did not describe how accuracy results were distributed across her participants sensitive to aspiration, except in regard to the aspiration assessment with noise, which is not relevant to the present study. Further consistent with Saz's results, aspiration sensitivity was directly correlated with dialect contact. The effects of dialect exposure on improved perception of those dialect features has been supported by numerous studies (Tauroza and Luk, 1997; Schmidt, 2009; Sumner and Samuel, 2009). The present study found that all three participants who responded "plural" to AS stimuli, reported noteworthy contact with aspirating dialects in their LHQ. Interestingly, LexTale proficiency scores were not a reliable indicator of aspiration sensitivity since the participant with obvious learning effects had an average LexTale score, but the participant with the highest LexTale score showed weak learning effects.

Of the leniting conditions, only the AS condition showed indication of perceptual learning, or adapting to a novel speech sound after short exposure. This was most evident in Participant 9 who responded "singular" to AS stimuli in the first 7 trials, then only responded "singular" for 2 of the remaining trials. Effects were limited in Participant 3, who had more evenly distributed "singular" responses, though it seemed to trial off towards the end of the block. No clear effects were seen in Participant 2, who had randomly distributed "singular" responses. Given all these participants had significant dialect contact, results may speak to probabilistic inference theories that highlight the role of using prior linguistic experiences to generalize and make predictions about novel speech sounds (Kleinschmidt and Jaeger, 2015; Norris et al., 2016). Participant 9 was the only one who reported a significantly large amount of time listening to Spanish per week, even compared to Participants 3 and 2. Since this person had the strongest learning effect, it seems consistent with theories that her larger linguistic experience served as a reference bank for her to accurately interpret aspiration in the span of a 30-minute experimental task.

Overall, the results of this study indicate that low proficiency Spanish speakers, upon initial contact with native L2 speakers, may find dialectal variation uninterpretable. While it is possible for low proficiency learners to interpret, and maybe even adapt to, aspirated variants, this does not reflect the majority of participants. Schmidt (2018) suggests that participants may need to reach a type of proficiency "threshold" for perceptual acquisition to occur and/or more dialect exposure, as indicated by this study's LHQ correlations to accuracy. Relatedly, it may be possible for low proficiency learners, who already to meet these criteria, to exhibit perceptual learning shortly after initial contact with native L2 speakers.

Nevertheless, the limitations of this study must be acknowledged. With 14 participants, sample size was an obvious factor that prevented extensive analyses of the AS condition, the most promising result. Indication of trial order or block order effects were also hard to visualize due to sample size. In the future, data collection should continue to expand sample size for better statistical analyses. Next, the study intentionally limited feedback to participants during the visual world experiment so any apparent learning effects would be more organic. However, Kriengwatana et al. (2016) perform study supporting the value of feedback to adapt to linguistic variation. If the same experiment were to be continued to increase sample size and better understand the significance of the study, researchers may want to consider feedback for the DL condition, which was most challenging for participants at this proficiency level. Specifically, before doing this, researchers should also check DL stimuli for the presence of phonetic changes (anticipated by File-Muriel and Brown, 2011) by testing native Spanish speakers, with prior exposure to deleting dialects on the same tokens from this experiment. If there is a notable distinction, this phonetic cue should be analyzed in the waveform. Alternatively, it may be interesting to compare L2 learners' performance on DL stimuli for blocks with and without feedback. Feedback should not be provided for the AS condition, if the experiment were to continue. This is because a larger sample size may reveal other factors that explain why certain participants are sensitive to aspiration. It also provides more data to visualize trial order or block order effects for those individuals.

Conclusion

The results from this thesis demonstrate that low proficiency Spanish learners interpret singular, full /s/, and deletion stimuli categorically, in a way that is predictable by speech perception models and L1 phonotactics. Learners consistently respond in a way that supports L1 influence on L2 acquisition as the canonical sibilant was interpreted as a valid Spanish /s/, but any lenited variant was not accepted as a valid /s/. Closer evaluation of accuracy results reveled two types of low proficiency learners in this sample: Those who recognized aspirated stimuli as "singular" and those who recognize it as "plural". The latter subset provided possible support of perceptual learning and demonstrated a direct correlation between L2 dialect exposure and L2 dialect perception. Overall, these results do not suggest many low proficiency L2 leaners will be able to adapt to L2 variants upon initial contact with a native L2 speaker, unless the learner had sufficient dialect exposure. Future efforts are recommended to test deletion tokens against native Spanish speaker, with exposure to such dialect feature, and to continue this study to better understand if L2 learners can adapt to leniting Spanish variants.

