PUERTO RICAN ISLAND ENGLISH (PRIE):
ON THE EMERGENCE OF A NEW DIALECT OF AMERICAN ENGLISH

By

Cristopher B. Font-Santiago

A dissertation submitted in partial fulfillment of

the requirements for the degree of

Doctor of Philosophy

(English)

at the

UNIVERSITY OF WISCONSIN – MADISON

2021

Date of final oral examination: 3/26/2021

This dissertation is approved by the following members of the Final Oral Committee:

Thomas Purnell, Professor, Department of English
Joseph Salmons, Professor, Department of Language Sciences
Eric Raimy, Professor, Department of English
Rajiv Rao, Associate Professor, Department of Spanish & Portuguese
Catherine Stafford, Associate Professor, Department of Spanish & Portuguese
Fernando Tejedo-Herrero, Associate Professor, Department of Spanish & Portuguese
Paul Reed, Assistant Professor, Department of Communicative Disorders, University of Alabama

ACKNOWLEDGMENTS

It has been an incredible journey on my way to be the first in the family to earn a Ph.D., and there are so many folks to whom I am grateful that mentioning them all would make this section quite extensive, so I will keep it brief. First of all, I would like to give my co-advisors, Joe and Tom, a special shoutout for your support throughout the years that made all this possible. I would also like to thank my committee members, Eric, Cathy Stafford, Rajiv, Fernando, and Paul, for all their feedback on and contributions towards the making of this project.

Nevin, Cathy Fleck, Liz Dayton—may you rest in peace—NanVan, Nick, Rosita, and everyone else at the English Department at UPRM, I have not forgotten your support throughout my undergraduate years. It was your outstanding dedication to your students and research that reaffirmed my interest to pursue an advanced degree in the first place.

My colleagues, who I had the chance to meet here at UW – Madison and elsewhere, thank you for all the laughs and feedback throughout these last 6 years. My close friends, Bekki, Alan, Ash, Jonathan, Joseph, and my partner for nearly a decade, Brian, thank you for supporting me through thick and thin, for sponging a good chunk of that infamous graduate school stress, and for simply being there.

To my Padrino, Leo, a veteran and Borinqueneer, who could not witness this moment, know that I keep you and all our cherished memories in my thoughts. May you rest in peace. Finally, Papi y Mami, thank you for everything, from raising and supporting me to encouraging me to pursue an education throughout my life. Even though we do not agree on everything, you have grown to become my best friends in adulthood as well. Los amo con todo mi corazón. Gracias. Thank you, everyone.

ABSTRACT

From the beginning of Puerto Rico's colonial history in 1493, and more recently, the transfer of political power over it from Spain to the United States in 1898, Spanish has been the dominant language of culture and policy. This history makes way for a unique linguistic context in which two colonial languages, Spanish—first and mainly—and American English, coexist today. Recent research suggests the emergence of simultaneous Puerto Rican Spanish-English bilinguals in younger generations of speakers (Fabiano-Smith et al., 2014). With the emergence of larger numbers of L1 English speakers on the island in the last few decades and the development of tightly knit communities with shared sociopolitical identities, it is reasonable to expect subsequent dialect formation as a form of marking SOCIAL IDENTITY, not as part of a generational language shift, but rather as a reflection of complex cultural interactions and the amelioration of language attitudes towards American English.

To begin documenting this phenomenon, this dissertation employs approaches and concepts including PERCEPTUAL DIALECTOLOGY, SIGNAL DETECTION THEORY, and ROOTEDNESS to explore the process of ENREGISTERMENT and NEW DIALECT FORMATION in Puerto Rican Island English (PRIE) as an emerging variety of American English. The research framework employed throughout this dissertation partitions the perceptual study into four phases: speaker identification, speaker selection, signal detection for inferential statistical analysis, and descriptive mental maps of listeners' representations of the speaker dialect groups. Phases 1 and 2 are designed to control the speaker selection process, given the range of variability in production that language dominance introduces when considering bilingual speakers. Phases 3 and 4 focus on the analyses of listeners' responses through a set of judgment tasks and a Mental Map Task, drawing on listeners' perception judgments to explore the status of PRIE.

Participants were divided into speakers and listeners. 21 speakers were selected from monolingual English, sequential bilingual Spanish-English and simultaneous bilingual English-Spanish speaker groups. The goal with this broader speaker recruitment is to compare the PRIE speakers to a variety of Spanish-English speakers from the mainland. A total of 338 adult naïve listeners, mainly from the Upper Midwest, responded to three surveys: (1) a Matched-Guise Box Task, (2) a set of Identification and AX Signal Detection Tasks, and (3) a Mental Map Task. Nonparametric statistical analyses were conducted on the Identification and AX Tasks and comparative descriptive analyses were conducted on the results in the Mental Map Task.

The findings in the signal detection tasks confirmed that listeners identified PRIE as distinct from the other speaker groups and that PRIE holds a similar perceptual status as the other simultaneous Spanish-English bilingual varieties of mainland American English. The results from the Mental Map Task indicated that PRIE did not yet appear as a form that was uniquely enregistered to Puerto Rico. Perceptions of PRIE patterned closer with that of the mainland simultaneous bilingual groups that were most associated to the English control speakers, suggesting that PRIE held a similar perceptual status to these other mainland Spanish-English bilingual dialects.

Puerto Rico has been almost entirely viewed as an L1 Spanish/L2 English-speaking community, even with the increasing sociocultural and sociopolitical influence of the mainland United States on the island. This dissertation is the first study to show the rise of a new variety of American English in Puerto Rico that can be heard by listeners in the mainland United States. The findings in this study begin to reshape the conversation on the role of English on the island and dispel some of the myths associated with a lack of knowledge about the linguistic diversity and the changing linguistic landscape of Puerto Rico.

Keywords: Puerto Rico, English, dialect, new dialect formation, perception, bilingualism

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

CHAPTER 1

INTRODUCTION

For over a century, Puerto Rico and the mainland United States have shared a complicated sociopolitical relationship, ranging from short-lived civil unrest in the early to mid-20th century to a complex cultural interaction today. It was not until recent decades that the number of L1 English speakers began to notably increase in younger generations of bilingual Puerto Ricans, partly from the influence of return migrants from the mainland and in greater part from those born and raised in the island, like myself (Fabiano-Smith et al., 2014; U.S. Census Bureau, 2019a). The rise in the number of English speakers, alongside an increasing openness to a Puerto Rican identity that does not require Spanish among younger speakers in Puerto Rico (Zentella, 2003, pp.249-250) and overall more accepting attitudes toward English (Vicente Vélez, 2000; Falcón, 2004; González-Rivera & Ortiz López, 2018), provides the ideal conditions to observe the potential emergence of Puerto Rican Island English (PRIE) as a new L1 English dialect of American English—largely in connection to the sustained sociocultural and ecopolitical influence of the United States over the island.

Both language emergence and NEW DIALECT FORMATION are motivated by a combination of elements, such as national as well as individual sociopolitical identity and sociohistorical factors (Wolfram & Schilling-Estes, 2006, pp.28-64). In the case of Spain, for instance, Castilian Spanish was not recognized as a separate language for years. In that time, speakers believed they spoke a different dialect of the same language, "Latin," even though there was plenty of surface level evidence that indicated substantial divergence from other speakers of "Latin" elsewhere in Europe (Janson, 2004, pp.90-91). In combination with national identity, it was not until the writing revolution, where the oral variety was codified in writing and began competing with

Latin, that Spanish rose to a level of explicit linguistic awareness and its speakers claimed it as their own (Janson, 2004, p.92). This was the case regardless of whether outsiders believed speakers in the Spanish region spoke the same way as other speakers of "Latin" elsewhere.

From a more contemporary perspective, speakers in the Basque Country speak both Basque and Spanish and are simultaneous bilingual speakers of both languages. Yet, even though Basque is, in great part, their cultural language, they also have a recognized variety of Spanish that has its unique features (Ciriza, 2010). I postulate that the kind of surface-level divergence in dialect formation that one would expect does not generally take place unless there is: (1) an identity associated to the language, (2) historicity, and (3) a claim over the connections between (1) and (2). The desire to create a contrast with other dialects of the same language (e.g., Labov's (1972) findings in Martha's Vineyard) is part of the catalyst that fuels NEW DIALECT FORMATION and feature divergence. A similar kind of intentional linguistic boundary setting can be observed in the spelling movement of the original American settlers from Britain, where they underscored linguistic divergence orthographically to assert American English as a distinct national language (Whelan, 2002). A language or dialect, a political distinction in and of itself, does not surface to a conscious level of linguistic awareness until its own speakers claim it as part of their identities, regardless of the number of recognizable or subliminal features that constitute their speech. In the analogical and historical context of this discussion is where the emergence of PRIE in Puerto Rico is situated.

On a personal note, growing up bicultural and bilingual in rural Puerto Rico highlighted an identity and language use gap that developed between monolingual and bilingual social circles, where both English and Spanish were our primary means of written and spoken communication. English has always been an integral part of my identity and my upbringing for

as long as I can remember. My parents often underscored the importance of learning English as a way to access better educational and career opportunities. To me and my bilingual peers, however, it became a connection to a powerful nation that represents access and opportunity. The same rings true for many other Puerto Ricans (Vicente Vélez, 2000, p.38). American English, from the sheer breadth of its global influence to the social and career opportunities it presents to Puerto Ricans, plays an important role. This is the story of many young bilinguals on the island today, who in many ways represent a generational departure from what has been the linguistic norm in Puerto Rico for centuries—one that remains prevalent in the minds of many Americans from the mainland.

Historically, in Puerto Rico, Spanish has been the language of culture and internal exchange, while English was reserved as an alternative language for official purposes and external exchange. This configuration has changed over the years, but it has remained in this official status for decades (Bischoff, 2017, p.287). Furthermore, until recently, Spanish-English bilingualism in Puerto Rico was attested to be reserved for an elite class of Puerto Ricans (Pérez Casas, 2008). However, as it has been noted and as this study shows, the usage of English in Puerto Rico is no longer reserved for an elite class. The results from the pilot study for this dissertation indicated that simultaneous Puerto Rican Spanish-English non-elite bilinguals were largely perceived as native speakers of English.[1] I argue that the emergence of PRIE as a broader phenomenon in the Puerto Rican populace is largely due to a combination of the changing linguistic landscape on the island and the acceptance of a bilingual Puerto Rican identity among some younger bilinguals in Puerto Rico, a region where English use is steadily increasing.[2]

[1] These results are contextualized and discussed in the introduction of the methodology for the present study in Chapter 3.
[2] This development is discussed in Chapter 2, which presents previous research and the most recent census data at the time that this dissertation was written.

It is important to underscore that this observation is not part of a monolithic shift in language use, but rather a continued diversification, where Spanish and English coexist to different extents depending on communities and the individual—all within a range of monolingual Spanish, sequential Spanish-English bilingual, and simultaneous Spanish-English bilingual speakers. The emergence of PRIE in younger simultaneous bilinguals on the island represents a paradigm shift in the understanding of English language usage on the island. As a relatively new and, thus, undiscussed phenomenon, the development of PRIE as a dialect for a subset of bilingual Puerto Ricans on the island emphasizes a notable gap in existing conversations and research—one that this study begins to address from the perspective of World Englishes, NEW DIALECT FORMATION, and ENREGISTERMENT in the unique colonial linguistic context of Puerto Rico.

1.1 Purpose of Study

With the previous discussion in mind, the goals of this dissertation are to observe the emergence of PRIE as a dialect of American English through a combination of methods in PERCEPTUAL DIALECTOLOGY and statistical analyses. PERCEPTUAL DIALECTOLOGY provides us with a guideline to understand how outsider listeners group PRIE speakers, and in what ways these listeners mark PRIE speakers as members of a same linguistic group. As stated earlier, dialect recognition and codification do not happen until patterns of speech rise to a level of awareness that attracts attention to them, and PERCEPTUAL DIALECTOLOGY provides a means to an end, identifying whether PRIE has been enregistered in the minds of listeners from the mainland and to what extent it may compare to other parallel Spanish-English bilingual varieties.

