US Latinidad and the Spanish Heritage Community:

Perception, Production and Identity

by

Cecelia Staggs

A dissertation accepted and approved in partial fulfillment of the

requirements for the degree of

Doctor of Philosophy

in Linguistics

Dissertation Committee:

Melissa Baese-Berk, Chair

Guillem Belmar Viernes, Core Member

Julie Sykes, Core Member

Sergio Loza, Institutional Representative

University of Oregon

Summer 2026

© 2026 Cecelia Staggs

DISSERTATION ABSTRACT

Cecelia Staggs
Doctor of Philosophy in Linguistics
Title: US Latinidad and the Spanish Heritage Community: Perception, Production and Identity

The purpose of this dissertation is to provide a deeper understanding of the nuances involved in how the languaging habits of Spanish heritage speakers in the United States are evaluated at the institutional level, and how identity is subsequently assigned to heritage speakers and learners when evaluated by their peers. This evaluative process is heavily influenced by the ideologies and social expectations within the Latinx community regarding proficiency, linguistic ability in the heritage language, and "sounding native." In addition, this dissertation addresses the expectations placed around both grammatical and phonological language production in research and educational spaces, analyzing how these structural standards dictate the external validation and social categorization of these individuals.

Particular attention is paid to beginner-level heritage learners, who usually fall within later generations, given that these expectations remain insistent and powerful even as groups more susceptible to language loss begin to outnumber the earlier generations, who typically exhibit the linguistic skills that community members, researchers, and educators anticipate. The dissertation begins by challenging previous arguments claiming that heritage acquisition is incomplete, suggesting instead that the error lies within methodological approaches that rely on monolingual and second language (L2) norms. The dissertation then evaluates the conceptual boundaries surrounding the labels heritage *speaker* and heritage *learner*, exploring how proficiency in the Spanish language influences variables such as social perception and community acceptance. Finally, the dissertation concludes by challenging researchers to be

thoughtful in their terminology and descriptions, given the lasting impacts our findings have on education and policy.

This dissertation includes previously published and unpublished coauthored material.

TABLE OF CONTENTS

DISSERTATION ABSTRACT **3**

TABLE OF CONTENTS **5**

CHAPTER 1: INTRODUCTION **10**

1.1 The history of Spanish in the United States 20
1.2 Multilingualism in the US 25
1.3 Spanish as a Heritage Language in the United States: Ideologies and Attitudes 31
1.4 Spanish Heritage speakers: Definitions and terminology 36

CHAPTER 2: A COMPARATIVE ANALYSIS OF LEXICAL DIVERSITY IN ORALLY PRODUCED SPANISH AND ENGLISH SPEECH **42**

2.1 Introduction 42
2.2 Background 43
2.2.1 Spanish heritage speakers *43*
2.2.2 Heritage acquisition *44*
2.2.3 Heritage production - Spanish *47*
2.2.4 Lexical diversity *49*
2.3 The current study 51
2.4 Methods 52
2.4.1 The corpus *52*
2.4.2 Linguistic and corpus analysis *54*
2.5 Results 54
2.5.1 Lexical diversity – English *54*
2.5.2 Lexical diversity – Spanish *57*
2.6 Discussion 60
2.7 Conclusion 65

CHAPTER 3: BUT… ARE YOU REALLY LATINO? WITHIN COMMUNITY PERCEPTIONS OF SPANISH HERITAGE SPEAKERS AND LEARNERS **67**

3.1 Introduction 67
3.2 Background 69
3.2.1 Spanish as a heritage language in the United States *69*
3.2.2 Spanish heritage speakers vs learners *72*
3.2.3 Standard Language Ideology *75*
3.2.4 Language ideologies and the Spanish heritage community *77*
3.3 The current study 81
3.4 Methods 82
3.4.1 Participants *82*
3.4.2 Task *83*
3.4.3 Stimuli *84*

3.5 Results ... 86
3.6 Discussion ... 89
3.7 Conclusion and future directions ... 91

CHAPTER 4: THE INFLUENCE OF SOCIAL INFORMATION ON SPEECH INTELLIGIBILITY WITHIN THE SPANISH HERITAGE COMMUNITY ... **93**

4.1 Introduction ... 93
4.2 Background ... 95
4.2.1. The Role of ideology and expectation in speech perception ... *95*
4.2.2 Language Ideologies and attitudes in the Spanish heritage community ... *97*
4.3 The current study ... 100
4.4 Guise validation - Methods ... 102
4.4.1 Stimuli ... *102*
4.4.2 Participants ... *102*
4.4.3 Procedures and analyses ... *103*
4.5 Intelligibility task - Methods ... 104
4.5.1 Stimuli ... *104*
4.5.2 Participants ... *106*
4.5.3 Procedure and analyses ... *106*
4.6 Results ... 107
4.6.1 Guise Validation ... *107*
4.6.2 Intelligibility Task ... *108*
4.7 Discussion ... 111

CHAPTER 5: A COMPARATIVE ANALYSIS OF SPANISH HERITAGE SPEECH PRODUCTION: SCRIPTED VS. SPONTANEOUS SPEECH ... **115**

5.1 Introduction ... 115
5.2 Background ... 117
5.2.1 Spanish as a heritage language ... *117*
5.2.2 Pedagogical Approaches to SHL ... *118*
5.2.3 Attitudes and ideologies: their influence on HL maintenance and shift ... *122*
5.2.2 Heritage speech production ... *124*
5.3 The Current Study ... 128
5.4 Methods ... 130
5.4.1 Materials ... *130*
5.4.2 Analysis ... *131*
5.5 Results ... 132
5.6 Discussion and Conclusion ... 135

CHAPTER 6: GENERAL DISCUSSION AND CONCLUSIONS ... **137**

6.1 Summary of findings ... 137
6.1.1 Chapter 2 Findings ... *139*
6.1.2 Chapter 3 Findings ... *145*
6.1.3 Chapter 4 Findings ... *150*
6.1.4 Chapter 5 Findings ... *153*
6.2 Contributions to the field ... 155

6.3 Limitations & Future Directions 161
6.4 Conclusions: Call to action! 166
APPENDIX A **172**
APPENDIX B **173**
APPENDIX C **174**
APPENDIX D **181**
APPENDIX E **184**
REFERENCES CITED **186**

LIST OF FIGURES

Figure Page

1. Figure 1 English MTLD92 and MATTR11 scores grouped by speaker. 56
2. Figure 2 Spanish MTLD92 and MATTR11 scores grouped by speaker. 59
3. Figure 3 Ratings for fluency, accentedness and Latino identity for each guise type. 86
4. Figure 4 Ratings for fluency, accentedness and Latino identity. 87
5. Figure 5 Ratings for fluency, accentedness and Latino identity for each Speaker. 88
6. Figure 6 A spectrogram of the word porfiaban produced with a pure approximant. 131
7. Figure 7 A spectrogram of the word porfiaban produced with a tense approximant. 132
8. Figure 8 Pure approximants and tense approximants shown by speech style 134
9. Figure 9 Pure and tense approximants grouped by phoneme type and task type. 135

LIST OF TABLES

Table Page

1. Table 1 Descriptive Statistics for MTLD92 and MATTR11 for the English data. 55

2. Table 2 Descriptive statistics for MTLD92 and MATTR11 for the Spanish data. 58

3. Table 3 Descriptive statistics by guise type for each construct. 107

4. Table 4 Descriptive statistics for manner of articulation across all phonemes 133

CHAPTER 1: INTRODUCTION

Spanish was introduced to the United States in ways similar to the English language. Hundreds of years of exploration followed by years of war and colonial rule forged the oppressive and hegemonic practices that ultimately granted these languages a superior status in the land. Despite both Spanish and English entering the territory in similar ways, further years of political strife led to English eventually becoming the dominant language in power in the US. As Potowski (2018) draws attention, the concept of a 'country' is fundamentally problematic given that, "The emergence of most modern nation states in the 19th century resulted from wars and annexations, and as a result, many populations of individuals with similar ethnic, religious, and/or linguistic affiliations were arbitrarily split by lines drawn on a map." (p. 2). This characterizes the expansion of the US in the aftermath of the Mexican American War, for example, which resulted in the sudden and abrupt distinction – with respect to citizenship status – between the Spanish speaking population occupying Mexico, and those who occupied the northern territory of Mexico that was ceded to the United States after the war (MacGregor-Mendoza, 1998). Events such as this have resulted in lasting social, cultural, and political impacts that have affected the population of Spanish speakers who already inhabited the land, and those who would come to inhabit it for years after.

Although Spanish has consistently been the most widely spoken language other than English in the US for many years, this has not prevented it from being met with resistance when discussing its legitimacy in spaces whether they be political, educational, media-based, and even personal spaces – e.g., the home where Spanish speaking parents are ridiculed for speaking Spanish to their children, with some courts even going so far as to compare it to forms of child abuse (Carter, 2018). Despite debate regarding the degree to which Spanish should be *embraced*

in the United States, its prolific and ubiquitous role in the country cannot be ignored (Beaudrie & Marrero-Rivera, 2024). The presence of Latinx and Spanish speaking communities is regularly increasing, and the language practices of these groups are a direct representation of the multilingual US Latinx experience. The language practices that characterize US Spanish are influenced by the role both Spanish and English play in their lives and developed despite the constant resistance to and degradation of languages other than English in the United States. Overcoming the multitude of social and systemic obstacles set in place to develop and sustain usage of a familial or heritage language to any extent in a country well known for generally favoring and promoting monolingual English practices is challenging in its own right for Spanish speaking families and communities. Persistent societal indoctrination of a monolingual English standard has led, and continues to lead, to copious amounts of external pressures placed on immigrant families making it very difficult to maintain the heritage language – i.e., a minoritized language in a respective country spoken by individuals who have an ancestral or cultural connection to the language (Fishman, 2001)[1].

Notable exceptions to the promotion of monolingualism appear when multilingualism is commodified or acquired later in life as a second or foreign language (Wiley & Lukes 1996, Carter, 2018). Thus, the existence of systemic ideologies that place heritage languages in a position of circumstantial value leaves both the heritage language and those who speak them – i.e., heritage speakers – in vulnerable positions. Confining the perceived benefit of languages other than English to specific groups (e.g., non-heritage learners) and/or for specific purposes (e.g., when a 'native' speaker expert is preferred) has belittled heritage speakers and restricted their ability to use, maintain, identify personally with, and embrace their heritage language with

[1] This is just one of the many ways the heritage community has been defined. More definitions and terminology will be introduced in the literature review section.

confidence. Various difficulties and hurdles experienced in language maintenance often result in the heritage language being completely lost in families by the third generation – a phenomenon so common it has been called the *three-generation model of language loss* (Fishman et al., 1966, Fishman, 1972; Veltman, 1983).

This trend of generational language attrition – although common and relatively consistent – manifests differently and for different reasons in each family and for each heritage speaker. The Spanish heritage community in the United States is fundamentally diverse, influenced by multiple varieties of Spanish interacting with each other as well as the dominant language, English. In addition to the general variation introduced from different varieties of Spanish, there are large amounts of variation in these families regarding levels of proficiency and fluency in both productive and receptive abilities in the heritage language (Rothman et al., 2016). While the *three-generation model* outlines a typical pattern surrounding the anticipated linguistic ability in the heritage language for each generation, this trend is generalized in nature and therefore does not account for variable circumstances – e.g., within generation variation when one child can exhibit strong productive and receptive skills while the other exhibits only receptive skills. An additional circumstance that the three-generation model is unable to account for is the expansion in the availability of heritage language learning programs and the number of students enrolled in them (Beaudrie & Fairclough, 2012; Beaudrie & Marrero-Rivera, 2024). More and more we encounter instances where this model can be considered 'interrupted' or 'reversed' through the presence of these programs and the deliberate decisions from individuals to enroll in them in an effort to develop and/or maintain their heritage language. Therefore, it has not gone unrecognized in the field of heritage research how formidable and at times seemingly implausible a task it is to describe and define the community of heritage speakers through a

singular definition that best acknowledges the very nuanced nature it exhibits. The many ways in which the field has approached defining the community is further explained in the literature review portion.

This dissertation examines the perceptual and productive patterns that exist within and around Spanish as a heritage language in the United States, specifically exploring how they relate to the ideological framework within the Latinx community that equates Spanish proficiency with Latinx identity. To contextualize these dynamics, the literature review begins by focusing on the significant role the Spanish language has had in the United States, recounting its history since its initial imposition on the region and examining how the language, as well as its community of speakers, has been susceptible to language loss and influenced by ongoing contact with English, thereby shifting and fluctuating over time. Crucially, although Spanish and its speakers have been minoritized within the US context, it remains a European colonial language and has served a similar hegemonic role to English by historically enacting political dominance and forced cultural assimilation across Latin America and other territories ruled by Spain. Outlining this history not only elucidates how Spanish became a subordinated language in the US, but also reveals how this historical transition in status continues to shape perceptions, alongside the unique productive patterns and languaging habits of US Spanish.

Building upon this historical foundation, the section then examines how the entrenchment of English monolingualism and nativist ideals in the United States has impacted both bilingual and multilingual practices, and the accompanying public perception of these linguistic behaviors and the communities who employ them. The within-community perceptions examined in this dissertation, as well as the external perceptions that exist, are a direct result of the role the English language has had in the minoritization of Spanish. While these dynamics affect

multilingual communities broadly, the scope of this dissertation centers primarily on how Spanish is structurally positioned as a threat to national unity, thereby influencing perceptions of and within the US Latinx community.

Consequently, particular attention is paid to the attitudes and ideologies that reinforce monoculturalism by framing English acquisition as a prerequisite to achieving and maintaining a unified nation. These beliefs have had a considerable impact in shaping language policy and multilingual educational practices throughout the US. Both have played an active role in hindering and suppressing the use and maintenance of Spanish and other heritage languages, creating a system that marginalizes non-English languages, devaluing them as familial languages but incentivizing them as foreign and second languages when acquired by native monolingual English speakers (Carter, 2018). Therefore, I will outline how language policy in the US has operated in tandem with the educational system by first developing voter-backed mandates and measures that contribute to heritage language loss and negatively target multilingualism in minoritized communities, and then utilizing the educational system as a primary facilitator to consecrate monolingual ideals. Indeed, the education system is often where individuals are first introduced to institutionalized ideologies such as those that aided in the establishment of the linguistic regime that cements the dominant status of English and reifies the negative perceptions around Spanish and its speakers (Flores & Rosa, 2015).

Through the racialization of language, this linguistic regime indirectly, albeit intentionally, establishes an ethnoracial hierarchy that sanctions implicit discrimination and can enable its explicit forms against minoritized communities (Flores & Rosa, 2015). The systemic devaluing of Spanish has encouraged and normalized the process of language loss and attrition by ascribing it to the necessary demands of societal integration and social cohesion in the United

States. The systemic pressure placed on the Latinx community to succumb to language shift has influenced both the internal community perceptions and the productive patterns I observe and examine. However, even with the normalization of language loss, the relationship between Spanish language proficiency and Latinx identity has not disappeared despite the imposition of English acquisition in the United States, making it a complex process of ongoing identity negotiation amongst Latinx.

The literature review concludes by synthesizing the many definitions that have been employed to describe the US Spanish heritage community. Throughout this section, I underscore the challenges and implausibility of developing a singular definition that accounts for the heterogeneity within the community due to the immense variation in Spanish language proficiency, national origins, cross-dialectal variation, and individual social experiences and upbringings. I further describe the role the education system has played in complicating the development of a definition and emphasize the need to embrace the term 'heritage language learner' as a way to account for variation in proficiency while including those who aim to reconnect with their Latinx heritage via language acquisition.

This dissertation contains two chapters that have been co-authored. The first co-authored chapter is Chapter 2 which is currently under review for the *Heritage Language Journal* and was written along with my advisor Dr. Melissa Baese-Berk. Chapter 4 is published in the special issue Investigating the Effect of Listening Conditions on Speech Intelligibility for the journal *Languages* and was also written with Dr. Melissa Baese-Berk, as well as Dr. Charlie Nagle. Chapters 3 and 5 are prepared for the potential of future publications but have not been submitted to journals.

The analysis chapters of this dissertation adopt a top-down approach by delving into larger-scale language production questions that aim to better understand Spanish heritage production more broadly and ending with more fine-tuned language production analyses[2]. Chapter 2 begins with a broader look at the language produced by heritage speakers by looking at their production in both English and Spanish. This study was motivated by the historical tendency to describe heritage speech as incompletely developed and inadequate (Silva-Corvalán, 1994; Montrul, 2002; Montrul & Bowles, 2009), and to describe heritage speakers as not knowing either language well or being *languageless* (Rosa, 2016) with the general view that heritage and nonstandard language is a deficient byproduct of poor multilingual practices (Rosa & Flores, 2023). These assumptions led to many studies making direct comparisons between heritage speakers and L1 English monolinguals, L1 Spanish speakers, L2 Spanish learners, etc. in ways that position these other language communities as the prototypical speaker whose language usage patterns should be directly emulated by heritage speakers. However, recent acknowledgement of a heritage variety of Spanish has led to pushback against claims arguing that these differences are traces of incomplete acquisition and instead argue they are directly influenced from the unique acquisition environment in which heritage speakers develop language (Pascual y Cabo & Rothman, 2012). Along these lines, the analyses in Chapter 2 serve to challenge the previously established norm by making observational claims regarding similarities or differences in production rather than claims that aim to highlight perceived deficiencies. Using lexical diversity scores I aimed to determine the extent to which monolingual English, English dominant Spanish heritage, and L1 Spanish-L2 English speakers exhibit comparable or variable

[2] The audio that either provided the data or the stimuli for each chapter were all borrowed from the Archive of L1 and L2 Scripted and Spontaneous Transcripts and Recordings (ALLSSTAR) developed at Northwestern University (Bradlow, n.d.). This corpus, and how it was used, is further explained in each respective chapter.

lexical productions in spontaneously produced oral speech. Lexical diversity scores are often used to approximate vocabulary size with higher scores equating to greater variety in the speech (Fusté-Herrmann et al., 2006). Results from this study support the idea that heritage production is a direct result of their acquisition environment and the constant contact between the languages in their lives. These noted differences should be understood for what they are rather than mislabeled as inadequate. Furthermore, findings in this study pave the way for a deeper discussion regarding the tools and instruments we as researchers employ to measure multilingual production. Historically speaking, heritage speakers have been the square peg awkwardly misplaced in the round hole. While using instruments designed for monolingual speakers or L2 learners has in fact measured something, the bigger question of whether these measurements are accurately interpretable has begun to arise more and more and is addressed further in this chapter.

Chapter 3 introduces the role that ideologies and attitudes have on speech perception by analyzing how varying social information about a speaker can impact listeners' perceptions of their speech, and their perceptions of the speaker's Latinx identity. This chapter introduces the key distinction between heritage *speakers* and *learners.* Previous research in the field mostly conflates these two terms using them in nearly interchangeable ways. However, given the increase in heritage language learning programs and the impact they are having on the ways we have previously understood models of heritage language maintenance, it is crucial to acknowledge the growing community of individuals who have chosen to enroll in Spanish language courses. Therefore, for the purposes of this study a narrow definition for heritage speaker was employed which describes *speakers* as individuals who acquired Spanish as an L1 in early childhood and underwent a shift in language dominance to English upon entering primary

school, while *learners* are individuals who have an ancestral or cultural connection to the heritage language and acquired it in an academic setting, whether that be through World Language classes or bilingual education (see Valdés, 2001). The Latinx community and Spanish speaking communities at large have long been known to associate proficiency in Spanish with the Latinx identity, so much so that it can be a deciding factor for some when determining in-group associations. These ideologies are centered around certain expectations and pressures placed on US Latinx and has had the residual effect of decreased confidence for heritage speakers when claiming their Latinx identity. The expectation to speak Spanish is, at times, extended to require a 'correct' and formal variety of Spanish (Showstack, 2017). Therefore, individuals that speak non-standard varieties of Spanish, limited Spanish, or do not speak Spanish at all, risk having their Latinx identity questioned by others (Rosa, 2019). These repeated experiences lead to the development of insecurities around identity (Potowski, 2012) and motivated the decision to examine how the role of education might influence within community perceptions. Findings in this study indicate the potential for a shift in expectations regarding speaker proficiency, accentedness, and fluency when making judgements regarding Latinx identity, which can likely be attributed to the fact that second-generation Latinx and later are outnumbering first-generation and immigrant groups (Abdi, 2011; Funk & Hugo Lopez, 2022).

Chapter 4 of this dissertation expands upon the third by further examining the influence of ideologies and attitudes on speech perception, also using a matched-guise methodology to distinguish between heritage *speaker* and *learner* identities. This study, however, incorporates intelligibility measures to determine if differing social information impacts to any degree how intelligible a particular heritage speaker or learner is perceived to be by others within the heritage

community. An additional step was taken to determine whether individual bias may play a role in listeners' perception by analyzing how listener attitudes interacted with and/or influenced their perception of others. Results in this study elucidate that despite being a good indicator of how perception may play out in real time interactions, expectations do not always determine how a listener may perceive or process language. The general descriptive trend showing that listeners expected to notice a difference in Spanish ability between guise types did not actually impact intelligibility, nor did listeners' attitudes regarding the importance of knowing and maintaining Spanish.

The final analysis chapter, Chapter 5, attempts to understand if there are certain features within the audio stimulus that may be triggering perceptual patterns amongst heritage participants leading to findings that do not replicate previous ideas around the influence of social information. The talkers used in chapters three and four may exhibit productive patterns that trigger certain expectations we are not yet capable of ascertaining. For this reason we focused on the talkers' productions of /bdg/, which are lenited in the intervocalic position most often to become pure approximants (Rao, 2015). Given that the realization of these phonemes is known to correlate with listeners' perceptions of accentedness (Schoonmaker-Gates, 2015) across both English and Spanish, I found this to be an interesting starting point. We also included the variable of speech style given that literacy rates in Spanish throughout the heritage community are typically not very high, unless they were educated in Spanish as well; therefore, likely triggering some disfluencies while the talkers were reading the scripted stimuli. Overall, we found insignificant differences in production of the Spanish stops between talker groups and speech styles, however, we were able to notice descriptive trends. Talkers from the previous two studies who were perceived as aligning more with the speaker guises patterned similarly to

monolingual Spanish speakers while the talkers who aligned more with learner guises had inconsistent patterning in their stop productions.

1.1 The history of Spanish in the United States

The Spanish language has long had a place in what is now the United States since its first introduction as a colonial language in the 16th century when Ponce de Leon arrived in Florida (Potowski & Carreira, 2010). In fact, the first established European settlement was under Spanish rule making it the first European language spoken in the territory. Many historical events following this initial introduction would lead to a drastic shift in the socio-political influence of Spanish from the once powerful language that epitomized colonial rule, to an inferior immigrant or minoritized language (Lozano, 2022). However, the acquisition of new land to the United States territory and subsequent waves of mass migration has allowed Spanish to maintain the robust presence it still has today in this country. Throughout this diachronic shift in social prestige, the Latinx and Hispanic populations have continued to grow – mostly fueled by the country's expansion westward – establishing their collective position as a dominant part of society and eventually making the United States home to the second largest Spanish speaking population in the world, only behind Mexico (Beaudrie, 2012; Pascual y Cabo, 2016).

One of the most significant increases in the presence of Spanish and its speakers in the United States was following the Mexican American War in 1848, which was ultimately resolved in two stages. These two stages collectively resulted in the United States acquiring all of California, Nevada, and Utah, as well as parts of Arizona, Colorado, Texas and Wyoming, in addition to a sudden influx of new citizens – many of whom had originally resisted joining the quickly expanding US. Thus, the Mexican citizens who resided in these areas suddenly became citizens of the United States in a fashion most appropriately summarized in MacGregor-Mendoza

(1998) in which she implores us to remember, "… that many of the inhabitants in the Southwest regions of the Union had not come to America willingly, rather, it is more accurate to state that America had come to *them*" (p. 57). During this process of integration into the Union, those who were previously Mexican citizens were guaranteed to maintain land ownership rights and religious and cultural freedoms as they navigated inhabiting a new country with new policies and expectations (Kloss, 1977). These promises unfortunately were unfulfilled in very tangible ways. Not only did most of the formerly Mexican citizens have to sell their land due to staggering legal fees or lose it to squatters during the Gold Rush, but the promise and guarantee of linguistic and cultural freedom was, in retrospect, laughable given the intentional efforts made by the US government to hinder and even completely eliminate these freedoms, as outlined below.

The deliberate efforts to not only discourage the maintenance and growth of the Spanish language but to also reduce its usage are most obviously seen when looking into how boundaries between state lines were drawn, as well as when certain territories were granted statehood. State lines were drawn in the Southwest region in ways that ensured an English-speaking majority population, and in instances where this was not readily possible, there were delays in granting statehood until an English-speaking majority was achieved. States like Arizona and New Mexico spent roughly 60 years as territories for no other reason than the high population of Spanish speakers and the perception that English was still a foreign language in these territories (MacGregor-Mendoza, 1998).

Another notable era in US politics follows the Spanish American War of 1898, which ultimately resulted in Spain ceding control of Puerto Rico to the United States – in addition to other colonial territories who were at the time operating politically in Spanish. However, it was not until the Jones-Shafroth Act of 1917 that the island officially became a territory granting

statutory citizenship to its inhabitants (Library of Congress). The 1920's saw a rapid influx in migration from the island of Puerto Rican to the Northeast US mainland so much so that East Harlem in the early 1930's was no longer recognized for its large Jewish population but rather its Puerto Rican and Hispanic population (Lynch, 2018). Another boom in migration was recorded in the post-World War II decade with approximately 50,000 Puerto Ricans entering the mainland annually (Zentella, 1997).

Concurrently, the early 1940's during World War II was also met with massive growth in the US Spanish speaking population when an agreement was reached with the Mexican government to institute the *Bracero* program. This program invited around 250,000 Mexican citizens over a five-year span to come to the US to mitigate work shortages during the war (Lynch, 2018; Potowski & Carreira, 2010). Massive growth in the Mexican population was not accepted with open arms – especially coming not long after the Great Depression, when Mexicans were often accused of taking American jobs. These previously existing hostilities compounded with the increased patriotism at the time of the war led to many negative sentiments towards the invited Mexican population and their languaging habits.

A similar trend was seen in the 1980's when Cuba underwent what is referred to as a massive refugee exodus between April and September, bringing in approximately 125,000 Cubans known as the *Marielitos* (Pedraza, 2004). The general perception of and attitudes toward this group of Cuban refugees were negative and riddled with stereotypes, despite research demonstrating that claims of criminality and danger were unfounded (Aguirre, 1994). The *Balseros* in the 1990's also introduced a notable increase in the presence of the Cuban population in the United States and Spanish speakers in areas outside of the generally anticipated Southwestern region (Ackerman, 1996).

In fact, the Spanish language and the cultures and communities that accompany it have grown to be so ubiquitous that although the US Southwest, New York, and Florida are most commonly known for their significant Hispanic and Latinx populations, communities previously viewed as being devoid of a substantial or notable Latinx presence are experiencing some of the largest growth – such as North Dakota, South Dakota, Montana, and New Hampshire (Beaudrie & Marrero-Rivera, 2024). While Mexico remains the largest contributor of Spanish speaking immigrants to the United States, the latter portion of the 20th century saw an increase in immigration from the Dominican Republic, Venezuela, and countries throughout Central America, ultimately contributing to the increase in Latinx populations outside of the regions historically known for their presence (Haner & Lopez, 2023).

This sustained growth in the Hispanic/Latinx community has made it not only the largest ethnic group in the country, but also the fastest growing ethnic group. Numbers have jumped from 50.7 million people in the 2010 census to 62.1 million in the 2020 census, representing roughly 18.7% of the overall population. Considered to be a benchmark of population and linguistic growth by the U.S. Census Bureau, the number of states reporting Hispanic densities above 10% grew from including only Colorado and the four border states in 1980, to including 17 states in the 2010 census (Jenkins, 2018) and now 28 states according to the 2024 census (U.S. Census Bureau, 2024). Furthermore, in the ten years between the 2010 and 2020 census, not a single state saw a decline in its Hispanic/Latinx population, and current projections predict a steady, continual growth that will anticipatorily lead to a Hispanic/Latinx population of *more* than 111 million people by 2060 (Beaudrie & Marrero-Rivera, 2024).

Therefore, the most pressing question for many who have examined this history became how – despite its original role as a colonial language that imposed assimilation, its long-standing

place in the country predating even that of English in some regions, and its continual usage facilitated by far more than just immigration – Spanish underwent such an extreme fall from grace as it pertains to its status as a spoken language in the United States. It cannot be ignored that these historical events did not develop in a vacuum. Each moment in history that is marked by a growth in the Latinx community coincides with both covert and overt efforts from the US government to reinforce and perpetuate ideals that exalt English as the superior language and reify negative sentiments toward those who speak Spanish, and by extension any language other than English (Lawton, 2013). Although extensive historical analyses have led sociolinguists to argue that, "far from being exotic or foreign, Spanish is a thoroughly American language." (Lozano, 2022, p.33; see also Alonso, 2006; Lozano, 2018; Fuller & Leeman, 2020), repeated actions taken by the US government have allowed for increased distance between the widespread social perception of the place Spanish has in this country and the historical reality of the circumstance. Spanish soon became associated with immigration, and the migrants who speak it or speak with a Spanish accent are deemed outsiders who are "non-citizens, inherently illegal, poor, uneducated, criminally inclined, [and] refusing to learn proper English" (Rosa et al., 2022, p.98).

These deeply internalized negative ideologies about Spanish stem from years of their foregrounding and the continual careful work by even the earliest of policymakers to establish them systemically. While the US government is not unfamiliar with employing and allowing the usage of Spanish and other minoritized and/or foreign languages when it is convenient and advantageous (Nichols, 1989), this limited and restricted usage has allowed for the rich and critical history of the Spanish language to be misrepresented, almost to the point of erasure (Urciuoli et al., 2022). While Spanish is the largest minoritized language in the US, any language

other than English has long been susceptible to scrutiny at both the social and political level. The United States has always ensured that the English language sustains a superior status relative to that of any other language that may be heard within the country and has, therefore, been a driving cause for the ways in which multilingualism in the United States is approached.

1.2 Multilingualism in the US

The role of English throughout the history of the United States has always been that of a resource employed for assimilatory purposes to devalue and delegitimize minoritized languages. More often than not, it has been forcibly pushed upon communities that do not already speak it – whether they be the native population who originally inhabited the land, immigrants to the US, or those who became 'overnight citizens' in the aftermath of war and land expansion. Traces of English Only policy date back to the 18th century and continue to be a constant and persistent theme deeply ingrained in even contemporary US policy (Balestra et al., 2008; Rosa & Flores, 2017). Essentially, the promotion of a single unifying language has long been a foundational quality of US colonialism and governance, relying almost entirely on what sociolinguists recognize as the ideology of *linguistic purity* where one language is viewed "…as naturally corresponding with one ethnonational lineage, forging a primordial bind between that language and the spirit of the nation" (Dick, 2011, pp. 230-231).

Early historical accounts are riddled with these sentiments, expressing language as an integral facet of national identity (Wiley, 2014) where the "language" referred to is that of a *standard* variety of English spoken by the ethnic and social group in power. These sentiments are ultimately centered on the existence of a prototypical, idealized "native speaker" who serves as both the model for any minoritized group identified as "other" (Bonfiglio, 2010) and as the justification for the many actions taken to assimilate and/or disenfranchise them. The

racialization of language throughout history refers to the social process of mapping race onto language in an effort to use language as an inherent quality that deems the dominant group as undoubtedly superior and others as inferior without having to directly use race as the basis for this hierarchical status (Rosa, 2016; Urciuoli et al., 2022). This process has been highly successful and has allowed for countless attacks on minoritized communities to be carried out via attacks on their language usage habits and practices. Therefore, not to be fooled by the faux sense of hubris when the US boasts its "melting pot" metaphor, the goal of early leaders and policymakers has always been the advancement of a single superior culture unified by a single tongue.

Although it is well acknowledged that assimilation has served as a common thread woven throughout US history, a key distinction must be made between the types of assimilatory processes employed and how they have differentially impacted communities that speak minoritized languages. Behavioral assimilation is identified as the favored approach adopted throughout the larger sociopolitical culture of the US. This process employs educational practices simply for the purposes of ensuring subjugation, while the alternative process, Structural Incorporation, is reserved for communities deemed worthy enough for social and economic integration and full participation in society (Weiss, 1982; Wiley, 2014). The Spanish language and its speakers have historically fallen under the group expected to undergo a form of behavioral assimilation rather than structural incorporation, thereby ensuring it remains ideologically perceived as an inferior language associated with inferior cultures.

Early policy in the country originally allowed for limited usage of Spanish to facilitate a smooth transition process in instances such as with the mass acquisition of land from Mexico. However, this was short-lived and does not imply that the intention of eliminating the Spanish

language was not there from the very beginning. Rather, the permitted usage was solely meant to benefit the US government and not intended to necessarily support Mexican settlers (Lozano, 2022). For example, when the US acquired what is now the Southwest and granted immediate citizenship to the Mexican citizens residing in these territories, this came with the exclusion of the Indigenous communities in areas like New Mexico and California. The US government then relied on the Spanish speaking population to aid in diminishing and eliminating Indigenous control of the lands, thus leading to Spanish being permitted in social and political institutions (Balestra et al., 2008). Despite this, traces of English Only frameworks were still prominent when considering the many steps taken to ensure an English majority population in each newly integrated territory.

It was by the end of the 19th century that clear anti-bilingual measures began rapidly appearing and extended into the 20th century, leading to the passage of laws like the Nationality Act of 1906 which required English speaking ability to obtain naturalization (Carter, 2018). Anti-bilingual and anti-immigration sentiments were exacerbated amid heightened patriotism in the post-World War I era and most of the legislative frameworks developed strongly prioritized national identity. Language became a clear target, and efforts to further promote monolingualism as a means of achieving a unified nation were so successful that by the early 1920's there were 34 states that had passed legislation deeming English the official language (Wiley, 2014).

This restrictive status-quo persisted until the mid-20th century; indeed, it was not until the 1960's and 70's that legislation lessened its attacks on multilingualism and began extending some protective rights toward immigrant populations and individuals with limited English-speaking abilities. The passage of the Civil Rights Act of 1964 made it illegal to discriminate against someone on the basis of national origin – and by extension, individuals who speak a

language other than English. The subsequent passage of the Bilingual Education Act in 1968 also established a state and district requirement to address the needs of English learners (Gándara & Orfield, 2012). While these federal initiatives may have temporarily provided a false sense of hope to minoritized communities in this country, states in the Southwest with exceptionally high rates of Spanish speakers actually began introducing more restrictive policies to limit Spanish usage and attack multilingualism.

Within this Southwestern landscape, the state of California is one of the states that – while well known for its large immigrant and Spanish speaking populations – has been particularly active in ensuring the subordination of these communities via the passing of legislation and policies often directed toward education and language rights. It was in the late 1980's that the state denied the extension of the previously established Chacone-Moscone Bilingual-Bicultural Education Act ending the requirement for districts to provide bilingual and bicultural education to non-English speaking students (Johnson & Martínez, 2000). In 1994, Prop 187 was passed by voters which denied undocumented immigrants certain rights including access to public education – an initiative that would disproportionately impact the Mexican community, although this was deemed unconstitutional and never actually implemented. The most notorious policy passed in California, however, is likely Prop 227 which was passed in 1998 and completely banned bilingual education while enforcing English immersion for non-English speaking students in public schools. The circulating rhetoric at the time attempted to mask the oppressive monocultural undertones of this proposition by blaming low English literacy rates and high drop-out rates among Latinx on an apparent ineffectiveness of bilingual education.

This ban lasted for 18 years until the passage of Prop 58 – originally known as Education for the Global Economy initiative and subsequently renamed Language Education, Acquisition

and Readiness Now – which removed previous English-only requirements and reintroduced bilingual education opportunities. On the surface the passage of Prop 58 seems like a step in the right direction moving away from previously enforced hegemony toward a society that embraces other cultures; however, a deeper dive into the actual phrasing of the proposition along with voter motivations potentially reveals otherwise. Studies have found that the motivations behind this bill do not seem to reflect an actual desire to increase cultural acceptance, civil rights, or greater equity throughout society. Rather, they reflect neoliberal ideals to develop economically and globally advantageous skills amongst students, reducing language to a resource (Muñoz-Muñoz et al., 2023; Katznelson & Bernstein, 2017; Kelly, 2018). Further, scholars argue that carefully curated language may have played in a role in influencing voters given that the bill makes mention of English proficiency – known to be important in a state that values English-only ideals – and shifts from using the term bilingualism to multilingualism as a way to better suit the outlined narrative of "language as a global resource" (Citrin et al., 2017).

Similarly to California, the state of Arizona is another state in the union well known for its historic tendency to attack immigrant and Spanish speaking populations through years of systemic oppression aimed at furthering an English speaking agenda. Since the passage of the state's original Official English legislation in 1988, Arizona has been considered among the strictest and harshest with regard to its treatment of languages other than English, with particular focus on Spanish (Combs & Jaime-Diaz, 2022). This early legislation included a comprehensive ban on the usage of any language other than English in the public sector – a mandate so restrictive that it was determined to be unconstitutional in 1997 by the U.S. Supreme Court (Cashman, 2009). Despite this judicial ruling, governmental and societal support for English-only frameworks persisted, and in late 2000 Proposition 203 was passed, which banned bilingual

education for non-English speaking students and enforced English immersion (Gándara & Orfield, 2012). The circulating rhetoric surrounding Prop 203 highlighted the importance of English by positioning it as granting access for immigrant children and focusing on the economic and global advantages it might afford both them and society as a whole (Bernstein et al., 2025).

Continued support led to English once again being voted as the official language of Arizona in 2006, less than 10 years after the previous federal ruling against it. In doing so, the ban on the usage of languages other than English in the public sector was reinstated, with law enforcement and emergency services being the sole exceptions to this ban (Cashman, 2009). Even when hints of bilingual education were reintroduced to the state, this was done in a manner that can only be perceived as a snub towards minoritized communities (Yamagami, 2012). The bill passed to implement the Critical Languages and Economic Development Pilot Program included explicit guidelines to exclude those considered native speakers of the targeted languages and include only "native" speakers of English (Kelly, 2018). This parameter clearly signaled that enrollment was reserved for individuals who fit the prototypical category of the monolingual English Speaker in the United States. Mirroring the motivations of Proposition 58 in California, the reintroduction of bilingual and multilingual education in Arizona represents an economic and global interest – with some focus on national security – while the core principles of equity, cultural acceptance, and civil rights remain of little importance to the government (Kelly, 2018). Thus, the discussion of multilingualism in the United States is a difficult one muddied by consistent attempts to belittle and almost erase cultures that do not align with the hegemonic majority, while endorsing the adoption of minoritized languages and cultures only when commercially or geopolitically advantageous and convenient.