Appendix A

EsPal Lexical Data for Target Nouns

Key: *Frq* = frequency *Syll* = number of syllables *SyllA* = accented syllable

Phon = phonetic transcription *Adj* = corresponding adjective in trial

Item	Frq	LogFrq	Syll	SyllA	Gen	Phon	Familiarity	Adj
ballena	8.08	0.96	3	2	F	baLena	5.65	grande
banco	83.77	1.93	2	1	M	baNko	6.37	pequeño
bañera	10.99	1.08	3	2	F	baJera	6.64	limpia
barco	115.70	2.07	2	1	M	barko	5.55	blanco
barra	17.38	1.26	2	1	F	baRa	5.93	corta
bata	8.75	0.99	2	1	F	bata	5.64	cara
bebida	31.58	1.51	3	2	F	beBiDa	6.49	rica
bicho	9.62	1.03	2	1	M	biCo	6.15	rojo
billete	17.69	1.27	3	2	M	biLete	6.51	verde
bloque	8.12	0.96	2	1	M	bloke	5.15	pequeño
boca	135.27	2.13	2	1	F	boka	6.52	grande
boleto	13.45	1.16	3	2	M	boleto		barato
bolsillo	32.45	1.52	3	2	M	bolsiLo	6.19	pequeño
bomba	70.37	1.85	2	1	F	bomBa	5.20	negra
bombero	5.14	0.79	3	2	M	bomBero		valiente
botella	47.24	1.68	3	2	F	boteLa	6.42	pequeña
bruja	30.70	1.50	2	1	F	bruxa	5.36	vieja
bufanda	4.24	0.72	3	2	F	bufanda	6.08	naranja
caballo	89.28	1.96	3	2	M	kaBaLo	5.60	blanco
cama	177.77	2.25	2	1	F	kama	6.52	blanda
cámara	94.31	1.98	3	1	F	kamara	6.23	barata
camino	324.42	2.51	3	2	M	kamino	5.97	largo
capilla	7.17	0.91	3	2	F	kapiLa	5.15	roja
cara	269.74	2.43	2	1	F	kara	6.92	bonita
carne	86.11	1.94	2	1	F	karne	6.51	condimentada
carro	37.29	1.58	2	1	M	kaRo	5.48	nuevo
carta	109.36	2.04	2	1	F	karta	6.61	nueva
cartera	20.92	1.34	3	2	F	kartera		vacía
chaqueta	31.24	1.51	3	2	F	Caketa	6.68	naranja

chica	391.16	2.59	2	1	F	Cika	6.53	placentera
chico	334.19	2.53	2	1	M	Ciko	6.31	placentero
clínica	21.41	1.35	3	1	F	klinika	5.63	médica
coche	216.99	2.34	2	1	M	koCe	6.64	lento
coco	10.04	1.04	2	1	M	koko	6.72	duro
cohete	10.47	1.06	3	2	M	koete	5.66	rojo
columna	18.24	1.28	3	2	F	kolumna	5.95	dura
coma	26.36	1.44	2	1	F	koma	5.93	negra
cómoda	15.05	1.21	3	1	F	komoDa	6.23	vieja
contrato	45.77	1.67	3	2	M	kontrato	5.13	privado
copa	58.51	1.77	2	1	F	kopa	6.02	delicada
corbata	17.21	1.26	3	2	F	korBata	5.55	rayada
cortina	6.74	0.89	3	2	F	kortina	5.66	corta
criatura	30.55	1.50	3	2	F	krjatura	5.01	peluda
cuadro	20.68	1.34	2	1	M	kwaDro	6.32	blanco
cuchillo	47.19	1.68	3	2	M	kuCiLo	6.34	puntiagudo
cuello	63.65	1.81	2	1	M	kweLo	6.02	fuerte
cuenta	378.83	2.58	2	1	F	kwenta	6.01	cara
cuento	47.47	1.69	2	1	M	kwento	6.47	fantástico
cuerda	27.91	1.46	2	1	F	kwerDa	6.07	larga
cuerpo	218.60	2.34	2	1	M	kwerpo	6.22	fuerte
cuna	8.26	0.97	2	1	F	kuna	4.72	baja
dedo	43.86	1.65	2	1	M	deDo	6.38	gordo
diablo	79.95	1.91	2	1	M	djaBlo	5.86	negro
diamante	12.28	1.12	3	2	M	djamante		falso
diente	10.58	1.06	2	1	M	djente	6.31	blanco
droga	37.17	1.58	2	1	F	droGa	6.21	farmacéutico
ducha	24.63	1.41	2	1	F	duCa	7.00	limpia
fábrica	4.42	0.73	3	2	F	faBrika	5.66	nueva
falda	9.49	1.02	2	1	F	falda	6.34	morada
firma	38.08	1.59	2	1	F	firma	5.66	famosa
flauta	4.00	0.70	2	1	F	flauta	5.46	pequeña
flecha	9.25	1.01	2	1	F	fleCa	5.65	puntiaguda
foto	105.69	2.03	2	1	M	foto	6.63	vieja
frente	142.19	2.16	2	1	F	frente	6.29	grande
fruta	14.01	1.18	2	1	F	fruta	6.33	dulce
fuego	180.95	2.26	2	1	M	fweGo	6.76	caliente
fuente	43.16	1.65	2	1	F	fwente	6.34	rica
galleta	8.73	0.99	3	2	F	gaLeta		dulce
gallo	6.96	0.90	2	1	M	gaLo	6.24	pequeño