While a yet unrecognized PRIE dialect may have overtly distinguishing features, like most recognized American English dialects, it is also possible that, given the present

sociohistorical context, these features have not just emerged yet. Therefore, the presence of a subliminal accent, much like Wisconsin English in college-aged speakers (Schuld et al., 2017), is also possible. This dissertation observes that PRIE exists in simultaneous Puerto Rican Spanish-English bilinguals as a byproduct of the coexistence of English and Spanish and the ongoing amelioration of language attitudes towards English on the island.

I argue that dialect enregisterment depends on two conditions: that outside listeners notice it as belonging to an insider group of speakers, and that the speakers of the dialect claim it as their own, at which point more salient features can begin to develop to solidify in-group identity and create linguistic contrast. SOCIOPHONETICS informs potentially emerging features, helping us understand how acoustics diverge or coalesce while playing a role in what outsiders perceive. SOCIOPHONETICS also helps us identify whether cohesive dialect formation has begun and, if so, to what extent and in what direction. Both PERCEPTUAL DIALECTOLOGY and SOCIOPHONETICS play a role in dialect enregisterment and codification that cannot be understated, but the object of research must first be recognized before it is explored. This project, therefore, begins to address the research gap by drawing from the former to create a foundation to explore PRIE via the latter at a future time.

1.2 Chapter Roadmap

With the discussion above in mind, this study focuses on observing and analyzing the naïve perception judgments of English listeners mainly from the Upper Midwest (i.e., listeners with relatively little to no exposure to Spanish and/or the English of bilingual Spanish-English speakers), relative to what constitutes a canonical variety of American English to them. Because this group of listeners is unaware of the development of PRIE, positive evidence from these listeners in support of PRIE's perceptual salience as a dialect would present a strong case for its

current emergence and development in simultaneous Spanish-English bilinguals from Puerto Rico.

Chapter 2 introduces additional background information on the history and the current linguistic context of Puerto Rico, relative to the island's sociocultural and ecopolitical history with, first, Spain, and then the United States. These observations are contextualized within the development of bilingualism in Puerto Rican speakers and previous research on the importance of research on perceptions and language attitudes thereof. This discussion also includes a literature review that motivates the theoretical framework of this study, drawing from perspectives in PERCEPTUAL DIALECTOLOGY, NEW DIALECT FORMATION, and ROOTEDNESS.

Chapter 3 details the methodological framework for this study, which includes four phases, controlling for confounds in perceptual studies, while addressing the main research questions that focus on determining the status of PRIE in the minds of listeners from the mainland and in relation to other parallel bilingual dialects. The results from the first two phases focus on speaker selection, and are thus, also included as part of the methods in Chapter 3.

Chapters 4 and 5 present and discuss the results for the latter two phases that focus on answering the research questions presented in Chapter 3 through adapted approaches from SIGNAL DETECTION THEORY and PERCEPTUAL DIALECTOLOGY, respectively. Finally, Chapter 6 summarizes the overall findings and contextualizes them within the literature review from Chapter 2 and the expectations outlined in Chapter 3. The final chapter also outlines this study's contributions to the field, limitations, and future directions.

CHAPTER 2

BACKGROUND & RESEARCH CONTEXT

This chapter explores the emergence of PRIE in Puerto Rico as a phenomenon that largely follows the paradigm of post-colonial World Englishes. The theoretical framework for this project focuses on PERCEPTUAL DIALECTOLOGY, NEW DIALECT FORMATION, and ROOTEDNESS as core concepts for this exploration.

The first section of this chapter provides a historical and contemporary overview of demographic changes in Puerto Rico, starting from its incorporation into the United States as a territory in 1898 to the present day. The second section explores the role of Spanish-English bilingualism in the emergence of PRIE and the specific circumstances through which the current linguistic landscape in Puerto Rico has become possible, as opposed to the historical understanding of Puerto Rico as an almost exclusive L1 Spanish/L2 English-speaking region. The third section expands on the role of bilingualism as a catalyst for the emergence of Englishes around the world, particularly from the perspective of the redefinition of sociopolitical identities and shifting language attitudes. The fourth section investigates previous approaches to PERCEPTUAL DIALECTOLOGY and establishes its function in this project as a tool that provides insight to language and/or dialect enregisterment.

2.1 A Sociohistorical and Sociolinguistic Overview of Puerto Rico

NEW DIALECT FORMATION can be typically observed over generations of speakers after a speech community settles into a region (Kerswill & Trudgill, 2005 in Schuld et al., 2017, p.1). Even though this phenomenon can still be observed in many regions of the United States, such as the Pacific Northwest (Ingle et al., 2005) and in perceptions of Wisconsin English speakers (Schuld et al., 2017), the emergence of PRIE differs in that its development also represents

language enrichment. Since the beginning of Puerto Rico's colonial history in 1493, and more recently, the transference of political power over it from Spain to the United States in 1898, Spanish has largely been the dominant language of culture and policy. Since then, there have been many attempts by the US government to shift Puerto Rico from an L1 Spanish to an L1 English region, such as through local government reforms in 1917 (Rivera, 2019).

Kerswill identified a series of sociohistorical parameters through which dialect and language emergence can be explored: migration, geography, sociocultural factors, and language contact, among others (2005). These factors provide a foundation to understand the context of the linguistic landscape of Puerto Rico, where PRIE is emerging.

2.1.1 Population and Migration

Kerswill underscores that migration has "profound sociolinguistic consequences" as population and demographic changes affect the intracommunity relationships between the sociolinguistic groups in a society (2005, p.30); that is, whenever there are changes in the composition of the speakers within a larger dialect group, changes follow as older and more recent speakers orient themselves and their language use relative to the changes taking place. It is, therefore, important to establish a sociohistorical, economic, and linguistic reference point to better understand the relationship between language use in the island and its shifting population of speakers. Travel between Puerto Rico and the mainland mainly takes place through airports, and the recent diaspora of Puerto Ricans often leads to transient residents, some of whom relocate between the two regions and others who visit (Pousada, 2010).

Since Puerto Rico became a territory of the United States in 1898, migratory flows from the island to the mainland have waxed and waned. Following the patterns of today, migration from the mainland to the island is at a net negative of about -40,000 (Rodríguez, 2000, p.3). Even

though Puerto Rico became a territory of the United States in 1898, it was not until the mid-20th century that the first major wave of Puerto Ricans came to the mainland. The first wave occurred in two steps: the first in the 1950s, which focused on east Harlem in New York City, and the second in the 1960s, where the amount of Puerto Ricans migrating to New York City increased, alongside a further population increase diffusion in other mainland states, particularly in states such as Florida, Illinois, Connecticut, and Pennsylvania (Rodríguez, 2000).

Since the economic crisis in the late 2000s, and more recently, the natural disaster that was Hurricane Maria, a third notable wave of migrants from Puerto Rico to the mainland appears to be taking place (Montalvo & Laughlin, 2017). Puerto Rico has, since, experienced another steady decline in population over the last several years, as illustrated in Figure 2.1, largely motivated by searches for gainful employment and better overall economic opportunities.

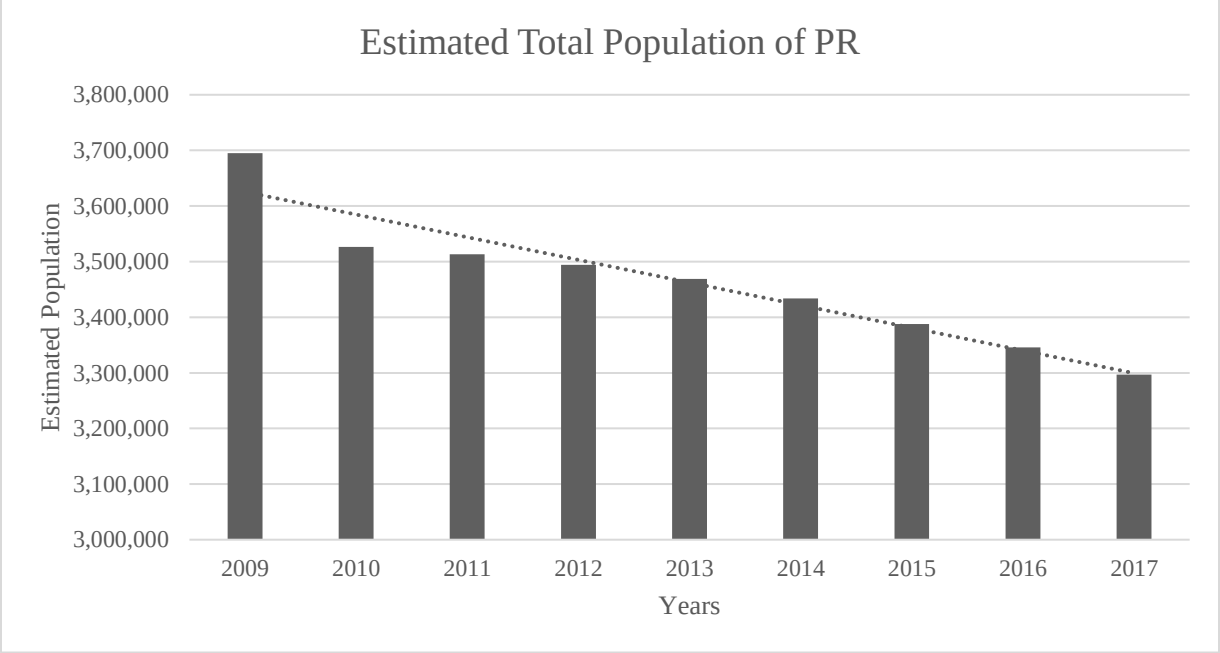


Figure 2.1: Estimated population of Puerto Rico from 2009-2017 (U.S. Census Bureau, 2019a).

The increased sociocultural interaction with the mainland U.S., in part due to the increasing number of Puerto Ricans on the mainland with family still on the island, partly motivates the rise of concurrent Spanish-English bilingualism in Puerto Rico.

2.1.2 Bilingualism in Context

As the field has developed, new definitions of 'bilingualism' (Nash, 1968; Hickling, 2000; Argyri & Sorace, 2007; among others) have been proposed over recent decades. Initially, studies, such as Nash (1968), treated bilingualism as an umbrella term to refer to any speaker who spoke more than one language: "Bilingualism is here defined as the habitual use of two languages" (p.1). However, this definition of bilingualism fails to capture important aspects, such as language dominance, Age of Acquisition (AoA), among a plethora of other factors. A speaker who acquires language(s) in their childhood exhibits a much different outcome from one who acquires an L2 during their late teenage or adult years. Since then, definitions of 'bilingualism' have become more nuanced. For instance, in Argyri & Sorace's study on crosslinguistic influence in bilinguals, language dominance and proficiency were key in participant selection for "early bilinguals" (2007, p.86); other factors, such as the parents' L1 and language history, are also considered important in this then-emergent methodological paradigm.

Hickling (2000) evaluates the utility of distinguishing between SIMULTANEOUS and EARLY/LATE SEQUENTIAL BILINGUALISM, where the former term is observed when a speaker is exposed to both languages at or before around the age of three, and the latter term at a later point in their childhood years or beyond. Hickling underscores that this distinction is not only necessary, but important because the order in which one acquires languages determines whether they develop simultaneously (SIMULTANEOUS BILINGUALISM) or whether the L2 is nested within an already-developing L1 (EARLY SEQUENTIAL BILINGUALISM) (2000, p.40). LATE SEQUENTIAL BILINGUALISM happens after an L1 is either almost or completely acquired; these speakers are referred to as L2 learners.