1.3 Spanish as a Heritage Language in the United States: Ideologies and Attitudes

Originating in Canada in the late 1970's the term *heritage language* eventually gained wide acceptance among researchers and academics in the United States beginning in the 1990's (Cummins, 2005; Montrul, 2010). The term has since been employed to refer to immigrant, indigenous, minoritized languages in the country (Beaudrie & Fairclough, 2012; Potowski, 2018). Spanish as a heritage language has gained particular focus given the prolific nature of the Spanish speaking community in the United States and the long history of socio-political oppression that has led to difficulties in language maintenance or complete language loss. The Latinx community in the US has been susceptible to attacks and marginalization due to ideologies throughout the country that place immense value on assimilation to the dominant culture through the acquisition of English, while simultaneously demonizing the Spanish language as a threat to national unity (Lawton, 2013) and even as a threat to the general well-being of children raised in Spanish speaking homes (Carter, 2018).

Kroskrity (2005) defines language ideologies as beliefs or feelings about languages as they are used in their social worlds; however, the dissemination and adoption of language ideologies have been accurately described as a tactic used for social and political control over groups deemed linguistically distinct from the group in power (Martínez, 2006). Homing in on these differences has been a key factor throughout US history in establishing a linguistic hierarchy that places English at the top and has often served to justify the discriminatory treatment of the Latinx community. Therefore, the role of language ideologies has been to assist in the development of this linguistic hierarchy via attempts to direct culture change and to alter identities that differ from the dominant and covertly, yet purposefully, resulting in an ethnoracial hierarchy in the process (Kroskrity, 2005). Acquiring English has been posed throughout history

as the ability that opens up endless opportunities and paves the way for social mobility, forcing a choice that makes assimilation appear more attractive and desirable despite there being no other option.

Those that don't speak English are, in turn, ridiculed and perceived as hopelessly disadvantaged, unable to fully participate in society or make a meaningful impact. In fact, many throughout the country view the continued usage of Spanish as a direct snub and insult to the English-speaking dominant majority, inciting fears that manifests in the form of *language panic*. Language panic is a phenomenon that expands from previous definitions around moral panic, which occurs when public discourse becomes centered around a particular social phenomenon or problem in a way often described as obsessive and alarmist (Cameron, 1995). Thus, within the paradigm of language panic, technical issues surrounding language usage are hyper fixated on (Hill, 2001). This alarmist approach tends to manifest through actions like English-only mandates and restrictive bilingual education policies (Cashman, 2009), dialect dissing – which describes attacks against US varieties of Spanish and Spanglish (Zentella, 2003) – and the usage of mock Spanish (Martínez, 2006).

Essentially, the clear purpose of inciting sentiments of language panic is to reinforce ideologies that elevate English and whiteness as superior and dominant, while perpetuating negative stereotypes that belittle and subjugate the Spanish speaking community (Hill, 2001). In doing so, public anxieties have been directed at the Spanish language, and consequentially at those who speak it. These directed and focused public anxieties result in the community and their language being viewed as problematic in and of themselves, and as the underlying cause for society's problems at large (Cashman, 2009). For example, the pursuit of policies like Prop 227 in California discussed earlier, was made successful through the societal belief that Mexican

immigrants, and Spanish speakers more generally, are criminals, violent, and uneducated (Yamagami, 2012) framing them as inherently threatening to the social and political foundation of the country (Zentella, 1997). Therefore, acquiring English is viewed as an opportunity afforded to them to dispel their problematic ways and assimilate into mainstream American society. Given that public perception views the pressures placed on the Spanish speaking community as stemming from an altruistic place rather than a racist one, the development of discriminatory policies directed at the Latinx community are justified and argued to be devoid of anti-immigrant sentiments (Yamagami, 2012).

In the process of expecting the acquisition of English, minoritized and heritage languages have then been relegated to the familial domain and restricted to private use. The expectation that heritage languages will be developed in the home does not represent an acceptance of heritage languages nor does it imply that the US finds value in heritage language maintenance. Instead, it places the responsibility entirely on the family to continue generational usage of the heritage language. When faced with the stark reality of public perception of non-English languages, the task becomes nearly impossible. The difficulty these families experience in maintaining their heritage language is at direct odds with the ideology rampant throughout the US and regurgitated by influential leaders and media sources without regard: that Spanish speakers are lazy and simply prefer not to, or completely refuse to, learn English (Zentella, 1997; Huntington, 2004). Pressures to assimilate have been found to lead to the internalization of these ideologies among US Spanish speakers, frequently rendering them highly anxious to acquire English quickly. However, the motivations behind this swing between a desire to be able to pursue avenues like work and education and participate in society (Zentella, 1997), and a desire to simply avoid public discrimination and racism (Cobas & Feagin, 2008).

The simple usage of Spanish in public has led to aggressive attacks and mistreatment, making its use at times dangerous and a risk (Rosa et al., 2022) – ultimately facilitating the rapid process of language loss. Overall, there is a general impression that there is not a safe and accepted space to use the Spanish language that doesn't either pose a threat to the Anglo-dominant majority or lead to suspicions that those choosing to speak Spanish are doing so out of defiance or a lack of interest in learning English. As a result, research has demonstrated that Spanish speakers in the US do indeed experience a shift to English monolingualism, and the community at large displays trends of language shift, better yet loss, at the same pace or even quicker than European immigrants (e.g., European communities who immigrated in the post-World War era) (Rumbaut et al., 2006) highlighting how profoundly misconceived the ideology is that asserts stating they are not learning or refusing to learn English.

There is a plethora of research demonstrating how detrimental the nativist ideals of the United States have been to the survival of minoritized languages (Portes & Schauffler, 1994; Portes & Hao, 1998; Alba et al., 2002) leading to the nation earning a reputation of a *linguistic graveyard* (Portes & Rumbaut, 2006; Rumbaut, 2009; Rumbaut & Massey, 2013). Contrary to public opinion, Spanish has not been immune to this seemingly inevitable fate. This systemic decline is conceptualized as the *three-generation model of language loss* (Fishman et al., 1966, 1972; Veltman, 1983), a framework through which researchers have long analyzed the influence of assimilation on the paradigmatic shift toward English in minoritized families. In this model, first-generation immigrants acquire some English but still prefer the comfort of their mother tongue, by the second generation there is greater preference for English but maintained usage of the minoritized language at home, and finally by the third generation a clear preference towards English is shown and only English is spoken in the home. This preference often develops as the

perception of the usefulness of the heritage language begins to decrease with each generation. Given that society presents the notion that Spanish is useless, problematic, and even a hinderance, there no longer is a "need" for it once everyone in the family can communicate in the majority language.

The US does, however, at times assign value to heritage languages although the assigned value is not due to a desire to embrace multiculturalism and diversity. Rather it is motivated by neoliberal ideologies surrounding multilingualism that focus the purpose of language as a way to connect globally and advance economically (Katznelson & Bernstein, 2017). Even more, the benefit of embracing multilingualism and the usage of non-English languages is seldom reserved for actual heritage speakers and is more often reserved for Anglo L2 learners (Schmid, 2001). Researchers have long studied this asymmetrical treatment of minoritized languages, defining phenomena such as *differential bilingualism* (Aparicio, 1998) or the *double figuration* of Spanish (Carter, 2018). Both concepts refer to the trend wherein multilingualism is viewed as a hinderance to minoritized populations or heritage speakers – in a manner that prevents full inclusion in society and *Americanization* – while, on the other hand, it is celebrated as an economic and global resource for those within the Anglo majority.

The ideological polarity around multilingualism that is introduced by the commodification of Spanish and other subordinated languages is exacerbated when considering that even attempts to view heritage multilingualism as equally beneficial to Anglo multilingualism are consistently thwarted. For example, within workplace dynamics, the same employers who seek out multilingual candidates, have been known to fire those same individuals for their usage of Spanish when it is used on tasks unrelated to work (e.g., when co-workers are communicating amongst themselves rather than with a customer in need of assistance) (Carter,

2018). Therefore, it is not enough that heritage speakers are expected and forced to lose their home language; they must also constantly observe instances in which native English speakers can openly communicate in Spanish, or any minoritized language they choose, and are in turn praised for their knowledge and ability to contribute as a globally oriented members of society. Importantly, even amid the many challenges encountered in language maintenance and high rates of language loss, the Latinx community consistently pushes back against this systemic devaluation, perceiving attacks toward Spanish as direct personal attacks on their heritage and identity (Johnson & Martínez, 2000). Therefore, the willingness to learn English and avoid discrimination does not necessarily represent or rely on a desire to abandon or lose a connection to the heritage culture.

1.4 Spanish Heritage speakers: Definitions and terminology

The Spanish heritage community in the United States is consistent in at least one factor – variation – allowing us as researchers to always expect high variability whether it be regarding language proficiency, background, and/or experience. This variability remains a constant given expansive manner in which the term "Spanish heritage" has been used to describe a wide array of individuals within the Spanish speaking and Latinx community in the US. Over the years, institutional and scholarly attempts to formulate a definition best suited to describe this population have differed with regard to how they conceptualize the three variables above; however, all definitions agree that a familial and cultural connection to the heritage language is essential. As a result, definitions for heritage speakers have been categorized by researchers into broad and narrow paradigms, with narrow definitions being those that implement strict parameters regarding language proficiency.

The formation of these two categories has been guided by two of the most well-known and well-accepted definitions in the field. Valdés (1997) popularized the reliance on proficiency ultimately paving the way for more narrow definitions surrounding Spanish heritage speakers. This framework defines a heritage speaker as an individual who is raised in a home where a non-English language is spoken and who is, to some degree, bilingual in the heritage language and English (Valdés, 2001). The broad and general definition was developed by Fishman (2001) describing Spanish heritage speakers as individuals who were born in the United States and have an ancestral and cultural connection to the Spanish language.

Later definitions developed tend to stem from these two original definitions by either placing primary focus on speaker proficiency or the individual's relationship to their heritage. However, as has been acknowledged in previous works (e.g., Beaudrie & Fairclough, 2012; Wiley, 2001) each of these definitions, including the ones provided by Valdés and Fishman, attempt a notable feat in trying to describe such a variable community in a singular way. Researchers instead must make conscious and deliberate decisions when adapting and refining definitions and these decisions depend on the individual or group being described *and* the one describing the population, often reflecting and reinforcing their goals and interests. It is further argued that while the various definitions provided in the past have been appropriate for the specific contexts and communities of heritage speakers they are describing, they have difficulty being extended further than that (Benmamoun et al., 2014). This can result in conclusions lacking in external validity or findings that are overrun with a series of caveats.

Ultimately, definitions describing Spanish heritage speakers in the US fluctuate in degrees of conservatism regarding who 'qualifies' as a heritage speaker with one of the most conservative and narrow views often describing the heritage population as, "asymmetric

bilinguals who learned language X – the 'heritage language' – as an L1 in childhood, but who as adults, are dominant in a different language" (Benmamoun et al., 2014, pg. 260). Conversely, the broadest of definitions simply describes heritage speakers as individuals with an ancestral or cultural connection to the heritage language with a nondescript level of command and proficiency in that language. The narrow definition does not adequately capture this linguistic variation, and its overuse without recognition of the inherent within-community diversity, risks conveying the notion that the term *heritage* is representative of a monolithic population. On the other hand, the inclusive nature of the well-established definition provided by Fishman effectively mitigates exclusionary concerns, while the lack of detail regarding speaker proficiency allows the researcher to operationalize the metric in a manner that most appropriately suits the group of heritage participants. Still, the lack of granularity in the broad definitions and the limitations of the narrow definitions have not been taken for granted, as they present distinct practical challenges when considering student placement in language courses and programs.

The term *heritage language learner* has been used in education and language acquisition research to refer to individuals who *are identified as* heritage speakers and have enrolled in language learning courses. Emphasis is placed on the action of them being identified by others given that the nature of heritage courses tends to exclude those with limited proficiency in the heritage language (Beaudrie & Ducar, 2005). Courses and programs designed for Spanish heritage speakers are becoming more common throughout the country even in regions not typically recognized for their large Spanish speaking population (Beaudrie, 2012). However, historically speaking, this availability has been scarce. Even with this notable increase these programs are still not a regular offering across institutions; thus even heritage speakers who exhibit high command of Spanish have been known to enroll in non-heritage programs – such as

foreign language or second language courses – and sometimes do so intentionally due to the endurance of standard language ideologies that place languages within a social hierarchy that aims to ultimately reinforce the self-proclaimed power of the dominant majority (Ducar, 2012). For this reason, the term *Spanish heritage language learner* has been applied across various studies to describe, "anyone who is learning the language of their ancestors in a cultural context in which it is not the official language of power" (Sánchez-Muñoz, 2016). This definition, therefore, includes those who are enrolled in mixed classes – which are comprised of both L2 learners and heritage learners in the same classroom – in addition to those enrolled in bilingual education courses, heritage courses and courses referred to as Spanish for Native Speakers.

A browse through the available literature on this population reveals that the many studies adopting this terminology assume a degree of proficiency in the language ranging between beginning and advanced levels of receptive and productive abilities. However, beginning level language learners are significantly less researched and are sometimes even erased altogether when considering who is allowed to enroll in, or who is placed in, programs designed for those that have an ancestral and cultural connection to the language. Dating back 20 years to the early 21st century, predictions arose that advanced level heritage learners would become far less common on university and college campuses, leading to an overall increase in students with beginning level heritage language abilities or very limited abilities altogether (Carreira, 2003, 2004). Again, these individuals *usually* encapsulate the third and fourth generations, given the way in which language loss manifests in this country, and represent potentially the most highly debated subgroup when considering who gets included in the already umbrella term of *heritage*.

Researchers have even been found to question the legitimacy of those who lack any linguistic ability in the heritage language claiming they should not be regarded as 'true' heritage

speakers when considering their space in the realm of language acquisition (Polinsky & Kagan, 2007). The erasure of this portion of the community from even research settings is particularly concerning, leading to attempts to disentangle the complicated web that is woven when considering linguistic heritage vs cultural heritage. Scholarship highlights that individuals with limited linguistic ability are often excluded from heritage language programs, even when they are socially recognized as heritage (Beaudrie & Ducar, 2005). Those questioning the legitimacy of these individuals have opted to develop differentiating terminology such as *culturally motivated learners* (Polinsky & Kagan, 2007) or *learners with a heritage motivation* (Van Deusen-Scholl, 1998; Van Deusen-Scholl, 2003). The resistance in identifying them as heritage speakers lies in the drawn parallels between their acquisition process and that of L2 learners of Spanish (Wiley, 2001). Essentially, these individuals would not exhibit the expected advantages that are usually afforded to others in the heritage community who have had substantial prior exposure to Spanish throughout childhood.

Although even this has been found to lack complete validity when considering *childhood overhearers* or individuals who did grow up hearing the language, but this exposure never extended past simple background noise – meaning they never attempted to speak the language, and they were not directly spoken to in the language (Au et al., 2002). Even the reduced act of only overhearing the language was found to lead to advantages in pronunciation and phonology when compared to L2 learners, however, this advantage does not extend to the morphosyntactic domain (Knightly et al., 2003; Au et al., 2008). While this does present an interesting caveat when contemplating course and programmatic development, one's linguistic ability – while indeed incredibly interwoven and linked to cultural identity – is not by any means the only defining factor and excluding this portion of the population risks further perpetuating feelings of

insecurity that influentially impact issues of identity, inclusion, and exclusion (Beaudrie & Fairclough 2012; Wiley, 2001).

CHAPTER 2: A COMPARATIVE ANALYSIS OF LEXICAL DIVERSITY IN ORALLY PRODUCED SPANISH AND ENGLISH SPEECH

The following chapter is a co-authored chapter that has been submitted to the *Heritage Language Journal* for publication and is currently under review as an independent manuscript. I was primarily responsible for the conceptualization, methodological design, carrying out the investigation, data curation, statistical analyses, and writing and developing the original draft. Dr. Melissa Baese-Berk contributed substantially to this work by supervising the work, assisting in each stage of the research process, and guiding the edits and revisions for the manuscript.

2.1 Introduction

For the past decade there has been an influx of research dedicated to better understanding the acquisition process and linguistic systems of heritage Spanish speakers in the United States (Lynch, 2008; Montrul, 2009; Rao & Ronquest, 2015; Cuza & López Otero, 2016). Overall, the Spanish heritage community is receiving greater support through both the increase in both research and education programs focused on attending to the specific needs of heritage learners. Early research efforts to inform these programs had initially argued that heritage speakers' linguistic differences were a direct result of incomplete acquisition or developmental deficiencies (Polinsky & Kagan, 2007; Polinsky, 2008). However, more research is challenging this narrative and forcing recognition of the influence of environmental differences as well as differences in input and exposure to both Spanish and English that may lead to the noted differences in production (Pascual y Cabo & Rothman, 2012). This research is drawing upon the understanding that multilinguals and heritage speakers do not operate through separate and distinct linguistic systems, nor do they represent individual monolinguals in a single mind (Grosjean, 1989; Kroll

et al., 2015). Rather, these systems interact with each other representing a continuous system containing linguistic information from each of their respective languages (Urciuoli, 1985; Vogel & García, 2017; Toribio & Duran, 2018). Therefore, it is crucial to examine both Spanish and English productions when looking at heritage speaker competency to truly make sense of results and interpret findings.

While these differences are beginning to be understood as indicative of a heritage variety of language instead of a "deficient" form of a standardized language, much of the research to date has focused on grammatical features (Zyzik, 2014; Toribio & Bullock, 2016), leading to limited knowledge regarding lexical development. Given that lexical knowledge has garnered limited attention, this study will focus on examining lexical differences between Spanish Heritage speakers, English monolingual speakers, and L1 Spanish-L2 English speakers in both English and Spanish. Lexical diversity, while unable to provide a detailed, deep dive into specific lexical decisions made by speakers, appears to serve as an appropriate mechanism for an initial comparison between populations given that it is directly influenced by language input levels and their shift over time (Jacobson & Walden, 2013). Through this initial examination we are able to compare heritage speakers with their counterparts (i.e., often compared Spanish speaking communities like L1 monolingual speakers or L2 Spanish speakers, as well as monolingual English speakers).

2.2 Background

2.2.1 Spanish heritage speakers

The Spanish heritage population in the United States is ever-growing as many Latinx communities continue to undergo generational language attrition, sometimes resulting in complete loss of the heritage language(s) in families (Cashman, 2009; Silva-Corvalán, 2004).

This generational language attrition is a result of sociopolitical factors that have led many families of Latinx descent to feel pressure to assimilate to the dominant culture of society – Mainstream American society – and in doing so promote acquisition of the majority language within their families (Leeman, 2012). This study will focus on Spanish as the primary heritage language of interest, though it is important to acknowledge that many within the heritage community may also be undergoing identical experiences in additional familial languages.

As a result of factors driving generational language attrition, heritage speaker communities are highly variable with regard to Spanish proficiency, in addition to the linguistic and cultural backgrounds that shape their production. Therefore, it is often difficult to develop a singular definition that appropriately encapsulates the community as a whole. However, for the purposes of this study, we define a heritage speaker as an individual that speaks a heritage language, where the heritage language is defined as a minoritized language in a given country, with which a speaker has an ancestral or historical connection (Beaudrie & Fairclough, 2012).

2.2.2 Heritage acquisition

The continued increase in speakers described as heritage in the United States has inspired much of the research geared towards investigating the unique language acquisition process within this community. Many studies have concluded that heritage speakers can best be defined as having undergone incomplete acquisition of certain grammatical structures (Montrul, 2008). In addition, for more than a decade, research has often chosen to approach this topic by describing results as being representative of incomplete acquisition when comparing heritage speakers to their native peers by referring to their command of their heritage grammar as unstable and underdeveloped (Benmamoun et al., 2010). On the other hand, their command – or perceived lack in command – of grammatical features like tense, aspect, and mood have been

viewed as being representative of having faced gaps in their knowledge when compared to L2 learners (Montrul, 2009; Rodriguez & Brandl, 2019). These findings have inspired the development of new research that challenges this approach and proposes that using the term incomplete to describe the acquisition process of heritage speakers fails to address the unique experiences they encounter when developing language (Kupisch & Rothman, 2018).

Instead of framing heritage acquisition as incomplete acquisition of Spanish and English, it may be more appropriate to frame this process as complete acquisition of a contact variety of both their heritage language – in this case, Spanish – and their dominant language – English. These contact varieties are a result of the differences in acquisition environment and exposure, and of the continuous interaction between both linguistic systems within their repertoire. A study conducted by Potowski, Jegerski and Morgan-Short (2009) takes a deeper look into this phenomena and postulates that the error lies within the attempt to classify US Latinx as showing signs of incomplete acquisition. Their conclusions support the above argument in favor of complete acquisition of a contact variety of the heritage language. However, the fact that heritage acquisition is indeed complete can be overlooked when attempting to analyze their language production by comparing them to monolingual, native Spanish speakers as opposed to the more appropriate demographic with the most input influence: first-generation immigrants (Pascual y Cabo & Rothman, 2012).

By understanding that the acquisition process heritage speakers encounter in Spanish classrooms is, for some heritage speakers, that of acquiring a second dialect (i.e., that of a standardized, monolingual form) rather than a process of reacquisition or an attempt to fill ‘gaps’, educators can begin to promote more positive language attitudes in classrooms. Greater awareness of the variation that exists within the community and where this variation stems from

can allow heritage speakers to have a deeper understanding of the language they bring with them to the classroom in a way that invites confidence rather than uncertainty and insecurity. Pascual y Cabo and Rothman (2012) further support the argument of acquiring a second dialect by claiming that heritage grammars are in fact native grammars, but distinct from monolingual ones. They state that we as researchers, "should embrace documented [heritage speaker] differences for what they are, unique outcomes of different realities" (453-454). Much of the discussion is centered on recognizing that the linguistic input heritage speakers receive is different from both L2 speakers who acquire language later in life and native speakers, making it reasonable to observe certain differences (e.g., lexical, grammatical, phonological, etc.). Similarly, heritage speakers' production of English can also be viewed as that of a contact variety when considering that the acquisition of their dominant language is unlikely to occur independently and cut off from any influence of their heritage language if they grew up and interact in multilingual environments. This notion further highlights the necessity to analyze all linguistic resources available to each speaker when interpreting heritage results and attempting to make direct comparisons to either monolingual English or monolingual Spanish speakers.

As research on the Spanish heritage population and their acquisition process has continued to expand, educational systems across the country have taken notice of the need to address identified differences in production (e.g., differences in morpho-syntactic features, differences in lexicon due to English influence, etc.) by developing classes and/or programs focused on embracing the cultural background that makes these students different from their native and second-language learner counterparts. As a result, studies have shown that heritage classes are no longer unique and restricted to regions in the country with well-established Spanish speaking communities, rather they are now available in more than half of the fifty states

and Washington D.C., representing a large expansion and increase in accessibility (Beaudrie, 2012). Due to the rapid growth in available programs and focused instruction, heritage speakers are now able to receive curated curriculum that acknowledges both their background and experiences, while also supporting their language development.

2.2.3 Heritage production - Spanish

Despite early research focused on bilingual and heritage language acquisition often framing their findings as perceived deficits, or aspects of incompleteness (Silva-Corvalán, 1994; Montrul, 2004; Montrul & Bowles, 2009), most current research argues in favor of recognizing the existence of an independent heritage variety influenced by the language acquisition environment and speaker experiences (Pascual y Cabo & Rothman, 2012; Kupisch & Rothman, 2018). Additional work has been done to further examine patterns within heritage Spanish and has been arguing for alternative explanations and terminology surrounding observed differences between heritage groups and monolingual controls (Paradis, 2023; Rothman et al., 2023; Wiese et al., 2022).

Although research is shifting toward embracing terminology that better reflects the Spanish heritage community, work in this area has often ignored the importance of vocabulary and lexical knowledge, and instead prioritized deeper understanding of grammar development (Fairclough & Garza, 2018). Therefore, the current section aims to describe current literature examining heritage production by focusing on both lexical knowledge and morphosyntactic production to provide insight into the characteristics associated with heritage Spanish in the United States.

Of the available literature analyzing lexical knowledge within the Spanish heritage community, most of it has focused on the influence of enrollment in Spanish language programs

and how extended enrollment impacts vocabulary development – with emphasis on development of formal and academic registers (Fairclough, 2006; Moreno-Fernandez, 2007). Findings demonstrated that increased Spanish language instruction had a significant impact on production of these formal registers by reducing English transfer, reducing the usage of colloquialisms or stigmatized forms, and ultimately resulting in patterns more similar to monolingual Spanish production (Garza, 2013). While these studies demonstrate how heritage production can develop through intervention-based methods, it fails to recognize or acknowledge patterns in production as they have naturally developed and may exist within the communicative world. While few studies have targeted questions regarding natural heritage Spanish production, findings show that heritage speakers are likely to make use of lexical creations based on English words using Spanish phonology (Sánchez-Muñoz, 2010), and that many high frequency calques are undergoing lexicalization within the US Spanish variety (Fairclough, 2013). These findings further support the notion that heritage speakers are likely to develop a variety of their own due to their unique acquisition environment.

In addition to demonstrating differences with regard to production in informal contexts, heritage speakers have been shown to produce morphosyntactic features – like the expression of subjunctive mood – following unique patterns demonstrating that heritage speakers do in fact diverge from native monolingual Spanish speakers with regard to their usage patterns. Results often show trends of underproduction in expected subjunctive contexts, overproduction in expected indicative contexts, or complete avoidance of the structure altogether (Giancaspro, 2019). While proficiency has been shown to be a significant factor predicting production of subjunctive mood, frequency of verbs, exposure to certain verbs, and verbal regularity were all interpreted as being influential factors with regard to production of the subjunctive mood leading

to the postulation that the lexicon plays a highly influential role in heritage speaker variability in mood production (Giancaspro et al., 2022). López Otero, Hur & Goldin (2024) further examined the role of language exposure as a factor modulating syntactic production by examining clitic climbing – where clitic climbing in Romance languages refers to instances where a clitic pronoun can detach itself from an infinitive verb and 'climb' to attach to the inflected verb – amongst simultaneous, and sequential bilingual Spanish heritage speakers. They found that age of acquisition only influenced receptive knowledge, while exposure and use of Spanish impacted the ability to distinguish between both Spanish clitic types more. They claim, "the findings in [their] study suggest that the acquisition of syntactic knowledge in which optionality is involved, as is the case for clitic climbing, is primarily shaped by rich and sustained language exposure and use in different contexts" (p. 18). Therefore, observed differences are more and more being attributed to differences in input, exposure, and other multi-faceted social factors, rather than incompleteness. This increased acknowledgment in the role of input and language background experiences, paired with the lack of knowledge associated with heritage production in informal settings and independent of intervention, highlights the need for expanding and reexamining current approaches.

2.2.4 Lexical diversity

Lexical diversity is often used to estimate speaker proficiency (Elabdali et al., 2022) and has been argued to be influenced by both the amount of input speakers receive, and the way in which the amount of input may fluctuate over-time throughout developmental stages (Jacobson & Walden, 2013). When considering these key features, it invites the opportunity to further explore how this measurement may serve as an appropriate way of analyzing language usage within the heritage community. Heritage speakers inherently experience variable degrees of

exposure to Spanish, the amount of input they receive varies throughout their lifetime, and, as a community, they vary greatly in the variety of Spanish spoken in the home. Furthermore, lexical diversity is strongly correlated with speaker competency and is a mechanism employed for approximating vocabulary size of a speaker, where greater lexical diversity exhibited by an individual is often interpreted as being indicative of greater vocabulary size and greater variety present in the speech (Fusté-Herrmann et al., 2006). The variation in input within the heritage community highlights a need for examining how this mechanism can be informative for heritage production.

Due to the value and need for analyzing lexical diversity being clear, indices used for measuring lexical diversity have been continuously developed overtime and their validity tested. Previous studies have relied on indices like Token-Type Ratio (TTR) – which divides the average number of unique words by the number of total running words (Templin, 1957). However, this metric proved to be sensitive to factors related to text-length. Therefore, additional indices such as Moving Average Token Type Ratio (MATTR; Covington & McFall, 2010) and the Measure for Textual Lexical Diversity (MTLD; McCarthy & Jarvis, 2010), were developed to mitigate these analysis issues. MATTR serves to correct for text-length effects by gathering the TTR values across a given designated window (usually set to 50 words for written texts and adjusted to 11 for oral texts) and averaging them. MTLD stabilizes across various lengths by calculating how quickly TTR values stabilize and averages the number of words required for stabilization in a given text. This 'point of stabilization' can best be understood as reaching the moment in which "neither the introduction of repeated [token] types nor even a considerable string of new types can markedly affect the TTR trajectory" (McCarthy & Jarvis, 2010, p. 386).

To date most studies analyzing lexical diversity have focused primarily on L2 English learners and written modalities with even the existence of automated tools uniquely designed to analyze written L2 English (Kyle et al., 2021; Zenker & Kyle, 2021). Despite this, recent work has shown that lexical diversity indices and measurements can be appropriate tools to determine the diversity present within transcribed orally produced speech as well, and potentially even more stable across task types given that oral speech tends to contain more repeated function words while written texts tend to include more content words (Biber et al., 2004). To confirm these claims, Kyle, Sung, Eguchi & Zenker (2024) aimed to determine whether lexical diversity indices can be extended to analyzing orally produced speech. Ultimately, they determined that MTLD and MATTR both served to be a reliable and valid methods for measuring lexical diversity in oral tasks.

2.3 The current study

The goal of this study was to examine heritage speakers' patterns of production in both English and Spanish by examining lexical diversity. Heritage speakers have been found to develop differences in production when compared with monolingual English or Spanish speakers in various linguistic domains (i.e., phonologically, grammatically, and lexically). Studies have increasingly shown the importance in analyzing these features with specific focus on heritage patterns, to further shed light on the role that background, upbringing, and experiences have on language acquisition. The overall goal of considering these additional factors is to prevent continued comparisons between heritage speakers and monolingual speakers for the purpose of setting an ultimate acquisition goal that reveals deficiencies of heritage speakers; and to promote analyses that aim to recognize the Spanish heritage language as a unique variety. The following questions have guided this study:

1. To what extent do L1 monolingual English speakers, English-dominant heritage Spanish speakers, and L1 Spanish-L2 English speakers vary with regard to their lexical diversity in spontaneously produced oral speech in English?

2. How do L1 Spanish-L2 English speakers compare to English dominant Spanish heritage speakers with regard to their lexical diversity in spontaneously produced oral speech in Spanish?

The L1 monolingual English group and the L1 Spanish-L2 English group served only to inform our findings within the Spanish heritage group. These groups allow us to determine what similarities or differences exist when these groups are compared to heritage speakers; however, they do not serve as a benchmark for heritage speakers. The approach here is to support the current directions of the field by recognizing that any differences in heritage production are indeed legitimate, and to further add to this field by examining the lexicon. Given that lexical diversity is limited in its ability to make cross-linguistic comparisons, the heritage productions in Spanish were compared only to the Spanish produced by the L1 Spanish group and the heritage speakers' productions in English will be compared to both the English produced by the L1 monolingual English group and the L2 English speakers.

2.4 Methods

2.4.1 The corpus

The corpus used to conduct the analyses was the Archive of L1 and L2 Scripted and Spontaneous Transcripts and Recordings (ALLSSTAR) developed at Northwestern University (Bradlow, n.d.). This is a multilingual corpus consisting of audio recordings from more than 120 talkers across more than 20 different language backgrounds. Talkers completed tasks eliciting spontaneous speech in a question-and-answer session (See APPENDIX A) and a story telling

task where they narrated a comic strip (See APPENDIX B). Audio was collected in both the L1 and L2 languages for each talker, with exception to the English participants who recorded only their L1.

We utilized data from 11 speakers from each language background: English, Spanish, and Spanish Heritage. The English participants in the corpus were identified as native English speakers between the ages of 18-22 who were raised monolingually with some identifying the development of additional languages later in life. The Spanish participants were between the ages of 23-35 and developed Spanish as a first language and English much later in life. Some of the Spanish participants, like the English participants, identified acquiring additional languages later in life as well. The Spanish Heritage participants in the corpus were all between the ages of 18-22. Speakers reported having acquired Spanish as a first language and English as a second language between the ages of 5-8. Detailed information regarding the home variety for the participants is not provided, however, some demographic information was collected during elicitation. Participants represent multiple Latinx backgrounds and identities. All participants reported having undergone a shift in language dominance from 100% usage of Spanish in the home during early childhood to less than 20% usage of Spanish after early childhood. None of the speakers involved in elicitation reported having taken Spanish language learning courses in high school[3].

[3] See Blasingame (2018) for more background information about the Spanish heritage speakers in the ALLSSTAR corpus. Given there is no information with respect to participants' variety of Spanish, the current study is unable to specify the variety of heritage Spanish being investigated.

2.4.2 Linguistic and corpus analysis

To conduct analyses on the speech provided by the ALLSSTAR corpus, the audio was first transcribed and manually tagged for certain speech phenomena like repetitions (e.g., 'so I bought, I bought a shoe'), blends or truncated speech (e.g., 'gonna', 'cause'), or self-corrections (e.g., 'so he bought, she bought…'). After tagging the speech for these phenomena, the data was then cleaned and preprocessed so lexical items could be properly tagged for part of speech. The English data was analyzed using the freely available online Tool for the Automated Analysis of Lexical Diversity (TAALED) (Kyle et al., 2021), which allowed for the automated analysis of the transcribed text using the chosen indices. This study focuses on Moving Average Token Type Ratio 11 (MATTR11) – which uses the window size (i.e., 11 words) that is recommended for spoken language – and the Measure of Textual Lexical Diversity 92 (MTLD92) – a version of the original index that is shown to be most appropriate when analyzing oral productions. The original Python code for TAALED was then adjusted for analyses of the Spanish data. To do so, Spanish language models were loaded to serve as a reference corpus for the transcribed speech. After adjusting the language models, the same steps used for the English data were followed.

2.5 Results

2.5.1 Lexical diversity – English

Before beginning analyses, the number of tokens were averaged by group to determine the amount of speech produced within each group, which ultimately informed our interpretations. As may be expected, the English group produced the most speech across both task types (M = 219.970) and the Spanish group produced the least amount of speech (M = 127.424), while the heritage group's average token production fell in between the two groups (M = 134.030). To

further examine the data, descriptive statistics for each speaker group were gathered on the primary indices of interest MTLD92 and MATTR11; a summary of these results is shown in Table 1.

Table 1 *Descriptive Statistics for MTLD92 and MATTR11 for the English data.*

Speaker type	Task	M MTLD92	SD MTLD92	M MATTR11	SD MATTR11
English	Q & A	8.910	0.435	0.399	0.059
	Storytelling	13.048	1.168	0.911	0.026
Spanish Heritage	Q & A	8.677	0.724	0.434	0.058
	Storytelling	12.667	1.329	0.909	0.022
Spanish	Q & A	8.828	0.844	0.388	0.080
	Storytelling	11.929	1.299	0.888	0.031

When looking at MTLD92 scores, English speakers on average had the greatest lexical diversity across tasks, while Spanish Heritage speakers had greater scores than the Spanish speakers in the storytelling task but slightly lower scores in the question-and-answer task. It is not surprising to find that the English participants have yielded greater lexical diversity scores across task types because they are the only monolingual speakers in this study. The remaining results are more peculiar, however, when considering that neither the heritage group nor the Spanish group consistently exhibited greater lexical diversity across tasks. The English group and the Spanish Heritage group are the participants that have had the greatest and most consistent exposure to English, while the Spanish group acquired English much later in life, leading us to expect that both groups would outperform the Spanish group across task types.

While we do not find the anticipated trend when looking at MTLD92 scores, we do see this trend in the MATTR11 results. The Spanish Heritage group and the English group on average demonstrated greater lexical diversity when compared to the Spanish group. There is some fluctuation between which group yielded greatest lexical diversity scores across tasks

between the English and Spanish Heritage groups, but the most salient result is that both English dominant groups outperformed the L2 English group (i.e., the Spanish group). The fact that the Spanish Heritage speakers are English dominant and report communicating in English approximately 80% of the time, insinuates that it is likely they may perform similarly when compared to monolingual English speakers, but we should still anticipate some discrepancies. Figure 1 shows the results for both indices split by task type. When looking at the figure, there is a distinct difference in scores between the question-and-answer task and the story telling task, which may be a result of average length of texts given that the question-and-answer prompts had a minimum time of five minutes. Furthermore, the MTLD92 scores demonstrate the inconsistency in Q&A scores, however, a more anticipated trend for the storytelling task. The MATTR11 scores also demonstrate a less consistent trend across tasks with the heritage group yielding higher diversity scores in the Q&A task, but not in the storytelling task.

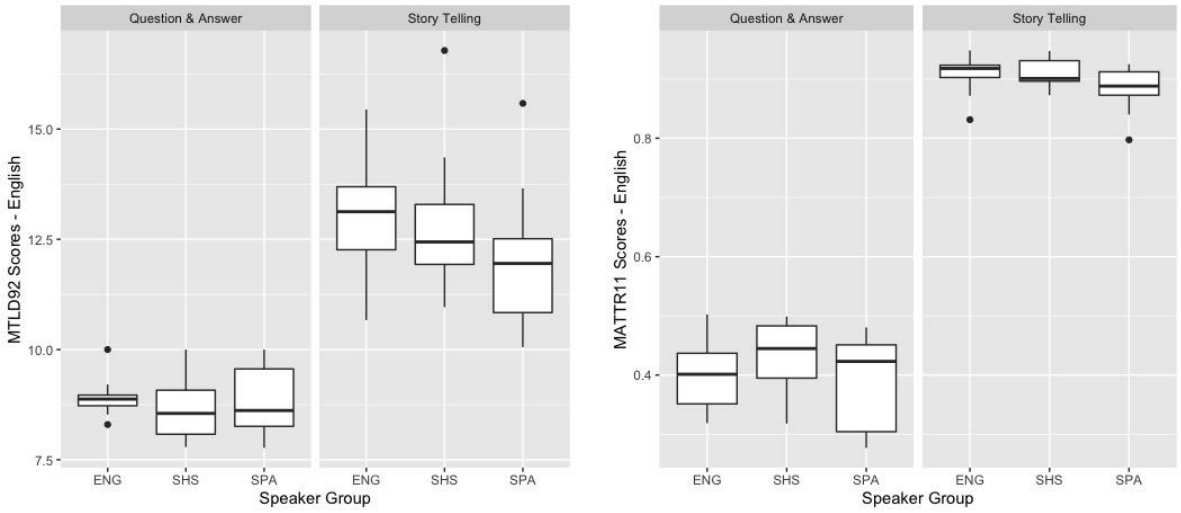


Figure 1 *English MTLD92 and MATTR11 scores grouped by speaker and separated by task type.*

After reviewing the descriptive statistics and visualizing the English data, Linear Mixed-Effects Models were run to determine if the observed differences are significant. To do so, the role that speaker group may have on lexical diversity scores was analyzed, while considering the potential differences that may be introduced from task type (Question and Answer vs. Story Telling). Task type was contrast coded to help control for the fact that the story telling tasks were much more restricted in the potential vocabulary that may be used, given that each speaker described the same comic strips leaving little room for much variation between participants due to the characters and scenery being described. Speaker group was also contrast coded to ensure that the Spanish heritage group was analyzed as the reference group and the appropriate comparisons were made for the most interpretable results.