garganta	24.98	1.41	3	2	F	garGanta	6.46	dolorosa
gato	54.97	1.75	2	1	M	gato		blanco
globo	11.39	1.09	2	1	M	gloBo	5.56	rojo
guardia	68.51	1.84	2	1	F	gwarDja	5.74	valiente
guitarra	16.38	1.24	3	2	F	GitaRa	5.66	cara
juego	215.16	2.33	2	1	M	xweGo	6.51	divertido
lago	37.52	1.59	2	1	M	laGo	5.96	grande
lágrima	3.36	0.64	3	1	F	laGrima	5.50	pequeña
lámpara	10.54	1.06	3	1	F	lampara	6.15	radiante
lechuga	3.09	0.61	3	2	F	leCuGa	6.68	verde
lengua	43.95	1.65	2	1	F	leNgwa	5.61	grande
letra	21.24	1.35	2	1	F	letra	6.25	vieja
libro	160.29	2.21	2	1	M	liBro	6.86	viejo
llama	259.78	2.42	2	1	F	Lama	5.07	peluda
llave	82.00	1.92	2	1	F	LaBe	6.74	metálica
lobo	28.49	1.47	2	1	M	loBo	6.13	fuerte
luna	84.49	1.93	2	1	F	luna	5.96	radiante
madre	613.28	2.79	2	1	F	maDre	6.95	paciente
mago	20.13	1.32	2	1	M	maGo	5.83	magnífico
maleta	26.21	1.43	3	2	F	maleta	5.84	llena
manga	7.27	0.92	2	1	F	maNga	6.12	larga
mango	4.25	0.72	2	1	M	maNgo	5.40	naranja
mano	284.06	2.45	2	1	M	mano	6.51	grande
mapa	41.20	1.63	2	1	F	mapa	5.94	norteamericano
médica	21.52	1.35	3	1	F	meDika	4.80	concentrada
médico	125.78	2.10	3	1	M	meDiko	6.54	viejo
mercado	45.10	1.66	3	2	M	merkaDo	6.01	grande
metro	25.23	1.42	2	1	M	metro	5.58	rápido
mochila	10.74	1.07	3	2	F	moCila	6.28	llena
moneda	19.89	1.32	3	2	F	moneDa	6.22	metálica
mono	34.99	1.56	2	1	M	mono	6.10	peludo
montaña	40.63	1.62	3	2	F	montaJa	6.37	grande
monte	17.46	1.27	2	1	M	monte	6.25	pequeño
moto	18.97	1.30	2	1	M	moto	6.86	rápida
muchacho	126.88	2.11	3	2	M	muCaCo	5.24	placentero
mundo	698.35	2.84	2	1	M	mundo	6.54	pequeño
museo	23.28	1.39	3	2	M	museo	6.18	viejo
nave	130.88	2.12	2	1	F	naBe	5.32	grande
nido	11.98	1.11	2	1	M	niDo	6.01	pequeño
niña	157.95	2.20	2	1	F	niJa	6.90	linda