SEQUENTIAL BILINGUALISM covers all language acquisition that occurs after an L1 is almost entirely or entirely acquired; many of these speakers are often referred to as 'L2 learners' as well. However, 'L2 learners' is often employed to refer to either adult learners of an L2 language, who already acquired an L1 (Montrul, 2014) or younger learners who acquire a target language in institutional settings (Carstens, 2016). In effect, anyone who speaks more than one language can be broadly considered bilingual, and anyone who is accurately described by the 'L2 learners' terminology can also be considered a sequential bilingual to different extents.

Nevertheless, it is important to note that age of exposure (AoE) and AoA are not the sole factors in determining future language proficiency or dominance. As is in the case of many HERITAGE SPEAKERS, who are typically sequential learners, they may be exposed to only one language in their early childhood but receive no community support for their L1 or they switch out of it entirely in preference for an L2. The resulting outcome is a sequential bilingual who is highly dominant in their L2 over their L1; this phenomenon is discussed at length in Benmamoun et al. (2013), who characterize HERITAGE SPEAKERS narrowly as asymmetrical bilinguals who acquired an L1 in their childhood but speak a different language as adults.

Depending on AoE, AoA, and the linguistic environments in which a child is raised, and/or a learner acquires (a) language(s), different language dominance outcomes are possible. To summarize, SEQUENTIAL BILINGUALISM takes place when a learner fully acquires a first language and then learns a second; EARLY SEQUENTIAL BILINGUALISM can occur soon after a first language is acquired and LATE SEQUENTIAL BILINGUALISM much after. Dominance and identity play notable roles in a speaker's ultimate language outcome at a given point in their lives, as can be exemplified in heritage language speakers, who most of the time are sequential bilinguals.

SIMULTANEOUS BILINGUALISM can emerge from different environment TYPES, where the speaker has >1 L1, while proficiency can still depend on dominance and other factors.

With the understanding that simultaneous bilinguals can have multiple L1s, the role that these speakers would play in the potential emergence of PRIE in a bilingual region cannot be understated. Simultaneous bilinguals, especially those with balanced or more English-dominant proficiencies within a historically Spanish-dominated region, are crucial to observing the potential emergence of PRIE.

Furthermore, Figure 2.2 (Pew Research Center, 2014) compares the population of Puerto Ricans living on the island versus in the mainland over the last four decades. Juxtaposing the information from both Figures 2.1 and 2.2 provides a clearer picture on the correlation between migration to the mainland and the declining population trend. Of course, even with a larger amount of migration to the mainland, there are those who return or relocate to the island from the mainland. Based on statistics from the 2000 census, Pousada (2010) observes that around 6% of the population in Puerto Rico was born in the mainland United States, while around 3% (Duany, 2001 in Pousada, 2010) were circular/return migrants. Since then, the number of Puerto Rican residents born in the mainland United States has remained consistent, at slightly under 5%, according to 2017 estimates (U.S. Census Bureau, 2019a). Given this discussion, a direct influence from mainland varieties of Puerto Rican Englishes appears unlikely.[3]

[3] Following the findings from previous work on NEW DIALECT FORMATION and DEMOGRAPHIC DETERMINISM in Trudgill et al. (2000), it has been claimed that, for the transmission, adoption, and maintenance of outside and/or emergent dialect features in a community, the speakers of a variant usually represent a majority of the speech community. Even if that holds true, current island wide census data do not indicate that the necessary influx of speakers needed has been observed for this kind of effect to be expected in Puerto Rico.

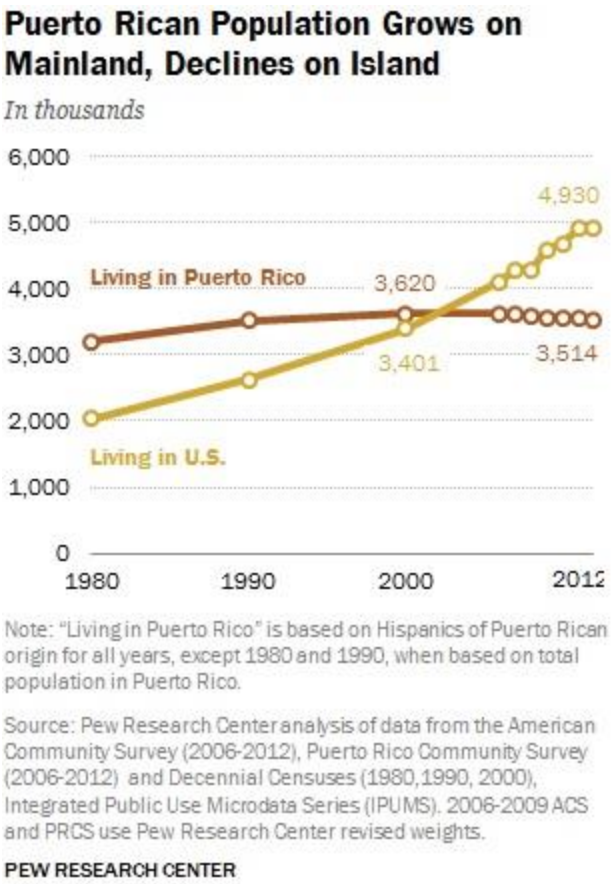


Figure 2.2: Puerto Ricans in the Island versus the Mainland (Pew Research Center, 2014).

Even with a declining population and a relatively low proportion of migrants from the mainland, there has been a consistent increase in L1 English speakers over the last decade (see Figure 2.3 below). A large portion of these respondents skew younger; almost a third of these speakers are between the ages of 5-17, with another third between the ages of 18-64, and the rest above 64 years of age (U.S. Census Bureau, 2019).

Given the data so far, it is possible that these speakers of an emerging English are isolated from speakers of mainland Puerto Rican Englishes. The overarching implication is that it is unlikely that there is a major influence of mainland Puerto Rican Englishes on the overall population of the island.

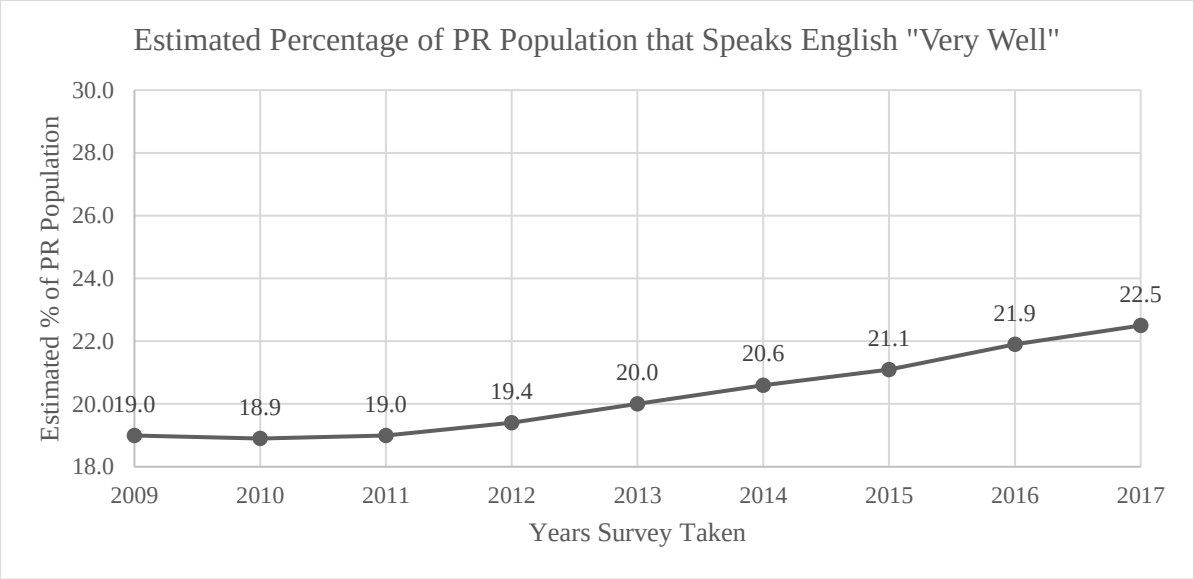


Figure 2.3: Estimated percentage of the population in Puerto Rico that speaks English "Very Well" (U.S. Census Bureau, 2019a).

The population density map (Figure 2.4 at 1,088.2 people/mi^2) from the last official indicates what are largely three major metropolitan areas: (1) the San Juan Metro Area in the north, (2) the greater Mayagüez area in the west and southwest, and (3) the Ponce metropolitan area in the south. Each of these areas boast a national or an international airport, facilitating both greater chances for population conglomeration and a higher ratio of transients.

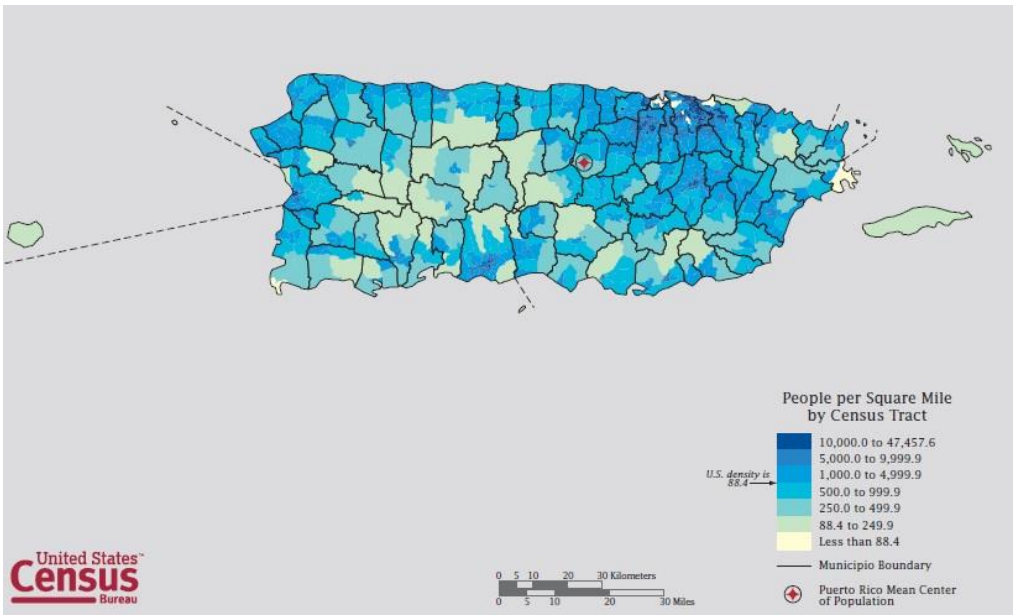


Figure 2.4: Population density map of Puerto Rico (U.S. Census Bureau, 2010).

2.1.3 Geography

Although Puerto Rico is almost always referred to as an island, it is technically an archipelago, consisting of several smaller islands, two of which are notably populated municipalities, Vieques and Culebra to the east. The other islands, such as the Mona island to the west, are not notably populated and are mainly used as hubs for research centers.

Wolfram & Schilling-Estes highlight that geographical barriers, such as mountains, lakes, and rivers, are important in observing dialect development and diffusion as well as settlement patterns (2006). Of these, there are two major geographical barriers in Puerto Rico: (1) the chain of mountains in the center of the main island, and (2) the Atlantic Ocean, separating the main island from its two municipalities, Vieques and Culebra, to the east (see Figure 2.5). Observing Figures 2.4 and 2.5 together reveals that the major metropolitan areas named above circumvent these geographical barriers. As such, it is expected for there to be more linguistic isolation in (1) and (2), with a correspondingly lower amount of English usage.