To compare the two mixed-effects models for MTLD92 scores, an analysis of variance was run to establish whether adding speaker group as an independent variable significantly impacted results. Results were not found to be significant at alpha = 0.05 with p = 0.071. The same comparison was run for MATTR11 scores and yielded different results with speaker group showing to significantly impact the model at an alpha level of 0.05 with p = 0.0495. Thus, while it appears that the level of diversity in orally produced speech varies in some way across indices, this difference is manifesting differently given that results were found to be significantly different for MATTR11 but not for MTLD92 scores when looking at the English data.

2.5.2 Lexical diversity – Spanish

As with the English data, the average number of tokens produced by each speaker group was evaluated first. The Spanish group on average produced more speech across task types (M = 339.606) than the Spanish Heritage group (M = 277.455). Upon first glance at the descriptive statistics shown below in Table 2, it appears that the Spanish group also shows greater diversity

with regard to the lexical choices they make when comparing both MTLD92 and MATTR11 scores. Despite the Spanish group rendering higher scores, they seem to be more variable, when looking at MTLD92 scores, due to the much higher standard deviations across tasks. This difference in variation appears to level out for the MATTR11 scores, with the Spanish group and the heritage group demonstrating similar degrees of variation for this lexical diversity index.

Table 2 *Descriptive statistics for MTLD92 and MATTR11 for the Spanish data.*

Speaker type	Task	M MTLD92	SD MTLD92	M MATTR11	SD MATTR11
Spanish Heritage	Q & A	14.124	1.604	0.918	0.016
	Storytelling	12.871	1.446	0.908	0.027
Spanish	Q & A	15.630	2.045	0.928	0.017
	Storytelling	15.007	2.482	0.930	0.022

The greater diversity observed in MTLD92 scores is not necessarily surprising given that the amount of input and exposure to Spanish in the monolingual group may be inherently more varied than the Spanish heritage group who is gaining most of their exposure from social and familial experiences, rather than through education or more formal communicative circumstances. However, this noted difference does not hold across indices making it difficult to interpret the discrepancy. Potential explanations for these differences will be further disentangled in the discussion section. The data shown in Table 2 is visualized below in Figure 2 to allow for further understanding of the distribution of scores and the variance within each speaker group.

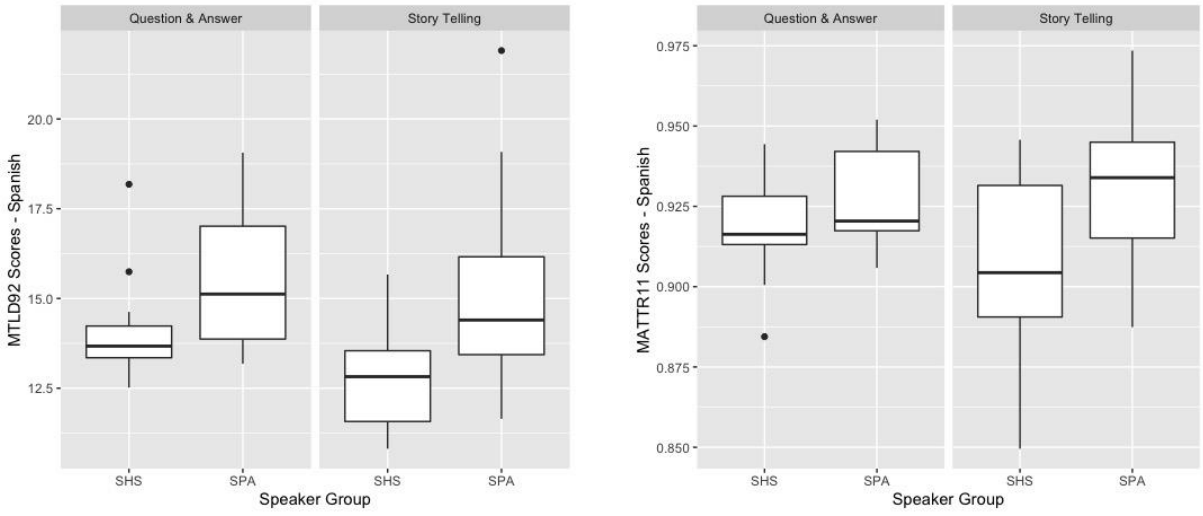


Figure 2 *Spanish MTLD92 and MATTR11 scores grouped by speaker and separated by task type.*

When looking at both MTLD92 and MATTR11 scores for the Spanish data, the Spanish group demonstrates greater lexical diversity across both task types. However, unlike the English data, the question-and-answer task yields greater diversity scores than the storytelling task demonstrating a difference in production between the two groups when communicating in Spanish and English.

Following the same analysis as with the English data, Linear Mixed-Effects Models were run to determine the influence speaker group may have on lexical diversity scores. Results showed that adding speaker group as a variable to the model for MTLD92 scores significantly impacted results at alpha = 0.05 with $p = 0.004$. When comparing results for MATTR11 score we again see that adding speaker group to the model significantly impacted results with $p = 0.005$. This slightly differs from the English results in that the Spanish group consistently rendered greater diversity scores while the English results were more conflicting.

2.6 Discussion

The goal of the current study was to further our understanding of the potential similarities and differences with regard to lexical diversity when comparing the Spanish heritage community with both monolingual English speakers and L1 Spanish-L2 English speakers. The first research question aimed to examine the variation in lexical diversity scores in spontaneously produced English, between monolingual English speakers and Spanish heritage speakers, with the Spanish speakers serving as a control. We found that observed differences between lexical diversity scores for speaker groups are only significant when looking at MATTR11 scores but not MTLD92 scores. Findings are further muddied when considering that the Spanish Heritage group yielded the lowest scores of all groups when looking at the question-and-answer scores for MTLD92; however, this was the only instance in which the Spanish group yielded higher scores for the English data. This result is surprising considering we would expect the Spanish group to score the lowest across task types given that they have acquired English much later in life and have had markedly less exposure and input when compared to the other two groups. When comparing indices (MTLD92 and MATTR11) for the English data, there is a fluctuation between the English group and the Spanish Heritage group for lexical diversity scores. This data may seem unclear, yet I interpret it to support the well-established argument recognizing that despite heritage speakers exhibiting differences in production when compared to their monolingual and L2 counterparts, these differences do not necessarily indicate deficiencies in development (Kupisch, 2013; Meisel, 2013). The second research question examines the comparison between lexical diversity scores in spontaneously produced oral speech from Spanish heritage speakers and L1 Spanish-L2 English speakers. Results showed that the Spanish group yielded significantly greater lexical diversity scores across both indices. This is not the dominant

language for the heritage speaker group; therefore, it is not very surprising that they would score lower than the Spanish group. Language dominance, combined with the fluctuation in input levels between each of the languages, may have ultimately affected how diverse the Spanish Heritage group's lexical usage patterns are (Parodi, 2008).

It is also worth noting previous research which has examined the role that social domains and semantic fields have on lexical knowledge within heritage and bilingual communities. Findings show that diversity in each respective language varied based on the domain and field with which each language was most used (Fairclough & Garza, 2018). Therefore, if speakers used Spanish more commonly in the home they had greater productive vocabulary in Spanish on related topics, and if they used English at school, they demonstrated greater English vocabulary in topics related to academics. There is growing acceptance for the argument that bilinguals should develop skills differently in each language because the skills being developed are a direct result of a speaker's needs and purposes while communicating, which may not be the same for each language they acquire (Grosjean, 2016; Treffers-Daller, 2019). Keeping this in mind, the tasks completed for the ALLSSTAR corpus can cross various domains depending on the participant and how they chose to address the prompt. The question-and-answer task, for example, asked participants to speak for five minutes with a set of six questions to guide them. These questions mostly touched on topics related to their upbringing, culture and family, however, some participants ended up discussing academics or work and crossed between multiple semantic fields throughout the recording. Without direct knowledge on how language usage varied based on social domain for the participants in this study it is difficult to ascertain how this may have influenced results, although it is fair to assume this impacted, to some degree, the differences in scores and may provide a possible explanation for the results found in this

chapter. The L1 Spanish speakers acquired language in a monolingual environment well before acquiring English as an L2, therefore, we can deduce that this group would be less constrained when switching between semantic fields while communicating in Spanish. See APPENDIX A and APPENDIX B to view the comic strips and interview questions for the question-and-answer task.

In addition to the role that semantic field may have played in limiting participants, potential concerns arise from the methodological approach used for gathering data. The data analyzed in this study was transcribed from a preexisting corpus that asked participants to complete tasks and perform in monolingual mode. Given that in multilingual environments heritage speakers can employ resources from each respective language within their system, the seemingly simple act of restricting multilingual participants to a monolingual mode could have influenced lexical diversity scores due to potential extra effort needed to constrain and restrict the system for monolingual production. For bilingual and multilingual speakers, each language in their repertoire develops together as a single and dynamic system and represents, "how [multi]lingual people fluidly use their linguistic resources – without regard to named language categories…" (Vogel & García 2017, p. 4). Therefore, the nature of the tasks themselves may lead to inaccurate representations of participants' abilities. Enforcing a monolingual mode ultimately may have made participants unable to access the full array of linguistic resources available to them, and, asking them to avoid a multilingual mode where they are potentially most comfortable could have led to stress or other confounding factors. While stress is already a contributing factor in any in-lab research study given that participants are in non-natural settings with a moderator, additional external factors (e.g., language mode) have also affected why the language analyzed from the heritage speakers is unlikely to be an accurate depiction of how they

naturally communicate. Language mode is arguably even more of a concern for the heritage group than the other bilingual group, the Spanish group, given the sociocultural nature of translanguaging and the tendency for L2 classrooms to discourage non-monolingual practices (Otheguy et al., 2015).

Ultimately it is difficult to speculate how results may have varied had language not been controlled for in this way, and even more challenging because lexical diversity is a language specific metric making it incapable of allowing for cross-linguistic comparisons. Lexical diversity being a metric that is correlated with speaker competency and heavily influenced by the input an individual receives makes it appear as an ideal metric for heritage production. However, the differences in lexical diversity scores and the inconsistent results found beg the question of whether this tool is suitable to adequately measure bilingual and multilingual production as it exists outside of classroom environments. Lexical diversity tools are not the only instruments being used that, upon deeper dive, may not be meeting our needs as researchers when it comes to producing results that are truly interpretable and meaningful for this kind of data. For example, native speaker norms are often portrayed as the standard benchmark for which speakers of other identities and backgrounds should attain. Heritage and multilingual speakers are often inappropriately compared to monolingual groups and many of the assessment methods and tools used to date represent these practices (Gathercole et al., 2013). But when we take a more nuanced look into the language production that is supposedly setting language norms, there is actually a lot of variability within L1 development and production that is ignored and overlooked making the establishment of these norms even less clear (Hulstijn, 2015; Treffers-Dellar, 2019). While it seems obvious that even native, monolingual speakers would vary in terms of production – whether it be due to factors like region, age, gender, race, etc. – we as researchers

still consistently take this for granted when establishing expectations for any group that falls outside of the monolingual and/or native category. Therefore, while our goal may be to discover between group differences, when these differences do manifest, how do we interpret results if we are not properly and accurately contextualizing them within our understanding of the underlying cognitive processes of multilingualism, and within the greater sociocultural landscape of our participants? Essentially, it is one thing for us as researchers to acknowledge the continuity of the multilingual system, but another entirely to thoughtfully incorporate this fact into experimental and methodological designs. In acknowledging these limitations we also introduce the responsibility to fill this major gap in the field as well as a need to reevaluate previous claims that may have been overextended. Some studies to date have attempted this feat by employing instruments that target overall vocabulary production rather than language specific productions (Pearson et al., 1993; De Houwer et al., 2014; Haman et al., 2015), but there is yet to be an expectation within the field as whole to prioritize methods and assessment tools that capture the entire linguistic system of multilingual speakers.

Overall, these findings provide support for the idea that differences in heritage speakers' productions should be understood as being demonstrative of the differences in their linguistic development. This reinforces that we shouldn't expect to see identical lexical diversity scores across speaker groups and reiterates the idea that the language produced by heritage speakers is in fact different from monolingual speakers or speakers who acquired language in a more traditional L2 setting. Greater interaction in multilingual environments is likely to lead to exposure to different vocabulary and at differing levels of frequency than individuals who find themselves in predominantly monolingual environments. However, we are still confined in the scope of our interpretations and do need to acknowledge limitations as a result of not being able

to directly compare between languages, and from not fully capturing the linguistic richness from the multilingual groups due to the restriction of language mode.

2.7 Conclusion

By examining lexical diversity scores in spontaneously produced speech between Spanish heritage speakers, monolingual English speakers, and L1 Spanish-L2 English speakers, we can see that although there are significant differences between their scores, it can likely be assumed that the heritage speakers are still producing diverse speech despite our inability to fully encapsulate that diversity. These initial findings suggest that heritage differences should not be perceived as faults, rather they should be seen as being a result of having acquired contact varieties of Spanish and English, and as a result of the instruments being used to collect measurements. In acknowledging these very influential factors we can help to shift attitudes towards heritage production in a way that prevents linguistic insecurity that develops in heritage communities from judgments made regarding their production. While the lack of speaker background information and the use of monolingual mode limit the scope of this study regarding claims that can be made about lexical usage in the Spanish heritage community, these constraints provide an avenue for future research. Future studies should embrace the sociocultural differences in multilingual acquisition and avoid perpetuating the implementation of inappropriate tools.

This chapter, although coauthored as an independent manuscript aligns with the broader dissertation by providing insight into some of the external pressures that exist around heritage language production. Specifically, this section thoroughly addresses how the use of labels – such as “deficient” and “incomplete” – has been inappropriately employed and overused, demonstrating that evaluative judgments regarding heritage languaging habits operate at *both* the

institutional and within community level. These results focus on the institutional judgements, showing that they are made even when inadequate methods and models are employed to measure production. Consequently, this methodological mismatch can lead to claims based on data that fail to capture the full linguistic repertoire of heritage speakers, reflecting an unclear understanding of the origins of these observed differences. The following chapter expands upon this work by exploring *within* community evaluative judgments and examining how they are influenced and potentially altered when the social information about the speakers is manipulated.

CHAPTER 3: BUT… ARE YOU REALLY LATINO? WITHIN COMMUNITY PERCEPTIONS OF SPANISH HERITAGE SPEAKERS AND LEARNERS

3.1 Introduction

As a result of the growing Spanish Heritage community in the United States, there has been an influx of research dedicated to better understanding the unique needs, backgrounds and linguistic systems of this community (see, for instance, Lynch, 2008; Valdés & Geoffrion-Vinci, 2011; Montrul, 2013; Rao & Ronquest, 2015; Amengual, 2018; Beaudrie et al., 2019). This work has demonstrated that heritage speakers represent a very heterogenous group of individuals, with language proficiency, language background and cultural upbringing being variable across individuals. Years of systemic linguistic repression has led to the development of many negative sentiments in the US towards the Spanish language and those that speak it, resulting in almost complete loss of the heritage language by the third or fourth generation in Latinx[4] families (Cashman 2009). The implementation of various policies discouraging bilingualism (e.g., Proposition 227 in California and Proposition 203 in Arizona), combined with the social pressures to assimilate to a monolingual society have led to loss of the heritage language in families – for the purposes of this chapter the heritage language in question is Spanish but this can be true of any minoritized, familial language (MacGregor-Mendoza, 2000). In an attempt to restore their heritage language, many Latinx in the United States are beginning to enroll in courses specifically designed to meet their needs as a heritage speaker (Ducar, 2012). As the availability of courses offered to those with a heritage connection to Spanish is increasing, more

[4] The majority of this chapter, like the rest of the dissertation, adopts the term Latinx to refer to the community; however, maintains the use of Latino when referring to the data given that this was the term used at the time of data collection.

and more US Latinx are taking advantage of the opportunity to develop and/or strengthen their ability to communicate in their heritage language (Beaudrie, 2012; Ducar, 2012). This has given rise to the need for a term used to distinguish certain groups within the Spanish heritage community: Spanish heritage *learner* (SHL). Spanish heritage speakers (SHS) and learners differ primarily in that, according to most definitions, Spanish heritage learners have relied more on language courses to develop their productive and receptive abilities while Spanish heritage speakers usually learn Spanish as their first language and then undergo a shift in language dominance once they attend primary school (Octavio, 2018). These two groups are encapsulated by the umbrella term Spanish heritage, highlighting just one aspect of the heterogeneity of this community.

The mixed exposure to Spanish varieties amongst heritage speakers and learners – both formal and informal forms – has led this community to develop certain linguistic features that may index a nonnative identity to others in the Hispanic/Latinx community. As a result, many Spanish heritage speakers have to navigate the perceptions and ideologies about their language use and identity not only from their Anglo peers, but from within the Spanish speaking community as well. Given that each community of speakers develops its own set of language ideologies that sustain certain linguistic and sociolinguistic practices (Martinez, 2006), those that identify as Spanish heritage have to navigate ideologies from within both the Spanish speaking community and the English-speaking population in the US. Many of these ideologies within the Spanish speaking community are centered around certain expectations and pressures placed on US Latinx that often affect their confidence in claiming their Latinx identity. One ideology that is often adopted by Spanish speakers is that a Latinx identity is dependent on an ability to speak Spanish with proficiency – and at times, this attitude can be extended to require a 'correct' and

formal variety of Spanish (Showstack, 2017). Therefore, individuals that speak non-standard varieties of Spanish, limited Spanish, or do not speak Spanish at all, risk having their Latinx identity questioned by others, which often leads to the development of insecurities around this identity (Potowski, 2012).

Some research has focused on the presence of these ideologies within the Latinx community and in the educational domain in particular (Abdi, 2011; Ducar, 2012; Randolph Jr., 2017), still, little research has worked to better understand the subtle nuances around the way these ideologies actually come into effect and how other US Latinx actually perceive the speech produced by Spanish Heritage speakers and learners. For this reason, in the current study we aim to better understand the influences that fluency in Spanish, or perceived non-native productions may have on how Spanish speakers perceive Spanish produced by heritage speakers, and how this may influence who is classified as Latinx. Given the many ideologies that associate Spanish proficiency with Latinx identity, we hope to better understand how these perceptions may be influenced and/or altered when participants are provided with different biographical information about the speaker's upbringing and relationship to the heritage language. We anticipate that the background information about the speakers will be influential enough to outweigh perceptions directly associated with the linguistic content of the speech.

3.2 Background

3.2.1 Spanish as a heritage language in the United States

The Latinx population in the United States of America has been a significant part of society since the country's acquisition of the Southwest. Yet, despite their prominent role in society, consistent efforts throughout history to assimilate minoritized populations and promote monolingual ideals has led to a trend of generational decrease in linguistic ability in Spanish for

families of Latinx descent, and in some cases a complete loss of this heritage language by the third or fourth generation (Cashman, 2009; Silva-Corvalán, 2004). A heritage language is defined as a minoritized language in a given country, with which a speaker has an ancestral or historical connection, in this instance the language is Spanish (Beaudrie & Fairclough, 2012). Many Latinx families face difficult circumstances when attempting to maintain their heritage language and share it with their children, in a society that does not always respect or embrace the expression of multiculturalism and the benefits of multilingualism (Zentella, 2002). As a result of the large population of individuals of Hispanic descent and pressures to assimilate, there is an emerging population of Spanish Heritage Speakers (SHS) and Learners (SHL) in the US.

The Latinx community is not unfamiliar with hostile and negative encounters directed towards their culture and language, often leading them to feel pressures to assimilate to the Anglo dominant culture in society and ultimately preventing them from confidently maintaining and expressing their heritage. Many social and political factors like the implementation of English-only policies and the promotion of monolinguistic ideals, have led those of Latinx descent to feel obligated to learn English and decrease their usage of Spanish to combat discrimination. Furthermore, in an effort to help their children achieve equal opportunities, Latinx families have also been discouraged from teaching Spanish to their children and encouraged to enforce English-only policies in their home to avoid judgements from external entities that accuse them of causing 'harm' to their children by speaking Spanish (Martínez, 2012). This relationship between monolingual and nationalistic ideals have paved the way for many forms of linguistic repression, like the previously referenced policies that are aimed at limiting and discouraging bilingualism (MacGregor-Mendoza, 2000). The development of these policies in turn only further perpetuates the belief that bilingualism leads to children being

cognitively confused and hinders an individual's ability to be successful and fully contribute to society (Zentella, 2002). These beliefs and attitudes put the culture and language of Latinx families in a very vulnerable position as they attempt to navigate maintaining their heritage among pressures to assimilate to the dominant majority. Research has shown that the influx in growth in the Latinx population can actually be attributed to the increase in birthrate rather than an increase in immigration, indicating that second-generation Hispanics will emerge as the largest group (Abdi, 2011). In addition, Potowski & Carreira (2010) explain that the high Latinx birthrates in the US are currently outnumbering immigration rates. This further highlights that the Spanish Heritage community is rapidly growing and gaining an even larger presence, however their ability to maintain their culture and language is still greatly affected by their inability to express themselves freely in society as result of negative ideologies towards the Spanish language (Potowski & Carreira, 2010).

The institutionalization of these ideologies has led to the development of a system of values and beliefs that ultimately shapes the way society views the practices of multilingual communities and the individuals that belong to these communities (Leeman, 2012). Therefore, Spanish speakers in the United States are often confronted with strong beliefs which devalue and demonize their language and culture at both a social and political level (Zentella, 1995), resulting in the internalization of these ideologies and attitudes in a very personal way, preventing many Latinx from feeling confident in their identity and in claiming their heritage (Potowski, 2002). This linguistic repression has led these individuals to lose an important part of their culture, and as a result gave rise to a new set of ideologies amongst US Latinx. The brief sociopolitical history described in this section has led to a growth in the Spanish heritage population in the

United States and has led to a complex interaction of ideologies around language usage within the Latinx population.

3.2.2 Spanish heritage speakers vs learners

As the large presence of the Latinx population in the United States and as the community of individuals identifying as Spanish Heritage continues to grow, more and more studies are beginning to focus on better understanding the backgrounds and needs of these individuals. The aforementioned generational language attrition is a result of many sociopolitical factors that have led many families of Latinx descent to feel pressure to assimilate to the dominant culture of society, and in doing so promote acquisition of the majority language within their families (Leeman, 2012). These societal pressures combined with the distinct acquisition environment of heritage speakers has led them to develop a contact variety of Spanish that also has distinct lexical and grammatical features when compared to the standard variety of Spanish (Pascual y Cabo & Rothman, 2012). The distinct acquisition environment for those in the heritage community combined with the fact that many in the heritage community often are exposed to their own variety of Spanish sheds light to the inherent heterogeneity when considering not only plethora of social backgrounds that exist under the umbrella term "US Latinx", but also the heterogeneity that exists within their linguistic backgrounds from both a variety and proficiency perspective.

As more studies begin to emerge this variation becomes even more apparent, inspiring a subset of research currently focused on how to best describe and define those that belong to it. However, not to be forgotten a singular definition is difficult to navigate and has been highly debated in the field. Researchers are often tasked with synthesizing a definition that best suits their population of interest due to the inability to describe the community in a one-size-fits-all

way. This is heavily acknowledged and not taken for granted in the field. Also well recognized is that the consistent difference between many of the heritage definitions used depends greatly on the importance placed on linguistic factors, like the individual's ability to speak the heritage language, when determining who classifies as heritage (Carreira & Kagan, 2011; Polinsky & Kagan, 2007). For this reason, the definitions described below have been intentionally chosen to most appropriately describe the community of Spanish Heritage included for analyses in this study.

As has been addressed previously in this dissertation, in broad terms a heritage speaker can be defined as an individual that speaks a heritage language and has an ancestral and cultural connection to the heritage language (Fishman, 2001; Beaudrie & Fairclough, 2012). However, more narrowly speaking they can be described as, "asymmetrical bilinguals who learned language X – the 'heritage language' – as an L1 in childhood, but who as adults, are dominant in a different language" (Benmamoun et al., 2014, p. 260). For the purposes of this study the narrow definition will be adopted to describe heritage *speakers* – with the emphasis on the word speaker being important in this case and will be further explained.

Due to the high variation in experiences amongst those within the community, the field has recognized two distinct groups: heritage *speakers* and heritage *learners*. Definitions used to describe heritage language *learners* are just as variable as those aiming to describe heritage speakers, for similar reasons. However, they agree on one key aspect: the role of education in the facilitated acquisition of the heritage language. Heritage learners are individuals who have a cultural and ancestral connection to the heritage language and have enrolled in language learning courses. Reasons behind their enrollment in language courses include but are not limited to: wanting to acquire their heritage language in general, wanting to acquire a more standard variety,

and/or wanting to 'relearn'/ 'improve' their heritage language. Motivations have been found to be anything from the more intrinsic and/or integrative (i.e., a desire to learn Spanish out of enjoyment or to connect more with their community), to the more extrinsic and/or instrumental (i.e., viewing language as a resource for educational or professional opportunities) (Ducar, 2012). They vary greatly between individuals, are not mutually exclusive, and are also known to shift within a person throughout their lifetime (Gardner & Tremblay, 1994). Regardless of the reason and/or motivation, heritage learners are influenced by academia, whether that be World Language classes, bilingual education, or heritage classes (Montrul, 2010; Octavio, 2018).

The growing usage of this terminology, however, has naturally inspired discussions regarding the various learner profiles that enter the Spanish classroom and elucidates even more the lack of agreement within the research community surrounding the 'qualifications' for who is deemed a heritage speaker. The individuals questioned most tend to lack even receptive abilities making their inclusion in the overall heritage community, from a research perspective, even more contentious. Their legitimacy as a heritage *learner* is often questioned upon enrolling in Spanish classes given that their acquisition process is considered to be more similar to that of an L2 learner (Wiley, 2001). However, for identity and inclusion reasons their cultural connection to the heritage language cannot be ignored, even with the lack of heritage language proficiency (Beaudrie & Ducar, 2005). It is for this reason that the current study focuses on this often ignored and overlooked portion of the heritage learner community.

In order to create a clearer divide between the two groups of heritage *speakers* and heritage *learners*, this study intentionally narrows the heritage learner definition to describe: those within the heritage community that may fall within a later generation that is more susceptible to language loss (i.e., the third or fourth) and thus exhibit very limited receptive

and/or productive abilities; or, those that may have had some exposure in the home from relatives but lack of maintenance of the heritage language led to the language not being acquired in childhood. Regardless of circumstance heritage learners are still raised within the culture and traditions of the Latinx community, however, they must rely primarily on an educational setting for acquisition of their heritage language (Rothman et al., 2016). This more fine-grained distinction between the two groups was adopted for the purposes of creating the guises used for the study which will be further explained in the methods section of this chapter.

3.2.3 Standard Language Ideology

Standard language ideology (SLI) is a belief recognized as bias towards promoting a single pure, correct manner of speaking a given language, and that any deviation from this standard form is considered representative of lower status (Showstack, 2017). This variety is considered an idealized abstraction of language due to the fact that it relies on the assumption that language is static rather than dynamic and should be impervious to influences from languages in contact and the general diversity and creativity amongst speakers (Leeman, 2012). In addition, the standard variety is often modeled after the speech produced by the upper and middle classes, furthering its perceived its level of prestige. The perceived level of prestige and social status of the standard variety often creates many social divisions amongst different social groups due to the many social ideologies associated with nonstandard varieties that often attribute lower economic status or education level to those that speak them (Villa, 2002). Standard language ideology is both covertly and overtly promoted in various aspects of society as a way to make judgements on other language varieties based on socioeconomic status, education, and various other social factors, but arguably the most notable, and less covert of settings, can be found in the education system where students are constantly confronted with

expectations of the 'right' way to speak (Leeman, 2012). These expectations placed on individuals during their education only helps to further disseminate the SLI, as well as further the authority of the elected standard variety. As a result, many speakers internalize the very ideologies that are used to criticize them, resulting in the persistence of the standard (Cashman, 2009).

As described above, Spanish heritage speakers are already confronted with negative ideologies regarding their bilingualism, causing them to have to navigate not only the presence of SLI in English, but in Spanish as well. Octavio (2018) highlights the pressures that heritage speakers and learners feel to hide their linguistic identity in order to be better accepted by society – this unfortunate dilemma that heritage speakers are often confronted with will be discussed in greater detail in section 2.4. This is further complicated as educational institutions across the country begin to recognize the need for programs and classes designed for heritage speakers. Many institutions promote instruction of the standard variety of Spanish and gather their materials from the Real Academia Espanola (RAE – Royal Spanish Academy). The RAE is an entity located in Madrid, Spain that has been overseeing official policy regarding the Spanish language for the past 300 years and ultimately determining the rules that influence roughly 450 million Spanish speakers in the world (Zentella, 2017). The RAE serves to further enforce the authority of the Castilian variety of Spanish as the correct and most formal variety, consequently resulting in the development and reinforcement of many insecurities within US Latinx, and Latinx more generally as well. The promotion of the Castilian variety conveys the idea to US Latinx that the Spanish language is not theirs and it's the responsibility of a governing entity to explain to them how to speak their language, only perpetuating the insecurities that many Latinx feel around language usage (Helmer, 2013). Standard language ideology is just one of the ways

in which US Latinx are often faced with pressures and expectations around the way in which they speak, however even more intricacies arise when considering the additional expectations present within their own community regarding proficiency and accentedness when speaking Spanish.

3.2.4 Language ideologies and the Spanish heritage community

Given the very nature of language ideologies, Martínez (2006) argues, "that every language group constructs a set of beliefs about language in order to serve its own interests" (p. 9). We can see this taking place in the United States with English and the promotion of both monolingual ideals that pose minoritized languages as a threat, and ideals that elevate the abstract 'standard' variety of US English. In Spanish this can be seen through its own set of ideologies which have been institutionalized through the governing body of la Real Academia Española (the Royal Spanish Academy) and its dogmatic mission to maintain lexical and grammatical purity in Spanish for all of its speakers worldwide (Zentella, 2017). Given the coexistence of ideologies belonging to multiple communities, US Latinx find themselves in what Leeman (2012) refers to as a double-bind; a result of being marginalized by their Anglo peers for their relationship to Spanish and being ostracized by Hispanophones for speaking Spanish 'wrong', or for their inability to speak Spanish at all (p. 48).

Furthermore, Potowski & Carreira (2010) argue that "the most studied foreign language in the nation is criticized when spoken by those who actually speak it natively" (p. 79) and this judgement is coming from multiple different communities and groups. Thus, US Latinx often encounter and express interethnic hegemonic ideologies regarding certain varieties of Spanish by other Hispanophones; however, these beliefs are often driven by social factors associated with the group that speaks a certain variety rather than the actual linguistic features themselves. This

has long been recognized as a process in which language is racialized leading to the mapping of race onto language, whereas language serves as a proxy to achieve goals centered around racism, ultimately aiming to perpetuate the established hierarchical social strata that arbitrarily elevates the group in power, usually the white majority (Urciuoli, 2008; Rosa, 2016). Zentella (2002) details this by explaining that Spanish spoken in the Caribbean shares many linguistic similarities, yet Spanish spoken by upper-class, light-skinned Cubans is criticized far less than Spanish spoken by working class, dark-skinned Dominicans and Puerto Ricans. These strong prejudicial judgements often become deeply internalized, resulting in the perpetuation of these negative stereotypes, even by those directly affected by them (Peace, 2024). Thus, even speakers who themselves speak what would be considered a nonstandard variety of Spanish have been known to uphold notions of standard language ideology. These beliefs around standardness often lead Spanish speakers to feel pressured to either acquire a more prestigious variety of Spanish through education or abandon the language altogether.

In addition to negative ideologies associated with nonstandard Spanish varieties, there are also many strong beliefs that arise among US Latinx surrounding the variety they speak that has been influenced by contact between English and Spanish. This variety is a clear result of the differences in acquisition environment from a Spanish monolingual, however heritage speakers are still often accused of corrupting both English and Spanish through their language mixing (Zentella, 2002). Code-switching, lexical borrowing and other linguistic phenomena are key features defining Spanish spoken by many US Latinx and later passed on to future generations of heritage speakers, yet those that acquire these features in their speech are described as being incapable of mastering either language (Anderson & Toribio, 2007; Carvalho, 2012).

These strong beliefs on the linguistic ability of Spanish heritage speakers furthers the pressures placed on them and further demonstrates the complex position US Latinx find themselves in. Latinx families faced many social and political pressures to conform to a monolingual society; however, those within the Spanish heritage community still battle linguistic insecurity due to the belief amongst Latinx communities that a Latinx/Hispanic identity is dependent on language proficiency, as well as the ability to speak the 'correct' and formal register (Leeman, 2012; Showstack, 2017). Therefore, those that maintained Spanish may often be at risk of being criticized for speaking Spanish 'wrong', and those that speak with low proficiency, or did not acquire their heritage language may struggle with confidently claiming their Latinx identity (Abdi, 2011). A commonality between various types and backgrounds of heritage speakers is that these ideologies often serve as motivating factors for them to pursue Spanish language courses. Their desire to 'correct' their Spanish and/or restore their connection with their heritage and strengthen their understanding of their identity as a Latinx are common motivators for those in the Spanish Heritage community when enrolling in Spanish courses (Ducar, 2012; Potowski, 2002).

In response to years of linguistic repression that have led to generational language loss in Latinx families, institutions across the country have begun to recognize the affects that these policies and ideologies have had on Latinx communities. This recognition has resulted in an increase in the availability of programs designed for heritage students (Beaudrie, 2012). The intent of these programs has generally been to help students strengthen and/or continue the development of their heritage language, with some programs providing curriculum that also focuses on helping them restore a connection with their culture as well (Beaudrie, 2011).

Unfortunately, when students enter the heritage classroom, they are often confronted with ideologies and attitudes regarding their home variety, their proficiency, and other factors that may lead them to feel insecure about their Spanish (Beaudrie, 2012). Often, heritage learners upon entering the classroom have a level of awareness that their home variety is not the standard or formal register; however, there are still expectations – whether at the familial level or the systemic – to acquire this variety and use it in the appropriate contexts (Leeman, 2005). Therefore, despite the desire to improve their Spanish serving as a motivating factor, the pressures and insecurities around levels of proficiency and nonstandard varieties that arise in the classroom can often prevent students from participating and being successful (Helmer, 2013). These judgements and beliefs are often expressed in the classroom by both teachers and students (Randolph Jr., 2017) which can have profound effects on speakers' confidence and sense of self.

An example of the effects wrought by perpetuation of these negative ideologies towards the linguistic characteristics of heritage students is described by Abdi (2011). The study profiles a mixed heritage and foreign language class in which a particular student, born to a Canadian father and a Honduran mother, was judged by her fellow classmates and teacher as not being Hispanic or Latina due to her insecurities in participating and using her oral skills in the classroom. This case demonstrates, that despite the student having grown up with a Latina mother and having acquired relatively strong receptive skills and oral skills, she was still not accepted as Latina. This exclusion was driven not by her actual level of proficiency, but by her lack of confidence in openly practicing her heritage language in the classroom. Experiences like the one are not uncommon, leading many Spanish speakers to feel uncomfortable and insecure about their own language abilities, and have motivated the current study.

3.3 The current study

The aim of this study is to better understand the many social perceptions that exist amongst members that identify with the Latinx community that ultimately affect and may prevent certain individuals from confidently claiming their own Latinx identity. Furthermore, the goal of this study is to determine the influence of different factors like perceived nonnative accent when speaking Spanish and perceived level of proficiency or fluency, have when participants receive limited visual information about the speaker and altered biographical information. The research questions for the study are as follows:

1. How influential is the type of guise (i.e., speaker or learner) in driving participant ratings for fluency, accentedness, and perceived relationship to the Latino/Hispanic identity?
2. Do ratings in fluency and accentedness also affect the ratings for the speaker's identity?
3. How influential is the audio stimulus in driving participant ratings for fluency, accentedness and perceived relationship to the Latino identity, compared to the guise type?

It is expected that the speaker guises will be more likely to receive higher ratings for their relationship to the Latinx identity, despite potentially receiving lower scores for fluency and accentedness. In addition, we expect that, when controlling for guise, the lower a participant rates a speaker in accentedness and proficiency level, the more likely that speaker is to be perceived or labeled as 'less Hispanic' or less likely to receive high ratings for Latinx identity, therefore, the audio stimulus itself should become less influential than the guises.

3.4 Methods

The current study analyzes within community perceptions for heritage speakers and learners. To better understand speakers' and learners' perceptions, participants were asked to participate in a matched-guise study to extract participants perceptions regarding the way in which other individuals identifying as Latinx speak Spanish. The task and stimuli are discussed in further detail below.

3.4.1 Participants

Participants in the current study were recruited using two methods. The first recruitment method utilized the online platform Prolific. We were looking to recruit individuals that self-identified as Hispanic or Latino, have some familiarity with the Spanish language (enough to be able to listen to and comprehend the audio recordings in the study) and are US citizens and/or have been living in the United States for a significant period of their life – this varied based on factors such as age.

The pre-screening functions available through Prolific allowed reduce the participant pool to individuals that identify as current US citizens, Latino/Hispanic, and claim English as one of their first/early languages. Participants that passed the pre-screening through prolific were then further looked at to determine if there were any necessary exclusions that needed to be made. There was only one participant that was removed for stating that they were born in the United Kingdom and did not move here until adulthood. This participant was excluded due to the differences in sociopolitical culture between the US and the UK resulting in differences in ideologies associated with the Spanish language and those that speak it in the United Kingdom. The second method used to recruit participants was through the University of Oregon Spanish Heritage Program. These participants were not included in the current analysis, but this data will

be explored in future analyses of the data. After excluding this participant from the Prolific data and not including the University students a total of 50 participants were admissible for analysis.