niño	272.40	2.44	2	1	M	niJo	6.51	lindo
novia	123.26	2.09	2	1	F	noBja	6.07	bonita
número	222.10	2.35	3	1	M	numero	6.51	negro
padre	819.27	2.91	2	1	M	paDre	6.57	paciente
página	28.28	1.47	3	1	F	paxina	6.60	blanca
pájaro	30.95	1.50	3	1	M	paxaro	6.51	colorido
palabra	183.06	2.26	3	2	F	palaBra	6.69	negra
pantalla	25.24	1.42	3	2	F	pantaLa	6.24	pequeña
papa	60.08	1.79	2	1	F	papa	6.29	frita
paquete	33.02	1.53	3	2	M	pakete	5.99	grande
parque	43.63	1.65	2	1	M	parke	6.31	bello
pata	24.85	1.41	2	1	F	pata	6.44	pelada
pato	13.86	1.17	2	1	M	pato	5.59	pequeño
pavo	14.46	1.19	2	1	M	paBo	4.71	grande
pecho	40.99	1.62	2	1	M	peCo	6.13	fuerte
pelota	32.87	1.53	3	2	F	pelota	6.52	dura
perro	156.54	2.20	2	1	M	peRo	6.64	pelado
persona	282.75	2.45	3	2	F	persona	7.00	placentera
piano	24.88	1.41	2	1	M	pjano	5.86	caro
pie	119.14	2.08	1	1	M	pje	6.42	fuerte
piedra	48.11	1.69	2	1	F	pjeDra	6.42	dura
pierna	54.41	1.74	2	1	F	pjerna	6.48	fuerte
piloto	39.11	1.60	3	2	M	piloto	5.19	valiente
pingüino	3.83	0.68	3	2	M	piNGwino		pequeño
pintura	31.74	1.52	3	2	F	pintura	5.03	colorida
pirata	12.83	1.14	3	2	F	pirata	5.57	valiente
planta	31.47	1.51	2	1	F	planta	6.60	verde
plato	26.94	1.45	2	1	M	plato		fino
pluma	13.77	1.17	2	1	F	pluma	5.81	negra
pollo	40.68	1.62	2	1	M	poLo	6.36	blanco
pregunta	209.60	2.32	3	2	F	preGunta	6.36	negra
premio	37.81	1.59	2	1	M	premjo	6.25	metálico
príncipe	55.03	1.75	3	1	M	prinsipe	5.55	guapo
puente	60.09	1.79	2	1	M	pwente	5.44	colgante
puerta	334.18	2.53	2	1	F	pwerta	6.78	dura
puerto	24.63	1.41	2	1	M	pwerto	6.20	tecnológico
punta	22.21	1.37	2	1	F	punta	5.72	puntiaguda
punto	231.74	2.37	2	1	M	punto	6.48	negro
radio	79.89	1.91	2	1	M	RaDjo	6.31	vieja
rata	25.28	1.42	2	1	F	Rata	6.31	blanca

rayo	22.70	1.37	2	1	M	RaHo	5.90	radiante
regalo	95.23	1.98	3	2	M	ReGalo	5.97	caro
regla	33.08	1.53	2	1	F	ReGla	6.27	corta
reina	80.46	1.91	2	1	F	Reina	5.51	bonita
rueda	18.71	1.29	2	1	F	RweDa	6.34	dura
ruta	28.92	1.48	2	1	F	Ruta	5.82	rápida
tabla	12.88	1.14	2	1	F	taBla	5.83	dura
tanque	22.45	1.37	2	1	M	taNke	5.28	grande
tapa	10.93	1.08	2	1	F	tapa	5.63	metálica
tarjeta	57.35	1.77	3	2	F	tarxeta	6.48	casera
tarta	14.85	1.20	2	1	F	tarta	6.24	dulce
teatro	43.56	1.65	3	2	M	teatro	5.92	viejo
techo	40.94	1.62	2	1	M	teCo	6.52	duro
tele	26.71	1.44	2	1	F	tele	6.50	vieja
templo	25.45	1.42	2	1	M	templo	5.64	magnífico
tienda	100.92	2.01	2	1	F	tjenda	6.61	barata
tigre	17.04	1.26	2	1	M	tiGre	5.54	rayado
toalla	13.26	1.15	3	2	F	toaLa	6.60	naranja
tobillo	8.24	0.97	3	2	M	toBiLo		fuerte
tormenta	39.39	1.61	3	2	F	tormenta	6.59	dura
torre	34.85	1.55	2	1	F	toRe	5.81	famosa
tortuga	11.10	1.08	3	2	F	tortuGa	6.29	verde
traje	113.38	2.06	2	1	M	traxe	6.42	caro
triángulo	4.34	0.73	3	1	M	trjaNGulo	5.92	blanco
trofeo	6.23	0.86	3	2	M	trofeo	3.72	metálico
trompeta	4.63	0.75	3	2	F	trompeta	4.74	metálica
tronco	7.11	0.91	2	1	M	troNko	5.43	duro
vaca	20.66	1.34	2	1	F	baka	6.12	moteada
vaquero	14.26	1.18	3	2	M	bakero	5.57	valiente
venda	8.08	0.96	2	1	F	benda		pequeña
ventana	80.71	1.91	3	2	F	bentana	6.82	limpia