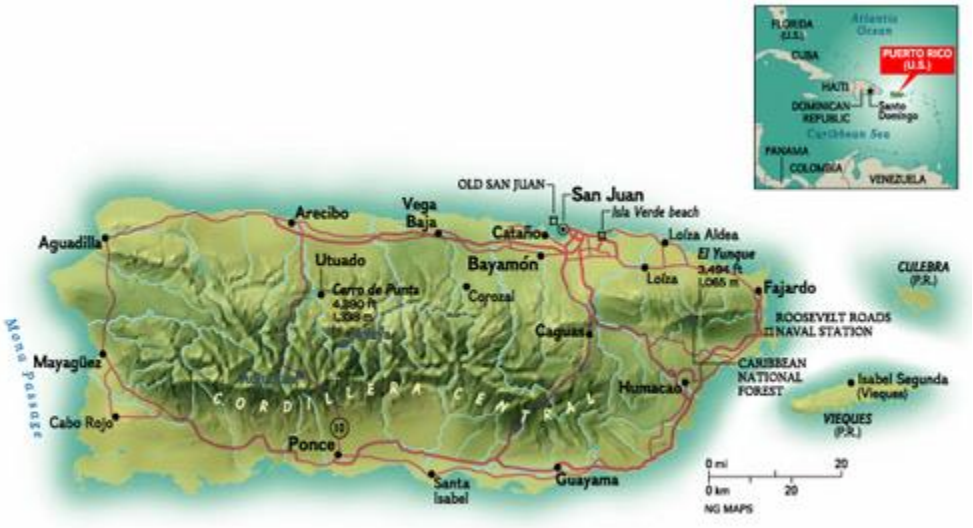


Figure 2.5: Geographic map of Puerto Rico (Rivera, 2019).

2.1.4 Socioeconomic Factors

Much like the mainland, the urban areas of Puerto Rico are loci of change due to more language contact. One of Puerto Rico's biggest economic driving forces is tourism, and these metropolitan areas have been tailored to accommodate it. However, other areas of Puerto Rico are focused on agriculture, especially in the central mountains, which are often referred to by locals as the coffee-producing zone.

To further observe this relationship, Figure 2.6 provides census estimates on languages spoken at home "other than English" (U.S. Census Bureau, 2019b). The map essentially translates to the percentage of people who speak Spanish at home. Closely following what has been observed thus far, there is a lower percentage of Puerto Ricans who only speak Spanish at home (i.e., a higher percentage who speak English) in the metropolitan areas, and a corresponding higher percentage of Spanish speakers in rural zones and the central mountains.

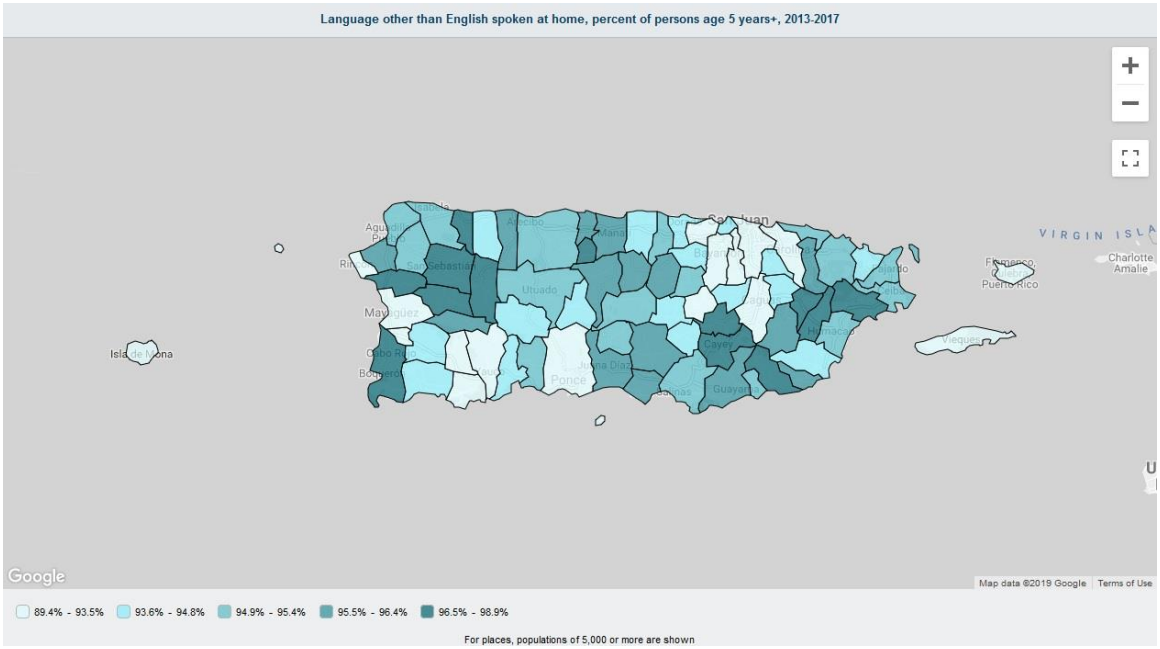


Figure 2.6: Language other than English at home; Age >5 years (U.S. Census Bureau, 2019b).[4]

[4] The color scale provided by the U.S. Census Bureau for these maps is off scale. The lowest end of the color scale starts at 89%, and the highest end of the scale ends at 99%. This distribution is expected for a highly bilingual region.

Figure 2.7 summarizes educational attainment (B.A. or higher) by percentage of the population in each municipality. Further reinforcing previous observations on geography and economic ecology, the higher proportion of the educated population resides in the initially defined metropolitan areas. As Antonio Barreto underscores, for many Puerto Ricans, English is seen as a valuable resource (2000), an instrumental one that represents opportunity for succeeding in higher education and in finding gainful employment on the mainland. Of course, for many other Puerto Ricans, English also serves as a means through which one can express identity in different social contexts, such as in worship, work, raising children, marriage, among others (Pousada, 2010, p.5).

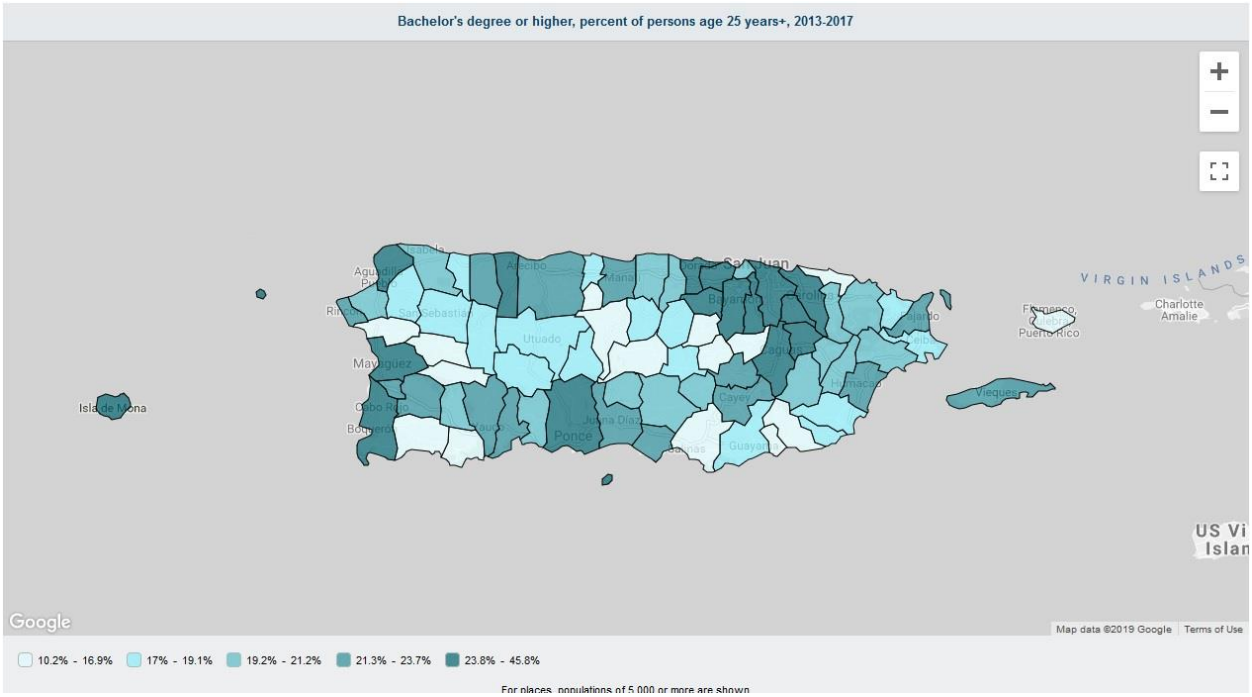


Figure 2.7: Educational attainment in Puerto Rico from 2013-2017 (U.S. Census Bureau, 2019b).[5]

[5] The color scale provided by the U.S. Census Bureau for these maps is off scale. The lowest end of the color scale starts at 89%, and the highest end of the scale ends at 99%. This distribution is expected for a highly bilingual region.

2.1.5 PRIE

Alongside the sociohistorical and socioeconomic factors discussed above, recent research suggests the emergence of simultaneous Puerto Rican Spanish-English bilinguals in younger generations of speakers (Fabiano-Smith et al., 2014)—whose English is termed PRIE in this study. With the emergence of native L1 English speakers in the island in the last few decades and the development of tightly-knit communities with shared sociopolitical identities, it is reasonable to expect subsequent dialect formation as a form of marking SOCIAL IDENTITY, not as part of a generational language shift, but rather as a reflection of cultural assimilation and the amelioration of language attitudes towards American English as a result.

While studies have been conducted on the performative Spanish-English bilingualism of Puerto Ricans (e.g., codeswitching/translanguaging) in the mainland (Poplack, 1978 for North Philadelphia Puerto Rican English; and Zentella, 1997 for the of English New York Puerto Ricans, often referred to as Nuyoricans), there is a paucity of data on the status of PRIE as an emerging dialect of simultaneous bilinguals in Puerto Rico.

2.2 New Dialect Formation and Bilingualism

From the beginning of Puerto Rico's colonial history in 1493, Spanish has been the dominating language of culture and, for the most part, policy. Since then, in the sense that Puerto Rico is officially a political territory of the United States, Puerto Ricans have also found themselves at odds, historically, with the mainland regarding language policy; there have been a number of failed attempts by the US government to implement English as the native language in Puerto Rico, such as through local government reforms in 1917 (Rivera, 2018) and the 1968 Bilingual Education Act (Smallwood Ramos, 2020). This history makes way for a linguistic context in which two colonial languages, Spanish—first and mainly—and American English, coexist today.

It is a cornerstone of this dissertation that ameliorative attitudinal changes towards American English (Vicente Vélez, 2000; Falcón, 2004) are pivotal in the emergence of PRIE. These changes are facilitated by a generational shift and increasing language contact in tandem with sociopolitical and socioeconomic assimilation (Domínguez-Rosado, 2015). Before examining NEW DIALECT FORMATION and PRIE, it is imperative to connect the linguistic situation in Puerto Rico to the concepts of Reed's concept of ROOTEDNESS (2016 and 2018) and the conversation on World Englishes, particularly Indian English as a potential analogue to PRIE, facilitated by Kachru's THREE CIRCLES (1990) model.

Kachru's model (1990) builds on observations of English usage and norms throughout the world. From the perspective of linguistic norms, the Inner Circle of Englishes is defined by those L1 English-speaking countries that are perceived as independent and norm-defining forces, traditionally associated to the monolingual English-speaking populations of the world; the United States, the United Kingdom, Australia, and New Zealand exemplify the Inner Circle. In relation, the Outer Circle can be defined as varieties of Englishes pertaining to countries where it functions either as a lingua franca or a region where the formation of an independent dialect, with its own usage norms, is in process; Singapore, Sri Lanka, and Bangladesh are examples of the Outer Circle. Finally, and in contrast with the two innermost circles, the Expanding Circle of Englishes encompasses the rest of the world, where English is largely considered a foreign language. The THREE CIRCLES model goes beyond the mere identification of the modalities of Englishes spoken around the world; it also serves to codify the sociopolitical relationship between the English-speaking countries of the world.