3.4.2 Task

Individuals participated in an online matched-guise perception task in which they were asked to listen to pre-recorded Spanish audio files of both scripted and spontaneous speech. Each audio recording was accompanied by a brief biography describing the speaker's background. In addition to the biography, there was a picture of a plain silhouette for each speaker. Participants were asked to first read the biography provided and informed that they would only be able to listen to each recording once (See APPENDIX D to review the social guises used). After listening to the recording, participants were prompted to respond to three questions regarding the speaker. The first question asked them to rate the speaker on a Likert scale 1-9 how fluent they perceived the speaker to be (1 being non-native, 9 being native/advanced), the second question asked how accented they perceived the speaker's Spanish to be (1 being 'sounds non-native', 9 being 'sounds native') and the final question asked how they perceived the speaker's relationship to the Hispanic/Latino identity (1 being not Hispanic/Latino or no relationship, 9 being of full Hispanic/Latino descent).

After completing the matched-guise portion of the study, participants were asked to respond to a language background questionnaire. In the background questionnaire they were asked questions designed to better understand their language dominance and identity as a United States Latinx. Questions asked participants to report information such as: age at which they began speaking English and Spanish; self-reported proficiency in reading, writing, speaking and listening in both English and Spanish; comfort level speaking English and Spanish; parents' country of origin; languages spoken in the home and with peers; and more. In addition,

participants were asked to respond to a brief questionnaire to gather some of their attitudes around the relationship between Spanish and the Latinx identity. The data gathered from the questionnaires will be used for future analyses but were excluded to maintain a targeted analytical focus on the initial questions. See APPENDIX C for the full language background questionnaire used for the experiment.

3.4.3 Stimuli

The stimuli used for the study were borrowed from the publicly available Archive of L1 and L2 Scripted and Spontaneous Transcripts and Audio Recordings Corpus (ALLSSTAR) developed at Northwestern University (Bradlow, n.d.). The corpus contains audio recordings from Spanish Heritage speakers ages 18-22 that all reported having been born in the United States but acquired Spanish at birth as their first language. They each reported English as their second language, having learned it between the ages of 5 and 8 years old. Furthermore, these talkers reported a notable shift in language dominance throughout their lives with Spanish having gone from 100% usage in the home during early childhood to less than 20% after early childhood. None of the eleven talkers reported taking Spanish classes during high school and there was no information provided regarding language courses in higher education. For a more detailed description of the Spanish Heritage talker backgrounds from the ALLSSTAR corpus, see Blasingame (2018).

A total of four speakers were chosen for the stimuli – two male and two female. Given that the metadata provided by the ALLSSTAR corpus did not contain specific proficiency scores, speakers were selected based on their ability to read or describe a scenario with some fluency– meaning those that could read through the story with minimal interruptions and disfluencies that may affect overall comprehension of the listeners. In addition, all Spanish Heritage speakers in

the corpus received intelligibility scores from native Spanish speakers that used Spanish exclusively during childhood (see Blasingame 2018 for more details regarding the intelligibility task). All Spanish heritage speakers, with the exception to one speaker who was not used for this study, had Spanish intelligibility scores ranging 0.84 – 0.95 demonstrating that they were all relatively easy to comprehend and understand by native Spanish speakers.

The scripted audio chosen for the study was a recording of speakers reading the story The North Wind and the Sun. For the spontaneous speech selected, speakers were provided with a comic-like series of images displaying interactions between two pigs and were asked to create a story narrating the images they saw. The survey was then prepared for randomization so that participants were exposed to different guises for each speaker. In order to randomize guise and speaker exposure, a total of eight guises were created with four of them describing a heritage learner background and four guises describing a heritage speaker. A set of four guises containing two learner guises and two speaker guises were used for the North Wind and the Sun audio. Three blocks were formed to align the speakers with different guises. The first block aligned Speaker 101 with guise 1, Speaker 100 with guise 2, Speaker 105 with guise 3 and Speaker 107 with guise 4. To create the remaining two blocks, audio recordings were shifted so that the first speaker of the second block was Speaker 100 and for the third block the first was Speaker 105. To randomize the presentation of stimuli for the spontaneous speech the remaining four guises were used. The same process was used to ensure that each speaker was presented with a different guise for each block. This was to ensure that each participant had a chance of being exposed to each speaker in either learner guise or a speaker guise.

3.5 Results

To analyze the participants' ratings, an analysis of variance was run to determine the main effects and test for the presence of an interaction.

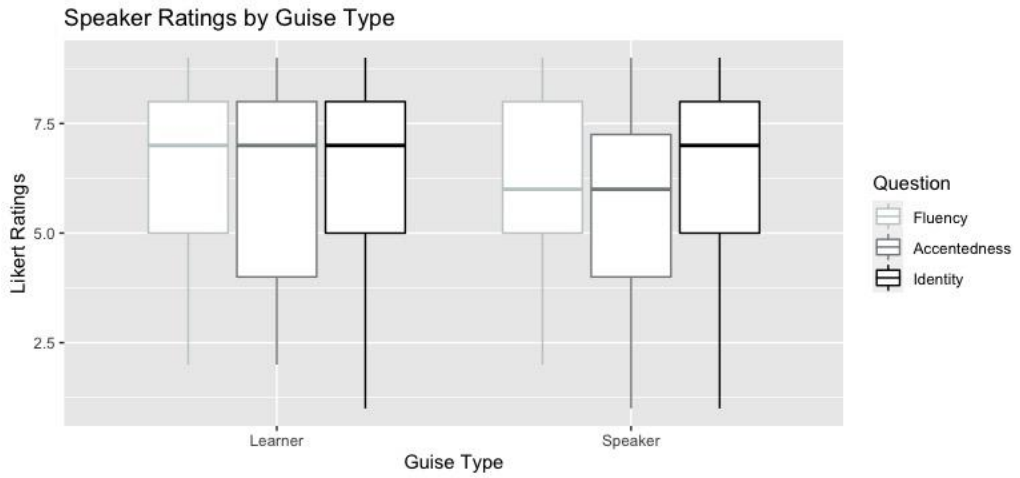


Figure 3 *Ratings for fluency, accentedness and Latino identity for each guise type across all speakers.*

Figure 3 shows that participants rated the learner guises as more proficient and native-like, as well as being more strongly associated with the Latinx identity. An analysis of variance showed a main effect of question ($F(2,63) = 7.642$, $p = .001$) but not guise type ($F(1,15) = 3.598$, $p = .058$). This demonstrates that the difference in mean ratings for each question were found to be significant at an alpha level of 0.05; however, guise type did not significantly influence ratings. The interaction between guise type and question type was also not found to be significant ($F(2,5) = 0.668$, $p = .513$). The learner guises yield very similar mean ratings with a preference to label them as 'more Latinx', fluent and native sounding for each question while there appears to be more variance in the ratings for the speaker guises across questions.

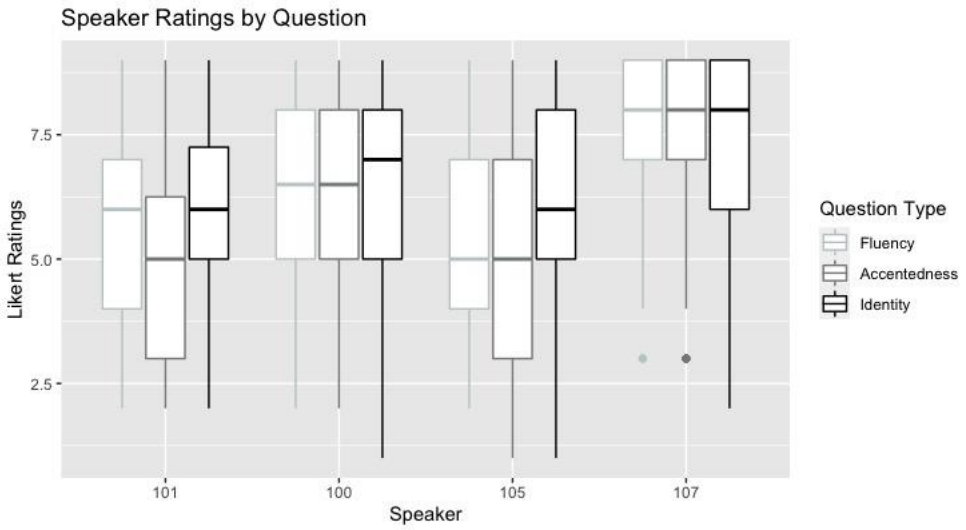


Figure 4 *Ratings for fluency, accentedness and Latino identity displayed for each Speaker.*

In addition, it appears that participants were willing to rate speakers relatively highly when considering their relationship to the Latino identity, despite potentially giving low ratings for a perceived non-native accent. This trend can be seen in Figure 4. Speakers 101 and 105 were both rated fairly similarly to the other speakers for Latino identity but received low ratings for fluency and accentedness (where high ratings represent native-like fluency and accent). This trend is interesting, as these speakers have less consistency in their scores. Speaker 105 was still rated very highly by participants on the identity question; however, the average ratings for the language related questions were fairly low for fluency and accentedness. Indeed, Speaker 105 actually had the lowest average rating for accentedness. A similar trend can be found with Speaker 101, who received the second lowest average accentedness rating but was still rated similarly to other speakers on the identity question. Figure 4 also shows that Speakers 100 and 107 were favored by participants as being highly fluent, having a native-like accent and having a strong relationship with the Latino identity. An analysis of variance showed a main effect of speaker ($F(3,866) = 86.351$, $p < .001$) and question ($F(2,63) = 9.381$, $p < .001$) further supporting

these notable differences in the figure for speaker ratings. There was also a significant interaction found between speaker and question ($F(6,81) = 4.060$, $p = .001$) showing that the speaker did influence ratings.

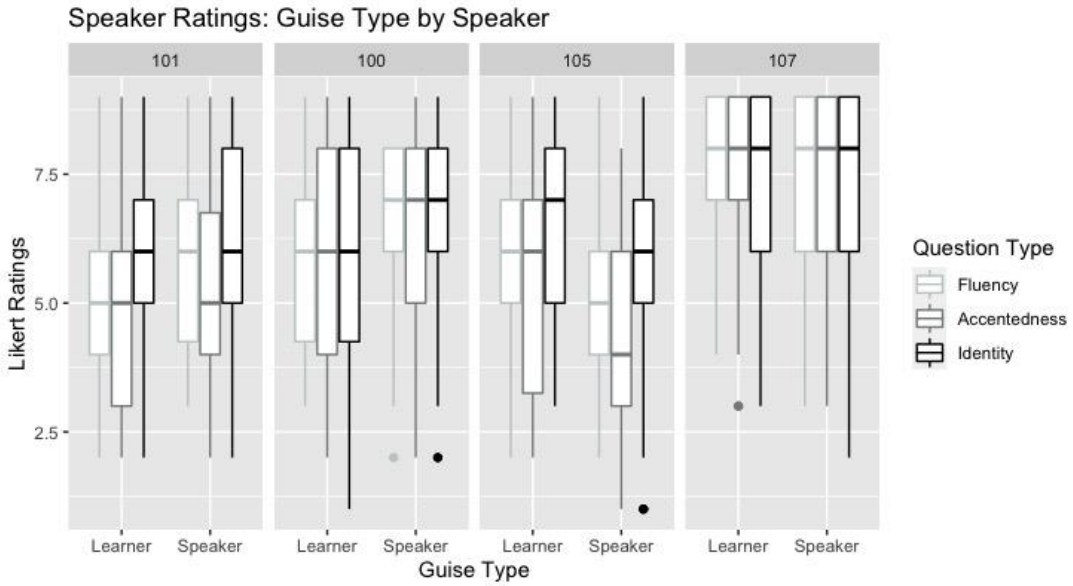


Figure 5 Ratings for fluency, accentedness and Latino identity for each Speaker, grouped by guise type within each Speaker.

Figure 5 provides a deeper look into the guise ratings for each individual speaker. Speakers 100 and 107 were favored by participants as being both more native-like in their language abilities and more Latino. That is, they received higher average ratings for the speaker guises. On the other hand, Speakers 101 and 105 had a higher average rating for the learner guises. This trend is most distinct for Speaker 105 where this preference for the learner guise can be clearly seen in the third panel in Figure 5 (i.e. all "learner" boxes are higher than the "speaker" boxes). Despite these observations, guise type was again found to be insignificant ($F(1,1) = 0.263$, $p = .608$), however, the interaction between speaker and guise was found to

significant ($F(3,125) = 12.467$, $p < .001$), showing that participants rated certain speakers differently depending on the type of guise the speaker was presented with.

3.6 Discussion

The results shown above are unable to support the first hypothesis indicating that speaker guises would be perceived as more strongly associated with the Latino identity. Rather, they show that participants felt relatively comfortable associating each guise type with the Latino identity. This potentially highlights the trends mentioned above stating that the second-generation Latinx population would become more prominent (Abdi 2011), therefore, potentially making some of the learner guises feel very salient for some of the participants. The learner guises referenced second or third generation individuals that were taking Spanish language courses to strengthen their skills, which potentially represents a very relatable narrative for some of the participants given the research showing the increase in availability in courses designed for heritage speakers. This could be an explanation for why the speaker guises were not given higher overall ratings as expected, but a deeper look into the demographic data for the participants is necessary to ascertain this claim.

The second hypothesis was also unsupported given that participants were still willing to give high ratings for the identity question, despite low ratings on accentedness and fluency. These results suggest that there is potentially an outside factor driving ratings given that the factors that Spanish speakers are seemingly judged the most on – proficiency and accentedness – did not actually appear to influence whether or not they were perceived as having a strong relationship to the Latino identity. Furthermore, the low ratings for the language-related questions are interesting when considering the intelligibility scores provided by the ALLSSTAR corpus. Speakers 101 and 105, who received the lowest ratings for fluency and accentedness,

both received some of the highest intelligibility scores (0.94 and 0.93) given by Native Spanish participants. Speaker 107, on the other hand was consistently rated the highest by participants for all three questions, yet has the lowest intelligibility score out of all four of the Speakers. A further analysis into aspects of their speech like word choice and acoustic features could provide insight into these discrepancies.

The third hypothesis also yielded null results given that the audio stimuli was actually found to be more significant than guise type in influencing ratings. However, the interaction between guise type and Speaker does show that the guise did to some extent effect the ratings Speakers received. The two Speakers that were rated as most native-like and received the highest scores for the identity question are also the two Speakers that received higher ratings for the speaker guises than the learner guises, while the two Speakers that received the lowest ratings overall received higher ratings for the learner guises. This potentially demonstrates that the participants' may have already expected that Speakers 100 and 107 would align more with a speaker guise (i.e., had more exposure in the home, potentially spoke Spanish as an L1 but underwent a change in dominance, etc.) given their high fluency and native sounding accent, therefore a learner guise may have come across as misleading or unexpected thus yielding lower ratings. For Speakers 101 and 105, however, we may be seeing that the guise is doing more work in driving participants' rating if the Speaker is perceived to be less native-like and less fluent. Meaning, if participants perceive certain aspects of the Speakers' speech that they believe to be indicative of a non-native accent they may be more willing to maintain high ratings for the identity question, given that these guises specify that these Speakers needed to acquire the majority of their Spanish speaking abilities in the classroom due to limited home exposure or limited productive abilities. Ultimately, we do see that the guise type is influential but how

influential the guise is, appears to be reliant on the individual speaker and particular features of their speech that have not yet be determined.

3.7 Conclusion and future directions

Overall, the results of this study provide insight into the factors driving Spanish speaking listeners' perceptions of heritage speakers and have revealed a plethora of avenues for future research we plan to carry out. Despite the existence of ideologies that mark a Latinx identity as exclusive to proficient Spanish speakers of the standard variety, US Latinx appear to be beginning to embrace more inclusive ideals for those that choose to identify as Latinx. Specifically, participants demonstrated that they were willing to rate all Speakers across guise types as being relatively strongly associated with a Latino identity even if they gave low ratings for speaker fluency and accentedness. This highlights that as the Latinx population continues to grow and as the second generations and later begin to account for a large majority of the community in the US, we may be seeing a shift in the beliefs that exist amongst US Latinx in particular.

This community of speakers may be beginning to recognize the diversity that exists within the Spanish heritage community making them more willing to rate a variety of biographies as 'Latino'. In addition, some of the participants could have found commonalities between their own upbringing and that described in the guises potentially making them more likely to give high ratings for those guises. For this reason, we intend to conduct future analyses taking into consideration the background information for participants to help elucidate the potential effects that salience between the guises and participant upbringings may have had on ratings. For this analysis, we intend to group participants using factors like their self-reported linguistic ability, language dominance, and age, to determine the influence that these properties

may have on ratings. In order to better understand the findings from this study, we also hope to compare ratings for spontaneous versus scripted speech to see if there are significant differences in ratings based on stimuli type. Given that the rating for the guises were dependent on the individual speaker, there could be potential differences, like word choice or speaking rate, that arise in spontaneous speech and influence ratings – i.e., the inclusion of colloquial words indicative of certain varieties may influence ratings, or long pauses while speakers think of how to continue the story plot may lead to low fluency ratings.

Finally, as mentioned, we have also collected data from students enrolled in the University of Oregon Spanish Heritage program which were not included in this analysis. We intend to include this data in further analyses to better understand how Spanish heritage speakers' speech production is judged by other individuals that identify as Spanish heritage and are currently enrolled in heritage classes. We hope to look further into the effect that may arise when specifically focusing on individuals enrolled in heritage courses that are currently or have taken effort to strengthen their Spanish skills and connection to their heritage through university courses. We anticipate that there will be effects that depend on the pedagogical method adopted by the institution that may ultimately influence the ideologies students have gained and/or lost from enrollment in the program, in turn shaping and informing the way their heritage students view varieties of Spanish and those that speak them. The results from these future analyses may help to better inform heritage programs as they approach navigating the presence of ideologies in the classroom.

CHAPTER 4: THE INFLUENCE OF SOCIAL INFORMATION ON SPEECH INTELLIGIBILITY WITHIN THE SPANISH HERITAGE COMMUNITY

The following chapter is a co-authored chapter with Dr. Melissa Baese-Berk and Dr. Charlie Nagle that was published in the special issue Investigating the Effect of Listening Conditions on Speech Intelligibility of the journal *Languages* in September 2022. I was primarily responsible for the conceptualization, methodological design, carrying out the investigation, data curation, and development of the original writing and draft formation. Dr. Baese-Berk and Dr. Nagle supervised the work, assisted in each stage of the research process, and guided the edits and revisions for the manuscript. Additionally, Dr. Nagle carried out the formal analyses.

From Staggs, C., Baese-Berk, M., & Nagle, C. (2022). The influence of social information on speech intelligibility within the Spanish heritage community. *Languages*, *7*(3), 231.

4.1 Introduction

Speech perception is heavily influenced by social factors that prompt assumptions about standard language, prestige, and social expectations of language production (Hay & Drager, 2010; Kutlu, 2020; Rubin, 1992). Listeners exploit both linguistic and paralinguistic information—information regarding the speaker themselves rather than acoustic information—in the speech signal, which can ultimately trigger both implicit and explicit biases that can alter speech processing (Babel & Russell, 2015; Chappell, 2020). Many studies to date have focused on the interaction between language ideologies, social expectations, and paralinguistic information in the perception of varieties of English, including nonstandard and non-native varieties. This body of literature has shown that implicit biases do trigger some cognitive mechanism that can result in reduced intelligibility, although there is debate regarding the

underlying mechanism—whether it be reduced listening effort (Rubin, 1992) or a mismatch in expectation (McGowan, 2015).

While these findings are relatively well-established in the literature dedicated to English and its varieties, fewer studies have analyzed the influence of ideologies and social expectations within the United States Latino[5] community, including how they may play a role in shaping listeners' judgements about and perception of the speaker. Furthermore, even less is known about these effects with regard to the perception of Spanish heritage language, that is, Spanish produced by heritage speakers and learners who were born in the United States and have an ancestral and cultural connection to the Spanish language (Fishman, 2001). Spanish heritage speakers in the United States often find themselves in what has been described as a "double bind" resulting from being marginalized by both their Anglophone peers for speaking "impure" English and for their connection to Hispanic/Latino culture, and by their Hispanophone peers for their Spanish proficiency, or lack thereof (Leeman, 2012). For example, the usage of non-native pronunciation patterns and nonstandard language often contributes to a sense of othering and may lead to attitudes related to languagelessness (Rosa, 2016). This creates a complicated balancing act where heritage speakers must navigate the social pressures from the English-speaking and Spanish-speaking communities, both of which are integral to their lives. Therefore, it is important that researchers not only study English speakers' perceptions of Spanish-accented English, but also Spanish speakers' perception of Spanish produced by heritage Spanish speakers, to better understand both sides of the double bind.

To address this gap, we investigated how giving listeners information about the socio-cultural background and upbringing of the speaker influences their perception of that individual's

[5] Given that this chapter is a previously published manuscript, the usage of 'Latino' is unchanged throughout.

speech. In this study, we generated descriptions targeting heritage backgrounds, making a special effort to better encapsulate the heterogeneity of the Spanish heritage community by distinguishing between speakers (i.e., individuals who acquired Spanish as a first language but underwent a shift in language dominance from Spanish to English during primary school) and learners (i.e., individuals who have an ancestral and cultural connection to Spanish but acquired most of their skills in a foreign or world languages course). The terms "bilingual" and "heritage" have often been employed monolithically, rendering a static portrait of a diverse group of individuals. Such a characterization fails to acknowledge the variability in both their social lives and linguistic experiences; we therefore developed the speaker and learner guises to explicitly address a portion of this variability. In the sections that follow, we describe the current state of research on the role of ideology and expectation in speech perception, as well as the language ideologies and attitudes that exist within the U.S. Latino community.

4.2 Background

4.2.1. The Role of ideology and expectation in speech perception

Language exists within an inherently social world and as a result, the sociocultural reality that speakers inhabit is known to play a role in speech perception (Kang & Rubin, 2009; Kutlu, 2020; McGowan, 2015; Niedzielski, 1999; Rubin, 1992). Previous research has shown that listeners exploit socio-indexical cues in the speech stream to make higher level linguistic and social judgments. These judgments have, in turn, been found to lead to linguistic profiling and perceptions of heightened accentedness and reduced intelligibility (Giles & Trudgill, 1983; Williams et al., 1999), demonstrating that "our beliefs about accents have social consequences for how we hear others and judge them as authentic speakers of the language" (Levis, 2020, p. 18).

Early studies focusing on the role of social information on speech perception were grounded in theories of language ideology and bias in communication between native and non-native speakers of English. For instance, Rubin (1992) found that when listeners were presented with an Asian face, they perceived the speech as more accented and transcribed it less accurately than when presented with a Caucasian face. Overall, Rubin argued that social information can influence the amount of effort listeners invest in decoding the speech signal. This early work was the basis of theories of reverse linguistic stereotyping, according to which social factors related to a speaker's group membership lead to altered perceptions of their language style and proficiency. For instance, listeners often develop expectations about the speaker based on their nationality, and these expectations have been shown to influence listeners' processing of the speech and their perception of the speaker (Kang & Rubin 2009). These findings have been corroborated in research on a range of varieties of English. Niedzielski (1999) determined that speakers' perceived nationality was crucial to whether they were perceived as speaking standard American English or Canadian English by listeners from Detroit, MI. The role of implicit racial bias has also been found to be influential in speech perception and accentedness judgements, with minority visual stimuli (i.e., minority faces) and minority varieties of English being viewed as more accented regardless of the actual speech stimuli presented (Kutlu, 2020). Running through this body of work is the idea that socioindexical information can activate biases that affect listeners' perception of speech.

Other researchers have argued that the degree of congruence between listener expectation and reality can lead to increased or decreased processing demands, which could account for the effect of socioindexical information on speech perception—however, as described, this finding is not consistently agreed upon in the literature. In other words, according to these researchers,

listeners are not subconsciously rejecting their end of the communicative burden due to the activation of implicit bias, but rather they experience increased processing demands when exposed to stimuli that run counter to their expectations (Walker & Hay, 2011; Van Engen & Peelle, 2014). Babel and Russell (2015) investigated the effects of paralinguistic information on speech perception by analyzing social associations between speech and ethnicity in a multicultural, multilingual, urban context. The results from this study showed that even in a diverse area where the Asian community is prominent, Chinese Canadian voices were perceived as more accented and less intelligible. Furthermore, this effect was even stronger when the audio stimuli were paired with visual information. To explain these findings, the authors suggested that there was a greater processing cost in incongruent audio–visual conditions—conditions where an Asian face was presented with audio from a White Canadian speaker or conditions where a White face was presented with audio from an Asian Canadian speaker—than in congruent conditions. In short, there are two explanations that could account for the effect of social information on speech processing: implicit bias shaping perceptual processing, and a mismatch in expectations driving differential processing costs. Although there is a lack of clear consensus in this area, it seems likely that both ideology and expectation affect perception.

4.2.2 Language Ideologies and attitudes in the Spanish heritage community

The Spanish heritage community in the United States is heterogenous, insofar as it encompasses individuals with vastly different linguistic abilities who come from a range of sociocultural backgrounds. Some heritage speakers have relatively strong productive and receptive skills, others have weak productive skills (or none at all) but strong receptive skills, and still others have limited or no ability in both skills. Despite such variability, the common

factor that connects heritage speakers is that they identify culturally and emotionally with their heritage language (Fishman, 2001).

Language ideologies develop around a set of constructed beliefs with the purpose of furthering social objectives in both language communities (Martínez, 2006), and, as a result, heritage speakers struggle to find acceptance in either community. Because of the situated, contextually defined nature of language ideologies, it is necessary to focus specifically on the ideologies that exist within the U.S. Latino community, rather than the Spanish-speaking community more generally, because such situated, local ideologies are most likely to influence the perception of Spanish as a heritage language in the United States.

Heritage speakers, along with other bilinguals, are confronted with ideologies that often conflate issues related to language proficiency and language variety (e.g., who speaks the "correct" variety of Spanish) with social issues such as ethnic identity (Potowski, 2012; Showstack, 2017). Given the dynamism of the U.S. Spanish-speaking community, where second-generation Latinos are beginning to outnumber Spanish-speaking immigrants (Abdi, 2011), it should come as no surprise that language ideologies are in flux. For example, U.S.-born Latinos seem to place less importance on proficiency than Latinos who have immigrated to the U.S. (Potowski, 2012). Despite this potential change in attitudes and ideologies, the relationship between language and identity remains complex. Proficiency in Spanish is still found to contribute to a personal sense of ethnic identity, but even advanced proficiency in Spanish does not guarantee community acceptance (Rosa, 2016). For this reason, heritage speakers face conflicting circumstances. On the one hand, they feel connected to their heritage through their acquisition process, but on the other, they feel compelled to achieve a certain linguistic standard in Spanish to gain community acceptance. As Rosa (2016) explained, heritage speakers are often

viewed as having a linguistic deficit in both languages and as a result are susceptible to accusations of languagelessness from both their Anglophone and Hispanophone peers, further reinforcing the double bind that these individuals face. In fact, heritage speakers face scrutiny and judgment toward their Latino identity by other Latinos simply for speaking English and having a connection to it (Rosa, 2019). Researchers are currently exploring these complex dynamics, but more work is needed to better understand the role they play in speech perception.

Listeners have certain expectations about the way a speaker "should" speak or sound that are formed based on external factors such as speaker appearance or identity. Furthermore, it is often the case that "… listeners' immediate reactions to speaker Hispanicity can trigger linguistic assumptions, prompting differential expectations… [and] these unconscious biases can result in the proliferation of linguistic violence toward heritage speakers, regardless of the phonetic variants they employ" (Chappell, 2020, p. 49). Simply put, heritage speakers face judgment based on both their linguistic production in Spanish and the way they are perceived by the listener, which is rooted in notions of identity, belonging, etc. In one illuminating study, Chappell (2020) investigated the extent to which listeners' perceptions of speakers' Hispanicity, intelligence, confidence, and other social attributes may be altered based on how the speaker produces the orthographic <v> and the orthographic <b> in Spanish, which are both graphemes of the phoneme /b/ and can be realized as either [v], [b], or [β], depending on the phonological environment and dialectal characteristics. Overall, their findings indicated that ideologies and attitudes differentially affected each speaker and prompted different sets of judgements and attitudes based on both speaker realization of the variant and listener assumptions of speaker Hispanicity. Such results indicate the need to analyze the extent to which these expectations may impact speech processing, especially the extent to which the differential

assumptions that listeners develop based on the information they receive about the speaker influence their perception.

In this study, we therefore distinguished between heritage speakers and learners. We defined heritage speakers as "asymmetrical bilinguals who learned language X—the 'heritage language'—as an L1 in childhood, but who as adults, are dominant in a different language" (Benmamoun et al., 2014, p. 260). For the purposes of this study, the key difference between speakers and learners is related to the role of education in the acquisition of primary productive skills. Heritage learners, like heritage speakers, have a cultural and ancestral connection to the heritage language; however, as the term suggests, most of their productive skills in the heritage language were acquired in an academic setting, whether that be through language classes or bilingual education (Montrul, 2010). We further narrowed the heritage learner definition to describe individuals who are either third- or fourth-generation Latinos, who are most susceptible to language loss and thus tend to be either receptive bilinguals or have limited receptive and productive abilities. We also included individuals who may not fall within these groups but still align in terms of their language skills (e.g., second-generation Latinos who, despite being expected to have strong skills in the heritage language, actually have limited receptive and productive skills). Thus, the primary distinction between speakers and learners, as conceptualized in this study, is that certain factors have led heritage learners to rely on external resources to acquire most of their heritage language skills.

4.3 The current study

In this study, we examined the effect that social information has on speech intelligibility, as well as how the listeners' own identity may impact their perception of the speaker. We manipulated social and linguistic stimuli using a matched-guise methodology to better

understand how different Latino social guises may prompt certain expectations on the part of the listener that can, in turn, affect speech processing. We aimed to understand how social variables—both attributed to the speaker and the listener—may impact speech intelligibility. This study was therefore driven by the following questions:

1. Do social expectations influence whether heritage Spanish speakers are perceived as more or less intelligible when paired with different social guises (i.e., heritage *speaker* vs heritage *learner* guise)?
2. Do listener background variables affect the intelligibility of heritage Spanish speech, and do these variables moderate any potential guise effects?

We expected social guise to influence participants' intelligibility scores, potentially leading to the heritage *speaker* guises receiving higher scores and heritage *learner* guises receiving lower scores regardless of the audio stimulus. This hypothesis is driven both by the results from our validation study (reported below) indicating that U.S. Latinos may expect heritage *speakers* to be more fluent, native-sounding, and Latino than *learner* guises, and previous findings showing that social guise can affect speech processing depending on the extent to which the speech characteristics match listeners' expectations—which are rooted in a preexisting concept of how each guise "should" sound when speaking Spanish. Thus, we reasoned that speaker intelligibility might not be as objective as one might expect, but rather could be influenced by social factors. We also developed an exploratory hypothesis that the listener's own identity could moderate the guise manipulation, such that, for instance, listeners who identified Spanish as their L1 might exhibit a stronger guise effect than listeners who identified Spanish as their L2.

4.4 Guise validation - Methods

Before analyzing the influence of the social guises on speech intelligibility, we first conducted a study to determine the effectiveness of the social guises we designed. These results serve as an initial comparison of participants' expectations for heritage speakers and heritage learners and help inform the results described in this paper.

4.4.1 Stimuli

The stimuli used for this study were developed to highlight the key social and language background differences between the two guise types. Both speakers and learners have a cultural and ancestral connection to the Spanish language; however, they differ with respect to when and where they acquired their heritage language. For the purpose of this study, heritage speakers are asymmetrical bilinguals who underwent a shift in language dominance from Spanish to English in early childhood, often around the time they entered primary school. Heritage learners, on the other hand, acquired most of their Spanish skills in the classroom through foreign or second language classes, often later in life than heritage speakers. Some heritage learners may enter the classroom as receptive bilinguals—they can understand Spanish, but they do not speak it—whereas others enter the classroom with some exposure to Spanish but lack both productive and receptive skills in the language. These characteristics were used to guide the creation of the guises for this study (APPENDIX D).

4.4.2 Participants

We used the online platform Prolific to recruit 35 raters who were (1) U.S. citizens, (2) identified as Latino/Hispanic, and (3) claimed English as one of their early or first languages. All participants reported being born in the United States, with the exception of one participant who was born in Chile and moved to the United States prior to entering primary school. To shed light

on the social and linguistic environment in which the participants were raised, participants were asked to provide information on their parents' birthplace and languages spoken. Parents were from the United States (23), Mexico (37), El Salvador (4), Chile (2), Guatemala (1), Cuba (1), Honduras (1), and Brazil (1). Most were Spanish–English bilinguals (39), but some spoke only Spanish (19) or English (10). One parent was bilingual in Spanish and Portuguese, and another was trilingual in English, Spanish, and Portuguese.

4.4.3 Procedures and analyses

Participants were asked to complete an online survey where they rated their expectations for heritage speaker and learner guises on three questions. First, participants were presented with a short biographical description providing information about a fictional individual's language background, including their family and upbringing. After reading the description, participants were asked to rate on a 9-point Likert scale how fluently they expected the person to speak Spanish (1 = beginner/non-native and 9 = advanced/native-like)1 and how accented they expected the person to sound (1 = sounds non-native and 9 = sounds native), based only on the information provided. Participants were also asked to rate their perception of the individuals' relationship to the Latino identity (1 = not of Latino descent/no relation and 9 = of full Latino descent). After rating each of the eight social guises, participants completed a language background questionnaire, language experience questionnaire, and language attitudes questionnaire. Although the survey was entirely in English and a specific level of Spanish proficiency was not required, we collected self-reported proficiency scores. Participants reported their abilities in Spanish and English speaking, listening, reading, and writing using the scale 1 = poor, 2 = needs work, 3 = good, 4 = very good, and 5 = native speaker command. Participants

reported on average that listening was their strongest skill in Spanish (M = 3.86, SD = 1.29) and writing was their weakest skill (M = 2.77, SD = 1.21).

4.5 Intelligibility task - Methods

4.5.1 Stimuli

The pre-recorded audio stimuli used for the task were taken from the Archive of L1 and L2 Scripted and Spontaneous Transcripts and Audio Recordings Corpus (ALLSSTAR) developed at Northwestern University. The Spanish heritage speakers in this corpus were between the ages of 18 and 22 and were born in the United States, and each reported Spanish as their first language and English as their second language, acquired between the ages of 5 and 8. They also reported a distinct shift in language dominance from 100% use of Spanish to less than 20% use after early childhood, and none reported taking Spanish language courses. For a detailed description of the participants who recorded the audio stimuli, see Blasingame (2018).

We selected four Spanish heritage speakers, two male and two female, from the Corpus: Speaker 100, Speaker 101, Speaker 105, and Speaker 107. The ALLSSTAR corpus provides self-reported proficiency scores for reading, writing, speaking, and listening (See APPENDIX E). The scale was 0 = none to 10 = perfect. Speaker 100 reported a 9 for both speaking and listening ability in Spanish, and Speaker 101 reported an 8 for both skills. Speaker 105 reported the lowest scores, with a 5 for speaking and a 7 for listening, and Speaker 107 reported the highest, with a 10 for both skills.

While there is no information on the dialects of Spanish spoken by the speakers, there are metadata on the languages spoken by each speaker and the country/countries they have lived or spent a significant amount of time in. Speaker 100 reported speaking Spanish, American English, and Portuguese, with Spanish and English being their familial languages and the languages they

used with friends and at school, while Portuguese was used only with classmates and teachers. They did not report spending a notable amount of time in a country other than the United States. Speaker 101 reported speaking Spanish with family, friends, and co-workers, and Quechua mostly while in Peru with extended family. They reported speaking English while in the United States with siblings, co-workers, classmates, and teachers. Speaker 105, a speaker of Spanish, French, and American English, did not report spending a significant amount of time in a country other than the United States. They reported using Spanish and English with family, coworkers, and roommates and French with their spouse and friends. Finally, Speaker 107 reported speaking American English, Spanish, French, and Italian. They used English with coworkers, friends, classmates, and teachers and Spanish with family, extended family, and housemates, and at school. They also indicated that they used French and Italian at school and Italian with extended family. Speaker 107 reported spending a significant amount of time in Argentina.

We sampled 30 sentences per speaker from a total of 120 hearing noise test (HINT) sentences provided in the ALLSSTAR corpus. HINT sentences were initially designed to be of similar length and syntactic structure and roughly equal intelligibility (Nilsson et al., 1994); however, it should be noted that these stimuli were not developed expressly for use with talkers from a variety of language backgrounds. In spite of this, we do not believe there are any stimuli that are notably more or less difficult. We masked each sentence in speech-shaped noise at SNR level −5 to simulate a real-world listening scenario, which often includes moderate background noise. Each speaker was presented once with a speaker guise and once with a learner guise for a total of 8 guises, four *learner* and four *speaker*.

4.5.2 Participants

For this study we recruited 50 new listeners (i.e., listeners who did not participate in the guise validation study) using Prolific. Participants were pre-screened to ensure that they (1) identified as Hispanic and/or Latino, (2) were born in the United States or had lived most of their life in the United States, and (3) had strong enough receptive abilities in Spanish to be able to listen to the Spanish stimuli and complete the study. Receptive skills for each of the listeners were confirmed through the language background questionnaire where participants self-reported their Spanish abilities in listening, speaking, reading, and writing on a scale of 1 = very low proficiency to 9 = native/native-like proficiency. Although the scales used for self-reported proficiency were not identical, we still compared listeners' scores with raters' scores from the guise validation study. Overall, we can see that raters and listeners compare, with Spanish listening being their strongest skill (M = 7.96, SD = 1.33) and writing being their weakest (M = 6.40, SD = 2.18), which is common in the heritage community. Prolific pre-screening allowed us to restrict the participant pool by current country of residence, ethnicity, first language, and fluent languages. We instructed participants to use headphones for the entirety of the study, which we verified by asking them to indicate whether or not they had done so upon completion of the task. After excluding participants who did not use headphones during the task (n = 5), data from 45 participants were available for analysis.

4.5.3 Procedure and analyses

Participants completed a matched-guise transcription task. The task was structured such that participants were presented first with the social guise followed by 15 sentences from a single speaker. Each set of 15 sentences varied across talkers; that is, each talker was presented producing a unique set of 15 sentences. Each sentence was presented to participants only once,

after which they were immediately prompted to transcribe the entire sentence. Participants were not constrained for time while transcribing.

Upon completion of the transcription and ratings tasks, participants were asked to respond to a language background questionnaire and language attitudes questionnaire to gain insight into their beliefs regarding the relationship between the Spanish language and Latino identity. Participants were asked to share information such as the age at which they began speaking English and Spanish; their self-reported proficiency in reading, writing, speaking, and listening in both English and Spanish; their comfort level speaking English and Spanish; the extent to which they thought speaking Spanish is integral to the Latino identity; and so on.

As part of a larger study on social information and speech perception, we also collected scalar, listener-based ratings on speaker proficiency, accentedness, and relationship to the Latino identity. In this chapter, we focus on the intelligibility data.