Appendix B

Language History and Dialect Contact Questions

1. Participant number
2. Age
3. Gender
4. Highest level of completed education
5. Area(s) of study for all post-high school education. Example: Major(s), minor(s), dissertation topic.
6. Country of residence
7. Native language
8. Second language
9. If you started learning your second language through school, at approximately what stage of your education did this begin?
10. If you did not start learning your second language through school, please describe at what age and how your started learning.
11. Rate your ability to perform the following skills in your second language on a scale of 1 to 5. (1 = low performance; 5 = high performance)
    a. Write (1 = “I write only very simple sentences”; 5 = “I write extensive papers with few grammatical errors”)
    b. Read (1 = “I read only very simple sentences”; 5 = “I read extensive papers with few grammatical errors”)

c. Speak (1 = “I can recite simple colors, numbers, object names, etc.”; 5 = “I have near-native speech and can discuss desires, ideas, politics, etc.”)

d. Listen to formal speech (1 = “I’m unable to understand what is spoken”; 5 = “I fully understand what is spoken”)

e. Listen to conversational speech (1 = “I’m unable to understand what is spoken”; 5 = “I fully understand what is spoken”)

12. Rate the extent to which you think you have a foreign accent in your second language on a scale of 1 to 5. (1 = “I sound like native speakers of the language”; 5 = “I sound like a foreigner to native speakers of the language”)

13. Rate how frequently you use your NATIVE language in the following contexts on a scale of 1 to 5. (1 = “Never”; 5 = “Always”)

    a. At home
    b. At school
    c. At work
    d. In media (video games, movies/TV, videos, music, leisurely reading, etc.)
    e. With native ENGLISH-speaking friends
    f. With native SPANISH-speaking friends

14. Rate how frequently you use your SECOND language in the following contexts on a scale of 1 to 5. (1 = “Never”; 5 = “Always”)

    a. At home
    b. At school
    c. At work
    d. In media (video games, movies/TV, videos, music, leisurely reading, etc.)

e. With native ENGLISH-speaking friends

f. With native SPANISH-speaking friends

15. If you use your SECOND language with native ENGLISH-speaking friends, what proportion of the conversation is in your second language on a scale of 1 to 5. (1 = “Only a few words”; 2 = “All of it”)
16. If you use your SECOND language with native SPANISH-speaking friends, what proportion of the conversation is in your second language on a scale of 1 to 5. (1 = “Only a few words”; 2 = “All of it”)
17. Estimate how many total hours you spend LISTENING to your second language per week. This could be in class lectures, videos, music, etc.
18. Do you notice any sounds produced **by your Spanish professors at Penn State** that are different than how you learned Spanish previously? If yes, what did you notice?
19. Do you notice any sounds produced **by the speakers in this experiment** that are different than how you learned Spanish previously? If yes, what did you notice?
20. Do you think the speakers in this experiment came from the same or different place? Please explain your reasoning.
21. Indicate how often you interact with friends, family, teachers/professors, or colleagues in the past 5 years from the following countries on a scale of 1 to 5. (1 = “Never”; 5 = “Very often”)
    a. Spain – North/Central
    b. Spain – South
    c. Caribbean (Dominican Republic, Puerto Rico, Cuba)
    d. Ecuador

e. Peru – Coast (including Lima)

f. Peru – Interior (Andes)