Despite its over 400-year-old history with English, one that started with the imposition of English as a colonial language that was eventually assimilated as a language of over 50 million

Indian English speakers (Mukherjee, 2010, p.167), India was once traditionally placed in the Outer Circle of World Englishes. As Mukherjee (2010) observes, Schneider's (2003) "model of the evolution of post-colonial Englishes" (to be discussed in more detail in relation to PRIE further on in this section) elucidates the process by which a colonial language, English in this case, becomes an L1 variety (pp.167-168). The process requires two factors: (1) for there to be a change in the construction of a linguistic identity, and (2) for there to be a change in the interactions between settlers and the indigenous population.

Since attention was brought to Indian English in the 20th century, the conversation on its status as an English variety has shifted, from one that framed Indian English as an Expanding Circle English to an Inner Circle English. This conversation was built on research over decades that reported on the emergence of Indian English (Kachru, 1990; Baker & Eggington, 1999; Chand, 2009; among others) as its own variety, with a longstanding linguistic history, its own norms, and its own L1 speakers. Figure 2.8, below, is an updated visualization of Kachru's Circles of Englishes model, which now includes India in the Inner Circle.

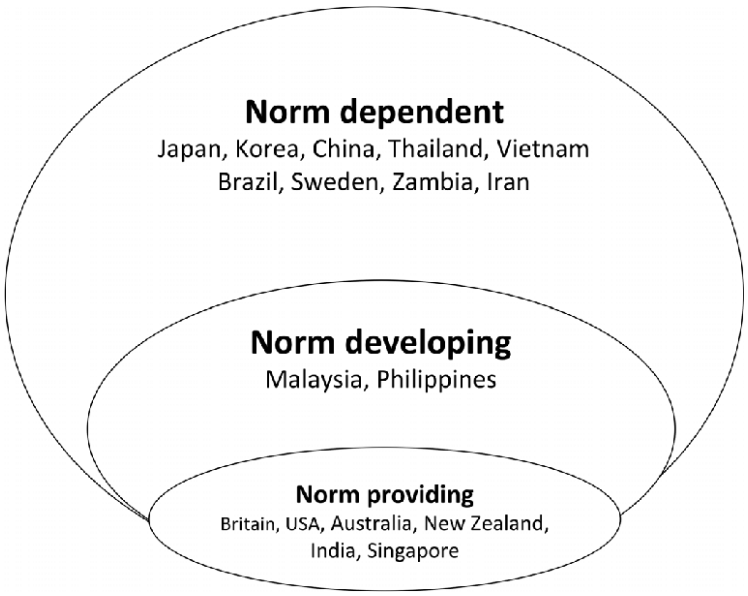


Figure 2.8: A more recent rendition of the Three Circles Model (Haswell, 2013) that observes Indian English as a Norm-Providing/Inner Circle English.

In this dissertation, the case of the development of Indian English serves as a model for the discussion of the development of PRIE, one where a colonial relationship gave rise to a culture of bilingualism, which in turn, allowed for the development of a local variety of what once was the colonial language. It is, then, important to ask when exactly a variety begins to emerge, and when we treat it as such, distanced from the mantle of influence of the norm-providing Englishes. Describing the development of Indian English, Mukherjee (2010, p.168) provides a summary of Schneider's model (2003), which outlines five general phases in the development of post-colonial Englishes:

Phase I – Foundation: In this initial phase, the English language is transported to a colonial territory.

Phase II – Exonormative Stabilization: There is a growing number of English settlers/speakers in the new territory, but the language standards and norms are still determined by the input variety and are, thus, oriented towards [the norm providing variety].

Phase III – Nativization: The English language becomes an integral part of the local linguistic repertoire, as there is a steady increase in the number of competent bilingual L2 speakers of English from the indigenous population.

Phase IV – Endonormative Stabilization: After independence, English may be retained as a/an (co-)official language and a medium of communication for a more or less wide range of intra national contexts (e.g. administration and the press, academia and education); in this phase, a new variety of English emerges with generally accepted local standards and norms.

Phase V – Differentiation: Once a New English variety has become endonormatively stabilized, it may develop a wide range of regional and social dialects.

While Schneider's (2003) model is broadly applicable to post-colonial settings, which Puerto Rico is not, it does function as a baseline to observe the relationship between American

English and Puerto Rico. There are several key differences between the mainland United States and Puerto Rico which complicate the analysis from this model's perspective:

1. Puerto Rico remains a territory of the United States.

2. The indigenous population of Puerto Rico, the Tainos, was systematically wiped out, with few remnants left of their language and culture. Generally, the current population is largely a mix of different waves of settlers from Spain, Africans from the transatlantic slave trade, and the descendants of Tainos.

3. There has not been a notable influx of monolingual English speakers from the mainland to the island.

4. Puerto Rico has been a territory of two different countries in its history: Spain from 1493-1898 and the United States from 1898 to the present.

These differences complicate parallel analyses to Indian English involving Phases II and III, due to the loss of the indigenous population and the fact that Puerto Rico remains a territory of the United States. Nevertheless, the model serves to establish parallels in the development of PRIE in the island. Particularly, English was still brought in as (the second) colonial language after hundreds of years of Spanish on the island (Phase I); the most influential English in Puerto Rico since then has been American English (Phase II); there has been a steady increase in bilingual speakers on the island (Phase III), most notably those younger speakers who demonstrate ameliorated attitudes towards American English (Domínguez-Rosado, 2015); English is an official language and functions as a medium of communication in media (popular culture, social media, news, etc.), intra-territorial contexts (academia and government affairs, particularly when directly concerning the United States), and in urban areas, where there are larger concentrations of English speakers (Phase IV). It is at this phase that I argue that PRIE has emerged, as Schneider's (2003) model projects, where PRIE has not yet been enregistered (Phase V), but has reached the ideal conditions, both in speakers' language attitudes towards American

English and at the level of institutional implementation for American English to gain enough clout to coexist with Puerto Rican Spanish.

As identity is inextricably tied to language use, much as in the case of Indian English, the emergence of PRIE appears to be guided by sociopolitical and sociocultural shifts in the orientation of the linguistic aspect of what it means to be Puerto Rican. Through a series of multigenerational ethnographic interviews, Domínguez-Rosado (2015) has observed a gradual redefinition of Puerto Rican identity from a "language one" culture (i.e., only Puerto Rican Spanish) to a bilingual culture comprised of the coexistence of Puerto Rican Spanish and American English, partially due to the rise of global multiculturalism and more flexible language attitudes towards English, particularly in younger generations. This redefinition of the roles that both English and Spanish play in what is locally defined as a Puerto Rican identity is central to the emergence of PRIE, as it represents an effective unbinding of the regional language norms, largely dominated by Spanish until recently.

In this sense, Puerto Rico is a unique case in that it has covertly existed as a member of the Outer Circle of English for over a century and is simultaneously a part of an Inner Circle country, the United States. However, in its present state of linguistic affairs (between, what appears to be, Phases IV and V), the case of PRIE is not much too different from the case Indian English several decades ago—a growth in bilingual speakers tied to a redefinition of a regio-cultural identity and language attitudes towards English.

The exploration of the role of place, then, and principally the importance of speakers' connections to place, in sociolinguistic performance cannot be understated when observing the connections between language and identity. More recent methodologies have begun exploring this connection in more depth. Reed developed a ROOTEDNESS METRIC (2016), a contextually

designed in-group measurement tool, which aids in quantifying a speaker's attitudinal orientation towards a region. Reed (2016 and 2018) found that speakers from Appalachia that were more ROOTED in Appalachia (i.e., had a stronger sense of identity with the region) exhibited much higher usages of both overt (/a/ monophthongization) and covert (rising pitch accent) Appalachian English features. His approach demonstrates that the degree to which speakers are oriented towards a place affects their sociolinguistic performance, not unlike Labov's landmark Martha's Vineyard study (1972), which centered on observing co-variation as an epiphenomenal effect of an alignment to a local identity.

The ROOTEDNESS METRIC further explicates attitudinally driven language performance. Given that the results of the pilot study indicate that higher English dominance is connected to more monolingual English-like performance and that the quantification of English dominance in both the LEAP-Q and the BLP questionnaires is heavily dependent on domain use and identity, it is expected that Puerto Rican Spanish-English simultaneous bilinguals are less ROOTED to Puerto Rico, at least in as much as it has been conceived as an L1 Spanish/L2 English-only region. Hence, the ROOTEDNESS METRIC is adapted to observe the connection of the amelioration of attitudes towards American English and the recent rise of L1 English (and thus, PRIE) in Puerto Rico, particularly in the younger generations of simultaneous bilinguals.

Much like Schneider's (2003) model of post-colonial Englishes, conversations around NEW DIALECT FORMATION usually revolve around the eventual product of progressive multi-generational language shift, such as the case for New Zealand English; Trudgill outlines five phases through which New Zealand English likely developed into the distinct variety that can be observed today (2008). Other studies employ synchronic analyses of incipient dialect formation, grounded in the enregisterment of previously neutral varieties of American English—for

instance, the enregisterment of Matanuska Valley Alaskan English as a variety influenced by migration from the Upper Midwest (Purnell, Raimy & Salmons, 2009) and of Wisconsin English by way of its younger speakers (Schuld et al., 2017). These studies explore a crucial element in NEW DIALECT FORMATION, the combination of historical factors, drift, and eventual enregisterment via feature selection.

In its broadest sense, NEW DIALECT FORMATION in the context of colonial and post-colonial Englishes can be defined as a variety that arises from the imposition of a colonial language, which, by process of adoption and adaptation, becomes a part of a redefined identity claimed by subsequent generations of speakers. In this development, the emergent variety may represent a shift or replacement of previous languages (American English, New Zealand English, etc.) or coexist with other dialects or languages as part of a redefined regional or national linguistic identity (such as Hawaiian English or Indian English, respectively); PRIE, as a potentially emerging variety of American English, is best described within the latter category.

While similar historical analogues for the development of PRIE exist in post-colonial English-speaking countries like India, the unique sociopolitical and sociolinguistic context of Puerto Rico underscores a number of aforementioned differences. These differences call for a reinterpretation of the political elements surrounding extant explanations for NEW DIALECT FORMATION, while applying compatible observations in Kachru's THREE CIRCLES (1990) and Schneider's post-colonial Englishes models to explain the phenomena associated to the emergence of PRIE as a variety that is simultaneously a development of an ongoing politically colonial and culturally assimilating relationship with the United States.

To summarize, Reed's (2016) ROOTEDNESS METRIC is employed in this dissertation to further analyze previous results from the pilot study of this dissertation in simultaneous Puerto

Rican Spanish-English speakers in the pilot of this study, where American English speakers from the mainland perceived these speakers as native English speakers. The ethnographic evidence on the amelioration towards English in Puerto Rico and the redefinition of the role of language in defining a Puerto Rican identity (Vicente Vélez, 2000; Falcón, 2004; Domínguez-Rosado, 2015) present a strong case to consider a favorable context in which PRIE can emerge.

2.3 Perceptual Dialectology, Language Attitudes, and Perception

Language attitudes have been at the center of discussions on the effects of the making and stereotyping of dialectal features as a means to codify social stratification. Labov (1972), Preston (2013), Moreno Fernández (2009), and González-Martínez (2008), among others, have observed that the perception and stigmatization of linguistic features occurs as an epiphenomenon of the preconceptions about people and cultures that we develop to understand language in a social context. This phenomenon, in turn, creates the ideal conditions for mismatches between production and perception. Language attitudes, as a theoretical concept, are not usually overtly defined in studies. As Moreno Fernández observes: "language attitudes are a manifestation of individuals' social attitudes" (2009, p.177). Language attitudes can, therefore, be defined as the linguistic behaviors towards a conglomerate of sociolinguistic factors that affect interactions between interlocutors, such as group membership, the amelioration or pejoration of dialects against a standard form, among others.