4.6 Results

4.6.1 Guise Validation

As a first step, we computed descriptive statistics for each guise type for fluency, accentedness, and identity constructs. As shown in Table 3, participants expected the speaker guise to be more fluent, less accented (higher scores indicate less accent), and to have a stronger connection to the Latino identity than the learner guise.

Table 3 *Descriptive statistics by guise type for each construct.*

Construct	Speaker Guise	Learner Guise
Fluency	5.01 (2.05)	3.33 (1.75)
Accentedness*	4.57 (2.37)	3.01 (1.96)
Latino Identity	5.99 (2.12)	4.59 (2.10)

* Accentedness was found to be the only variable that varied significantly across guise types.

To determine if these differences were statistically significant, we fit separate linear mixed-effects models, one per construct, with guise type as a fixed effect (contrast-coded; learner = −0.50, speaker = 0.50) and by-listener and by-item random intercepts. We used a likelihood ratio test to compare these models against a baseline, intercept-only model that did not include guise type. For fluency and Latino identity, adding guise type did not significantly improve model fit: for fluency, $\chi2(1) = 2.91$, $p = 0.088$, and for Latino identity, $\chi2(1) = 2.94$, $p = 0.086$. The estimated mean difference for fluency was 1.69 (95% CI = [−0.35, 3.73]), and the estimated mean difference for Latino identity was 1.40 (95% CI = [−0.26, 3.06]). In contrast, adding guise type to the accentedness model significantly improved fit: $\chi2(1) = 4.93$, $p = 0.026$. For accentedness, the estimated mean difference was 1.56 (95% CI = [0.24, 2.88]). Overall, then, from a descriptive standpoint, the guise type manipulation seemed to generate differential expectations for the speaker and learner guises. However, this difference only reached statistical significance for accentedness.

4.6.2 Intelligibility Task

Next, we evaluated the effect of guise type on intelligibility, which we operationalized as the proportion of words that were transcribed correctly (i.e., intelligibility = correctly transcribed words/total words in the utterance). Descriptively, intelligibility was high overall, but the mean for the learner guise (M = 0.74, SD = 0.29) was slightly higher than the mean for the speaker guise (M = 0.72, SD = 0.31). We followed the same procedure to assess significance. We fit models with and without guise type and used a likelihood ratio test to compare them. These models contained by-listener and by-item random intercepts. It was not possible to model by-talker random effects because only four talkers were included in the study. Therefore, to control for the effect of talker, we created a contrast-coded control covariate, which we integrated into

the models as a fixed effect. As in the previous models, we contrast-coded guise type. Adding guise type to the intelligibility model did not improve model fit, $\chi 2(1) = 1.00$, $p = 0.318$, and the estimated difference between speakers and learners was very small, estimate = −0.03 (95% CI = [−0.08, 0.03]). Thus, guise type did not affect the accuracy with which listeners transcribed speech.

We hypothesized that listener demographic variables could modulate the effect of guise type on intelligibility; that is, that guise type might only affect certain types of listeners, or listeners who show certain characteristics. To test this exploratory hypothesis, we fit a series of models focusing on the potential impact of three listener background variables: (1) whether listeners identified Spanish as their L1 or L2, (2) the extent to which they felt it was important for Latinos to speak and maintain their Spanish, and (3) the extent to which they felt that someone who identifies as Latino should know Spanish. The first variable was a binary response, whereas the latter two variables (speak and maintain Spanish) were scalar ratings given on a 7-point Likert scale (1 = strongly disagree and 7 = strongly agree). There were 20 listeners who identified Spanish as their L1, and descriptively, the mean rating listeners gave to both Likert-type items was high, indicating moderate overall agreement: maintain Spanish, M = 6.11 (SD = 0.99), and know Spanish, M = 5.11 (SD = 1.39). We followed the same modeling procedure, fitting three models, one per listener variable, to the intelligibility scores. As in the guise type model, these models included by-listener and by-item random effects and a contrast-coded control covariate for talkers. The primary effect of interest was the interaction between the listener background variable and guise type, which, if significant, would suggest that the listener background variable mediated the effect of the guise manipulation. We compared each model to

a baseline, intercept model without fixed effects for guise type, the listener variable, and the interaction term.

For the listener L1 variable, the model with the listener L1 by guise type interaction was a significantly better fit than the baseline model, $\chi2(3) = 10.41$, $p = 0.015$. However, estimates showed that the improvement in fit was due to the main effect of listener L1. Listeners who indicated Spanish as their L1 transcribed the stimuli significantly more accurately than listeners who self-identified as L2 speakers (estimate = 0.11, $p = 0.003$, 95% CI = [0.04, 0.19]). The interaction term with guise type did not reach significance (estimate = 0.01, $p = 0.314$, 95% CI = [−0.01, 0.04]). To confirm that listener L1 did indeed affect intelligibility, we respecified the model without the interaction term. This model, which included guise type and listener L1 but not their interaction, was a significantly better fit than the baseline model, $\chi2(2) = 9.40$, $p = 0.009$. The model with the Likert item focusing on the importance of maintaining Spanish did not improve fit significantly, $\chi2(3) = 3.67$, $p = 0.299$, nor did the model focusing on the importance of knowing Spanish, $\chi2(3) = 7.60$, $p = 0.055$. For the sake of completeness and parity, we fit one additional model for each predictor, again without the interaction term, to test the relationship between the listener background variable and intelligibility. These models confirmed the findings reported above: integrating the background variables related to the importance of maintaining and knowing Spanish did not significantly improve model fit, $\chi2(2) = 2.76$, $p = 0.251$ and $\chi2(2) = 3.05$, $p = 0.217$, respectively. Taken together, these results suggest that other than listeners' self-reported L1 or L2 status, the listener variables had little impact on intelligibility, and none of them moderated the relationship between guise type and intelligibility.

4.7 Discussion

Regarding our first research question, guise type did not affect intelligibility, despite the fact that the guise validation showed a descriptive trend toward a difference in expectations concerning heritage speakers' and learners' Spanish ability, a trend that reached significance for accentedness. This finding runs counter to Babel and Russell (2015), who reported that the presence of social information alone triggered differences in intelligibility. Listener sampling practices could account for these differences. Unlike previous research, our target population encompasses individuals who, on the whole, have much more experience with the accent in question than participants in previous studies. McGowan (2015), for example, compared groups of listeners based on prior experience with and exposure to the target non-native accent and found that increased experience with the accent significantly predicted performance on the intelligibility task. The participants in our task likely have had a substantial amount of experience with Spanish speakers in the U.S. who demonstrate varying degrees of fluency and accentedness; thus, listeners in our study likely performed well on the task due to high experience levels with "heritage accented" speech—a variety of Spanish being further investigated for systematic productive differences in phonetic and phonological features (Benmamoun et al., 2010). Increased exposure to a "heritage accent" would allow for less listening effort to be exerted because as listeners acquire more experience throughout their lifetime with a given accent, they become increasingly able to understand it. This improved understanding has also been documented for predictive language processing, where experienced listeners face fewer processing demands related to lexical access when processing non-native speech than unexperienced listeners (Porretta et al., 2020).

With respect to our second (exploratory) research question related to listener background variables, we did not find that the importance listeners ascribed to knowing and maintaining Spanish affected their ability to understand heritage Spanish, nor did it mediate the effect of guise type. This null finding is not especially surprising given that, as outlined above, listeners in this study were arguably accustomed to processing a range of heritage Spanish accents. In addition to any benefit listeners may have derived from their previous exposure to heritage Spanish, it is also important to consider the greater socio-political factors that may be at play. For example, listeners in this study may feel more connected to the specific social guises used here than those in previous research settings and may be aware of the difficulty that Latino families in the U.S. face to maintain their language. Given that language loss is common in Latino families, listeners' possible knowledge of the challenges that heritage speakers face may have mitigated any potential links between their personal beliefs and the guise manipulation. Thus, although listeners clearly valued maintaining Spanish and saw Spanish as central to the Latino identity, those attitudes did not appear to shape their ability to transcribe speech as robustly as is seen when native speakers perceive unfamiliar non-native speech (Rubin, 1992). At the same time, it is important to acknowledge that additional social factors related to language ideologies and attitudes within the Spanish-speaking community could also influence results. Spanish, as is the case with many languages, has a hierarchical social structure, such that certain Hispanic countries and varieties are regarded as more prestigious than others (Zentella, 2017). For the purposes of our study, we chose to include a variety of Spanish-speaking backgrounds when designing the social guises to encapsulate multiple heritage identities. However, we recognize that listeners' attitudes may have triggered a variation of biases that could have influenced results. Future studies may benefit from analyzing in further detail the relationship between

speaker and listener backgrounds and the potential effect this may have on intelligibility and overall speech processing.

In addition to the data presented here, we collected accentedness ratings and other subjective measures from the same listeners. A future direction is to directly compare these subjective measures to the objective measures reported in this manuscript. It is likely that subjective and objective measures may be differentially affected by social guise information. By comparing these subjective ratings with intelligibility scores, we can understand how each guise type may affect different aspects of speech perception and to what extent these perceptual factors interact with each other. Furthermore, future research would benefit from investigating the effect of social information on comprehensibility given that research has shown that high intelligibility does not entail the ability to comprehend the speech well (Munro & Derwing, 1995; Nagle & Huensch, 2020; Schmid & Yeni-Komshian, 1999). Comprehensibility differs from intelligibility in that intelligibility is a measure of actual understanding, irrespective of the effort involved, whereas comprehensibility reflects the individual's subjective evaluation of processing effort (Munro & Derwing, 1995). Research investigating the potential influences on comprehensibility could benefit heritage programs as enrollment in courses across the country becomes more common (Beaudrie, 2012). As described, the Spanish heritage community is inherently heterogenous, with a key difference being the variety of Spanish that speakers are exposed to in the home. Differences in home varieties combined with the influence of English make the heritage classroom a linguistically rich setting; however, this may have certain effects on speech processing other than intelligibility. For this reason, integrating comprehensibility into this line of research could help shed light on the overall impact of expectation and ideology on processing effort. To that point, one study showed that listeners are sensitive to social bias manipulation

when evaluating L2 speech, and that the strength and even the direction of that effect interact with listener background characteristics such as age (Taylor Reid et al., 2019). It is important to note that the failure to find an influence of guise on transcription ability in the current study may suggest that either intelligibility may be less susceptible to the influence of social information or that different populations of listeners may show different influences of social information during perception. Both possibilities should be investigated in detail in future work.

This chapter aligns with the broader dissertation by serving to follow-up on analyses in Chapter 3, which examines whether the role of education in the acquisition of the heritage language would lead to negative evaluations from heritage peers regarding group acceptance. Chapter 3 examines listener judgments regarding perceived accentedness and fluency as they relate to whether a speaker is perceived as "Latino" or not and how these judgements may shift when the individual is identified as either a heritage *speaker* or heritage *learner*. Building upon these perceptual evaluations, Chapter 4 then investigates whether the social biographical information can have trickle-down effects that can significantly impact how intelligible the speaker is perceived to be, in addition to potentially influencing how they are rated across the three established metrics. The next chapter, Chapter 5, extends the inquiry by focusing on specific production patterns as was done in Chapter 2. Results from Chapter 4 indicate that the social guises did not significantly impact intelligibility scores, as was originally hypothesized; therefore, Chapter 5 examines specific features of the speech stimuli that were used, determining if there were acoustic factors that influenced listener ratings and intelligibility independent of social factors.

CHAPTER 5: A COMPARATIVE ANALYSIS OF SPANISH HERITAGE SPEECH PRODUCTION: SCRIPTED VS. SPONTANEOUS SPEECH

5.1 Introduction

The current chapter aims to analyze Spanish produced by Spanish heritage speakers with particular interest in how production may vary based on speaking contexts. Due to societal pressures to assimilate and the role of educational systems in the US, the Spanish heritage community has long been susceptible to language shift and sometimes even complete loss of the heritage language (Zentella, 1995). As the language shifts in usage from the original immigrants in the family to first-generation and so on, there begins to surface productive differences that are being further investigated (Tran, 2010; Montrul, 2014; Pascual y Cabo, 2020). These differences stem from various sources related to acquisition environment and the type and amount of input heritage speakers receive in their life (Potowski, 2013). Throughout history, however, educational practices have promoted a reduction in usage of the heritage language to encourage acquisition of English (e.g., subtractive bilingualism), ultimately in an effort to increase the progress of loss. Other pedagogical practices (e.g., additive bilingualism) have attempted to embrace heritage languages but fall short when still proposing its inferior status (e.g., appropriateness-based models to be further discussed in the literature review) and imposing standard language ideologies around production (Leeman, 2005; Vogel & García, 2017). Therefore, when students enter the classroom, they are confronted with judgements regarding their language usage patterns and the usage patterns of their peers. However, newer approaches to pedagogy such as Critical Language Awareness (CLA) are being promoted as a way to allow

heritage speakers and learners to resist ideologies and attitudes that may accuse their Spanish of sounding poor or incomplete (Beaudrie et al., 2019).

Research on Spanish heritage language has often focused on better understanding the production of morphosyntactic features that are unique to heritage speech (Montrul & Perpiñán, 2011; Pascual y Cabo, 2015); however, more recent literature dedicated to better understanding Spanish heritage phonetics and phonology have begun to emerge in an effort to fill this gap in knowledge (Rao & Kuder, 2016; Rao & Ronquest, 2015). This growth in available research has led to the development of arguments in support of a 'heritage accent' in Spanish heritage speakers, yet the origin of this accent is still uncertain (Benmamoun et al., 2010). By having greater understanding of the development of potential differences of certain phonological features of the heritage variety/accent we can begin to incorporate the exposure of this to heritage learners – and even educators – as we embrace pedagogical environments that value the recognition and inclusion of variation, such as CLA (Leeman, 2005). Though this chapter is limited in scope given the small amount of audio data examined, we argue that it is equally valuable to educate listeners about the importance of recognizing individual differences in addition to the heritage accent. Also, given there is still limited understanding of the characteristics and features of this accent and what constitutes the key phonetic and phonological differences produced systematically within the heritage community, we see further value in acknowledging the legitimacy of individual differences.

In addition to the limited research available investigating the phonetic and phonological features of heritage Spanish, there is also limited research focused on production differences in spontaneous speech when compared to scripted speech for Spanish in particular. While speech

style has shown to demonstrate productive differences (Bradlow, 2002), studies focused on examining these differences in Spanish are more limited. For this reason, the current study will examine a small sample of Spanish heritage speakers borrowed from the ALLSSTAR corpus used in earlier chapters of this dissertation to analyze for potential differences in production of the Spanish spirants /β, ð, ɣ/ when examining spontaneous and scripted speech.

5.2 Background

5.2.1 Spanish as a heritage language

It is well established that the Latinx population in the United States represents the largest minoritized community in the country and the second largest Spanish speaking community in the world (Beaudrie & Fairclough, 2012). The presence that both the Spanish language and many Latinx cultures have in society has led to the growth in individuals that identify with the label heritage speaker. A heritage language is defined as a minoritized language in a given country, with which a speaker has an ancestral or historical connection, and a heritage speaker can be broadly defined as a speaker of a heritage language (Fishman, 2001). These definitions do, however, tend to vary by study given that individuals considered heritage in a particular setting, varies on the importance placed on linguistic factors like the ability to speak the heritage language (Carreira & Kagan, 2011). In addition, research has shown that second and third generation Latinx are emerging as the largest community within the overall US Latinx community furthering the variation that exists amongst upbringings and early language experiences. The birthrate amongst the Latinx population is continuing to grow to the point of outnumbering the immigration rate, therefore this increase in the Latinx population in the United States can be attributed to the increase in birthrate rather than an increase in immigration, further explaining the growth in second and third generation Latinx (Abdi, 2011; Potowski & Carreira,

2010). Both the increase in Latinx birthrates in the United States and the pressures US Latinx families feel to acquire English and assimilate to the monolingual ideals of society, has contributed to the generational language attrition these families undergo, often resulting in complete loss of the heritage language (Cashman, 2009; Silva-Corvalán, 2004). Therefore, the aforementioned variation that exists amongst definitions is representative of the inherent heterogeneity of this community both regarding their sociocultural backgrounds and their linguistic abilities.

5.2.2 Pedagogical Approaches to SHL

One of the factors leading to generational attrition of the heritage language is that the motivation and goal of bilingual education in the United States has been directly representative of the desire for minoritized families to adhere to English only expectations. Early educational practices were directly representative of this and followed a subtractive model that aims to completely replace familial and heritage cultural values (Baker, 2006) even going as far as to alienate students from their heritage culture in the process (Valenzuela, 1999). Following this approach, the designated role of the bilingual classroom is to facilitate acquisition of English, viewing Spanish as nothing more than a barrier to assimilation (San Miguel, 2004). Classroom and program environments that adopt subtractive models have zero regard for maintenance of the heritage language and are often so deeply entrenched with nativist ideals that educators in these spaces have expressed sentiments that maintenance of the familial language is not their responsibility nor their concern (Gainer & Larrotta, 2010; Randolph Jr., 2017). Subtractive methods in Spanish bilingual education have also been used to further support standard language ideology by devaluing nonstandard varieties of both Spanish and English, perpetuating the idea that there is a single correct way to speak a language and those that diverge from the standard are

in some way inferior (Randolph Jr., 2017). The devaluing of nonstandard varieties has detrimental impacts on heritage speakers sense of self and identity construction, triggering the formation of linguistic insecurities for those that fall within the heritage demographic. Not only do a large of majority of Spanish heritage speakers in the United States lack formal education in Spanish – which elevates the European variety of Castilian Spanish as the 'correct' variety – but they also bring with them influences of US Spanish and Spanglish (Zentella, 1995), as well as language skills that deviate from the native speaker model and reflect the acquisition of a contact variety of Spanish (Pascual y Cabo & Rothman, 2012). Thus, with regard to structured development and learning in a single language subtractive methods are intended to 'fix' and correct traces of any nonstandard 'vernacular variety of Spanish to the point of complete replacement and erasure, while simultaneously invalidating it and its speakers.

The field of Spanish as a heritage language has long contested subtractive-based models, and in response has introduced models based on additive methods, motivated by ideas of appropriateness. The goal of programs employing these models is to increase heritage learners' sociocultural awareness and expand their linguistic repertoire to include standard varieties of language, while maintaining their heritage language, variety and culture throughout the learning process. Appropriateness-based models were developed as a way to avoid complete replacement and elimination of nonstandard varieties and heritage languages and cultures by instead exposing heritage speakers to the idea that all language varieties are valid and legitimate in their own right, but some are 'more appropriate' for certain contexts and settings. The structure of these models gives off the impression that they embrace linguistic diversity by simply informing students that certain varieties are better suited for conversations and communication between family and friends, while the standard variety and/or dominant language is best when navigating

professional and academic settings (Leeman, 2005). Throughout the process of increasing learners' awareness of the various social expectations regarding context specific language usage these models actually result in reinforcing the subordination of nonstandard varieties and minoritized languages, and further cement the established cultural hierarchy that has always promoted the hegemonic status of English and heavily privileged those that constitute the Anglo white majority (Fairclough, 1992). Therefore, appropriateness-based models unfortunately – and despite their best intentions to introduce an alternative approach that no-longer views complete eradication of heritage and minoritized languages as the necessary outcome for English learners – still fall short in adequately addressing heritage learners needs by relegating their home language and even their home variety to an inferior status.

Similarly to classrooms and programs who have adopted subtractive-based approaches, those employing additive approaches still encounter the presence of negative attitudes which often stem from presence of administrators and teachers who lack sociolinguistic awareness and in turn unwittingly perpetuate their attitudes that don't always align with language maintenance or a value for multilingualism (Leeman & Serafini, 2016). In addition, the classroom is a host to students from different backgrounds and upbringings who come to the classroom with their own attitudes regarding language and a shared lack of awareness of the ways in which it has been tactically employed to influence status and power dynamics in the US socio-political landscape (Leeman, 2012). Thus, educational programs structured on appropriateness are not immune to the influence of language ideologies and can even further demonstrate the impervious nature they exude by perpetuating these beliefs in an environment claiming to embrace all varieties and languages as legitimate. In fact, they further preserve the deeply entrenched and institutionalized ideologies around language and power, resulting in the creation of yet another space where

heritage speakers are forced to acknowledge the supremacy of English and standard varieties, thereby being faced with the constant reminder of the lesser status held by both the Spanish language in the US and their home variety of Spanish in the larger Latinx community (Peace, 2024).

To combat this, Martínez (2003) and Leeman (2005) introduced Critical Language Awareness as an approach designed to increase students' awareness around the dynamics of language and power that have resulted in the institutionalization of ideologies devaluing minoritized languages and nonstandard varieties. The goal of this approach is to better understand how these concepts relate to ethnic identity and their ability to construct, perform and confidently embrace their own identity (Leeman & Serafini, 2016). Unlike appropriateness-based models, a CLA approach avoids the idea that certain varieties are more suited for certain social spaces and instead teaches students to value and appreciate nonstandard varieties and linguistic diversity in general (Beaudrie & Loza, 2020). Language development is built from their home variety or prior knowledge resulting in a restored sense of agency for the learner as they are finally able to see value in the linguistic knowledge they provide and begin to partake in acts of resistance towards the negative sentiments they have experienced (Beaudrie et al., 2021). In doing so, heritage learners can develop a stronger sense of linguistic security and ethnic identity as they navigate various social spaces. Though it is important to remember this approach is still relatively new regarding its implementation in heritage education and not yet fully embraced (Beaudrie et al., 2021), it does provide a hopeful trajectory for language education throughout the country. Studies have revealed the effectiveness of CLA-based practices in the classroom, reporting these interventions increased CLA in students and their overall tolerance towards variation (Hudgens Henderson, 2022), and has had a positive influence in students'

understanding of their own ethnic-racial identity (Gómez García, 2022). CLA approaches have even been found to improve levels of CLA in L2 learners demonstrating the benefits it can present to various learner backgrounds (Quan, 2020). This exposes an opportunity to include beginning-level heritage learners who experience similar identity struggles as anyone else in the community and then are also often excluded from heritage programs when they are unable to exhibit at least intermediate level placement skills (Carreira, 2004).

5.2.3 Attitudes and ideologies: their influence on HL maintenance and shift

The ubiquitous nature of both English only ideologies that formulate the core principles of subtractive methods, and standard language ideology which formulate the general foundation for additive methods, has led to the deep internalization of these sentiments making them a catalyst for the development of linguistic insecurity, and decreased self-esteem amongst heritage speakers (Peace, 2024). Continued disparagement of essentially every aspect of their language ability whether it be in their now dominant language English or in their heritage language Spanish, has been found to compound with experiences that belittle their Spanish ability due to limited proficiency and/or the presence of nonstandard, and have resulted in the many perceiving this as indicative of not possessing a command of either language, or what is referred to as languagelessness (Rosa, 2016). The treatment endured from heritage speakers when communicating in Spanish from both the Anglo community who does not value minoritized languages when used by minoritized populations, and those within the Latinx community who make accusations like, *hablas pocho* (you speak the vernacular) or "you're a *no sabo* kid" (US born Latinx who have limited Spanish proficiency), ultimately exacerbates the presence of previously mentioned feelings around linguistic insecurities and has led some to discontinue their study of the heritage language to avoid judgements in the academic environment regarding their

command and linguistic ability (Sánchez-Muñoz, 2016). Some have been known to inhibit their usage of Spanish even having been found to avoid advertising their linguistic ability at all costs, out of fear of ridicule for their nonstandard features (Ayala-Saracay, 2025). Finally, in many extreme cases, some learners decide to abandon their usage of the heritage language altogether. This decision usually arises after repeated experiences that ultimately drive some heritage speakers and learners to the discouraging realization that the US does not see a tangible benefit to maintaining the heritage language, nor is there a safe space for the heritage language outside of familial conversations (Schwarzer & Petron, 2005). However, not to be forgotten that these environments still invite ample opportunity for ridicule, albeit likely to be much less aggressive or targeted, the received ridicule from familial interactions is impactful in its own right and has been found to influence motivations behind enrollment in language learning courses. Not unlike the Anglo-speaking community, the nature of ideologies in the Spanish speaking community place immense value on standard language ideology, in turn promoting the acquisition and hegemony of Castilian Spanish throughout the Latinx diaspora at large. Meaning many minoritized families immigrate to the US having been predisposed to systemic ideologies that stem from the monopoly the Spanish government has on the administration of Spanish language education throughout the world (Zentella, 2017). Upon immigration to the US, these families are then further confronted with ideologies that not only denigrates the usage of their language altogether, but also degrades their variety of Spanish given its deviation from the textbook variety taught in formal language classes. These ideologies around the standard are so prolific that some native English, L2 learners of Spanish even going so far as to claim their Spanish should be regarded as better given that the formal environment in which they acquired Spanish afforded them the opportunity to learn the ‘correct’ variety (Leeman & Serafini, 2016).

5.2.2 Heritage speech production

The large representation in the United States from various different parts of the Latinx diaspora naturally, and not surprisingly, introduces a lot of variation into the heritage community. This variation develops not just from contact between English and Spanish, but also from contact between different varieties of American English and the many varieties of Spanish spoken in the United States. This contact occurs at both the between-language and within-language level, blurring the lines even further between what constitutes a linguistic boundary between languages and dialects when considering the language development of bilinguals, multilinguals and those that fall under the heritage category (Potowski, 2018). This unique acquisition environment provides the foundational basis for their acquisition process and influences heritage speakers' development, thus shaping the formation of a *heritage accent*.

Both the influence of contact between English and Spanish, and the contact of many different varieties of Spanish has led to clear distinctions in the acquisition process of heritage speakers and native monolingual speakers – who, when describing the prototypical profile of the idealized monolingual speaker, acquires their first and only language in an environment with consistent exposure to a version of the language that is devoid of nonstandard features (Rothman, 2007; Bonfiglio, 2010). Despite this, much of the research to date has often argued that heritage speakers develop a heritage accent *because of* incomplete acquisition and/or attrition of the heritage language, and have failed to recognize the social factors that, without a doubt, have played an impactful role in the acquisition process. Due to this oversight in the literature and the field in general, recent studies are beginning to acknowledge the undeniable significance that these differences in acquisition environment have had on molding the linguistic development of heritage speakers and learners. Increased recognition in the role of the language acquisition

environment has argued that heritage speakers actually do undergo 'complete acquisition' of a contact variety that simply put has been considered to diverge from that of the monolingual standard (Benmamoun et al., 2010, p. 6). This contact variety is a direct outcome of the shared influence of the heritage language and the dominant language, making it characteristic of the bilingual and bicultural environment in which heritage speakers are raised (Kupisch & Rothman, 2018). Therefore, recognizing the acquisition of a contact variety, entails the idea that using monolinguals to serve as exemplars for the expected speech production may be inappropriate given that heritage competence is unlikely to be identical to monolingual competence. In acknowledging these differences, the analyses in the remainder of this article which may use monolinguals as a comparison, aim to use this data as a basis for analysis, to highlight areas for potential investigation in the heritage language, rather than to highlight potential deficits in the heritage language.

Productive differences between spontaneous and scripted speech have been established and well-studied and these two different speech styles have been found to elicit both suprasegmental and segmental differences (Harmegnies & Poch-Olivé, 1992; Kim, 2020). However, as mentioned there is limited available research on these differences in Spanish more generally and heritage Spanish in particular, with the available studies often being limited to small participant sizes. This presents a clear gap in the research given that there is a notable preference towards formal and scripted laboratory elicitation of speech over the elicitation of spontaneously produced speech. While lab speech allows for initial phonetic and phonological analyses of a given language, it is not always able to provide the full understanding of certain phenomena. More generally speaking, spontaneous speech leads to heightened coarticulation and can even have notable effects on the overall vowel space in speech production. For example,

studies have found that conversational speech can lead to voicing assimilation of voiceless stops as a characteristic of hypoarticulated speech in Spanish (Torreira & Ernestus, 2011) as well as other trends that may not arise, or be less common, in scripted speech.

An additional problem presented by eliciting scripted laboratory recorded speech is that this may pose many issues for heritage speakers given that high literacy rates are not always common and may lead to disfluencies while reading (Ronquest & Rao, 2018), or have been found to introduce effects from orthography given that heritage speakers may not be able to distinguish pronunciations of 'b' and 'v', often as a result of influence from English that does not typically arise in spontaneous data (Rao, 2015). Efforts to fill this gap and address the need to investigate the effects of speech style have influenced studies aiming to understand not only these effects in native Spanish speech but also in heritage Spanish speech as well (Face, 2003; Rao, 2014).

A frequent starting point for analyses in the available studies taking this approach is to focus on the production of Spanish stops with particular focus placed on the voiced stops /b, d, g/. Literature in the field of Spanish linguistics has established the presence of spirantization in Spanish voiced stops causing them to change to fricatives or approximants in certain environments like between vowels (González, 2006; Hualde et al., 2011; Prado, 2011) with pure approximants being the most common realization for voiced stops cross-dialectically (Rao, 2015). Recent studies, however, have found that this weakening appears to be less consistent than previously thought with regard to speech produced spontaneously and conversationally. An initial examination of this weakening in spontaneous speech was investigated using a corpus containing data from seven different Spanish speaking countries and determined that the degree of weakening a stop undergoes can be influenced by factors like stress, word frequency, word

boundary, and whether the segment is part of a suffix – though these results were inconsistent across the three stop phonemes (Eddington, 2011).

As described above, heritage speakers tend to vary from native monolingual speakers in terms of production. While it is still difficult to identify the exact features that contribute to the heritage accent, studies still demonstrate that there are notable differences between heritage and native speech, often using these findings to judge and diminish heritage speakers for the way in which they speak Spanish (Leeman, 2012). These differences are being further examined through early studies on heritage speakers' production. Rao (2014) found many factors that lead to differences in the /b/ phoneme, like decreased exposure to Spanish (depending on the background of the individual), word boundary, stressed syllables and the reading task implemented, that all ultimately introduce and contribute to the variation in the realization of this particular stop phoneme, and all led to less target like productions. It was also found that in scripted speech /b/ and /d/ phonemes tend to be pronounced more native-like than /g/ in heritage Spanish speech, given that /g/ often contains less weakening than the other phonemes and the findings from the analyses were further supported with native-speaker perceptions (Au et al., 2002). Gaining an understanding of production patterns, however, is further complicated by Rao (2015)'s comparison of scripted and spontaneous speech that found that heritage speakers produced the /d/ and /g/ phonemes with more native-like realizations than /b/. Overall, the discrepancies in findings between studies highlight the plethora of variability within the heritage community and is often very difficult to fully account for and elucidate the struggles many researchers face in establishing a set of features for the heritage accent; however, these findings still provide a basis for the types of variation that may be present overall. A finding that appears to be consistent is that overall results are often linked to individual differences (Rao, 2015). This

supports the findings motivating further analyses in the current chapter that found that individual features of the speakers themselves appeared to be influencing participants' perceptions most.

5.3 The Current Study

Although we cannot yet ascertain the exact features or differences that form the heritage accent, it is still well known that within the Spanish speaking community many ideologies exist around the way in which heritage speakers produce language and communicate (Abdi, 2011; Leeman, 2012; Showstack, 2017). The exact features that listeners home in on, however, are more difficult to discern. Therefore, the current study aims to take a deeper analysis into the results from Chapter 3 to attempt to see if the heritage speakers used in our study exhibit similar or different phonological productions that may have influenced ratings.

The previous study aimed to better understand the social perceptions within the Latinx community that generally lead to the development of linguistic insecurities amongst heritage speakers, by determining the influence that different factors – like perceived nonnative accent when speaking Spanish and perceived level of fluency – have when participants receive limited visual information about the speaker and altered biographical information through a matched-guise methodology. Findings indicated that the guises were found to be minimally influential in driving participants' ratings, rather the most significant factor in determining the influence of a particular guise was actually dependent on particular features of the talkers' speech – as was found in Rao (2014). Two of the four talkers used for stimuli in the original study received higher ratings for their perceived 'Latino' identity when paired with guises that described them as having acquired Spanish at a very early age, while the other two talkers received higher ratings when paired with guises indicating that the majority of their speaking abilities were acquired later in life through school.

Given these intriguing findings, the current study will analyze the features of these particular talkers' speech to gain better insight into potential acoustic features that may be influential to others' perceptions. This study will focus on the production of the Spanish voiced stops /b, d, g/ in both the spontaneously produced speech and the speech produced from a scripted story. The research questions driving the study are outlined below.

1. Do the two talkers that received higher ratings for the guises indicating early acquisition of Spanish, pattern distinctively from other the other two talkers?
2. At first glance, does speech style appear to affect the degree of weakening present in the segment?

It is expected that given that two of the speakers received higher ratings for the speaker guises, they will also produce more instantiations of a pure approximant when /b, d, g/ are between vowels; however, we expect speech style to have an influence on production. Low literacy rates may have presented issues for the heritage speakers; therefore, we expect greater productions of tense approximants in the scripted data than the spontaneous – this hypothesis is still difficult to support, however, due to the limited information about the talkers' educational background in Spanish. Given the certain limitations to the current study – like the number of talkers – the current investigation will be a preliminary look into the data available in the corpus. It is understood that this brings challenges to validity, however, the goal of analysis is to inform previous results rather than to make large claims or generalizations regarding Spanish heritage speech production. The claims being made will be specific to the individual talkers given that many studies already indicate that individual differences appear to be significant.

5.4 Methods

5.4.1 Materials

The following analysis will use the same audio files from the original perception study discussed in Chapter 3. The audio files are from the publicly available Archive of L1 and L2 Scripted and Spontaneous Transcripts and Audio Recordings Corpus (ALLSSTAR) developed at Northwestern University (Bradlow, n.d.). The corpus consists of audio recordings from Spanish Heritage speakers ages 18-22. All talkers were born in the United States but acquired Spanish at birth as their first language and English as their second language. Talkers in the corpus acquired English between the ages of 5 and 8 years old and underwent a shift in language dominance throughout their lives having gone from 100% usage of Spanish in the home during early childhood to less than 20% after early childhood. None of the eleven talkers reported taking Spanish classes during high school and there was no information provided regarding language courses taken during higher education (see Blasingame 2018 for a detailed description of the Spanish Heritage talker backgrounds from the ALLSSTAR corpus). Four of the talkers from the corpus were selected for the perception study with two of them being male and two being female; therefore, these will be the talkers analyzed in the current study.

The audios being analyzed were chosen from a selection of different scripted stories and spontaneous narrations that the talkers were asked to record. The scripted materials read for the stimuli used were from the story *The North Wind and the Sun.* The spontaneous speech was recorded by providing the talkers with a comic series of images in which they were asked to narrate a comic of two pigs and form a story (See APPENDIX B).

5.4.2 Analysis

Following similar methodological approach to analysis as Rao (2015), each of the audio files were analyzed in Praat and marked as either a pure approximant, tense approximant, or stop. To avoid potential differences that may be introduced due to the large amounts of variation and potential influence of home varieties on heritage speaker production, analysis was focused to the environments which are most salient cross-dialectically for weakening of the voiced stops. For this reason, words containing either /b, d, g/, between vowels either within a single word or across word boundaries were included in analyses – therefore, word initial sounds were included if the preceding sound was a vowel. The spectrograms in Figures 6 and 7 below provide a visualization of the acoustic differences between pure and tense approximants.

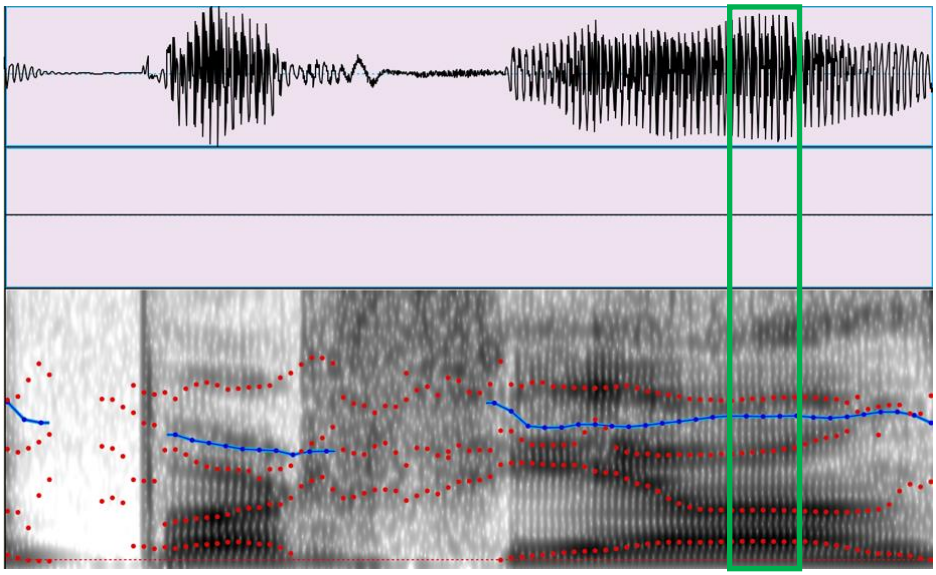


Figure 6 *A spectrogram of the word porfiaban produced with a* **pure** *approximant.*

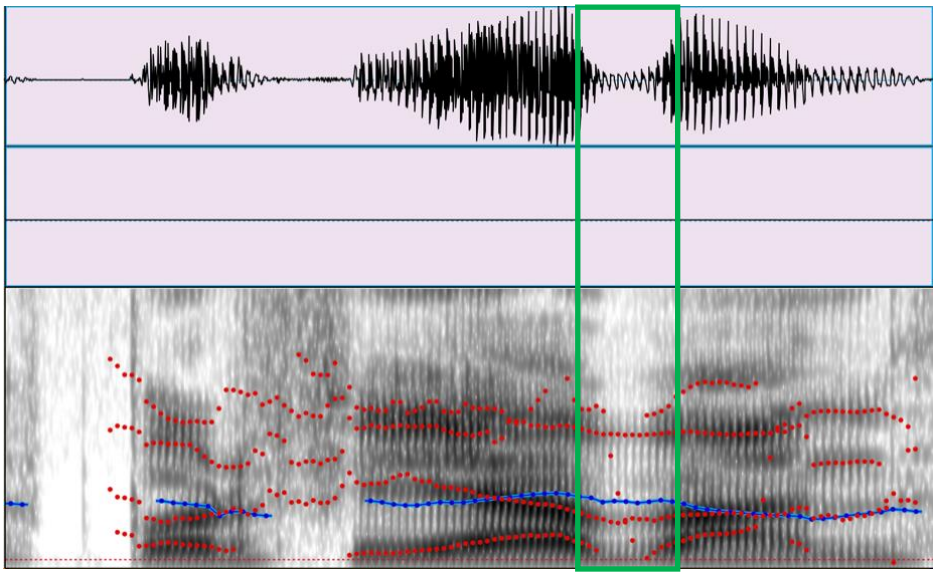


Figure 7 A spectrogram of the word *porfiaban* produced with a **tense** *approximant*.