g. Mexico

h. Costa Rica

i. Colombia

j. Argentine/Chile

22. If you have interacted with friends, family, teachers/professors, or colleagues in the past 5 years from any country NOT mentioned above, please list the country and rate how often you interact with these individuals. Use the format: Country, # (Use the rating scale from the previous question.)
23. Please list the number of Spanish classes you have taken **before university** and if the instructor was a native Spanish-speaker. If the instructor was a native Spanish-speaker, provide their country of origin.
24. Please list the number of Spanish classes you have taken **at Penn State** and if the instructor was a native Spanish-speaker. If the instructor was a native Spanish-speaker, provide their country of origin.
25. Describe any dialect differences in pronunciation that you are aware of in the sound highlighted in the following word. The "s" sounds in: *mi**s** gato**s***
26. Please comment below with any additional information that you think supplements the previous questions, may be interesting/important but was not included on this form, or helps us better understand your language history.

BIBLIOGRAPHY

Baker, W., and Smith, L.C. (2010). The impact of L2 dialect on learning French vowels: Native English speakers learning Quebecois and European French. *The Canadian Modern Language Review*, *66*, 711–738.

Bayley, R. (1996). Competing constraints on variation in the speech of adult Chinese learners of English. In R. Bayley and D. Preston (Eds.), *Second language acquisition and linguistic variation* (pp. 97–120). Amsterdam, The Netherlands: John Benjamins.

Best, C. T. (1995). A direct realist perspective on cross-language speech perception. In W. Strange (Ed.), *Speech perception and linguistic experience: Issues in cross-language research* (pp. 171–204). Timonium, MD: York Press.

Bradlow, A. R., and Bent, T. (2003). Listener adaptation to foreign-accented speech. In *Proceedings of the 15th International Congress of Phonetic Sciences* (pp. 2881-2884). Barcelona

Bradlow, A. R., and Bent, T. (2007). Perceptual adaptation to non-native speech. *Cognition*, *106*, 707-729.

Brown, K. (1968). Intelligibility. In A. Davies (Ed.), *Language testing symposium* (pp.180-191). Oxford, England: Oxford University Press.

Clarke, C. M. (2002). Perceptual adjustment to foreign accented English with short-term exposure. In *Proceedings of the 7th International Conference on Spoken Language Processing* (pp. 253- 256). Denver.

Clarke, C. M. (2003). Processing time effects of short-term exposure to foreign-accented English (Unpublished doctoral dissertation). University of Arizona, Tucson.

Bates, D., Maechler, M., Bolker, B., Walker, S. (2015) Fitting linear mixed-effects models using lme4. *Journal of Statistical Software, 67*(1), 1-48. doi:10.18637/jss.v067.i01.

Drummond, R. (2012). The Manchester Polish STRUT: Dialect acquisition in a second language. *Journal of English Linguistics*, *41*, 65-93.

Dupoux, E., and Green, K. (1997). Perceptual adjustment to highly compressed speech: Effects of talker and rate changes. *Journal of Experimental Psychology: Human Perception Performance*, *23*, 914- 927.

Eisenstein, M., and Berkowitz, D. (1981). The effect of phonological variation on adult learner comprehension. *Studies in Second Language Acquisition*, *4,* 75-80.

Eisner, A., and McQueen, J. M. (2005) The specificity of perceptual learning in speech processing. *Perception Psychophysics*, *67*(2), 224-238.

Erker, D. (2010). A subsegmental approach to coda /s/ weakening in Dominican Spanish. *Int'l J Soc Lang*, *203*, 9-26.

Escudero, P. (2005). Linguistic perception and second language acquisition: Explaining the attainment of optimal phonological categorization (Unpublished doctoral dissertation). Utrecht, The Netherlands: Utrecht University.

Escudero, P. (2009). Linguistic perception of "similar" L2 sounds. In P. Boersma and S. Hamann (Ed.), *Phonology in Perception* (pp. 151–190). Mouton de Gruyter.

Escudero, P., and Boersma, P. (2004). Bridging the gap between L2 speech perception research and phonological theory. *Studies in Second Language Acquisition*, *26*, 551–585.

Escudero, P., and Sisinni, B., Grimaldi, M. (2014). The effect of vowel inventory and acoustic properties in Salento Italian learners of Southern British English vowels. *J. Acoust. Soc. Am.*, *135*(3), 1577-1584.