PERCEPTUAL DIALECTOLOGY is positioned within FOLK LINGUISTICS, serving as a window to the social and attitudinal perspectives of speakers and listeners. PERCEPTUAL DIALECTOLOGY has existed within scholarship since the late 19th century, and it was gradually developed throughout the 20th century, particularly in the Netherlands and Japan (Preston, 2018, p.177). Preston (2013) discuses attitudes within the realm of FOLK BELIEFS, referring to the

influence of urban myths about language and the effects of social orientation on overall language perception; that is, that beliefs about a variety emerge from the relationship between social groups and the perceptions that are built upon that relationship (p.157). Within PERCEPTUAL DIALECTOLOGY, attitudes towards sociolinguistic groups can be defined by factors such as region, social class, linguistic features, and even age groups.

This dissertation considers the intersection of perceived dialect and geography as two such subfactors within the linguistic features that listeners notice. Particularly, the intersection of these two factors is built on Preston's (2013) observation that listeners interpret or filter the available background information (or lack thereof) on a speaker to attempt to socially orient themselves in relation to the speaker. This orientation is important for participants to determine that a group of speakers, PRIE speakers in this study, belong to the same speech community, a key indicator of potential dialect formation.

Perception, as exemplified in Nash's SURFACE POINTER MODEL (1996), is largely guided by a combination of sociolinguistic factors, where KINSHIP, COMMENSALITY, and COMMON CULT are the primary driving cultural forces behind group membership determination (p.25). From the perspective of PERCEPTUAL DIALECTOLOGY, perceptions of these cultural factors as indicative of group membership can be codified as linguistic features. However, with an absence of physically visible pointers or background information, as in this study, listeners have a limited number of tools through which they can interpret group membership.

Preston (1996) conducted a perceptual study on the North to South regional dialect sensitivity continuum in the United States. He recorded speakers of the same age, class, race, and gender reading a scripted text. The recordings were devoid of any dialect-specific features other than the speakers' phonetic features. Participants were then asked to match these speakers to a

regional continuum from Michigan to Alabama. The results revealed that perceptual distinctiveness broke along the lines of North versus South, rather than state by state, pointing to the existence of three macro dialects, the North, the Midlands, and the South. Wolfram & Schilling-Estes underscore that what guides these types of perceptual judgments are the prejudices against certain speech varieties, such as southern speech (2006, p.162), rather than a correlation of coexisting features along a continuum of speakers. These findings underscore the importance of recognizing the influence that socially constructed perceptions have on defining speakers' (folks') language attitudes and the categorization of social groups by means of established linguistic markers.

Participants have also been able to distinguish dialects from single linguistic cues. Through a series of four experiments, Purnell, Idsardi & Baugh's (1999) study investigated associations between linguistic features and group membership, particularly ethnicity. In one of the experiments, Baugh—a tridialectal speaker of African American Vernacular English, Chicano English, and Standard American English—called the same landlords in five different locales in time windows of at least 30 minutes between calls using the three different dialects to request apartment availability. Their findings revealed a "clear pattern of potential discrimination associated with the three dialects by geographic area" (1999, p.14). The landlords in this study appeared to have acted on preconceptions based on the dialect they perceived in a potential tenant, which, in turn, affected the availability of a rental offer. Other studies have found that speakers' SOCIAL IDENTITY, such as perceived nationality (Niedzielski, 1999) or sexual orientation (Mack, 2010), skew listeners' perceptions. These studies provide further evidence on the connections between production, perception, and SOCIAL IDENTITY.

Throughout these studies, it is evident that perceptions of the input that listeners receive are molded by their expectations according to the information on the speaker that they have available. In this sense, listeners are primed to classify and/or perceive certain linguistic features associated to a group because they expect them. It is, therefore, reasonable to accept that stereotypes and innate biases affect how listeners interpret speech input. However, an absence of marked features and/or background knowledge can also create a 'null effect' in listeners, wherein these 'null' (absent) features are interpreted as information in order to ascertain a speaker's social membership. This contrast is important in approaching PRIE as a potentially emerging variety of American English, one that has not yet been enregistered and is absent of STEREOTYPES.

Baird et al. (2018) investigated perceptions of Lexically-Specific Phonology Switches (LSPS) in Spanish-English bilinguals.[6] Through a Matched-Guise Task and a series of Perception Likert Scale tasks, 167 participants responded to recordings of two male Spanish-English bilinguals reading a scripted text, both with and without LSPS, on words of Spanish origin. Participants attributed more positive character qualities to the speakers when they read without LSPS; Baird et al. further suggest that having access to both Spanish and English phonologies is not "as stigmatized as Spanish-accented English" (Baird et al., 2018, p.79). The underlying implication in Baird et al.'s (2018) findings is that overtly codified associations to negative features easily surface with the corresponding SURFACE POINTERS, even in Spanish-English bilinguals that are fluent in both languages. Participants identified the STEREOTYPED LSPS, which prompted the association of negative social features to the stigmatized forms.

[6] The dialects of the bilingual speakers selected for this study were not specified.

The framework of PERCEPTUAL DIALECTOLOGY explains listeners' motivations to interpret linguistic features from a sociolinguistic perspective—both in relation to their own social groups and in understanding a speaker's dialect as a factor in determining group membership.

Nevertheless, most current studies in PERCEPTUAL DIALECTOLOGY (such as Preston's (1996) North to South and Montgomery & Stoeckle's (2013) generalized mental map tasks) have been focused on listeners' macro-level perceptions of speech stimuli. Their results, however, reveal that perceptions ultimately guide dialect categorization and the speakers' social orientation towards those dialects, regardless of the distinctiveness of intradialectal acoustic cues that are not perceived as differentiating (as opposed to those marked features that are).

In this dissertation, PERCEPTUAL DIALECTOLOGY is taken a step further and employed as a means to survey developing perceptions of potentially-emerging dialects—i.e., a new dialect category in listeners' dialect map of a language. This kind of approach is exemplified in Schuld et al. (2017), where selected acoustic cues were presented to listeners around the country and employed to survey the status of Wisconsin English as a developing regional dialect in college-aged speakers. As such, this research builds on the methods and approaches in the theoretical framework of PERCEPTUAL DIALECTOLOGY to explore ongoing dialect formation.

2.4 Summary

This chapter provides a baseline for the exploration of PRIE as an emerging variety of American English in Puerto Rico through the lens of current literature and research on PERCEPTUAL DIALECTOLOGY, NEW DIALECT FORMATION, and ROOTEDNESS. This theoretical framework functions as a point of departure to understand enregisterment, and thus, the perceptual salience, of PRIE as a historical circumstance of the complex sociolinguistic,

historical, and political relationship between Puerto Rico and the United States within the broader scope of World Englishes.

Given this background information, a set of hypotheses was determined to further explore the existence of PRIE:

> **H0**: There is *no* PRIE.
> **H1**: PRIE is emerging *in Puerto Rico*.
> **H2**: PRIE is *not* emerging *as a unique form, but rather, it is a subset of existing mainland American English varieties*.

Given the results from the pilot study for this research, where participants identified Puerto Rican Spanish-English bilingual speakers as native English speakers, it is unlikely that H0 is true. This dissertation explores the status of PRIE in relation to other varieties of American English. H1 is true if PRIE speakers are accurately identified as speakers of a same sociolinguistic dialect group, with the potential of finding its own set of subliminal or unmarked features. If so, from where, and how is it surfacing? H2 is true if PRIE speakers are speakers of different varieties of American Englishes and are identified as undifferentiated from the control groups, particularly the monolingual English control group.

In terms of expectations, H1 is expected to be true, given the combination of general shifts in language attitudes on the island, mirroring those observed in the emergence of Indian English. H2 is not expected to be true because of the current demographic trends on the island, particularly the combination of the nearly exclusive outflux of Puerto Ricans from the island, with little influx to counteract this new wave of diaspora, and a reduced influence of mainland Puerto Rican English varieties on the island relative to the whole population.

CHAPTER 3

METHODOLOGICAL FRAMEWORK

As discussed in Chapter 2, this study employs approaches from PERCEPTUAL DIALECTOLOGY, NEW DIALECT FORMATION, and ROOTEDNESS as the basis for its theoretical framework. This framework functions as a point of departure to understand ENREGISTERMENT, and thus, the perceptual salience, of PRIE as an emerging variety of American English due to the result of the complex sociolinguistic, historical, and political relationship between Puerto Rico and the United States, within the broader scope of World Englishes. As such, this retools previous approaches in PERCEPTUAL DIALECTOLOGY to explore NEW DIALECT FORMATION as a product of broader awareness of a variety and its eventual ENREGISTERMENT, in connection with the development of a new socio-regional identity of a group of speakers through shifts in language attitudes, with particular reference to younger simultaneous Spanish-English bilingual Puerto Ricans, for the purposes of this study.

One challenge is to ensure the validity and accuracy of the results by controlling for the effects of listeners' background knowledge and their sensitivity in identifying phonetic differences across the speakers' dialects. To this aim, this dissertation approaches listener selection and hypothesis testing in four separate phases. Phases 1 and 2 are designed to control the speaker selection process, given the range of variability in production that language dominance introduces when considering bilingual speakers. Phases 3 and 4 focus on the analyses of listeners' responses, providing closer attention to their language histories and sensitivity to dialect recognition.

The first section of this chapter introduces results from the pilot study for this dissertation. The second section introduces: (1) how these findings motivated the main research

questions for this dissertation, and (2) how this dissertation is structured to answer each research question, given that this research project is divided into four phases. The third and fourth sections each expand on the methods for Phases 1 and 2, which involve the initial speaker selection process. Finally, the last section provides a summary of the chapter. The specific methodologies for Phases 3 and 4, which are the main research tasks, are briefly described in this chapter and are covered in detail alongside their results and discussion in Chapters 4 and 5.

3.1 The Pilot Study: English Speaker 'Nativeness' Perceptions

The pilot study focused on investigating Upper Midwesterners' perceptions of simultaneous Spanish-English bilinguals' L1 English to determine whether they sounded like native speakers of English. The goal of this pilot was to ascertain whether native English speakers from the Upper Midwest identify a perceptual difference in the English of simultaneous bilingual speakers from Puerto Rico and, based on those findings, to determine whether there is a potentially emerging variety of English, unique to Puerto Rico.

3.1.1 Pilot Study Design

Three simultaneous Puerto Rican Spanish-English bilinguals, forming the pilot experimental group (pBS1, pBS2, and pBS3, in order of English dominance according to their results in the BLP[7]), were recruited for this study. All bilingual speakers were all born and raised in the southeastern coast of Puerto Rico and lived on the island for over 20 years before moving to the United States. The experimental speakers were compared against a monolingual General American English speaker (pEC) from the Upper Midwest and a late Spanish-English sequential bilingual (pSC) from Puerto Rico.

[7] To differentiate these speaker identification numbers from the ones in the main study, a 'p' for 'pilot' was added to each identification number.

The focus of this perceptual study was to examine whether listeners heard differences in simultaneous Puerto Rican Spanish-English bilinguals' English, based only on an audio clip of the speakers reading the first three paragraphs of "The North Wind and the Sun" (see §3.3.2 for more information). Because the experimental group consisted of overall-balanced simultaneous bilinguals, it was not expected that listeners would perceive the three simultaneous bilingual speakers in the experimental group as non-native English speakers. Rather, it was hypothesized that listeners' perceptions of the simultaneous bilingual speakers would pattern more like that of the monolingual English speakers, but with a lesser degree of certainty on the simultaneous bilinguals' status as perceived L1 English speakers. Contrastively, it was not expected that the perception of these simultaneous bilinguals' English would pattern like the sequential bilingual Spanish controls. To test these hypotheses for the purposes of providing a framework for this dissertation, a Speaker 'Nativeness' Judgment Task (discussed below in §3.1.2) and a Mental Map Task (discussed below in §3.1.3) were conducted in the pilot study.