To run statistics on the data, talkers were grouped by either speaker or learner guise labels to differentiate between the talkers that were rated more highly for early acquisition guises while the learner label represents the talkers that were rater more highly for guises indicating later productive acquisition of Spanish. Data were also grouped based on whether or not the token was produced in the scripted audio or the spontaneous audio, as well as by the phoneme and manner of articulation. Descriptive statistics were run on the data to gain an initial visualization of the data. After running descriptive statistics, an analysis of variance was run to determine if, despite the limited data, a significant effect still arises of guise type and speech style.

5.5 Results

An initial look at the data in Table 4 indicates that on average, there are slightly more productions of the tense approximants overall, however, this difference is very small with the

average production being 2.33 and the average pure approximant production being 2.12. The maximum production for tense approximants only outnumbered the production of pure approximants by one token and the production of stops was very limited. Only one talker produced stops between vowels in two of the tokens between both scripted and spontaneous speech.

Table 4 *Descriptive statistics for manner of articulation across all phonemes, talkers and speech style*

	Mean	SD	Median	Min	Max
Pure approximant	2.12	1.87	2	0	6
Tense approximant	2.33	2.24	2	0	7
Stop	0.08	0.28	0	0	1

The boxplots in Figure 8 below demonstrate that across the four talkers used for the stimuli, each of the voiced stops were more likely to become pure approximants when the talkers were speaking spontaneously and tense approximants when they were producing scripted speech. Despite this, an analysis of variance shows that for pure approximants, speech style was not found to be significantly influential in determining manner of production at alpha = 0.05, $F(1, 22) = 0.011$, $p = 0.916$. The production of tense approximants were also insignificant when controlling for speech style at an alpha = 0.05, $F(1, 22) = 1.677$, $p = 0.209$.

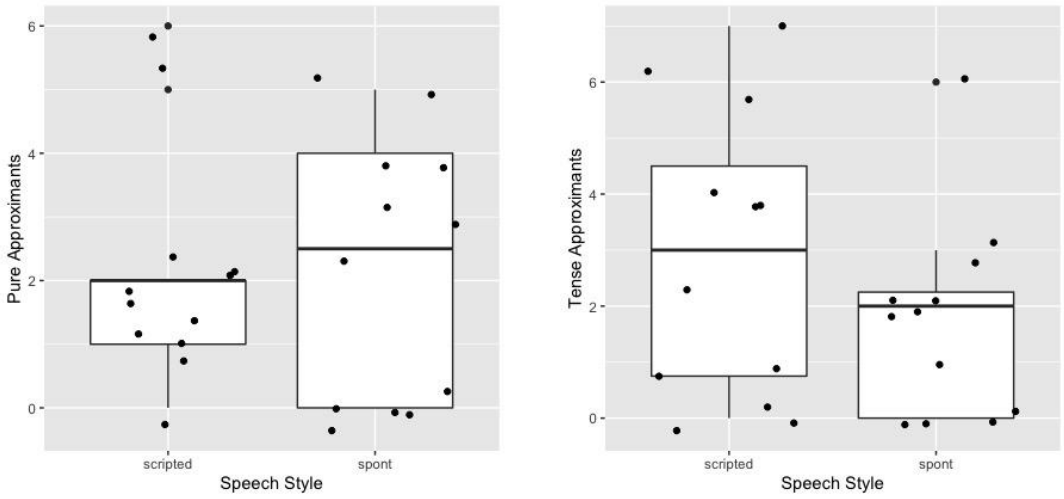


Figure 8 *Pure approximants and tense approximants shown by speech style*

Analyses of variance were also run to determine the influence that guise type has on productions of pure and tense approximants. Ultimately, results showed that neither pure nor tense approximants were significantly more common for either the speaker or learner guise. Despite these findings, the talkers that aligned with the speaker guises notably produced more pure approximants for the spontaneous speech style and more tense approximants for the scripted speech style while learner guises showed more overlap for pure approximant production across speech styles but still a preference for tense approximants in scripted speech.

When taking a closer look at the individual phonemes, we see that the /b/ phoneme appears to have greater productions of pure approximants across all participants in spontaneous speech while scripted speech appears to lead to greater productions of tense approximants. The /d/ phoneme showed a similar trend with much more drastic differences between speech styles for tense approximants. The /g/ phoneme, unlike the other two actually shows increased productions of tense approximants in spontaneous speech production rather than pure approximants. See the boxplots in Figure 9 below.

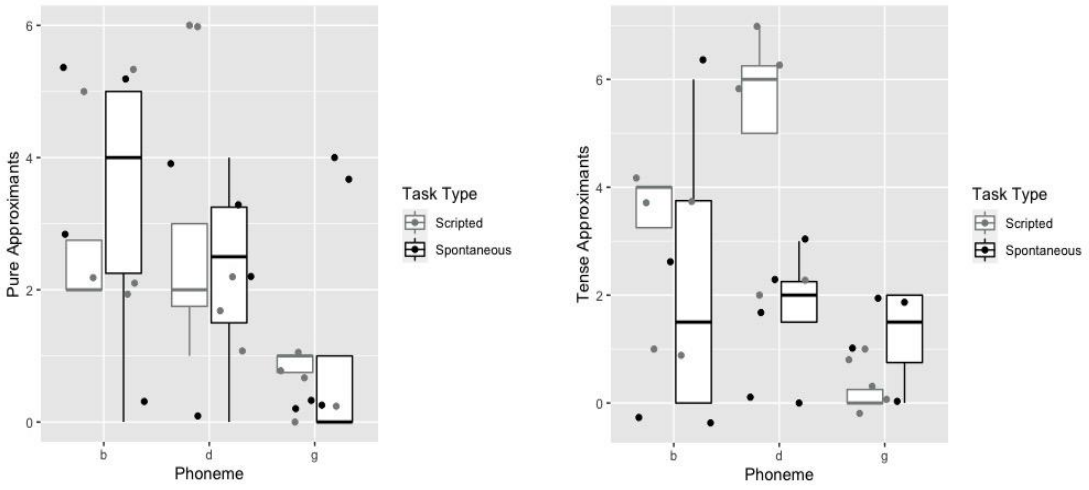


Figure 9 *Pure and tense approximants grouped by phoneme type and task type.*

5.6 Discussion and Conclusion

The results from the current study, though they do not render significant findings, still provide a stronger understanding of the previous findings associated with perceptions around the way each talker speaks Spanish. As expected, the speaker guises demonstrated a notable difference between manner of production of the voiced stops between each speech style despite the results not being significant. The talkers under the speaker guises were more likely to produce pure approximants when speaking spontaneously and tense approximants when producing scripted speech. This discrepancy in results could demonstrate potential literacy issues amongst the heritage speakers given that each talker showed some difficulty with the reading task, ultimately causing them to pronounce some sounds in a way that may not be completely natural to them. These results potentially explain why these two talkers may have received higher scores for relationship to the 'Latino' identity and for sounding 'native-like' when paired with the speaker guise rather than the learner guise. Given that cross-dialectically, pure approximants are the most common realization for voiced stops between vowels in Spanish (Rao, 2015), the two

talkers may have aligned more closely to native, monolingual patterns in terms of spontaneous speech production. The talkers linked to the learner guises show similar patterns to those linked to speaker guises, however, the differences are less distinct and there appears to be more overlap between pure and tense approximant usage.

These findings indicate that there should be further investigation into the influence of approximant production on social perceptions. Given that previous studies have already indicated that native speaker listeners are sensitive to stop productions and can determine which sound more 'native-like' or not (González-Bueno, 1997; Schoonmaker-Gates, 2015), the differences in approximant production could be driving some of the aforementioned ideologies that heritage speakers encounter within the Spanish speaking community regarding the way in which they speak. In addition, these results also invite further research to investigate the effects that frequency may play into the degree of weakening on a given segment. The talkers in this study all acquired Spanish as their first language, and although they underwent a shift in language dominance, they still had exposure to the language outside of academic settings. For this reason, certain high frequency words may have developed strong mental representations or exemplars in their mind, leading them to produce the more 'native-like' pure approximant for certain high frequency words, despite potentially having a tendency to favor tense approximant production overall. Therefore, considering word frequency may provide valuable insight into some of the discrepancies between production and social perceptions elucidated in this chapter.

CHAPTER 6: GENERAL DISCUSSION AND CONCLUSIONS

6.1 Summary of findings

The purpose of this dissertation is to provide a deeper understanding of the nuances involved in how identity is assigned to heritage speakers and learners of Spanish in the United States by their peers. This evaluative process is heavily influenced by the ideologies and social expectations within the Latinx community regarding proficiency, linguistic ability in their heritage language, and "sounding native". In addition, this dissertation addresses the expectations placed on the Spanish heritage community around both grammatical and phonological language production in social, research and educational spaces, thereby challenging previous misconceptions that frame heritage language production as inherently underdeveloped or deficient. Consequently, I have aimed to examine how these expectations around production, fluency, and accentedness shape ingroup perceptions among heritage speakers, specifically exploring how linguistic performance governs whether an individual is recognized as legitimately Latinx by their peers. Particular attention is paid to this process as it pertains to Latinx community members within the second generation and later, who frequently encompass beginner level heritage learners, given that they are the portion of the community most susceptible to language loss (Carreira, 2004). I focus on this portion of the community given that expectations around Spanish proficiency remain so insistent and powerful even as these groups who are more susceptible to language loss begin to outnumber the earlier generations who typically exhibit the linguistic skills that community members, researchers, and educators anticipate (Rothman et al., 2016).

This category within the heritage community is disproportionately impacted by attitudes and ideologies present in the Latinx community, given that language proficiency and competency

are often employed as a gatekeeping method – thus preventing community acceptance and inclusion in academic spaces (Ayala-Saracay, 2025). Many individuals of Latinx heritage experience restrictions on their social validation after enduring persistent accusations stating they are 'not Latino enough', not 'truly Latino', or even not 'truly' a heritage speaker. Feeling unaccepted by both their community and in academia, frequently complicates public validation of their Latinx identity, particularly when they lack a key feature associated with the performative aspect of speaking Spanish that is so deeply rooted in their heritage culture. Via the extrapolation of data examining actual production patterns and within-community listener perceptions, I emphasize the clear need to recognize the legitimacy of heritage varieties of Spanish. Ultimately, I underscore the legitimacy of those who may identify culturally with Spanish but present very beginner-level abilities or who did not acquire linguistic abilities until much later in life.

Further, this dissertation provides a deeper understanding of how 'accentedness' can interfere with how an individual is perceived by others in the heritage community. Researchers who investigate the phonetic and phonological systems within the heritage community have attempted to better understand the unique qualities of this system that shape the marked nature of the *heritage accent*, which can trigger peer judgements around language usage (Rao & Ronquest, 2015). To contextualize these dynamics, I describe the variation present within these productive patterns, framing them as potential barriers to community acceptance that can heavily impact whether speakers and learners are recognized as legitimate members of the Latinx community. This is complicated further when considering later generation Latinx individuals who enroll in Spanish classes to acquire their heritage language or to improve their established skills in an effort to connect more with their culture and thwart accusations questioning their authenticity.

This group of the heritage community presents us with the question of whether eventual late in life acquisition of the heritage language will grant community acceptance even though the majority of members within this category of heritage are unlikely to produce "native sounding" phonology once they do acquire the language or acquire greater command (Carreira, 2004). Essentially, while my data is descriptive of these production differences, I utilize this variation to evaluate the preexisting ideology around language ability, exploring how these patterns relate to an individual's external validation by the community at large. For example, those who speak the heritage variety are often judged for "improper" Spanish or an inconsistent presence of "native" features. Therefore, this dissertation considers how efforts to acquire the language later in life can still intersect with judgements by their heritage peers as they relate to who is recognized as being 'truly' Latinx due to existing perceptions around accentedness.

6.1.1 Chapter 2 Findings

Chapter 2 addresses the importance of validating heritage varieties and the features that characterize US Spanish by challenging the previous work claiming that heritage linguistic systems are a result of incomplete acquisition (Polinsky, 2008; Montrul & Perpiñán, 2011; Montrul & Foote, 2014). To do so, I examined spontaneously produced speech by monolingual English speakers, Spanish heritage speakers and L1 Spanish – L2 English speakers across two task types: a storytelling task where participants narrated comic strips and an open-ended Question-and-Answer task. I elected to measure production in both English and Spanish to highlight that the language development within the heritage community is not indicative of a systemic deficit, nor should we continue to regard linguistic differences as anything other than what they are, a product of different realities and experiences (Pascual y Cabo & Rothman, 2012).

While the data in this chapter is descriptive in nature, I position these findings to argue that traditional assumptions and conclusions around linguistic deficiency stem from inappropriate research designs within the field. Lexical diversity was used as the mechanism to measure production due to its ability to estimate speaker proficiency, given that it is influenced by the amount of input listeners receive and how this input fluctuates over time (Elabdali et al., 2022; Jacobson & Walden, 2013) – two factors that are highly variable and characteristic of the heritage experience. By utilizing these metrics descriptively, this dissertation critiques the flawed methodological frameworks that have inappropriately compared heritage speakers to monolinguals rather than evaluating them independently. The indices adopted in this study to measure lexical diversity include the Measure for Textual Lexical Diversity (MTLD 92) and Moving Average Token Type Ratio (MATTR 11). These indices were originally developed for written speech, however, have been adapted to adequately measure transcribed oral productions as well (Kyle et al., 2024).

The results from the English data in this study show that overall the monolingual English speakers produced the most speech across task types, followed by the Spanish heritage group (who identify as English dominant), and the group who produced the least amount of speech in English was the L1 Spanish group. When looking at lexical diversity scores for MTLD 92, the English group yielded the highest scores across both task types, while the heritage group had greater scores than the Spanish group for the storytelling task but not for the Question-and-Answer task. Despite these scores, linear-mixed effect models revealed that adding 'speaker group' as an independent variable to the model did not significantly impact results for MTLD 92. Scores for MATTR 11 show that the English group and the heritage group yielded higher scores than the Spanish group across both task types; however, there was a fluctuation between the

English group and the heritage group with the heritage group rendering higher MATTR 11 scores for the Question-and-Answer task but not for the storytelling task. Unlike MTLD 92, adding speaker group as a dependent variable to the model was found to significantly impact results. To examine lexical diversity scores in Spanish, we compared the Spanish heritage group and the L1 Spanish group. Results show that across both task types the Spanish group produced more speech overall than the heritage group and on average yielded greater scores for MTLD 92 and MATTR 11. Adding speaker group as a dependent variable to the model was found to significantly impact results for both indices, MTLD 92 and MATTR 11.

I interpret these findings to mean that differences in heritage production should not be perceived as faults, nor do they insinuate that heritage speakers lack adequate command of the languages that form their linguistic repertoire – in this case English and Spanish (Pearson & Fernandez, 1994; Wiese et al., 2022). The inconsistent differences between monolingual English lexical diversity scores and scores from Spanish heritage participants actually push back against the idea that heritage speakers lack adequate skills in both languages. The absence of a clear discrepancy demonstrates that monolingual English speakers' proficiency is not well above those in the heritage group. In addition, the significant differences between the SHS group and the L1 Spanish participants do not indicate that their Spanish is inadequate or deficient when considering the differences in acquisition environment and *type* of input these two groups receive in Spanish. Overall, the heritage speakers in this study still produced substantially diverse speech when considering scores for both MTLD 92 and MATTR 11, and none of the reported differences imply an inability to communicate or be understood in either Spanish or English. The heritage participants were perfectly capable of conveying concepts effectively in both languages. This then challenges us as researchers to better understand our goals and what we are actually

targeting if effective communication is not deemed sufficient for language proficiency standards and acceptance. Meaning, if we continue to focus only on how production compares to monolingual standards our goals for analysis become unclear.

The tendency to view the heritage variety as deficient and lesser, if even willing to consider it a legitimate variety at all, ignores the social and cultural influences that have long been known to impact language shift and change (Bhatt, 1995; Kachru, 2017). Bilingual and heritage speaker experiences and linguistic development do not mirror those of monolinguals, making the validity of direct comparisons questionable. Heritage Spanish development is influenced by multiple linguistic sources, including but not limited to the variety of Spanish spoken by the family, the Spanish they may have been exposed to in the community, and features of English (Rothman et al., 2016). Furthermore, this tendency underscores the flawed nature in how we have approached researching and understanding bilingualism in general, and thus heritage varieties.

The bilingual linguistic system represents a single repertoire that fluidly switches and navigates between named language categories (Li, 2018). Despite this, bilingual speakers and heritage speakers have been expected to develop two distinct and complete linguistic systems akin to the systems developed by monolingual speakers. This has been recognized as the 'monolingual view of bilingualism' (Grosjean, 1989) which relies on the idea of complete *balance* between the two languages and recognizes *true* bilinguals as individuals who are equally and fully fluent in each language (Bloomfield, 1933; Thiéry, 1978). However, true balance is unlikely and rare, and researchers have long been warned about making invalid comparisons in this way. The elected ignorance to the rarity of the ideal balanced bilingual has led some research to even go as far as adopting terminology such as *semi-lingual* (Zentella, 1995; Zentella, 2017)

to describe heritage communities. This misnomer completely ignores the abilities Spanish heritage speakers have in English and diminishes their heritage variety to the status of *nonlanguage*, insinuating that bilinguals and heritage speakers of Spanish in the US are incompetent members of society that do not possess adequate knowledge of any language at all (Rosa, 2016). The willingness of researchers and educators to completely delegitimize one's linguistic ability altogether – before even considering either the social reality or the cognitive processing that may influence linguistic differences – demonstrates just how difficult it can be to dismantle deeply ingrained ideologies around standard varieties of language that elevate the profile of native monolinguals at the expense of other backgrounds of speakers.

Instead, differences in bilingual production should be expected when compared to native monolingual production, or even L2 learner production. The observed differences in lexical diversity scores can likely be attributed to factors such as the domain specific nature of heritage language acquisition rather than deficiencies. Heritage speakers in the United States often develop skills in their heritage language as they relate to more familial and personal topics, and language skills in English as they relate to academic and professional contexts (Grosjean, 2016). Therefore, greater productive vocabulary for English is often found in formal social domains, while greater productive vocabulary of Spanish is concentrated to informal domains (Treffers-Daller, 2019).

There is, then, a glaring need to focus on the 'differential needs' and 'different social functions' each language in the bilingual repertoire serves, given that the two languages tend to distribute across social domains such that some domains are covered by one language, some domains are covered by the other language, and some domains may even be covered by both languages (Fishman, 1965; Grosjean, 2016). This distribution has been referred to as the

Complimentary Principle (Grosjean, 2016) which views bilingual communicative competence as equal but different from monolinguals and is succinctly described as, "Different aspects of life often require different languages" (p. 81). Studies have further argued that while single-language vocabularies for bilingual speakers may be smaller than those of monolingual speakers, the total vocabulary count when considering both languages is actually not different from monolinguals (Pearson & Fernández, 1994). Even knowing this, data collection methods and models used to measure bilingual linguistic ability in a monolingual mode are still common practice. These methods require heritage speakers to completely deactivate a given language avoiding any code-switching or lexical borrowings, and any traces of the additional language are thought to be signs of transfer or interference that require constraint to avoid (Weinreich et al., 1968; Grosjean, 2012). However, these traces of the additional language take place because complete deactivation is very rare and mostly impossible (Paradis, 1987), nor should it be expected. For the most part, both languages within the bilingual repertoire are *always* active (Grosjean, 2012).

This differs immensely from the monolingual English group in this study, who navigate every aspect of their daily life in one language and therefore do not experience the same challenges as the heritage group, who had to constrain their linguistic system when completing the tasks asked of them. Although the L1 Spanish speakers in the corpus used for this study are bilingual like the heritage speakers, the significant differences found between lexical diversity scores are likely because they did not acquire English until adulthood. Consequently, they are not constrained in the same ways that heritage speakers are in their Spanish usage. They are less likely to experience the domain specific restrictions that limit their Spanish to certain communicative contexts, similar to the monolingual English group.

Despite this, researchers still opt for direct comparisons between monolingual communities and heritage speakers using methods that are inadequate and restrictive. Given these flawed baselines, it is not surprising that, historically speaking, heritage Spanish has been viewed as deficient and incomplete. Heritage speakers continue to be compared inappropriately to populations and communities of speakers and the continuation of these practices will ultimately have detrimental effects to heritage language identity, and pedagogy. We as educators and researchers must ensure we are establishing and employing appropriate empirical paradigms when it comes to studying and describing heritage language. If we ourselves can't avoid making these mistakes in comparisons and overgeneralizations, then how can we expect the general population and the overall Latinx community – who typically lack overt sociolinguistic awareness – to be accepting of linguistic diversity and the experiences of heritage speakers? Therefore, if we can discontinue the mindset that any observed divergence from the native monolingual archetype represents errors in need of fixing (Seijas & Spino, 2023), we may begin to see the bigger picture and gain a greater understanding of the bilingual linguistic system as it exists in and of itself.

6.1.2 Chapter 3 Findings

Analyses in Chapter 3 shift focus from measuring production and challenging the research community to be mindful in the ways we approach our topics of inquiry, to evaluating within community perceptions that target the interaction between social identity, fluency, and accentedness. I sought to understand how deeply entrenched the ideology is that claims cultural identity is dependent on linguistic ability, describing the relationship between perceived Spanish proficiency and community acceptance. By measuring these evaluative behaviors, this study sought to determine if obtaining fluency in Spanish is perceived as the primary component

influencing group acceptance, or if a listener's awareness of an individual's language background is a salient variable that community members consider and evaluate. To target these perceptual questions, I employed a matched-guise methodology in which Spanish heritage participants were presented with audio stimuli from four different Spanish heritage speakers – who will now be referred to as heritage talkers to differentiate between them and the participants in the study. The talkers used for this study were all between the ages of 18 and 22 and acquired Spanish as their first language and English as their second language between the ages of 5 and 8. They now claim to be English dominant and did not report having enrolled in Spanish language courses.

Audio from each of the talkers was paired with a brief biography describing social background information that I prepared not necessarily related to the actual background of the talker themself. The same talker was presented twice, and the social background information was manipulated to ensure that each talker was presented once with a heritage *speaker* guise and once with a heritage *learner* guise. The heritage *speaker* guises in the study described the typical anticipated trend of acquiring Spanish as a first language and undergoing a shift in language dominance upon enrolling in elementary school. The *learner* guises in this study, however, were intentionally narrowed to represent the most stigmatized group within the heritage community who suffer the most scrutiny surrounding their status in the heritage community from both a social and educational perspective. These guises described individuals who enrolled in language courses having very limited or no command at all in Spanish, therefore, they may not depict the typical advantages expected of those within the heritage community (Van Deusen-Scholl, 2003). In fact, their acquisition process is expected to be more similar to that of L2 learners and they potentially would not qualify linguistically for a heritage program (Carreira, 2004; Beaudrie & Ducar, 2005).

After reading each biography and listening to the audio stimuli, participants were asked to rate on a Likert scale 1-9 how fluent they perceived the speaker to be (1 being non-native, 9 being native/advanced), how accented they perceived the speaker's Spanish to be (1 being 'sounds non-native', 9 being 'sounds native') and then how they perceived the speaker's relationship to the Hispanic/Latino identity (1 being not Hispanic/Latino or no relationship, 9 being of full Hispanic/Latino descent). Analyses focused on how influential the guise type itself was in driving ratings, how influential the audio stimulus was in driving ratings, and whether ratings around accentedness and fluency would influence how participants rated the identity question.

The learner guises were rated higher than the speakers guises for fluency and as sounding more native-like, in addition to also receiving higher ratings around their perceived Latino identity. There was a main effect of both question type and talker, however, the influence of the guise itself was not found to be significant nor was there a significant interaction between guise and question. Upon deeper dive into these effects, I found that talkers 100 and 107 received their highest ratings across questions when paired with the speakers guises while talkers 101 and 105 received their highest ratings for the learner guises. An additional unexpected finding revealed that participants were willing to rate the talkers as still having a strong relationship with the 'Latino' identity even when they gave the talkers low ratings on the fluency and accentedness questions. This is likely explained by the significant interaction between talker and guise type which revealed that certain talkers were indeed rated differently depending on guise type; however, again, I did not find that the guise in and of itself was as influential as originally anticipated.

These findings point to the potential for shifts in attitudes and ideologies that have relied so heavily on Spanish proficiency. Specifically, within this group of participants, both heritage *speaker* and heritage *learner* profiles were considered to have strong associations to the 'Latino' identity when judged by their peers. This descriptive pattern can be interpreted as a result of the change we are seeing in the demographic of US Latinx, suggesting that ideologies within this community may be shifting along with the changes in demographics given that US born Latinx have been found to place less importance on proficiency than Latinos who have immigrated to the U.S. (Potowski, 2012). The Latinx community now includes more second, third, and later generation individuals than first generation and immigrant Latinx, and the backgrounds described in this study were meant to reflect that shift. Furthermore, although not measured by the data, it is plausible that the participants in this study may have found some sense of relatability to these identities or at least to some extent an awareness that complete fluency and 'native' sounding Spanish is less common in general amongst the US Latinx population. If these data do not reflect a widespread shift in systemic ideologies, they may instead signal a tentative shift in more individualistic language attitudes, where those in the Spanish heritage community choose not to adopt these restrictive beliefs on a personal level. Because these ideologies remain ubiquitous in the broader Latinx community (Otheguy et al., 2007), it is unlikely that they are disappearing completely – especially given that immigration has not halted, nor is there an expectation that it will even as it slows (Funk & Hugo Lopez, 2022). Nevertheless, these findings suggest that at an individual level there may be pockets of understanding and acceptance that could hypothetically translate into larger social domains over time.

Even if heritage speakers are unable to overtly express ideas like generational language loss, or power dynamics related to this process, they are certainly capable of observing the reality

of the circumstances they and their peers encounter. This may have resulted in greater empathy and compassion for the different talkers and guises they were exposed to making them hesitant to completely deem learner profiles as not being 'Latino enough'. Upon learning about the backgrounds of the talkers, listeners may not want to maintain the same negative sentiments that they themselves have experienced from also being heritage speakers of Spanish. Also, a deeper dive into within community perceptions around the act of enrolling in Spanish classes would help to determine if the desire to reconnect with the Latinx heritage through language acquisition is viewed as admirable and therefore worthy of receiving high ratings on the identity question. Research on attitudes and motivations in the heritage language classroom has found that heritage language learners often describe integrative motivating factors as a reason for enrolling in Spanish classes, which simply put is the motivation to feel more accepted by their community (Ducar, 2012); however, little is known about how this is perceived by others in the community.

Previous studies may dismiss this idea being proposed given that research in classroom spaces have found that heritage speakers and learners often do make judgements regarding their peers' production, sometimes very harsh ones that reinforce national language ideologies (Leeman, 2012) and have even led to the rejection of certain heritage students as not being 'one of them' (Abdi, 2011). However, it should be noted that these sentiments are being expressed in an educational environment that is already highly policed regarding language and a place that both covertly and overtly perpetuates exclusionary ideologies. It may be the case that in these environments, heritage speakers convey more hostile and nationalistic ideologies than they might in a setting with less expectation around standardization and correctness. This was a formal research project, but the participants were simply asked to share their opinion about the audio and were made to understand there was not a right or wrong way to respond. This low-stakes

environment may have allowed for more lenient perceptual judgements, but this is indeed just speculation.

Although we can't ignore the role that expectations play in perceptual processing, which is why we purport that the audio stimulus did drive ratings for certain speakers. This is to say that there may have been confusion when a listener heard audio they perceived as 'sounding native' and rated high for fluency being paired with a *learner* guise type. Therefore, it might be the case that rather than judging the guise alone, the audio stimulus simply did not match the profile described, in listeners' opinions, making the ratings difficult for them to assign in general. Studies examining qualitative reports from heritage speakers have found that individuals within the Latinx community and the US in general do share social expectations related to someone's appearance and their Spanish ability. For example, Peace (2024) includes excerpts from heritage learners in university level courses in which one is reported to share, "También sobre cómo ser marron, la gente asume que sabes español. Yo he tenido miradas feas de la gente cuando digo, 'No hablo español.' [Also, being brown, people assume that you speak Spanish. I've had ugly looks from people when I've said, 'I don't speak Spanish.']" (p.71). While our study did not include visual information, there still may be potential that mismatch in the audio and the social information led to listeners' uncertainties or hesitancies, which is why we then focused on the role of social expectations and their influence on the perception of audio stimuli.

6.1.3 Chapter 4 Findings

Chapter 4 expands upon findings in Chapter 3 by targeting the question of how social information may impact speech intelligibility when the same interventions were employed. To verify our initial hypotheses stating that within community perceptions would differ between the heritage *speaker* guises and the heritage *learner* guises, I also included a component focusing on

listeners' expectations around the same questions we asked in Chapter 3 - fluency, accentedness, and relationship to the 'Latino' identity. This was included to confirm that the social guises alone do in fact influence and trigger certain attitudes independent of visual or audio information. Descriptive statistics confirmed this hypothesis and participants in the guise validation study showed a trend towards expecting the *speaker* guises to be more fluent, more native-sounding, and as 'more Latino' than the *learner* guises. However, the difference in ratings between guise types were only found to be significant for the accentedness ratings.

As mentioned, the next step in this study aimed to examine whether our guises may then impact listener intelligibility as has been previously reported in studies examining the effect of social information on native English listeners' perceptions of foreign accented English (Babel & Russel, 2015). To measure this a matched-guise methodology was again employed using the same talkers as in Chapter 3, and again presenting each talker once with a *speaker* guise and once with a *learner* guise. Participants were first asked to read a brief biography of the talker they were about to hear before beginning with the transcription task. The transcription task presented participants with 15 sentences for each guise, and intelligibility was measured as the number of correct words participants were able to transcribe in each sentence. After completing the transcription task they then filled out their ratings before proceeding to the next talker biography.

While heritage *learner* profiles showed a trend towards greater expectations around being less fluent, more accented, and less 'Latino' than *speaker* profiles, the average intelligibility scores were surprisingly slightly higher for the *learner* guises. However, this difference was not found to be significant, thus can be interpreted as not resulting in decreased intelligibility. This is a trend that diverges from findings claiming social information alone can lead to differences in

intelligibility even when audio stimuli are the same (Rubin, 1992). To discern whether listeners' demographic variables may have impacted results we looked further into three social background variables: (1) whether listeners identified Spanish as their L1 or L2, (2) the extent to which they felt it was important for 'Latinos' to speak and maintain their Spanish, and (3) the extent to which they felt that someone who identifies as 'Latino' should know Spanish. Findings indicated that listeners who reported Spanish as an L1 were significantly more accurate in transcribing speech, although the variables around knowing and maintaining Spanish did not have an effect.

There are a few factors we acknowledge are likely at play. This again highlights a need to reevaluate the current status of ideologies within the heritage community itself as well as need to better understand individual attitudes. To echo the point above, there is a unique relational aspect when carrying out within community heritage research that was not present in some of the previous studies examining effects of social information on speech intelligibility. The shared experiences of having their language variety deemed inferior if considered language at all (Vogel & García, 2017) may represent an underlying desire to avoid perpetuating them further. But also, this aspect of increased experience with the varieties they were exposed to may have led to the reduced influence of the social guise on intelligibility. Heritage speakers are likely to be highly familiar with US Spanish and heritage varieties of Spanish, and familiarity has been found to be a social factor that mitigates intelligibility effects (McGowan, 2015). Greater experience with increased variation might make their mental model of who a heritage speaker or learner 'should' be, and the background and experiences that go with that, more diverse than we had anticipated. Or, even more, this may not be a reflection of who they think 'should' or 'shouldn't' be heritage, but rather a simple understanding that the backgrounds they were exposed to, do in fact exist within the community.

As we can see there appears to be various ways in which social information interacts with speech perception and community acceptance within the heritage community that don't always necessarily align with trends found in the between community circumstances, unless familiarity is also high (McGowan, 2015). To revisit the influences at play that led to the learner guises themselves not being as distinctly and harshly judged as consistently as one might have expected when considering the rhetoric shared about judgements heritage speakers might make, it felt relevant to dig further into the question about features of the audio stimulus itself.

6.1.4 Chapter 5 Findings

Chapter 5 analyzed if the potential for phonological differences between the talkers that we can attribute to the observed differences in ratings between the guises. As observed in Chapter 3, two of the heritage talkers – talkers 100 and 107 – received their highest ratings for sounding more fluent, most 'native-like' and for the identity-based question when paired with *speaker* guises, while the other two talkers – talkers 101 and 105 – were rated lowest for these questions overall but received their highest ratings when paired with *learner* guises. Therefore, there may be something specific about the speech of each of these talkers that felt more aligned to certain guises than others. To examine this I adopted a similar methodological approach to Rao (2015) by focusing on the production of the voiced stops /b/, /d/, and /g/ in Spanish. Spirantization is a well recorded phenomenon in the Spanish language in which intervocalic stops undergo weakening, and the most common cross-dialectic realization is a pure approximant (Rao, 2015). I sought to determine if there were distinct differences between the talkers who were rated highest for *speaker* guises and the talkers who were rated highest for the *learner* guises. Furthermore, I took into account the role that speech style may have by differentiating between spontaneous speech production and scripted speech production. Due to literacy rates

often being low in the Spanish heritage community (Magaña, 2015), there was the potential that reading stimuli aloud may have led to within talker differences as well as any observed between talker differences.

Findings revealed that talkers 100 and 107, the *speaker* guise group, consistently produced more pure approximants, while talkers 101 and 105 had more realization of both tense and pure approximants across speech styles. Results were not significant though this is likely due to limited data. Despite this I still see value in some of the observations that were made regarding the talkers' productions. As was observed in Chappell (2019)'s sociophonetic study, "heritage speakers are highly adept at discerning social information based on variation to which they are regularly exposed" (p. 258). This addresses the point made when interpreting findings from Chapter 4 around the influence that familiarity with the variety in question might have on listeners' judgements. While Chappell (2019) focused on the social variable of gender as it pertains to productions of [v] between late immigrant speakers and heritage speakers, it is still worth acknowledging that results in her study demonstrate heritage speakers' implicit knowledge of phonological cues present in Spanish. Due to their immense exposure to various varieties of Spanish and this ability to extract social meaning from phonetic variation as found in this study, it is possible that participants garnered some social meaning from the audio stimuli that influenced their ratings more than the social guises were hypothesized to. The listeners providing ratings may have noticed these differences in phonetic realizations of the intervocalic stops. The production of stops and voice-onset time are very sensitive to listeners' perceptions regarding native-like speech and perceptions of a foreign accent (Schoonmaker-Gates, 2015). Considering this in conjunction with the fact that heritage speakers have been found to exhibit perceptual patterns similar to native monolingual Spanish speakers (Chappell, 2019), this definitely invites

an avenue for further investigation into how heritage listeners might attune to this feature when making perceptual judgements and how that may be influenced by a mismatch in their expectations around social information.

If we consider then that heritage speakers may exhibit perceptual patterns similar to that of native speaker patterns, yet heritage production does not always pattern to native-speaker production, we see a need to create more explicit awareness in the community around the variation that might form within the heritage Spanish and how it can shape the heritage accent. Current studies have supported the integration of dialectal and phonological variation into heritage curricula as a way to begin empowering heritage speakers, thereby allowing them to embrace pronunciation differences rather than stigmatizing them (Correa, 2011). In doing so, we may increase acceptance in the variation that does exist within US Spanish pronunciations as well.

6.2 Contributions to the field

First, it is important to revisit the key distinctions that are driving much of the work carried out in this dissertation, which lie within the differences between heritage *speakers* and heritage *learners*. For the purposes of these studies and inquiry, I intentionally chose to follow the definition proposed by Rothman et al. (2016) that describes heritage speakers as individuals who possess some degree of communicative competence in the heritage language. Within the field of Spanish as a heritage language (SHL), the term heritage *learner* was introduced to describe students who have an ancestral and cultural connection to the Spanish language, who enroll in language learning courses, and, more specific to SHL, rely on some level of naturalistic exposure and acquisition of the heritage language during childhood (Rothman, 2009). Research has tended to differentiate these groups by deeming those who do not enroll in language classes

as being representative of heritage language speakers "in the 'wild', naturalistic learners and users who are not seeking to learn their heritage language in the classroom" (Montrul & Polinsky, 2021).

In their work, Rothman, Tsimpli and Pascual y Cabo (2016) discuss how there has historically been a clear majority in the pedagogical field regarding the expectation around proficiency; specifically, heritage speakers and learners have been treated as synonymous with regard to the expectation that they 'should' have some degree of competency in Spanish that was learned in the home or a naturalized setting, usually without formal instruction. However, the authors mention there is not a unanimous agreement regarding how proficient an individual 'should' be to garner this classification (Rothman et al., 2016).The situation becomes even more opaque when attempting to group students based on their proficiency, given that there is also no complete agreement regarding who constitutes a beginner, intermediate, or advanced student. Often these standards are based on monolingual norms, or even norms for second or foreign language acquisition, which do not appropriately align with the heritage acquisition process (Rothman, et al., 2016).

The heritage language acquisition process should instead be recognized as the acquisition of a contact variety and then that of second dialect acquisition when students enroll in classes (Potowski et al., 2009). This acquisition of a second dialect is usually driven by models centered around appropriateness that aim to expand the heritage repertoire to include standard varieties, although this is typically done so in a manner that maintains the perception that the home variety is inferior (Flores & Rosa, 2016). Furthermore, the tendency to ignore this reality has led to the existence of work mislabeling those within the community. Rothman et al. (2016) highlight the well addressed fact that the degree of knowledge in the heritage language throughout the Spanish

heritage community is very inconsistent, reminding us that this variability is not different from native monolingual patterns, when it comes to variation, and of the factors that influence this variable knowledge:

> This fact is not at all surprising when one considers the continuum of exposure type (quantity and quality), their individual patterns of use of the HL, the level of literacy they have in the HL, the status of the HL in the society in which they live, or their access to other speakers of the HL. These differences do not pertain to monolinguals, at least not in the same way. While we typically do not use terms like intermediate and advanced levels to describe monolinguals, these terms are used in HS studies to equate their relative level as compared to matched monolingual norms. (p.19)

While this dissertation is not the first to address these concerns, it is meant to add to the field in such a way that it reminds us not to conflate previous ideas and concepts regarding how we discuss the Spanish heritage community, as well as how we understand their linguistic acquisition and development of their heritage language (HL). A crucial but only seldom adopted distinction is the one made between heritage *speakers* and *learners* in a way that positions the *learners* as individuals who still need to acquire their heritage language in general due to a lack of competence. On the other hand, *speakers* represent the population who already have a command of the language from exposure in naturalistic environments during childhood but underwent a shift in language dominance (Rothman et al., 2016). In fact, the important and necessary distinction between these groups that fall under the umbrella of heritage is so overlooked that even the UCLA National Heritage Language Resource Center (NHLRC) conflates these terms. The center states in its "About" statement that it represents, "… a federally funded language resource center whose mission is to develop effective pedagogical approaches

to teaching heritage language learners…" and then later in its "Mission" statement that "Heritage language speakers are bilinguals – often with two native languages" without acknowledging any difference between their use of learner or speaker. This institutional overlap ultimately illustrates that this is a shared concern across all heritage languages, not just Spanish.