File-Muriel, R. J. and Brown, E. K. (2011). The gradient nature of s-lenition in Caleño Spanish. *Language Variation and Change*, *23*, 223-243. doi:10.1017/S0954394511000056

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. (2005). Origins and development of the Speech Learning Mode [PowerPoint slides]. Retrieved from http://www.jimflege.com/files/SLMvancouver_updated.pdf

Fox, R. A., and McGory, J. T. (2007). Second language acquisition of a regional dialect of American English by native Japanese speakers. In O. S. Bohn M. J. Munro (Eds.), *Language experience in second language speech learning* (pp. 117–134). Amsterdam, The Netherlands: John Benjamins.

Greenspan, S. L., Nusbaum, H. C., and Pisoni, D. B. (1988). Perceptual learning of synthetic speech produced by rule. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *14*, 421-433.

Hammond, R. M. (2001). The sounds of Spanish: Analysis and application (with special reference to American English). Somerville, MA: Cascadilla.

Henriksen, N. C., Geeslin, K. L., and Willis, E. W. (2010). The development of L2 Spanish intonation during a study abroad immersion program in Léon, Spain: Global contours and final boundary movements. *Studies in Hispanic and Lusophone Linguistics*, *3*, 113–162.

Hochberg, J. (1986) Functional compensation for /s/ deletion in Puerto Rican Spanish. *Language*, *62*(3): 609-621.

Izura, C., Cuetos, F., and Brysbaert, M. (2014). Lextale-Esp: A test to rapidly and efficiently assess the Spanish vocabulary size. *Psicológica*, *35*(1), 49-66.

Kleinschmidt, D., and Jaeger, F. (2015). Robust speech perception: Recognize the familiar, generalize to the similar, and adapt to the novel. *Psychological Review*, *122*(2), 148-203.

Kriengwatana, B., Terry, J., Chládková, K., and Escudero, P. (2016). Speaker and accent variation are handled differently: Evidence in native and non-native listeners. *PLoS ONE*, *11*(6), e0156870. doi:10.1371/journal.pone.0156870

Kuznetsova, A., Brockhoff, P. B., Christensen, R. H. B. (2017) LmerTest package: Tests in linear mixed effects models." *Journal of Statistical Software*, 82(13), 1-26. doi:10.18637/jss.v082.i13

Norris, D., McQueen, J., and Cutler, A. (2016). Prediction, Bayesian inference and feedback in speech recognition. *Language, Cognition and Neuroscience*, *31*(1), 4-18, doi: 10.1080/23273798.2015.1081703

Psychology Software Tools, Inc. [E-Prime 3.0]. (2016). Pittsburgh, PA. Retrieved from https://www.pstnet.com.

Tauroza, S. (1993). Recognizing words in continuous speech: How important are word-final consonants? *ELT Journal*, *47*, 211-218.

Terrell, T. (1979). word-final /s/ in Cuban Spanish. *Hispania*, *62*, 599-612.

Schmidt, L. (2018). L2 development of perceptual categorization of dialectal sounds. *Studies in Second Language Acquisition*, *40*, 857-882. doi:10.1017/S0272263118000116

Schwegler, A., Ameal-Guerra, A., and Kempff, J. (2010). *Fonética y fonología españolas*, *4*, 386-391. Retrieved April 9, 2019.

Smith, L. C., and Baker, W. (2011). L2 dialect acquisition of German vowels: The case of Northern German and Austrian dialects. *Poznan Studies in Contemporary Linguistics*, *47*, 120–132.

SR Research Experiment Builder 2.2.1 [Computer software]. (2019). Mississauga, Ontario, Canada: SR Research Ltd.

Sumner, M., and Samuel, A. G. (2009). The effect of experience on the perception and representation of dialect variants. *Journal of Memory and Language*, *60*, 487–501.

van Leussen, J. W., and Escudero, P. (2015). Learning to perceive and recognize a second language: The L2LP model revised. *Frontiers in Psychology*, *6*, 1–12.

Wilcox, G. (1978). The effect of accent on listening comprehension: A Singapore study. *English Language Teaching Journal*, *32*, 118-127.

Witteman, M. J., Weber, A., and McQueen, J. M. (2014) Tolerance for inconsistency in foreign-accented speech. *Psychon Bull Rev*, *21*, 512-519. doi: 10.3758/s13423-013-0519-8

Wolfram, W., Carter, P., and Moriello, B. (2004). Emerging Hispanic English: New dialect formation in the American South. *Journal of Sociolinguistics*, *8*, 339–358.