237 English-speaking adult listeners were surveyed using the online Qualtrics Survey Software—around 15% (39) of which identified as 'Hispanic,' while the rest identified as 'Non-Hispanic.'[8] After analyzing the results, the emerging pattern between listeners was between Hispanic versus Non-Hispanic participants. The groups shared many similarities, including speakers ranging from 'not proficient at all' to 'fluent.' The main categories distinguishing these groups were: (1) exposure to Spanish, (2) place of origin, and (3) years of formal education in Spanish. These divisions are important because they demonstrate the different perceptual inclinations in the judgment tasks, while the contrasts were minimal in the mapping task.

[8] The 'Hispanic' versus 'Non-Hispanic' distinction, as opposed to other applicable terms such as 'Latinx,' was adopted for the pilot study from the wording employed in the United States 2010 Census.

3.1.2 Pilot Speaker 'Nativeness' Judgment Task Pilot Results Summary

Table 3.1

Hit Rate (H), False Alarm Rate (FA), A-Prime (A′), and Non-Parametric Bias (B''_D) Averaged Across the Listener Subgroups for the Speaker 'Nativeness' Judgment Task

Listener Subgroups	**H**	**FA**	**A′**	$\mathbf{B''_D}$
Hispanic	0.72	0.33	0.68	0.16
Non-Hispanic	0.78	0.13	0.86	0.64

For this task, listeners were asked to identify whether they believed that the five pilot speakers were native speakers of English. The results were then evaluated using a non-parametric A-Prime analysis in Table 3.1. In A-Prime, a number of stimuli (in the case of this analysis, audio clips of speakers) is identified as the target to measure both the strength and accuracy of the signal for listeners; in other words, A-Prime tests how sensitive listeners are to a set of stimuli. An expected positive response is identified as Hit, an unexpected negative response as a Miss, an unexpected positive response as a False Alarm, and an expected negative response as a Correct Rejection. Those results are then used to calculate the Hit Rate (H), which tells us how accurate listeners were at correctly identifying a signal; False Alarm Rate (FA), which tells us how often participants incorrectly identified a distractor as a signal; and the A-Prime score (A′), which tells us the listeners' overall sensitivity in identifying signals from distractors (more information on A-Prime methods and analyses is provided in Chapter 4).

Hispanic and non-Hispanic listeners generally identified the simultaneous bilinguals as native English speakers at approximately the same rate (H=0.72, 0.78, respectively). Non-Hispanic listeners, however, were more accurate in identifying pSC as a nonnative English speaker (FA Rate = 0.13). The overall results from this analysis indicated that group membership stratified the results in this task. The simultaneous bilinguals were generally identified as native English speakers by both groups of listeners, patterning more closely with pEC. However, the

much higher FA Rate for pSC ultimately affected the accuracy of the group of Hispanic listeners. It appears that the Hispanic listeners were overly sensitive to pSC. This bias can be linked to pSC's marked accent as a potential codifier of social identity, which ultimately guided the discrepancies between the final A′ scores between the listener groups.

3.1.3 Pilot Mental Map Task Pilot Results Summary

Following the Speaker 'Nativeness' Judgment Task, listeners were provided with an opportunity to listen to the recording of the speaker again and then answer whether they believed could place the speaker on a map of the United States (all 50 states plus territories). If participants answered that they could, they were provided with a blank map with identified state-level political divisions and clicked on the state or region from which they thought the speakers originated. The results for this activity were visualized as scatter point maps and compared descriptively across weighted region placement.

It was expected that the simultaneous bilinguals would be associated with more Spanish-speaking or Spanish-influenced regions of the United States when compared to the English control. Historically, the southwestern crescent of the United States, ranging from eastern Texas to northern or central California, has been seen as a Spanish-English language contact region, one that is considered the locus of Chicano English (Fought, 2006, pp.79-80) and other Spanish-influenced L1 varieties of American English. The results from the Mental Map Task are summarized in Figure 3.1 below, where each dot represents a response from one participant by clicking on a blank map. The ovals are provided to emphasize areas of higher response density.

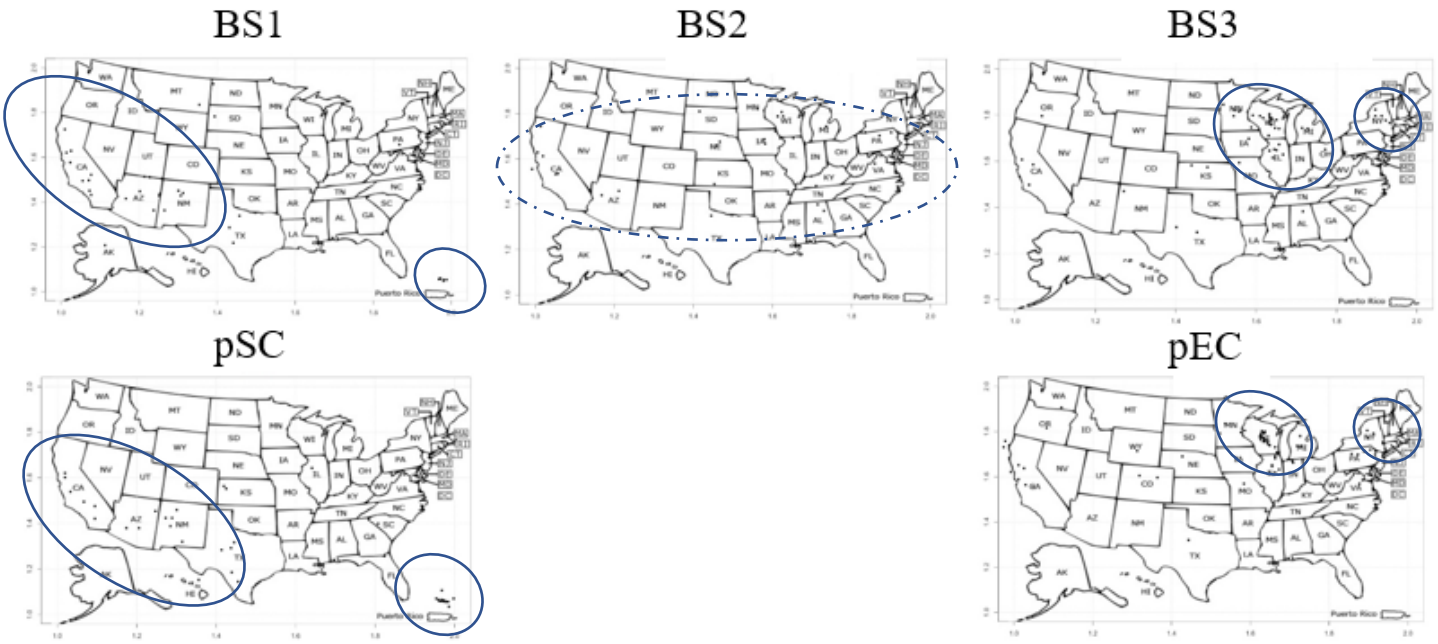


Figure 3.1: Mental Map results from least to most English dominant; the experimental speakers (BS1-3) are on the top row, left to right, from more Spanish-dominant to more English-dominant. The control speakers, pSC and pEC, are on the bottom row, from left to right.

Overall, the findings demonstrated a reinforcement of connections between speaker 'nativeness' judgments and the linguistic profiles that listeners developed based on their initial judgments. The bilinguals that were perceived more as non-native English speakers were placed more in regions of the United States that have been historically Spanish speaking or Spanish influenced, while the more English-dominant bilinguals were placed in the more historically anglophone regions, such as the Upper Midwest and the northeast. Unexpectedly, Puerto Rico was identified as one of the salient regions of origin for BS3 alongside pSC. BS3, as the most Spanish-dominant simultaneous bilingual, performed similarly to pSC, in terms of perceptual association to a dialect region. Conversely, listeners' perception judgments of the more English-dominant bilinguals (BS1 and BS2) patterned more closely to pEC and did not follow the same pattern that listeners ascribed to the more Spanish-dominant speakers.

3.1.4 Overall Findings from the Pilot Study

Alongside AGE OF EXPOSURE, language dominance also played a critical role in distinguishing between the simultaneous bilingual speakers. Noting that the language dominance score provided in the BLP quantifies how bilingual speakers both identify and conduct themselves in everyday linguistic interactions, the importance of this factor cannot be overlooked as a potential indicator of production and performance in a language.

With few exceptions, the overarching results from the tasks in this pilot supported the initial hypothesis that PRIE speakers would be perceived as L1 English speakers, but to a lesser degree than the EC speakers, as indicated in §3.1.1. Listeners rated the simultaneous Spanish-English bilinguals as native speakers of English, absent any background information, despite the idiolectal differences across the bilinguals with similar language histories. The observed response patterns suggest that these simultaneous Spanish-English bilinguals' English exhibit features that have yet to be associated with an overtly recognized variety.

The results from the pilot indicate an emerging variety of English in Puerto Rico, PRIE, which may contrast with previously studied mainland varieties of Puerto Rican Englishes (such as Nuyorican English). These results represented an important finding, as Puerto Rico has been almost entirely viewed as an L1 Spanish/L2 English-speaking region, even with the recently increasing influence of the mainland United States on the island over the decades.

3.2 Methodology

This dissertation builds on the results from the pilot study by improving on its methodological approaches and rigor, both by controlling more for listeners' preexisting language histories (exposure to Spanish speakers and sensitivity to regional dialect variation), while addressing the limitations of the pretested tasks in the pilot study. Therefore, the aim is to

explore whether there is an emerging variety of American English in the younger generations of Puerto Ricans living in Puerto Rico who are simultaneous Spanish-English bilinguals, given the sociopolitical relationship between Puerto Rico and the United States, alongside the amelioration of language attitudes towards English in these PRIE speakers.

To reiterate observations from Chapter 1, this study focuses on observing and analyzing the naïve perception judgments of English listeners mainly from the Upper Midwest (i.e., listeners with little to no exposure to Spanish and/or the English of bilingual Spanish-English speakers), relative to what constitutes a canonical variety of American English to them. Because this group of listeners is unaware of the development of PRIE, positive evidence from these listeners in support of PRIE's perceptual salience as a dialect would present a strong case for its current emergence and development.

With the results from the pilot study and the background information presented in Chapter 2 in mind, this dissertation adopts a multifaceted approach to data gathering and analyses, designed to begin to answer the following research questions:

1. Based on speakers' audio recordings and no other background information, can listeners identify PRIE speakers distinctly from other similar varieties of mainland American Englishes (Nuyorican, Chicano English, and Miami Cuban English)?
2. How do speakers of Puerto Rican Englishes (PRIE and Nuyorican) compare to other selected recognized varieties of mainland American Spanish-English bilingual varieties (Chicano English and Miami Cuban English)?
3. From a perceptual perspective, to what dialect region(s) of the United States are the selected American Spanish-English varieties associated vis-à-vis the Spanish and monolingual English control speakers?

To answer these questions, the study was divided into four phases. The first two phases of the study delimit the speaker selection process, while Phases 3 and 4 address the research questions with the speaker data gathered and screened for in Phases 1 and 2, respectively.