Even more, the majority within the field has indeed been challenged by other researchers within the field of SHL who seek to expand the definition with regard to its role in education specifically, focusing mostly on the linguistic ability of the speaker rather than relying strictly on the cultural and heritage side. Beaudrie and Ducar (2005) have been proposing that we include within heritage programs and classrooms:

> all individuals that experienced a relatively extended period of exposure to the language, typically during childhood, through contact with family members or other individuals, resulting in the development of either receptive and/or productive abilities in the language, and varying degrees of bilingualism. (p. 13).

They explain that the motivation for this is that while some students may not be Hispanic, they might *feel* it, and some examples provided for this type of individual were: those who were raised by a Spanish speaking caretaker or those who might have a spouse that speaks Spanish (Beaudrie, 2012). Essentially the argument against including *learners with a heritage motivation* is that their acquisition mirrors too closely to that of L2 learners; therefore, they feel an SHL class may be inappropriate and would not sufficiently meet their needs. Thus, from a social standpoint some researchers in the field think these learners classify as heritage learners, but not from a linguistic standpoint.

While findings in this dissertation do not explicitly demonstrate the need for this distinction, the within community perceptions do show that their inclusion in academic spaces is

warranted and can even be embraced by others in the heritage community. Findings demonstrating that both heritage *speaker* and heritage *learner* profiles are perceived by other heritage community members as connected to the Latinx identity show us how combined heritage culture classes – different from acquisition classes – can help to further bridge the divide presented from differences in Spanish proficiency. Further, this dissertation also contributes to the field by elucidating that as we incorporate new methods and models – as is the case with most of the models in corpus linguistics as they relate to heritage language research (Kisselev, 2021) – we need to make sure we are mindful of how we select and employ certain models and methodological frameworks. It is not safe to assume that because certain methodological practices or empirical models were effective at describing L1 or L2 acquisition and production, they will translate to be equally effective in heritage domains. Methodological frameworks must be developed to match the unique cognitive and sociohistorical profiles of the speakers being examined, rather than forcing them into inappropriate comparative frameworks.

Overall, the findings in this dissertation contribute to the field of Spanish Heritage Language (SHL) research by furthering our understanding of the impacts of improperly describing, comparing and even at times excluding backgrounds of speakers within the heritage community in research and academic spaces alike. In addition to this, the descriptive results remind us to be mindful when employing certain methodologies and models to ensure that we are actually obtaining meaningful results that can be interpreted properly and in a truly informative way. In doing so, we as researchers can then avoid deficit descriptors such as "incomplete" or "semi-lingual" that have been found to be only legitimate when comparing the heritage community to monolingual speakers.

Finally, a major contribution of this work is the inclusion of members from within the heritage community who acquired only limited Spanish in childhood or none at all. This portion of the community is often overlooked due to their limited linguistic abilities prior to enrolling in Spanish language courses; however, I argue this should not justify their exclusion. While I can understand the reasoning behind this from the language acquisition position, and I can only fathom how difficult it is from a programmatic and administrative standpoint to be able to develop heritage classes in the first place, I would like to say that the descriptive results in this dissertation can be interpreted to push back against this. If nothing else, there is a need to design a culture class that allows these students to partake in the heritage program in some way. The traditional, restrictive definition is limiting and can, although unintentionally, result in the perpetuation of language ideologies that have consistently othered this subgroup within the heritage community. From a solely formal linguistic viewpoint of language acquisition there may be a need to differentiate these groups, but it is my opinion that if we can include a group within SHL programs that *feel* Hispanic, I then also see a space for those who are Latinx and have a desire to reconnect with their heritage through their language.

The way this inclusionary framework would impact perceptions and attitudes within the community by carving space for these individuals in heritage programs could allow us as educators to be huge proponents in the process of deeming them legitimate members of the community. To echo sentiments expressed by others in the field who concur with their inclusion (Beaudrie & Ducar, 2005), the motivations combined with even some limited linguistic knowledge and cultural knowledge that this profile of heritage learner has is still valuable and we should not be the individuals determining otherwise. Therefore, this personal connection and

motivation may require us to reconsider exclusionary practices in education and research (Rothman et al., 2016).

Furthermore, to revisit the ideas around inappropriately judging the heritage language as incompletely acquired for those who are classified as speakers, the classroom is a site riddled with a plethora of ideologies from the Latinx community not just around who qualifies but also around the kind of Spanish being spoken. It seems as though members of the heritage community are constantly at the precipice of having their identity stripped from them should they make one wrong move. These attitudes also present negative effects on their willingness to communicate in Spanish; if we continue these ideas in the classroom, how are we helping heritage students if we are actually discouraging them from using Spanish? In essence, those who exhibit strong ideologies around the relationship between Spanish and a Latinx identity, can also end up being the reason that some US Latinx hide their Spanish ability due to ideals about standardization. This highlights the necessity for a shift in how we think about, research, and discuss the heritage acquisition process from both an empirical and educational aspect. These findings overall address certain questions around what it means to be a Spanish heritage speaker or learner in the United States as these individuals are constantly confronted with ideologies that pressure them to act a certain way and speak a certain way, without much regard to the historical, political, and sociocultural influences that led to the observed differences that have been examined.

6.3 Limitations & Future Directions

As the findings in this dissertation contribute to our understanding of the Spanish heritage community in the United States, the research carried out was not without its limitations, both as it pertains to the respective methods and approaches in each chapter, and also within the field as a whole. The limitations of the work presented here include small sample sizes for Chapters 2

and 5, which makes external validity a challenge. The usage of the ALLSSTAR corpus throughout was intentional for consistency purposes in the talkers being used and analyzed; however, this did limit the ability to make large overarching claims regarding the trends observed. Instead, many of the claims made here serve to initiate further conversations around the way we have historically discussed the Spanish heritage community in the United States.

In addition to concerns regarding small sample sizes, another limitation to the research is that the data were collected remotely. I did try to control for external noise by asking participants to go to a quiet room and use headphones, including a sound check so they could adjust the volume to their liking, and excluding any participants who seem to have rushed or taken an excessive amount of time, indicating potential distraction. While I cannot guarantee that each participant used the required resources to complete the task, and I cannot ascertain that they were in a quiet room devoid of distractions, there may have been a benefit to them not participating in a lab setting. Specifically, a remote environment potentially reduced participant stress and incidentally may have lowered the likelihood in making harsher judgements that mimic the institutionalized ideologies rather than those that reflect their own personal attitudes.

An additional limitation to this research is related to the social guises used for the ratings and intelligibility data. I intentionally presented the participants with a lot of variability in the backgrounds described in the guises so as to be representative of the variability within the Spanish heritage community. However, this design choice invites questions around whether this variability may have been more influential in driving the inconsistent ratings between speaker and learner guises. While this was an intentional and thoughtfully made decision, I do acknowledge that there are within community attitudes and ideologies around different varieties of Spanish – and the specific populations that speak each respective variety (e.g., Cuban vs

Puerto Rican vs Mexican) – exist independently and may have influenced listener ratings and findings in general. Furthermore, the guises do not account for the possibility of talkers having an additional heritage language within their family that was subsequently replaced with Spanish. Despite these limitations, the research presented remains relevant and valid in its contributions to the field.

A limitation present to the field as a whole is the general lack of bilingual or translanguaging data, which would have allowed for a larger picture of the linguistic repertoire of each of the participants in Chapter 2. This would have also very likely impacted the ratings in Chapters 3 and 4, considering what we know about stigmas associated with Spanglish or US varieties of Spanish (Parodi, 2008). This would have likely also impacted their phonological productions, as is observed in Chapter 5, which could have also had an impact on how each talker was rated. Allowing for a bilingual mode would have been particularly interesting to examine how this may or may not have led to a significant difference between perceptions of the guises or their intelligibility. While it is not possible to know for certain based on these data, the effects of expectation would have come to fruition. Laboratory based studies quite often restrict participants to a monolingual mode, which ultimately reinforces previously established and refuted ideas around bilinguals representing two monolinguals in one mind (Grosjean, 1989). The linguistic system for bilinguals and heritage speakers actually represents a single, dynamic system that does not conform to known politicized language boundaries (García & Li, 2014; Vogel & García, 2017). Moreover, bilingual speakers and heritage speakers tend to acquire skills in each language as it fits their needs; therefore, heritage speakers tend to have a greater productive vocabulary in their heritage language around topics related to family and community, while their English skills may be more advanced with regard to professional and academic

settings – unless they enrolled in Spanish classes to expand upon their abilities and the social domains in which use Spanish in (Grosjean, 2016). This is addressed in Parodi's (2008) analysis of the of the stigma associated with the California *Spanish Vernacular (SV)* spoken in Los Angeles, both in and out of the classroom. She mentions:

> [Spanish heritage speakers] have lexical gaps in all semantic fields. They solve this problem by paraphrasing, using body language, using English loan words, code switching, or changing topics. Their knowledge of passive vocabulary allows them to understand any monolingual statement or monolingual discourse, even if they cannot reply using vocabulary in the same sociolinguistic register. (p. 206).

It does not go unnoticed that Parodi is able to address the 'gaps' they may present while avoiding using language that describes this as problematic or compares these patterns to monolingual standards. This highlights even more that we are not seeing the full picture when we limit them to one language in a lab, especially in a corpus that is meant for longevity to allow many conclusions to be drawn from one large source of data, which is why scholars introduced translanguaging as a theoretical lens. Although this theory as it applies to language is relatively new, we in the fields of bilingualism and multilingualism at large have known for some time now that the linguistic system of these individuals is a multifaceted and nuanced one – with *one* being the key word – that does not comply with monolingual norms of language representation and production.

As is astutely pointed out by scholars in the field, Spanglish has actually developed to become symbolic of cultural hybridity. When heritage speakers are able to freely use this variety, they feel as if they can actually express their bicultural identity in a way that feels more genuine

and sincere to them, as if they are expressing themselves in a way actually representative of their daily lives (Betti, 2011; Ayala-Saracay, 2025). Thus, the research to date that intentionally confines and restrains the linguistic system of heritage speakers and learners is informing us of something – what production looks like when bilinguals are unable to access their full linguistic repertoire – although this does not address many of the questions that we as researchers are seeking out as they apply to their true linguistic competency.

Future directions for the work carried out in this dissertation include obtaining data that are actually representative of heritage speakers' tendencies to translanguage. It is very limiting and not entirely informative to restrict bilingual speakers to a monolingual mode; therefore, the best way to move forward with carrying out analyses and research in the field of Spanish as a heritage language, and heritage languages/multilingualism in general, is to allow participants to communicate in a way that feels natural to them. This ensures they have access to their full linguistic repertoire and communicative abilities. It would be interesting to see how results may differ if the talkers who provided the corpus data had not been restricted in such a way.

An additional avenue for further research would be to expand upon the phonological analysis of the language produced by the talkers who provided the stimuli for the perception studies. As mentioned, the results from Chapter 5 are limited in their scope due to focusing on a small number of talkers; however, because the results report a significant interaction between the talker and the guise type, it is clear that more needs to be understood regarding the acoustic features that may lead participants to rate talkers differently depending on the social guise. I speculate that a misalignment in expectations related to how each guise 'should' sound and the actual language they produced may be influential. However, without being able to tease apart the actual productive differences it is difficult to ascertain how much of an impact listeners'

expectations may have played. Based on the preliminary evidence provided in Chapter 5 revealing differences between stop productions, it is likely that additional phonological differences may also be present that could, in some way, influence how talkers are perceived.

6.4 Conclusions: Call to action!

Ultimately, this dissertation addresses the need for researchers and educators to actively inform the broader heritage and Latinx communities about the legitimacy of heritage language production. We as scholars can illuminate how the heritage variety and heritage accent represent a natural byproduct of bilingual development rather than linguistic flaws. In doing so, we can aid in the dismantling of internal gatekeeping practices and validate the ways in which individuals may reconnect with and express their cultural heritage. We must recognize the responsibility we bear within academia, ensuring we remain thoughtful and intentional regarding the way we discuss this topic and assign terminology.

Bayram et al. (2019) highlight this as a key need to abandon the usage of concepts such as incomplete, deficient, and attrited which lead to misconceptions that heritage grammars are anything other than simply different from monolingual ones. In their article, they remind us of the potential negative side effects we can introduce should this terminology reach more general audiences who typically have limited knowledge of the nature of the work being carried out. This is to say that if we within the research community continue to extend the notion that incomplete acquisition is an appropriate descriptor we then run the risk of these concepts being weaponized far past their original intentions; even if the argument is that it is necessary only on theoretical grounds and therefore innocently pejorative (Domínguez et al., 2019).

Support in favor of shifting our perceptions of the observed differences between monolingual speakers and heritage speakers has demonstrated that the only way ‘incomplete’

may be an appropriate descriptor for heritage speaker grammars is if we continue to make the fatal flaw of relying on adult monolingual grammars and acquisition trajectories – as is traditionally done in L2 practices – as our only baseline to draw conclusions from (Pascual y Cabo & Rothman, 2012; Kupisch & Rothman, 2018). It seems as if we can argue that for many years researchers have been prone to shoving the square peg of heritage acquisition into the round hole of L1 monolingual development, or even L2 developmental milestones, yet when the paradigmatic structure did not fit, it went unnoticed or unrecognized. As addressed in Chapter 2, this is flawed in nature given that relying on monolinguals as the baseline then takes for granted the plethora of variation that exists inherently within monolingual production as well. Making these comparisons essentially holds heritage speakers to a standard we ourselves don't even expect of monolingual speakers, and researchers have addressed how this then begs the question of whether we should attribute certain aspects of monolingual production, that are often markers of dialectal differences, as being incomplete as well, given their divergence from the 'standard' (Bayram et al., 2019).

Further inconsistencies are noted when considering that even those who are reluctant to relinquish the idea that heritage grammars are incomplete still recognize that they are representative of coherent grammars in their own right (Polinsky, 2008), which challenges the need for the deficit descriptor at all. Thus, within the field of heritage research, we must expand our own approach and understanding of heritage grammars as systems that naturally deviate from monolingual norms. This shift must be done in such a way that doesn't unintentionally perpetuate beliefs around languagelessness or linguistic purity. In initiating a shift at the empirical level, we can then extend awareness overall around heritage language development and the reasons why it may systematically and fundamentally differ. The ultimate goal lies in

enacting change at the heritage community level. As we become increasingly enlightened, we can utilize our knowledge to help educate and potentially facilitate a shift in the language attitudes and ideologies within the US Latinx community that have been cemented in for many years. The significant expectation placed on language ability within the Latinx community is not an ideology that will be uprooted overnight – especially as it pertains to attitudes around identity, belonging, and acceptance – yet this challenge doesn't mean we can't initiate and positively impact change.

There is also a direct need to ensure we are setting the appropriate example for inclusive educational practices – ones that can challenge ideologies which have previously led to hostilities and feelings of shame within the community. As the educational system has long been the space that introduces and reinforces negative ideologies (Cobas & Feagin, 2022), it can now serve as a space that dismantles them, allowing heritage speakers to regain their agency and confidence surrounding their Latinx identity. As we see a steady increase in university level heritage programs and support, we need to then ensure that these programs are well equipped to meet students' needs. As is often the case with heritage populations, one of the most integral needs is to provide them with a space where they can strengthen their ethnolinguistic vitality. A great starting point for enacting and setting a positive example for acceptance and inclusivity can begin with expanding our own definitions of heritage speakers and learners as they relate to program placement.

Relying on heritage profiles that expect some degree of proficiency in the heritage language results in essentially ignoring the portion of the community who do not possess linguistic ability, or who possess very limited ability, in their heritage language but who do possess a desire to connect further to their culture via heritage language acquisition. This critique

is not meant to overlook how difficult it is to accommodate this variation throughout the community that places heritage speakers on "a continuum, from rather fluent speakers, who can sound almost like competent native speakers, to those who can barely speak the home language" (Polinsky & Kagan, 2007, p. 371). Rather, it is simply to point out that if we continue to refer to certain portions of the community as not legitimately heritage speakers and exclude them in this way from heritage programs, how can we expect the general community to accept them as well?

It is important to emphasize, also, that I do not imply we should undo the years of work that have attempted to develop a space in education that is supportive and accepting of heritage speakers in general, nor is this meant to devalue heritage language maintenance. Instead, the hope is that the inclusion of these individuals will lead to a greater sociocultural awareness among the Latinx community about the process of generational language loss that has been intentionally forced upon them, and the role educational systems and governmental policy have historically played in accelerating this process. Gaining this awareness can hopefully support within community acceptance of Latinx individuals with limited or no ability in Spanish who acquired much of their skills later in life. This inclusion may afford those who may perpetuate these beliefs the opportunity to critically examine the nature of the ideologies they may be reinforcing and any personal attitudes they have around the relationship between Spanish proficiency and ability, and the Latinx identity.

In their 2005 article, Beaudrie & Ducar argue for the potential of including speakers who do not have a cultural connection to the language but did grow up speaking Spanish or with a substantial amount of prior exposure to it, so that they exhibit language usage trends similar to those of heritage speakers. While I understand this line of logic, making linguistic competency in the heritage language the only factor determining placement in heritage vs L2 classes, ignores

the rich cultural experiences these students may have had that didn't rely on their knowledge of Spanish. Further, not providing a space for beginner-level heritage learners then requires these individuals to enroll in classes designed for L2 learners, reducing the likelihood that they will be exposed to any sociolinguistic awareness around language varieties, language ideologies, or language maintenance and loss. As heritage researchers promote the implementation of Critical Language Awareness (CLA) approaches – which have been found to have positive impacts on students' ethnoracial identities in classrooms (Gómez García, 2022) – we are presented with a hopeful opportunity for the inclusion of these students.

Crucially, research acknowledges how difficult it is to accommodate the immense variation within the community (Carreira & Kagan, 2011; Rao & Kuder, 2016). Granted, many universities lack sufficient funding and institutional support to expand programs beyond only being able to provide a few heritage courses and even fewer programs provide sequences of courses (Beaudrie, 2012). However, it is still crucial to carve a space for anyone who identifies as heritage, even those who would place in beginner-level courses. Heritage speakers who display lower levels of linguistic competency are often the most susceptible to within community scrutiny and erasure from academic spaces. Consequently, they are never truly afforded the opportunity to grapple with their own insecurities or self-esteem issues around language and identity.

To reiterate, this is not to ignore the struggles that intermediate and advanced heritage learners have experienced to gain their space in education; rather it is intended to shed light on the need for, and importance of expanding our criteria for heritage programs in general, or even how we approach language acquisition on a larger scale. Developing a culture class that includes these learners and gives them agency to express their Latinx identity, background, and

experiences would also be a huge step in the right direction. Quan (2020) reported that even L2 students benefitted from critical language awareness, increasing their overall sociocultural awareness and rendering this an inviting potential option for underfunded or unsupported university heritage programs. Either way, we cannot simply forget about this group, leaving them high and dry. Even these subtle ways that we as researchers and educators perpetuate ideologies and expectations around language proficiency in the heritage community can have a large impact.

My final call to action in this dissertation will then be for policymakers, who wield systemic power and influence. As this dissertation aptly outlines, policy in the United States has always put English on a pedestal far above any other language and often employs English as a *proxy* for race, using discriminatory language practices to justify racial discrimination (Johnson & Martínez, 2000). The use of language as a proxy has been so successful that even the majority of the voting population – who have historically passed legislation in favor of assimilation and nativist ideals – truly believe they are doing a service to these families and individuals by giving them the opportunity to learn English with zero regard for their heritage language (Yamagami, 2012). This trend remains rampant even as contemporary policy attempts to embrace multilingualism (e.g., Prop 58 in California), given that it does so in such a way that positions languages other than English as only valuable to white Anglophones when viewed as a global and economic commodity (Carter, 2018). Therefore, I challenge policymakers to reconsider the role they have played in demonizing and delegitimizing minoritized languages throughout the United States. While multilingualism does afford the opportunity to connect globally, we must first prioritize systemic change that allows there to be value at the sociocultural level as well.

APPENDIX A

CHAPTER 2 - ALLSSTAR CORPUS QUESTION-AND-ANSWER INTERVIEW QUESTIONS

English Questions:

Tell us about the place or places where you grew up.

Tell us about your family or other people you are close to.

Tell us about a big time of celebration in your family or culture.

Tell us about some food or meals that are important in your family or culture.

Tell us your first impressions of your current home.

Tell us about what you like to do in your free time.

Spanish Questions:

Cuéntenos sobre el lugar o lugares en los que creció.

Cuéntenos acerca de su familia u otras personas cercanas a usted.

Cuéntenos de un gran momento de celebración en su familia o cultura.

Cuéntenos algo sobre la comida o las comidas que sean importantes en su familia o de en su cultura.

Cuéntenos sus primeras impresiones de su casa actual.

Cuéntenos de lo que le gusta hacer en su tiempo libre.

APPENDIX B

ALLSSTAR CORPUS COMIC STRIPS

Story 1 and Story 2





Story 3 and Story 4





APPENDIX C

CHAPTER 3 – LANGUAGE QUESTIONNAIRE

Age: 18 – 25 years
26 – 35 years
36 – 45 years
46 – 55 years
56+

Gender: Female
Male
Non-binary
Prefer not to say

What is your highest level of education completed?

Less than high school	High school graduate	Some College	2-year degree
4-year degree	Professional degree	Doctorate	

Country of origin (Please specify if not US)

United States Other _____

If you were not born in the US, how long have you lived in the US for? How old were you when you moved to the US?

Years in US:	Age at move:
<1 year	Infancy
1 – 5 years	1 – 5 years
6 – 10 years	6 – 10 years
11 – 15 years	11 – 15 years
16 – 20 years	16 – 20 years
20+ years	20+ years

Where are your parents/caregivers from?

Parent 1: Parent 2:

What languages do your parents/caregivers speak? Please list all below.

Parent 1: Parent 2:

What is your parents' highest level of education completed?

Parent 1:

Less than high school High school graduate Some College 2 year degree

4 year degree Professional degree Doctorate

Parent 2:

Less than high school High school graduate Some College 2 year degree

4 year degree Professional degree Doctorate

At what age did you first learn English and Spanish?

English: Spanish:

What languages did you **hear** in your home? (Select all that apply)

English Spanish Mixed Other _____

What languages do you use mostly with your parents/caregivers?

When they speak with you:

English Spanish Mixed Other _____

When you speak with them:

English Spanish Mixed Other _____

Did you watch TV/listen to the radio in Spanish?

Yes, both Just TV Just radio No/neither

Did you travel to Spanish speaking countries throughout your childhood/teenage years? If yes, please state where, when, how long and how often.

No Yes_____

Did your parents encourage you to speak Spanish?

Yes No

Did your parents correct your Spanish?

Yes No

At what age did you start to feel comfortable using the following languages?

English:	Spanish:
As early as I can remember	As early as I can remember
1 – 5 years	1 – 5 years
6 – 10 years	6 – 10 years
11 – 15 years	11 – 15 years
16 – 20 years	16 – 20 years
20+ years	20+ years
Not yet	Not yet

Rate your current overall language ability in **English**.

1 = understand but cannot speak

2 = understand and can speak with great difficulty

3 = understand and speak but with some difficulty

4 = understand and speak comfortably, with little difficulty

5 = understand and speak fluently like a native speaker

Rate your current overall language ability in **Spanish**.

1 = understand but cannot speak

2 = understand and can speak with great difficulty

3 = understand and speak but with some difficulty

4 = understand and speak comfortably, with little difficulty

5 = understand and speak fluently like a native speaker

Did you attend elementary school in the US? If no, please state where you attended elementary school.

Yes No _____

How many years of classes (grammar, history, math, etc.) have you had in the following languages (primary school through university)?

English:	Spanish:
0	0
1 – 2 years	1 – 2 years
3 – 4 years	3 – 4 years
5 – 6 years	5 – 6 years
7 – 8 years	7 – 8 years
9 – 10 years	9 – 10 years
11 – 15 years	11 – 15 years
16 – 20 years	16 – 20 years
20+ years	20+ years

On a scale from 1-5, rate your abilities in English and Spanish.

1 = poor 2 = needs work 3 = good 4 = very good

5 = native speaker command

English reading: English writing: English speaking:

English listening:

Spanish reading: Spanish writing: Spanish speaking:

Spanish listening:

How much time have you spent in a country/region where the following languages are spoken?

English:	Spanish:
0	0
1 – 2 years	1 – 2 years
3 – 4 years	3 – 4 years
5 – 6 years	5 – 6 years

7 – 8 years	7 – 8 years
9 – 10 years	9 – 10 years
11 – 15 years	11 – 15 years
16 – 20 years	16 – 20 years
20+ years	20+ years

In general, as a young adult which language do/did you prefer to use?

English Spanish It depends on whom I talk to Both

Do you feel Spanish is your native language or like a second language?

Native language Like a second language

Would you like to improve your Spanish skills?

Yes No

Why or why not?

In an average week, what percentage of the time do you use the following languages with **friends**? (Percentages should add up to 100%)

English:	0%	10%	20%	30%	40%	50%	60%	70%	80%	90%	100%
Spanish:	0%	10%	20%	30%	40%	50%	60%	70%	80%	90%	100%
Other languages:	0%	10%	20%	30%	40%	50%	60%	70%	80%	90%	100%

In an average week, what percentage of the time do you use the following languages with **family**? (Percentages should add up to 100%)

English:	0%	10%	20%	30%	40%	50%	60%	70%	80%	90%	100%
Spanish:	0%	10%	20%	30%	40%	50%	60%	70%	80%	90%	100%
Other languages:	0%	10%	20%	30%	40%	50%	60%	70%	80%	90%	100%

In an average week, what percentage of the time do you use the following languages at **school/work**? (Percentages should add up to 100%)

English:	0%	10%	20%	30%	40%	50%	60%	70%	80%	90%	100%
Spanish:	0%	10%	20%	30%	40%	50%	60%	70%	80%	90%	100%
Other languages:	0%	10%	20%	30%	40%	50%	60%	70%	80%	90%	100%

When you talk to yourself, how often do you talk to **yourself** in the following languages? (Percentages should add up to 100%)

English:	0%	10%	20%	30%	40%	50%	60%	70%	80%	90%	100%
Spanish:	0%	10%	20%	30%	40%	50%	60%	70%	80%	90%	100%
Other languages:	0%	10%	20%	30%	40%	50%	60%	70%	80%	90%	100%

When you count, how often do you count in the following languages? (Percentages should add up to 100%)

English:	0%	10%	20%	30%	40%	50%	60%	70%	80%	90%	100%
Spanish:	0%	10%	20%	30%	40%	50%	60%	70%	80%	90%	100%
Other languages:	0%	10%	20%	30%	40%	50%	60%	70%	80%	90%	100%

Please indicate how much you agree or disagree with each statement below.

Strongly disagree	Disagree	Somewhat disagree	Neither agree/disagree
Somewhat agree	Agree	Strongly agree	

I feel like myself when I speak English.

I feel like myself when I speak Spanish.

I identify with an English-speaking culture.

I identify with a Spanish-speaking culture.

Please indicate how much you agree or disagree with each statement below.

Strongly disagree Disagree Somewhat disagree Neither agree/disagree
Somewhat agree Agree Strongly agree

It is important for me to use (or eventually use) English like a native speaker.

It is important for me to use (or eventually use) Spanish like a native speaker.

I want others to think I am a native speaker of English.

I want others to think I am a native speaker of Spanish.

Please indicate how much you agree or disagree with each statement below.

Strongly disagree Disagree Somewhat disagree Neither agree/disagree
Somewhat agree Agree Strongly agree

It is important for Latinos to speak and maintain their Spanish.

When I meet someone that identifies as Latino, I think they should know Spanish.

Spanish is an important part of Latino identity.

APPENDIX D

CHAPTER 3 AND 4 – SOCIAL GUISES

Speaker Guise 1: The speaker you are about to hear was born in the United States to Mexican immigrant parents. Their first language is Spanish, and they began acquiring English during primary school. They still frequently use Spanish at home and with members of the community; however, English has become their dominant and most used language.

Speaker Guise 2: The speaker you are about to hear was born in the United States. Their father is a second-generation Cuban American, and their mother moved to North Carolina from Puerto Rico at age 10. Their father speaks very limited Spanish but can understand very well and their mother has maintained fluency since her move from Puerto Rico. The speaker acquired some Spanish speaking skills from their mother and grandparents but can understand Spanish much better than they can speak. Their dominant language is English.

Speaker Guise 3: The speaker you are about to hear was born in the United States to a non-Latina mother and a Cuban American father. Their parents spoke only Spanish in the home during their first 7 years; however, the speaker began to lose some fluency in Spanish after starting primary school. They now claim to speak mostly English with their parents but continues to use some Spanish.

Speaker Guise 4: The speaker you are about to hear is a second-generation Cuban American. Their mother was raised mostly speaking Spanish with her parents; however, their father was raised to speak only English. As a result, the speaker grew up with stronger listening skills from

listening to their mother and grandparents, but they lacked proficiency in their speaking abilities and have begun taking university Spanish courses to strengthen them.

Learner Guise 1: The speaker you are about to hear is a third-generation Mexican American. Their parents mostly understand Spanish and have limited speaking abilities. The speaker's speaking and listening abilities in Spanish are limited; however, they had some exposure growing up from their grandparents. The speaker tried to strengthen their skills in Spanish by enrolling in university courses; therefore, most of their abilities have been acquired in the classroom.

Learner Guise 2: The speaker you are about to hear was born in the United States to a U.S.-born mother and a Guatemalan immigrant father. The family primarily spoke English in the home, but the speaker was still exposed heavily to their father's cultural background. The speaker acquired most of their abilities from Spanish classes in high school and during their university studies.

Learner Guise 3: The speaker you are about to hear was born in the United States to Puerto Rican parents who moved to Arizona in their early adolescence. Their parents became English-dominant speakers after moving to Arizona and as a result, the speaker did not speak Spanish in the home. They grew up with some exposure to Latino pop culture, but it was limited. They began acquiring Spanish in high school.

Learner Guise 4: The speaker you are about to hear was born in the United States to a second-generation Mexican American mother and non-Latino, U.S.-born father. They grew up with some

Spanish exposure from their mother's parents, but overall, their Spanish skills were limited growing up prior to enrolling in Spanish language learning courses in high school and university.

APPENDIX E

CHAPTER 4 – SELF-REPORTED PROFICIENCY SCORES

ALLSSTAR Corpus heritage speakers*

	Spanish Speaking	Spanish Listening	Spanish Reading	Spanish Writing	English Speaking	English Listening	English Reading	English Writing
Speaker 100	9	9	8	7	9	9	8	7
Speaker 101	8	8	9	9	9	9	9	8
Speaker 105	5	7	6	6	8	8	8	8
Speaker 107	10	10	10	10	10	10	10	10

* The numbers in each column represent the actual score reported by each speaker.

Guise validation raters*

	Mean	SD
Spanish Speaking	3.20	1.39
Spanish Listening	3.86	1.29
Spanish Reading	3.29	1.32
Spanish Writing	2.77	1.21
English Speaking	4.91	0.28
English Listening	4.94	0.24
English Reading	4.94	0.24
English Writing	4.91	0.28

* The scale for the raters was different from the scale used for the intelligibility study listeners. This study's scale was 1 = *poor*, 2 = *needs work*, 3 = *good*, 4 = *very good*, and 5 = *native speaker command*

Intelligibility study listeners *

	Mean	SD
Spanish Speaking	7.27	1.85
Spanish Listening	7.96	1.33
Spanish Reading	7.49	1.66
Spanish Writing	6.40	2.18
English Speaking	8.82	0.53
English Listening	8.87	0.55
English Reading	8.91	0.42
English Writing	8.80	0.81

* The scale used for this study was 1 = *very low proficiency* to 9 = *native/native-like.*

REFERENCES CITED

Abdi, K. (2011). 'She really only speaks English': Positioning, language ideology, and heritage language learners. *Canadian modern language review*, *67*(2), 161-190.

Ackerman, H. (1996). The balsero phenomenon, 1991–1994. *Cuban Studies*, 169-200.

Aguirre, B. E. (1994). Cuban mass migration and the social construction of deviants. *Bulletin of Latin American Research*, *13*(2), 155-183.

Alba, R., Logan, J., Lutz, A., & Stults, B. (2002). Only English by the third generation? Loss and preservation of the mother tongue among the grandchildren of contemporary immigrants. *Demography*, *39*(3), 467-484.

Alonso, C. J. (2006). Spanish: The foreign national language. *ADFL BULLETIN*, *37*(2/3), 15.

Amengual, M. (2018). Asymmetrical interlingual influence in the production of Spanish and English laterals as a result of competing activation in bilingual language processing. *Journal of Phonetics*, *69*, 12-28.

Anderson, T. K., & Toribio, A.J. (2007). Attitudes towards lexical borrowing and intra-sentential code-switching among Spanish-English bilinguals. *Spanish in Context*, *4*(2), 217–240.

Aparicio, F. R. (1998). Whose Spanish, whose language, whose power? An ethnographic inquiry into differential bilingualism. *Indiana Journal of Hispanic Literatures*, *12*(1), 5-26.

Au, T. K., Knightly, L., Jun, S.A., & Oh, J. (2002). Overhearing a language during childhood. Psychological Science, 13, 238-243.

Au, T. K. F., Oh, J. S., Knightly, L. M., Jun, S. A., & Romo, L. F. (2008). Salvaging a childhood language. *Journal of memory and language*, *58*(4), 998-1011.

Ayala-Saracay, D. (2025). *Ni de Aquí, Ni de Allá: The Impact of the No Sabo Kid Label on Cultural Identity, Belonging, and Language Acquisition in the Latino Community* (Doctoral dissertation, Marymount University).

Babel, M. & Russell, J. (2015). Expectations and speech intelligibility. *The Journal of the Acoustical Society of America* 137: 2823–33.

Baker, C. (2006). Foundations of bilingual education and bilingualism. 4th ed. Clevedon, UK: Multilingual Matters.

Balestra, A., Martínez, G. A., & Moyna, M. I. (Eds.). (2008). *Recovering the US Hispanic linguistic heritage: Sociohistorical approaches to Spanish in the United States*. Arte Público Press.

Bayram, F., Kupisch, T., Pascual y Cabo, D., & Rothman, J. (2019). Terminology matters on theoretical grounds too!: Coherent grammars cannot be incomplete. *Studies in Second Language Acquisition*, *41*(2), 257-264.

Beaudrie, S. (2011). Spanish heritage language programs: A snapshot of current programs in the southwestern United States. *Foreign Language Annals*, *44*(2), 321-337.

Beaudrie, S. (2012). Research on university-based Spanish heritage language programs in the United States. *Spanish as a heritage language in the United States: The state of the field*, 203-222.

Beaudrie, S., & Ducar, C. (2005). Beginning level university heritage programs: Creating a space for all heritage language learners. *Heritage language journal*, *3*(1), 1-26.

Beaudrie, S., & Ducar, C. (2012). Language placement and beyond: Guidelines for the design and implementation of a computerized Spanish heritage language exam. *Heritage language journal*, *9*(1), 77-99.

Beaudrie, S., & Fairclough, M. (2012). *Spanish as a heritage language in the United States: The state of the field*. Georgetown University Press.

Beaudrie, S., Amezcua, A., & Loza, S. (2019). Critical language awareness for the heritage context: Development and validation of a measurement questionnaire. *Language Testing*, *36*(4), 573–594.

Beaudrie, S., & Loza, S. (2020). Spanish heritage education in the Southwestern United States: Fighting restrictive policies toward language maintenance in Arizona. In *Spanish across domains in the United States* (pp. 25-46). Brill.

Beaudrie, S., Amezcua, A., & Loza, S. (2021). Critical Language Awareness in the Heritage Language Classroom: Design, Implementation, and Evaluation of a Curricular Intervention. *International Multilingual Research Journal* 15: 61–81.

Beaudrie, S., & Marrero-Rivera, A. (2024). Research on university-based Spanish heritage language programs in the United States: The state of affairs in 2022. *Foreign Language Annals*, *57*(1), 10-26.

Benmamoun, E., Montrul, S., & Polinsky, M. (2010). *Prolegomena to Heritage Linguistics*. White Paper. Cambridge: Department of Linguistics, Harvard University.

Benmamoun, E., Montrul, S., & Polinsky, M. (2014). Defining an “ideal” heritage speaker: Theoretical and methodological challenges| Reply to Peer Commentaries. *Theoretical linguistics*.

Bernstein, K. A., Brown, L., Kaveh, Y. M., Yuhas, B., & Pazhouhi, S. (2025). The perils of policy workarounds: Lessons from Arizona's return to "English-only" policy. *Journal of Language, Identity & Education*, *24*(3), 742-748.

Betti, S. (2011). Spanglish in the United States: A legitimate expressive strategy? Modern Languages, (37), 33-53.

Bhatt, R.M. (1995) Prescriptivism, creativity, and world Englishes. *World Englishes* 14: 247–60

Biber, D., Conrad, S. M., Reppen, R., Byrd, P., Helt, M., Clark, V., Cortes, V., Csomay, E., & Urzua, A. (2004). *Representing language use in the university: Analysis of the TOEFL 2000 Spoken and Written Academic Language corpus*. TOEFL Monograph Series.

Blasingame, M. (2018). Early versus Extended Exposure in Speech and Vocabulary Learning: Evidence from Switched-dominance Bilinguals. Doctoral dissertation, Northwestern University.

Bloomfield, L. (1933). *Language*. New York: Holt, Rinehart & Winston.

Bonfiglio, T. P. (2010). *Mother tongues and nations: The invention of the native speaker* (Vol. 226). Walter de Gruyter.

Bradlow, A. (n.d.) ALLSSTAR: Archive of L1 and L2 Scripted and Spontaneous Transcripts And Recordings. Retrieved from https://speechbox.linguistics.northwestern.edu/allsstar

Bradlow, A. (2002). Confluent talker- and listener-related forces in clear speech production. In C. Gussenhoven & N. Warner (Eds.), Laboratory Phonology 7 (pp. 241-273). Berlin, Germany: Mouton de Gruyter.