ACADEMIC VITA
CAITLIN E. Y. REICHARD

EDUCATION

The Pennsylvania State University, Schreyer Honors College **University Park, PA**
Eberly College of Science | Bachelor of Science in Biology, Vertebrate Physiology May 2020
College of the Liberal Arts | Bachelor of Science in Spanish, Applied

Universidad Iberoamericana Puebla **Puebla, Mexico**
The Pennsylvania State University's six-week, faculty-led study abroad program *Summer 2018*

RESEARCH EXPERIENCE

MuGlab: The MultiGrammar Lab, Dr. Matthew Carlson, PhD **University Park, PA**
Research Assistant *Aug 2018 – May 2020*
- Investigated English-Spanish bilinguals' perception of Spanish dialectal variation for completion of an undergraduate honors thesis
- Aided in the research proposal, design, experiment programming, participant recruiting, and testing

Muscle Biology Lab, Dr. Gustavo Nader, PhD **University Park, PA**
Research Assistant *Jan 2017 – May 2018*
- Studied cellular mechanisms of skeletal muscle hypertrophy
- Performed western blots to evaluate the nuclear localization of key proteins in the mTOR signaling cascade

WORK EXPERIENCE

UPMC Pinnacle Hospital Systems, Clinical Information Management Systems (CIMS) **Harrisburg, PA**
Medical Scribe *May 2016 – May 2020*
- Logged 400+ clinical hours per year documenting information for the Electronic Health Record (EMR) in the patient's room with the attending physician
- Utilized medical terminology to chart the history of present illness, review of systems, and physical exam
- Trained newly hired medical scribes on medical charting and calling and communicating with healthcare professionals

CIVIC ENGAGEMENT/SERVICE

Straight Talks, Penn State Center for Sexual and Gender Diversity (CSGD) **University Park, PA**
Peer Educator *Spring 2020*
- Raised awareness and educated students about identity development, the LGBTQ+ community, and LGBTQ+ issues
- Developed professional presentation and training skills by serving on student panels and leading student ally learning workshops

ACADEMIC VITA
CAITLIN E. Y. REICHARD

Appalachia Service Project, Penn State Alternative Breaks **Guyan Valley, WV**
Volunteer *Mar 2019*
- Served 1 week (30+ hours) insulating the exterior and replacing subflooring of a trailer for a low-income family in rural, central Appalachia
- Participated in program activities and simulations to gain a deeper understanding of rural poverty and factors causing it

Remote Area Medical (RAM) **Gloucester, VA**
Volunteer *Nov 2018*
- Served 2 days (16 hours) with a mobile clinic to provide free medical and dental care to an underserved, rural community
- Sanitized the suction apparatus of the dental chair after extractions

Medical mission trip, MEDLIFE **Lima, Peru**
Volunteer *Dec 2016*
- Served 1 week (30+ hours) providing medical, educational, and developmental efforts to the impoverished *pueblos jóvenes* outside of Lima, Peru
- Utilized basic Spanish to instruct children how to properly brush their teeth
- Mixed cement for the construction of staircases and painted the staircases that would allow local citizens to seek land grants and better navigate the valley walls of their surroundings

LEADERSHIP EXPERIENCE

Eberly College of Science **University Park, PA**
Teaching/Lecture Assistant *Fall 2016 – May 2020*
- Provided tutoring, in-class assistance, and met with instructors about course development as a member of the instructional team for introductory courses
- Promoted to a senior position* for 4 of the 5 semesters served, which required the additional administrative tasks below
- Led exam review sessions of more than 100 students, resolved conflicts with students or TAs, and organized hiring and scheduling of the other TAs and distribution of course materials

Courses Assisted	Hours/Week	Semesters
*BIOL 110: Basic Concepts and Biodiversity**	6	Fall 2016-19
BIOL 133: Genetics and Evolution of the Human Species	4	Spring 2019
PHYS 251: Introductory Physics II	3	Spring 2020

HONORS AND AWARDS

Doris N. McKinstry Scholarship *Aug 2018*
Summer Health Professions Education Program (SHPEP) *June 2017*
Paterno Fellows Program *June 2017*
Schreyer Honors College Gateway Scholar *June 2017*
Small Project Grant in the College of Health and Human Development *May 2017*