Cameron, D. (1995). *Verbal Hygiene*. London, United Kingdom: Routledge.

Carreira, M. (2003). Profiles of SNS students in the twenty-first century. *Mi lengua: Spanish as a heritage language in the United States*, 51-77.

Carreira, M. (2004). Seeking explanatory adequacy: A dual approach to understanding the term "heritage language learner". *Heritage language journal*, *2*(1), 1-25.

Carreira, M., & Kagan, O. (2011). The results of the National Heritage Language Survey: Implications for teaching, curriculum design, and professional development. *Foreign Language Annals*, *44*(1), 40-64. doi:10.1111/j.1944- 9720.2010.01118.x

Carter, P. M. (2018). Spanish in US language policy and politics. In *The Routledge handbook of Spanish as a heritage language* (pp. 36-52). Routledge.

Carvalho, A. M. (2012). Code-switching: From theoretical to pedagogical considerations. *Spanish as a heritage language in the United States: The state of the field*, 139-157.

Cashman, H. R. (2009). The dynamics of Spanish maintenance and shift in Arizona: Ethnolinguistic vitality, language panic and language pride. *Spanish in Context*, *6*(1), 43-68.

Chappell, W. (2019). The sociophonetic perception of heritage Spanish speakers in the United States. *Recent advances in the study of Spanish sociophonetic perception*, *21*, 239-264.

Chappell, W. (2020). Perceptions of Spanish(es) in the United States: Mexicans' sociophonetic evaluations of [v]. In *Spanish in the US: Variation, Attitudes, and Pedagogy*. Edited by Scott M. Alvord and Gregory L. Thompson. New York: Routledge.

Citrin, J., Levy, M., & Wong, C. J. (2017). Politics and the English language in California: Bilingual education at the polls. *California Journal of Politics and Policy*, *9*(2).

Cobas, J. A., & Feagin, J. R. (2008). Language oppression and resistance: The case of middle class Latinos in the United States. *Ethnic and Racial Studies*, *31*(2), 390-410.

Cobas, J. A. & Feagin, J. R. (2022). Language Oppression and Resistance: The Case of Middle Class Latinos in the United States. In J. A. Cobas, B. Urciuoli, J. R. Feagin, & D. J. Delgado (Eds.), *The Spanish language in the United States: Rootedness, racialization, and resistance* (pp. 12–30). Routledge.

Combs, M. C., & Jaime-Diaz, J. (2022). Critical languages as capital and commodity in Arizona. *Kostas Skordoulis, Labrina Gioti*.

Correa, M. (2011). Advocating for Critical Pedagogical Approaches to Teaching Spanish as a Heritage Language: Some Considerations. Foreign Language Annals, 44(2), 308-320.

Covington, M. A., & McFall, J. D. (2010). Cutting the Gordian knot: The moving-average type–token ratio (MATTR). *Journal of quantitative linguistics*, *17*(2), 94-100.

Cummins, J. (2005). A proposal for action: Strategies for recognizing heritage language competence as a learning resource within the mainstream classroom. *Modern Language Journal*, 585-592.

Cuza, A., & López Otero, J. C. (2016). The acquisition of the semantic values of the Spanish present tense in L2 and heritage Spanish. *Revista Española de Lingüística Aplicada/Spanish Journal of Applied Linguistics. Published under the auspices of the Spanish Association of Applied Linguistics*, *29*(2), 462-486.

De Houwer, A., Bornstein, M. H., & Putnick, D. L. (2014). A bilingual–monolingual comparison of young children's vocabulary size: Evidence from comprehension and production. *Applied psycholinguistics*, *35*(6), 1189-1211.

Dick, H. P. (2011). Language and migration to the United States. *Annual Review of Anthropology*, *40*, 227-240.

Domínguez, L., Hicks, G., & Slabakova, R. (2019). Choice of words matters, but so does scientific accuracy: Reply to peer commentaries. *Studies in Second Language Acquisition*, *41*(2), 283-286.

Ducar, C. (2012). SHL learners' attitudes and motivations: Reconciling opposing forces. *Spanish as a heritage language in the United States: The state of the field*, 161-178.

Eddington, D. (2011). What are the contextual variants of [β, ð γ] in colloquial Spanish? Probus, 23, 1-19.

Elabdali, R., Wein, S., & Ortega, L. (2022). Can adult lexical diversity be measured bilingually? A proof-of-concept study. Bilingual Writers and Corpus Analysis, 121.

Face, T. L. (2003). Intonation in Spanish declaratives: differences between lab speech and spontaneous speech. Catalan journal of linguistics, 2, 115-131.

Fairclough, N. (1992). The appropriacy of 'appropriateness.'In N. Fairclough (Ed.), Critical language awareness (pp. 33-56). London: Longman.

Fairclough, M. (2006). Spanish/English interaction in US Hispanic heritage learners' writing. In Globalization and Language in the Spanish-speaking World: Macro and Micro Perspectives (pp. 76-93). London: Palgrave Macmillan UK.

Fairclough, M. (2013). Muestras de lexicalización en estudiantes hispanos bilingües en los Estados Unidos: Un estudio exploratorio. Revista internacional d'humanitats, 27, 105-114.

Fairclough, M., & Garza, A. (2018). The lexicon of Spanish heritage language speakers. *The Routledge handbook of Spanish as a heritage language*, 178-189.

Fishman, J. (1965). Who speaks what language to whom and when? *Linguistique,* 2, 67-88.

Fishman, J. A., Nahirny, V. C., Hofman, J. E., Hayden, R. G., Warshauer, M. E., Kloss, H., Lemaire, H. B., Christian, C. C., Christian, J., & Glazer, N. (1966). *Language loyalty in the United States : the maintenance and perpetuation of non-English mother tongues by American ethnic and religious groups*. Mouton & Co.

Fishman, J. (1972). *The Sociology of Language.* Rowley, MA: Newbury.

Fishman, J. (2001). 300-Plus years of heritage language education in the United States. In J. Peyton, D.A. Ranard, & McGinnis (Eds.). *Heritage Languages in America: Preserving a National Resource* (pp. 81-97). Washington, DC, and McHenry, IL: Center for Applied Linguistics and Delta Systems, Inc.

Flores, N., & Rosa, J. (2015). Undoing appropriateness: Raciolinguistic ideologies and language diversity in education. Harvard educational review, 85(2), 149-171.

Fuller, J. M., & Leeman, J. (2020). *Speaking Spanish in the US: The Sociopolitics of Language*. Bristol: Channel View Publications.

Funk, C. & Hugo Lopez, M. (2022). A Brief Statistical Portrait of U.S. Hispanics. Pew Research Center Science & Society (blog). June 14. Available online: https://www.pewresearch.org/science/2022/06/14/a-brief-statistical-portrait-of-u-s-hispanics/ (accessed on 25 July 2026).

Fusté-Herrmann, B., Silliman, E. R., Bahr, R. H., Fasnacht, K. S., & Federico, J. E. (2006). Mental state verb production in the oral narratives of English-and Spanish-speaking preadolescents: An exploratory study of lexical diversity and depth. Learning Disabilities Research & Practice, 21(1), 44-60.

Gainer, J. S., & Larrotta, C. (2010). Reproducing and Interrupting Subtractive Schooling in Teacher Education. *Multicultural Education*, *17*(3), 41-47.

Gándara, P., & Orfield, G. (2012). Why Arizona matters: The historical, legal, and political contexts of Arizona's instructional policies and US linguistic hegemony. *Language Policy*, *11*(1), 7-19.

García, O., & Li W. (2014). Translanguaging: Language, bilingualism and education. London: Palgrave Macmillan Pivot.

Gardner, R. C., & Tremblay, P. F. (1994). On motivation: Measurement and conceptual considerations 1. *The Modern Language Journal*, *78*(4), 524-527.

Garza, A. (2013). El registro académico en estudiantes hispanohablantes de herencia. Unpublished doctoral dissertation. University of Houston, Houston, TX.

Gathercole, V. C. M., Thomas, E. M., Roberts, E., Hughes, C., & Hughes, E. K. (2013). Why assessment needs to take exposure into account: Vocabulary and grammatical abilities in bilingual children. In V. C. Mueller Gathercole (Ed.), Issues in the assessments of bilinguals (pp. 20-55). Bristol: Multilingual Matters.

Giancaspro, D. (2019). Over, under and around: Spanish heritage speakers' production (and avoidance) of subjunctive mood. Heritage Language Journal, 16(1), 44-70.

Giancaspro, D., Perez-Cortes, S., & Higdon, J. (2022). (Ir) regular mood swings: Lexical variability in heritage speakers' oral production of subjunctive mood. Language Learning, 72(2), 456-496.

Giles, H. & Trudgill, P. (1983). Sociolinguistics and linguistic value judgements. In *On Dialect*. Oxford: Basil Blackwell.

Gómez García, E. (2022). Examining the Influence of Spanish Heritage Language Learners' Critical Language Awareness (CLA) Development on Ethnic Identity Formation. *Languages*, *7*(3), 190.

González, C. (2006). The phonetics and phonology of spirantization in North-Central Peninsular Spanish. Anuario del Seminario de Filología Vasca" Julio de Urquijo", 40(1-2), 409-436.

González-Bueno, M. (1997). Voice-onset-time in the perception of foreign accent by native listeners of Spanish. IRAL: International Review of Applied Linguistics in Language Teaching, 35(4), 251.

Grosjean, F. (1989). Neurolinguists, beware! The bilingual is not two monolinguals in one person. Brain and language, 36(1), 3-15.

Grosjean, F. (2012). An attempt to isolate, and then differentiate, transfer and interference. *International Journal of Bilingualism, 16*(1), 11-21.

Grosjean, F. (2016). The Complementarity Principle and its impact on processing, acquisition, and dominance. *Language dominance in bilinguals: Issues of measurement and operationalization*, 66-84.

Haman, E., Łuniewska, M., & Pomiechowska, B. (2015). Designing cross-linguistic lexical tasks (CLTs) for bilingual preschool children. Assessing multilingual children: Disentangling bilingualism from language impairment, 13, 196-240.

Haner, J., Lopez, M. H., (2023). 8 facts about recent Latino immigrants to the U.S. *Pew Research Center.* https://www.pewresearch.org/

Harmegnies, B., & Poch-Olivé, D. (1992). A study of style-induced vowel variability: Laboratory versus spontaneous speech in Spanish. Speech communication, 11(4-5), 429- 437.

Hay, J. & Drager, K. (2010). Stuffed toys and speech perception. *Linguistics* 48: 865–92.

Helmer, K. A. (2013). A Twice-Told Tale: Voices of Resistance in a Borderlands Spanish Heritage Language Class. *Anthropology & Education Quarterly*, *44*(3), 269-285.

Hill, J. (2001). "The racializing function of language panics". *Language ideologies: Critical perspectives on the Official English Movement, vol. 2* ed. by Roseann Dueñas González and Ildikó Melis, 245–267. Mahwah, New Jersey: Lawrence Erlbaum.

Hualde, J. I., Shosted, R., & Scarpace, D. (2011). Acoustics and Articulation of Spanish/d/Spirantization. In ICPhS (pp. 906-909).

Hudgens Henderson, M. (2022). Critical language awareness in the Spanish as a heritage language college classroom. *Languages*, *7*(3), 157.

Hulstijn, J. H. (2015). Language Proficiency in Native and Non-Native Speakers: Theory and Research. Language Learning & Language Teaching. Volume 41. John Benjamins Publishing Company.

Huntington, S. P. (2004). *Who are we?: The challenges to America's national identity*. Simon and Schuster.

Jacobson, P. F., & Walden, P. R. (2013). Lexical diversity and omission errors as predictors of language ability in the narratives of sequential Spanish–English bilinguals: A cross-language comparison. American Journal of Speech-Language Pathology, 22(3), 554-565.

Jenkins, D. (2018). Spanish language use, maintenance, and shift in the United States. *The Routledge handbook of Spanish as a heritage language*, *53*, 65.

Johnson, K. R., & Martinez, G. A. (2000). Discrimination by proxy: The case of proposition 227 and the ban on bilingual education. *UC davis L. rev.*, *33*, 1227.

Kachru, B. B. (2017). *World Englishes and culture wars*. Cambridge University Press.

Kang, O. & Rubin, D.L. (2009). Reverse Linguistic Stereotyping: Measuring the Effect of Listener Expectations on Speech Evaluation. *Journal of Language and Social Psychology* 28: 441–56.

Katznelson, N., & Bernstein, K. A. (2017). Rebranding bilingualism: The shifting discourses of language education policy in California's 2016 election. *Linguistics and Education*, *40*, 11-26.

Kelly, L. B. (2018). Interest convergence and hegemony in dual language: Bilingual education, but for whom and why?. *Language Policy*, *17*(1), 1-21.

Kim, J. Y. (2020). Discrepancy between heritage speakers' use of suprasegmental cues in the perception and production of Spanish lexical stress. Bilingualism: Language and Cognition, 23(2), 233-250.

Kisselev O. (2021). Corpus-Based Methodologies in the Study of Heritage Languages. In: Montrul S, Polinsky M, eds. The Cambridge Handbook of Heritage Languages and Linguistics. Cambridge Handbooks in Language and Linguistics. Cambridge University Press; 520-544.

Kloss, H. (1977). *The American bilingual tradition.* Rowley, MA: Newbury House.

Knightly, L. M., Jun, S. A., Oh, J. S., & Au, T. K. F. (2003). Production benefits of childhood overhearing. *The Journal of the Acoustical Society of America*, *114*(1), 465-474.

Kroll, J. F., Dussias, P. E., Bice, K., & Perrotti, L. (2015). Bilingualism, mind, and brain. Annu. Rev. Linguist., 1(1), 377-394.

Kroskrity, P. V. (2005). Language ideologies. *A companion to linguistic anthropology*, 496-517.

Kupisch, T. (2013). A new term for a better distinction? A view from the higher end of the proficiency scale. Theoretical Linguistics, 39(3-4), 203-214.

Kupisch, T., & Rothman, J. (2018). Terminology matters! Why difference is not incompleteness and how early child bilinguals are heritage speakers. International Journal of Bilingualism, 22(5), 564-582.

Kutlu, E. 2020. Now You See Me, Now You Mishear Me: Raciolinguistic accounts of speech perception in different English varieties. *Journal of Multilingual and Multicultural Development*, 1–15.

Kyle, K., Crossley, S. A., & Jarvis, S. (2021). Assessing the validity of lexical diversity using direct judgements. Language Assessment Quarterly 18(2), pp. 154-170.

Kyle, K., Sung, H., Eguchi, M., & Zenker, F. (2024). Evaluating evidence for the reliability and validity of lexical diversity indices in L2 oral task responses. Studies in Second Language Acquisition, 46(1), 278-299.

Lawton, R. (2013). Speak English or go home: The anti-immigrant discourse of the American 'English only' movement. *Critical approaches to discourse analysis across disciplines*, *7*(1), 100-122.

Leeman, J. (2005). Engaging critical pedagogy: Spanish for native speakers. *Foreign Language Annals* 38: 35–45.

Leeman, J. (2012). Investigating language ideologies in Spanish as a heritage language. In *Spanish as a Heritage Language in the United States: The State of the Field*. Edited by Sara Beaudrie and Marta Fairclough. Washington, DC: Georgetown University Press, pp. 43–59.

Leeman, J. & Serafini, E.J. (2016). Sociolinguistics for Heritage Language Educators and Students. In *Innovative Strategies for Heritage Language Teaching: A Practical Guide for the Classroom*. Edited by Marta Fairclough and Sara Beaudrie. Washington, DC: Georgetown University Press, pp. 56–79.

Levis, J. (2020). Revisiting the intelligibility and nativeness principles. *Journal of Second Language Pronunciation* 6: 310–28.

Li, W. (2018). Translanguaging as a practical theory of language. *Applied linguistics*, *39*(1), 9-30.

Library of Congress. (2023, December 14). *1917: Jones-Shafroth Act* in A Latinx Resource Guide: Civil Rights Cases and Events in the United States. https://guides.loc.gov/latinx-civil-rights/jones-shafroth-act.

López Otero, J. C., Hur, E., & Goldin, M. (2024). Syntactic optionality in heritage Spanish: How patterns of exposure and use affect clitic climbing. International Journal of Bilingualism, 28(3), 531-554.

Lozano, R. (2018). *An American Language: The History of Spanish in the United States* (Vol. 49). Univ of California Press.

Lozano, R. (2022). The Early Political History of Spanish in the United States. In J. A. Cobas, B. Urciuoli, J. R. Feagin, & D. J. Delgado (Eds.), *The Spanish language in the United States: Rootedness, racialization, and resistance* (pp. 33–47). Routledge.

Lynch, A. (2008). The Linguistic Similarities of Spanish Heritage and Second Language Learners. *Foreign Language Annals*, *41*(2), 252–381.

Lynch, A. (2018). A Historical View of US Latinidad and Spanish as a Heritage Language. In K. Potowski (Ed.), *The Routledge handbook of Spanish as a heritage language* (pp. 17–35). Routledge.

MacGregor-Mendoza, P. (1998). The criminalization of Spanish in the United States. *Language legislation and linguistic rights*, 55-67.

MacGregor-Mendoza, P. (2000). Aquí no se habla Español: Stories of linguistic repression in southwest schools. *Bilingual Research Journal*, *24*(4), 355-367.

Magaña, D. (2015). From pedagogy to communities: Issues within and beyond the Spanish heritage language classroom. *Studies in Hispanic and Lusophone Linguistics*, *8*(2), 375-388.

Martínez, G. A. (2003). Classroom Based Dialect Awareness in Heritage Language Instruction: A Critical Applied Linguistic Approach. *Heritage Language Journal* 1: 1–14.

Martínez, G. A. (2006). *Mexican Americans and language: Del dicho al hecho*. University of Arizona Press.

Martínez, G. A. (2012). Policy and planning research for Spanish as a heritage language: From language rights to linguistic resource. *Spanish as a heritage language in the United States: The state of the field*, 61-100.

McCarthy, P. M., & Jarvis, S. (2010). MTLD, vocd-D, and HD-D: A validation study of sophisticated approaches to lexical diversity assessment. Behavior research methods, 42(2), 381-392.

McGowan, K. B. 2015. Social expectation improves speech perception in noise. *Language and Speech* 58: 502–21.

Meisel, J. M. (2013). Heritage language learners: Unprecedented opportunities for the study of grammars and their development?. Theoretical Linguistics, 39(3-4), 225-236.

Montrul, S. (2002). Incomplete acquisition and attrition of Spanish tense/aspect distinctions in adult bilinguals. *Bilingualism: Language and cognition*, *5*(1), 39-68.

Montrul, S. (2004). Subject and object expression in Spanish heritage speakers: A case of morphosyntactic convergence. Bilingualism: Language and cognition, 7(2), 125-142.

Montrul, S. (2008). *Incomplete Acquisition in Bilingualism: Re-examining the Age Factor*. Amsterdam: John Benjamins.

Montrul, S. (2009). Knowledge of tense-aspect and mood in Spanish heritage speakers. International Journal of bilingualism, 13(2), 239-269.

Montrul, S. (2010). Current issues in heritage language acquisition. *Annual review of applied linguistics*, *30*, 3-23.

Montrul, S. (2013). Bilingualism and the heritage language speaker. *The Handbook of Bilingualism and Multilingualism*, 168-189.

Montrul, S. (2014). Structural changes in Spanish in the United States: Differential object marking in Spanish heritage speakers across generations. Lingua, 151, 177–196.

Montrul, S., & Bowles, M. (2009). Back to basics: Incomplete knowledge of Differential Object Marking in Spanish heritage speakers. *Bilingualism: Language and Cognition*, *12*(3), 363-383.

Montrul, S., & Foote, R. (2014). Age of acquisition interactions in bilingual lexical access: A study of the weaker language of L2 learners and heritage speakers. *International Journal of bilingualism*, *18*(3), 274-303.

Montrul, S., & Perpiñán, S. (2011). Assessing differences and similarities between instructed heritage language learners and L2 learners in their knowledge of Spanish tense-aspect and mood (TAM) morphology. *Heritage Language Journal*, *8*(1), 90-133.

Montrul, S., & Polinsky, M. (Eds.). (2021). *The Cambridge handbook of heritage languages and linguistics*. Cambridge University Press.

Moreno-Fernández, F. (2007). Anglicismos en el léxico disponible de los adolescentes hispanos de Chicago. In Spanish in contact: Policy, social and linguistic inquiries (pp. 41-58). John Benjamins Publishing Company.

Muñoz-Muñoz, E. R., Poza, L. E., & Briceño, A. (2023). Critical translingual perspectives on California multilingual education policy. *Educational Policy*, *37*(6), 1791-1817.

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

National Heritage Language Resource Center. (n.d.). *About the Center*. https://www.nhlrc.ucla.edu/

Nagle, C. L. & Huensch, A. (2020). Expanding the scope of L2 intelligibility research: Intelligibility, comprehensibility, and accentedness in L2 Spanish. *Journal of Second Language Pronunciation* 6: 329–51.

Nichols, S. P. (1989). *The official English movement in the United States with special reference to New Mexico and Arizona*. The University of New Mexico. Unpublished dissertation.

Niedzielski, N. (1999). The effect of social information on the perception of sociolinguistic variables. *Journal of Language and Social Psychology* 18: 62–85.

Nilsson, M., Soli, S. D., & Sullivan, J. A. (1994). Development of the Hearing in Noise Test for the measurement of speech reception thresholds in quiet and in noise. *The Journal of the Acoustical Society of America* 95: 1085–99.

Octavio, C. (2018). Hispanic children thrown into limbo: Language ideologies of Spanish heritage speakers and their English peers. Available from ProQuest Dissertations & Theses Global. (2102586304).

Otheguy, R., García, O., & Reid, W. (2015). Clarifying translanguaging and deconstructing named languages: A perspective from linguistics. Applied linguistics review, 6(3), 281-307.

Otheguy, R., Zentella, A. C., & Livert, D. (2007). Language and dialect contact in Spanish in New York: Toward the formation of a speech community. *Language*, *83*(4), 770-802.

Paradis, J. (2023). Sources of individual differences in the dual language development of heritage bilinguals. Journal of Child Language, 50(4), 793-817.

Parodi, C. (2008). Stigmatized Spanish inside the classroom and out. *Heritage language education: A new field emerging*, 199-214.

Pascual y Cabo, D. (2015). Issues in Spanish heritage morphosyntax. Studies in Hispanic and Lusophone Linguistics, 8(2), 389-401.

Pascual y Cabo, D. (2016). Charting the past, present, and future of Spanish heritage language research. *Advances in Spanish as a heritage language*, *49*, 1.

Pascual y Cabo, D. (2020). Examining the role of cross-generational attrition in the development of Spanish as a heritage language: Evidence from gustar -like verbs. *Linguistic Approaches to Bilingualism*, *10*(1), 86–108.

Pascual y Cabo, D. P., & Rothman, J. (2012). The (il) logical problem of heritage speaker bilingualism and incomplete acquisition. *Applied linguistics*, *33*(4), 450-455.

Peace, M. (2024). Heritage Spanish in the News: Understanding, Reevaluating, and Embracing One's Heritage Language on the National Stage. In *Advocating for Sociolinguistic Justice in the United States* (pp. 57-87). Routledge.

Pearson, B. Z., Fernández, S. C., & Oller, D. K. (1993). Lexical development in bilingual infants and toddlers: Comparison to monolingual norms. *Language learning*, *43*(1), 93-120.

Pearson, B. Z., & Fernández, S. C. (1994). Patterns of interaction in the lexical growth in two languages of bilingual infants and toddlers. *Language learning*, *44*(4), 617-653.

Pedraza, S. (2004). Los Marielitos of 1980: race, class, gender and sexuality. *Cuba In Transition*, 89-102.

Polinsky, M., & Kagan, O. (2007). Heritage languages: In the 'wild' and in the classroom. *Language and Linguistics Compass*, *1*, 368-395.

Polinsky, M. (2008). Gender under incomplete acquisition: Heritage speakers' knowledge of noun categorization. Heritage Language Journal 6(1): 40-71.

Porretta, V., Buchanan, L., & Järvikivi, J. (2020). When processing costs impact predictive processing: The case of foreign-accented speech and accent experience. *Atten Percept Psychophys* 82: 1558–65.

Portes, A., & Hao, L. (1998). Bilingualism and language loss in the second generation. *Sociology of Education*, *71*(4), 269-294.

Portes, A., & Schauffler, R. (1994). Language and the second generation: Bilingualism yesterday and today. *International migration review*, *28*(4), 640-661.

Portes, A. and Rumbaut, R.G. (2006) Immigrant America: A Portrait (3rd ed.). Berkeley: University of California Press.

Potowski, K. (2002). Experiences of Spanish heritage speakers in university foreign language courses and implications for teacher training. *Adfl Bulletin*, *33*(3), 35-42.

Potowski, K., & Carreira, M. (2010). Spanish in the USA. In K. Potowski (Ed.), *Language Diversity in the USA* (pp. 66–80).

Potowski, K., Jegerski, J., & Morgan-Short, K. (2009). The effects of instruction on linguistic development in Spanish heritage language speakers. Language learning, 59(3), 537-579.

Potowski, K. (2012). Identity and heritage learners: Moving beyond essentializations. Spanish as a heritage language in the United States: The state of the field, 179-202.

Potowski, K. (2013). Heritage Learners of Spanish. In K. L. Geeslin (Ed.), *The Handbook of Spanish Second Language Acquisition* (pp. 404–422). John Wiley & Sons, Ltd.

Potowski, K. (2018). Spanish as a heritage/minority language: A multifaceted look at ten nations. In *The Routledge handbook of Spanish as a heritage language* (pp. 1-14). Routledge.

Prado, A. S. (2011). Las aproximantes [β, δ, γ] del español en habla espontánea. Phonica, 7, 118-140.

Quan, T. (2020). Critical language awareness and L2 learners of Spanish: An action-research study. *Foreign Language Annals* 53: 897–919.

Randolph Jr, L. J. (2017). Heritage language learners in mixed Spanish classes: Subtractive practices and perceptions of high school Spanish teachers. *Hispania*, *100*(2), 274-288.

Rao, R. (2014). On the status of the phoneme/b/in heritage speakers of Spanish. *Sintagma: revista de lingüística: 26, 2014*, 37-54.

Rao, R. (2015). Manifestations of/bdg/in heritage speakers of Spanish. *Heritage Language Journal*, *12*(1), 48-74.

Rao, R., & Ronquest, R. (2015). The heritage Spanish phonetic/phonological system: Looking back and moving forward. *Studies in Hispanic and Lusophone Linguistics*, *8*(2), 403-414.

Rao, R., & Kuder, E. (2016). Research on heritage Spanish phonetics and phonology: Pedagogical and curricular implications. *Journal of New Approaches in Educational Research*, *5*(2), 99-106.

Rodriguez, E., & Brandl, A. (2019). Knowledge of Tense, Aspect and Mood in Heritage Language Speakers. Florida Linguistics Papers, 6(1), 64-75.

Ronquest, R., & Rao, R. (2018). Heritage Spanish phonetics and phonology. The Routledge Handbook on Spanish as a Heritage Language. London: Routledge, 164-177.

Rosa, J. D. (2016). Standardization, racialization, languagelessness: Raciolinguistic ideologies across communicative contexts. *Journal of Linguistic Anthropology*, *26*(2), 162-183.

Rosa, J. D. (2019). “I Heard That Mexicans are Hispanic and Puerto Ricans are Latino” Ethnoracial Contortions, Diasporic Imaginaries, and Institutional Trajectories. In *Looking like a language, sounding like a race*. Oxford: Oxford University Press, pp.71–101.

Rosa, J. D. (2019). “They’re Bilingual… That Means They Don’t Know the Language”: The Ideology of Languagelessness in Practice, Policy, and Theory. In *Looking Like a Language, Sounding Like a Race*. Oxford: Oxford University Press, pp. 125–43.

Rosa, A., Aranda, E., & Dotson, H. (2022). “You Are Not Allowed to Speak Spanish! This Is an American Hospital”: Puerto Rican Experiences with Domestic Discrimination 1. In *The Spanish Language in the United States* (pp. 96-110). Routledge.

Rosa, J. D. & Flores, N. (2017). Unsettling race and language: Toward a raciolinguistic perspective. *Language in society*, *46*(5), 621-647.

Rosa, J. D. & Flores, N. (2023). “Rethinking Language Barriers & Social Justice from a Raciolinguistic Perspective.” *Daedalus* 152, no. 3: 99–114.

Rothman, J. (2007). Heritage speaker competence differences, language change, and input type: Inflected infinitives in Heritage Brazilian Portuguese. *International Journal of Bilingualism*, *11*(4), 359-389.

Rothman, J. (2009). Understanding the nature and outcomes of early bilingualism: Romance languages as heritage languages. *International journal of bilingualism*, *13*(2), 155-163.

Rothman, J., Tsimpli, I. M., & Pascual y Cabo, D. (2016). Formal linguistic approaches to heritage language acquisition. *Advances in Spanish as a heritage language*, *441*, 13-26.

Rothman, J., Bayram, F., DeLuca, V., Di Pisa, G., Duñabeitia, J. A., Gharibi, K., Hao, J., Kolb, N., Kubota, M., Kupisch, T., Laméris, T., Luque, A., van Osch, B., Pereira Soares, S. M., Prystauka, Y., Tat, D., Tomíc, A., Voits, T., & Wulff, S. (2023). Monolingual comparative normativity in bilingualism research is out of “control”: Arguments and alternatives. Applied Psycholinguistics, 44(3), 316-329.

Rubin, D. L. (1992). Nonlanguage factors affecting undergraduates’ judgments of nonnative English-speaking teaching assistants. *Research in Higher Education* 33: 511–31.

Rumbaut, R. G., Massey, D. S., & Bean, F. D. (2006). Linguistic life expectancies: Immigrant language retention in Southern California. *Population and development review*, 447-460.

Rumbaut, R. G. (2009). A language graveyard? The evolution of language competencies, preferences and use among young adult children of immigrants. *The education of language minority immigrants in the United States*, *35*, 71.

Rumbaut, R. G., & Massey, D. S. (2013). Immigration & language diversity in the United States. *Daedalus*, *142*(3), 141-154.

San Miguel, G. (2004). Contested Policy: The Rise and Fall of Federal Bilingual Education in the United States, 1960-2001. United States: University of North Texas Press.

Sánchez-Muñoz, A. (2010). Different words for different contexts: Intra-speaker variation in Spanish as a heritage language. Spanish of the Southwest: A language in transition, 337-352.

Sánchez-Muñoz, A. (2016). Heritage language healing? Learners' attitudes and damage control in a heritage language classroom. *Advances in Spanish as a heritage language*, *49*, 205-217.

Schmid, P. M., & Yeni-Komshian, G. (1999). The effects of speaker accent and target predictability on perception of mispronunciations. *Journal of Speech, Language, and Hearing Research* 42: 56–64

Schmid, C. L. (2001). The Politics of Language: Conflict, Identity and Cultural Pluralism in Comparative Perspective. United Kingdom: Oxford University Press.

Schoonmaker-Gates, E. (2015). Measuring foreign accent in Spanish: How much does VOT really matter. In *Selected Proceedings of the 6th Conference on Laboratory Approaches to Spanish Phonetics and Phonology* (pp. 95-105).

Schwarzer, D., & Petrón, M. (2005). Heritage language instruction at the college level: Reality and possibilities. *Foreign Language Annals*, *38*(4), 568-578.

Seijas, J. M., & Spino, L. (2023). Written "corrective" feedback in Spanish as a heritage language: Problematizing the construct of error. *Journal of Second Language Writing*, *60*, 100989.

Showstack, R. E. (2017). Stancetaking and language ideologies in heritage language learner classroom discourse. *Journal of Language, Identity & Education*, *16*(5), 271-284.

Silva-Corvalán, C. (1994). The gradual loss of mood distinctions in Los Angeles Spanish. *Language Variation & Change*, *6*(3).

Silva-Corvalán, C. (2004). Spanish in the Southwest. *Language in the USA: Themes for the twenty-first century*, 205-229.

Taylor Reid, K., Trofimovich, P., & O'Brien, M. G. (2019). Social Attitudes and Speech Ratings. *Studies in Second Language Acquisition* 41: 419–42

Templin, M. C. (1957). *Certain language skills in children; their development and interrelationships.* University of Minnesota Press.

Thiéry, C. (1978). True bilingualism and second-language learning. In D. Gerver & H. W. Sinaiko (Eds.), *Language interpretation and communication* (pp. 145–153). Plenum Press.

Toribio, A. J., & Bullock, B. E. (2016). A new look at heritage Spanish and its speakers. Advances in Spanish as a heritage language, 49, 27-50.

Toribio, A. J., Durán, L. (2018). Understanding and Leveraging Spanish Heritage Speakers' Bilingual Practices. In K. Potowski and J. Muñoz-Basols (Eds.), The Routledge handbook of Spanish as a heritage language. (pp.284-298). New York: Routledge.

Torreira, F., & Ernestus, M. (2011). Realization of voiceless stops and vowels in conversational French and Spanish. Laboratory Phonology, 2(2), 331-353.

Tran, V. C. (2010). English gain vs. Spanish loss? Language assimilation among second-generation Latinos in young adulthood. *Social forces*, *89*(1), 257-284.

Treffers-Daller, J. (2019). What defines language dominance in bilinguals?. *Annual Review of Linguistics*, *5*(1), 375-393.

Urciuoli, B. (1985). Bilingualism as code and bilingualism as practice. Anthropological Linguistics, 363-386.

Urciuoli, B. (2008). Whose Spanish. *Bilingualism and identity: Spanish at the crossroads with other languages*, 257-277.

Urciuoli, B., Cobas, J. A., Feagin, J. R., & Delgado, D. J. (2022). Introduction: Language, racialization, and power. In J. A. Cobas, B. Urciuoli, J. R. Feagin, & D. J. Delgado (Eds.), *The Spanish language in the United States: Rootedness, racialization, and resistance* (pp. 3–11). Routledge.

U.S. Census Bureau. "Hispanic or Latino Origin." *American Community Survey, ACS 1-Year Estimates Detailed Tables, Table B03003*, https://data.census.gov/table/ACSDT1Y2024.B03003?g=010XX00US$0400000. Accessed on 19 Jan 2026.

Valdés, G. (1997). The Teaching of Spanish to Bilingual Spanish-speaking Students: Outstanding. *La enseñanza del español a hispanohablantes: Praxis y teoría*, 8.

Valdés, G. (2001). "Heritage Language Students: Profiles and Possibilities." In *Heritage Languages in America: Preserving a National Resource,* edited by Joy Peyton, Donald Ranard, and Scott McGinnis, McHenry, IL: Center for Applied Linguistics and Delta Systems.

Valdés, G., & Geoffrion-Vinci, M. (2011). Heritage Language Students: The Case of Spanish. In M. Díaz-Campos (Ed.), *The Handbook of Hispanic Sociolinguistics* (pp. 598–622). Wiley-Blackwell.

Valenzuela, A. (1999). *Subtractive Schooling: U. S. - Mexican Youth and the Politics of Caring* (1st ed.). State University of New York Press.

Van Deusen-Scholl, N. (1998). Heritage language instruction: Issues and challenges. AILA Newsletter, 1, 12-14.

Van Deusen-Scholl, N. (2003). Toward a definition of heritage language: Sociopolitical and pedagogical considerations. *Journal of language, identity, and education*, *2*(3), 211-230.

Van Engen, K. J., & Peelle, J. E. (2014). Listening effort and accented speech. *Frontiers in Human Neuroscience* 8: 577.

Veltman, C. J. (1983). *Language shift in the United States* (No. 34). Walter de Gruyter.

Villa, D. J. (2002). The sanitizing of US Spanish in academia. *Foreign Language Annals*, *35*(2), 222-230

Vogel, S., & García, O. (2017). Translanguaging. In G. Noblit & L. Moll (Eds.), Oxford Research Encyclopedia of Education.

Walker, A. & Hay, J. 2011. Congruence between 'word age' and 'voice age' facilitates lexical access. *Laboratory Phonology* 2: 219–37.

Weinreich, U., Labov, W., & Herzog, M. (1968). *Empirical foundations for a theory of language change*. University of Texas Press.

Weiss, B. J. (Ed.). (1982). *American education and the European immigrant: 1840–1940*. University of Illinois Press.

Wiese, H., Alexiadou, A., Allen, S., Bunk, O., Gagarina, N., Iefremenko, K., Martynova, M., Pashkova, T., Rizou, V., Schroeder, C., Shadrova, A., Szucsich, L., Tracy, R., Tsehaye, W., Zerbian, S. & Zuban, Y. (2022). Heritage speakers as part of the native language continuum. Frontiers in psychology, 12, 717973.

Wiley, T. G., & Lukes, M. (1996). English-only and standard English ideologies in the US. *Tesol Quarterly*, *30*(3), 511-535.

Wiley, T. G. (2001). "On Defining Heritage Languages and their Speakers." In *Heritage Languages in America: Preserving a National Resource*, edited by Joy Peyton, Donald Ranard, and Scott McGinnis, McHenry, IL: Center for Applied Linguistics and Delta Systems.

Wiley, T. G. (2014). Diversity, Super-Diversity, and Monolingual Language Ideology in the United States: Tolerance or Intolerance? *Review of Research in Education*, *38*(1), 1–32.

Williams, A., Garrett, P., & Coupland, N. (1999). Dialect recognition. In *Handbook of Perceptual Dialectology*. Edited by Dennis R. Preston. Philadelphia: Benjamins, pp. 345–58.

Yamagami, M. (2012). The political discourse of the campaign against bilingual education: From Proposition 227 to Horne v. Flores. *International Multilingual Research Journal*, *6*(2), 143-159.

Zenker, F., & Kyle, K. (2021). Investigating minimum text lengths for lexical diversity indices. *Assessing Writing*, *47*, 100505.

Zentella, A. C. (1995). The" Chiquitafication" of US Latinos and Their Languages, OR Why We Need an Anthropolitical Linguistics.

Zentella, A. C. (1997). Spanish in new york. The Multilingual Apple: Languages in New York City, 167-201.

Zentella, A. C. (2002). Latin@ languages and identities. *Latinos: Remaking America*, 321-338.

Zentella, A. C. (2003). "José, can you see?" Latin© Responses to Racist Discourse. In *Bilingual games: Some literary investigations* (pp. 51-66). New York: Palgrave Macmillan US.

Zentella, A. C. (2017). "Limpia, fija y da esplendor": Challenging the Symbolic Violence of the Royal Spanish Academy. *Chiricù Journal: Latina/o Literature, Art, and Culture*, *1*(2), 21-42.

Zyzik, E. (2014). Causative verbs in the grammar of Spanish heritage speakers. Linguistic Approaches to Bilingualism, 4(1), 1-33.