The experimental tasks were broken down by phases to control survey length, keeping the tasks under 30 minutes each, and consistency by recruiting a different pool of listeners; a consistent set of results from diverse participant pools strengthens data analyses. Phases 2-4 also included the same Demographics Questionnaire (see Appendix A for the full demographics questionnaire) at the end to determine which listeners qualify under the inclusion criteria for each study. The dissertation's methodological approach is structured as follows, by each phase and their associated task(s):

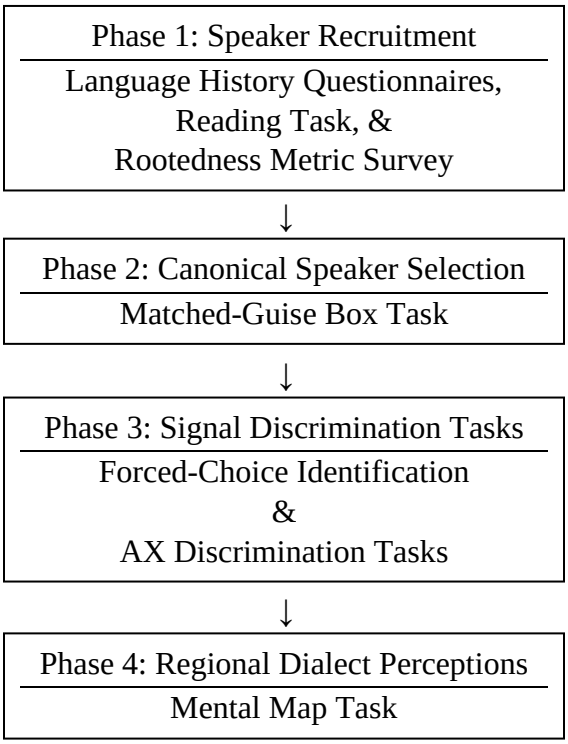


Figure 3.2: Structure of the methodology by phase with their relevant tasks.

To answer the first research question, a Yes-No Forced-Choice Identification Task was employed. This task allowed listeners to identify whenever they heard a speaker that they

believed was Puerto Rican or of Puerto Rican ancestry. To explore the second research question, an AX Discrimination Task was adopted. This task allowed further elaboration on the findings from the Identification Task by comparing just simultaneous Spanish-English bilingual groups against each other (PRIE, Nuyorican English, Chicano English and Miami Cuban English). The results from this task aid in observing how listeners group the members of each of the target bilingual speech groups with respect to each other. Both the ID and AX Discrimination tasks are part of Phase 3 of the study. A detailed explanation on the design and structure of the methodology for both the Forced-Choice ID Task and the AX Task is provided in Chapter 4.

For the third research question, a Mental Map Task, following Preston & Robinson (2005), is administered in Phase 4. The Mental Map Task improves on the piloted map task, described above, by allowing participants to select fixed regions. It is expected that participants identify similar speech samples in the speech signal so that samples align with regions based on similarity of phonetic features. A detailed explanation on the design and structure of the methodology for the Mental Map Task is also provided in Chapter 5.

3.3 Phase 1: Speaker Recruitment

Phase 1 draws from the methodological approach employed in the pilot study, with an additional set of measures—a more rigorous initial screening process and a second speaker selection process in Phase 2 to have a speaker pool that is as representative as possible to the median speaker of a bilingual group. That is, the aim in this two-fold process was to select the bilingual speakers that are as cohesively representative as possible of that speaker's group.

The goal of this perception study in terms of speaker recruitment is to compare the PRIE speakers, not only to control speakers, much like in the pilot study, but also to a variety of other simultaneous Spanish-English speakers from the mainland. These speakers' varieties should also

be either already enregistered or at a similar stage of development as PRIE. This approach allows for listeners the opportunity to further distinguish the PRIE speakers from a larger amount of variance, strengthening any saliency of PRIE that results from the data.

To that end, 33 speakers were interviewed: 24 were simultaneous Spanish-English bilinguals (six from each speaker group in Table 3.2) and nine speakers to function as control groups (3 from each speaker group, also in Table 3.2). To control for variance in speech production and prevent additional variables in the signal detection tasks in Phases 3 and 4, all of the selected speakers were males between the ages of 18-42.[9]

Table 3.2

Initial Number of Speakers Recruited in Phase 1 from Each Target Speaker Group

Speaker Group	Speaker N
Simultaneous PR Spanish-English Bilinguals (PRIE)	6
Simultaneous Nuyorican PR Spanish-English Bilinguals (NY)	6
Simultaneous Miami Cuban Spanish-English Bilinguals (CB)	6
Simultaneous Chicano Spanish-English Bilinguals (CH)	6
Monolingual English Control Speakers (EC)	3
Sequential PR Spanish-English Bilingual Controls (SCPR)	3
Sequential Mexican Spanish-English Bilingual Controls (SCMX)	3
Total Speakers Recruited	**33**

The Nuyorican, Miami Cuban, and Chicano simultaneous Spanish-English bilinguals all serve different purposes in identifying the perceptual salience of PRIE: Nuyorican English and Chicano English are both enregistered varieties of mainland American English, the former also being a variety that is ancestrally connected to Puerto Rican Spanish. Miami Cuban English is analogically closer to PRIE in that it is variety of American English, belonging to a group that is broadly bilingual and dominant in both languages. Miami Cuban English also appears to be

[9] See Poplack (1978, pp.97-102) for further information on degrees of stylistic variance in production based on gender differences.

undergoing NEW DIALECT FORMATION (Carter, López Valdez & Sims, 2020) that parallels PRIE in terms of its bilingual context and recency; this relationship is further explored in Chapter 6.

The six sequential Spanish-English bilinguals serve as the Spanish-dominant control group. Of those six speakers functioning as Spanish Control, three were selected from Puerto Rico and three from Mexico. The rationale for this selection is to provide analogous distractors for both the Chicano and PRIE simultaneous bilingual speaker groups. The monolingual English control speakers were all speakers of the same dialect of a city in the Upper Midwest, and they serve as counterweight distractors for the more English-dominant bilingual speakers, while also providing a baseline that listeners (Upper Midwesterners) were more likely to perceive as more familiar unmarked.

Finally, with the goal of controlling for the range of potential in-group variation that may come from bilingual speakers with different language backgrounds, the 24 simultaneous bilinguals belonging to those groups were further narrowed down to 12 (three from each group) in Phase 2. This process is discussed in further detail in §3.4.

3.3.1 The BLP and LEAP-Q Questionnaires

The speakers interviewed in Phase 1 were all administered the Bilingual Language Profile (BLP; Birdsong et al., 2012) and the Language Experience and Proficiency Questionnaire (LEAP-Q; Blumenfeld & Kaushanskaya, 2007). These tests allow for the operationalization of a speaker's language dominance on a scale. The BLP and LEAP-Q questionnaires[10] have been demonstrated to be efficient tools in quantifying dominance and experience-specific effects in language outcomes (e.g. Casillas & Simonet, 2018; Ramírez & Simonet, 2018; and Wong & Ng,

[10] Both questionnaires are available in multiple languages. The BLP is available in 28 languages (http://sites.la.utexas.edu/bilingual/using-the-blp/access-testing-materials/) and the LEAP-Q in 22 languages (https://bilingualism.northwestern.edu/leapq/).

2018 for the BLP questionnaire; as well as Cockcroft et al., 2017; Cornwell & Rafat, 2017; and Gonzales et al., 2019; for the LEAP-Q questionnaire—among many others for both questionnaires).

The results from these language profiles aid in more accurately determining the bilingual speakers' language proficiency and dominance. Both tests contextualize the results by providing a descriptively quantifiable method through which we can understand the complex language history of these bilingual speakers: language histories that correspond to different levels of dominance, and thus, proficiency and performance, which influence listeners' perceptions.

The BLP (Appendix B) provides an extensive quantification of speakers' sociolinguistic backgrounds by both language domains and overall dominance. An equally-weighted language score in a specific domain ranges from 0 to 54.5, and overall dominance is quantified between -218, indicating absolute Spanish dominance, and 218, indicating absolute English dominance; this global dominance score is obtained by subtracting the overall totals in both languages against each other. While the BLP successfully quantifies language dominance and distinguishes domain-specific variation, it does not account for usage as efficiently as the LEAP-Q. As such, this study triangulates the language profiles of all speakers to provide a more concise and thorough background of their domain-specific language proficiencies and overall dominance scores through the BLP, and usage/self-reported proficiency through the LEAP-Q (Appendix C).

Screening. As just noted, all of the speakers who were interviewed in Phase 1 were administered the BLP and the LEAP-Q questionnaires. Alongside quantifying the speakers' sociolinguistic backgrounds by language for contextual analyses of the perception task results, to qualify for Phase 1, those same BLP dominance scores were also used to screen speakers for the

inclusion criteria by speaker group. The BLP dominance score screening criteria are outlined below:

i. Simultaneous Spanish-English Bilingual BLP Qualification Score: ≤-50 to ≥70

ii. Late Spanish-English Sequential Bilingual BLP Qualification Score: >-50

iii. Monolingual English BLP Qualification Score: >200

These dominance score thresholds for the Phase 1 speaker selection process were guided by the results from the pilot study. The bilingual speakers that patterned closest to one another and less like the sequential speaker fell within the -50 to 70 score range, while the sequential bilingual that functioned as Spanish Control in the pilot (pSC) scored in the hundreds and the monolingual English speakers nearly scored the threshold of 218. Given the overall success of the pilot study, these scores were adopted as the baseline for Phase 1.

The full BLP score results for all speakers are included in Appendix D, and the results and analyses for the speakers selected for Phases 3 and 4 are discussed at length in the results subsection for Phase 2, §3.4, after outlining the canonical speaker selection process.

3.3.2 Data Collection

"The North Wind and the Sun" Passage. This passage, published by the International Phonetic Association (1999, p.44), has been employed in numerous previous studies (Kondo et al., 2015; Knight 2011; among many others), including the pilot study for this dissertation. It contains relevant English vowels and consonants in multiple phonetic environments, which makes the passage ideal to capture segmental variation in English, an important aspect in drawing perceptual distance across dialects for listeners. With the goal of targeting plosive voice onset time, coronal fricatives, and vowel realization in English, the pilot study had listeners react to speakers reading the first three out of the four lines in this passage. While this longer recording provided listeners with more input, the task was more time consuming, which was not

adaptable to this study, as there are many more speakers in each task. Consequently, only the first sentence of this passage, below, was used to create the clips for all of the perception tasks in Phases 2-4 (see Appendix E for the full passage).

"*The North Wind and the Sun were disputing which was the stronger,*

when a traveler came along wrapped in a warm cloak."

As Calamai (2015, p.291) observes in her first three points on common methodological assumptions in perception research on linguistic attitudes:

i. Naïve listeners can explicitly identify both linguistic and social categories from short speech samples;
ii. Very little speech is needed to discriminate among linguistic varieties/accents/dialects;
iii. Ethnic group affiliation is recoverable from speech; […]

In choosing to reduce the length of auditory input by providing clips of the speakers only reading the first line of the passage, the methodology in this study follows the traditional assumptions in prior perceptual studies, as those outlined above from Calamai (2015). This shorter segment of the North Wind and the Sun still provides sufficient input from all of phonetic environments and realizations that were originally targeted in the pilot study.

As initially discussed in Chapter 2, data were also gathered through an adaptation of the Rootedness Metric (RM) Survey. The RM discussion and analyses for the selected PRIE and PR Spanish speakers (post-screening) are provided in the discussion of Phase 2, in §3.4, after outlining the results of the canonical speaker selection process.

3.3.3 Procedures

For Phase 1, potential speakers were contacted through multiple social networks, known acquaintances, and direct e-mails. Interested speakers who identified as a member of one of the speaker groups for this study would then proceed to complete the Qualtrics surveys that included the consent form and the BLP questionnaire. Once they completed the questionnaire and their BLP qualification scores were calculated, following the metric outlined above in the Screening subsection, those speakers were contacted to arrange an online interview.

The interviews were audio-only, and they were conducted and recorded through the Skype voice chat software. All audio recorded through Skype[11] was output in .MP4 format. The audio was then processed through Audacity (Audacity Team, 2015) to apply noise reduction and remove excess audio from the recordings, resulting in the final clips for each speaker in .WAV format, all between 6-8 seconds in length.

All speakers completed a total of four tasks, two of which are planned to be used for future papers following this dissertation, the Carrier Phrase Task and the "Frog, Where Are You?" (Mayer, 1969) Picture Story Narration Task. The complete list of tasks in the order in which they were administered is provided below:

[11] Downloaded from https://www.skype.com/en/

1. BLP Questionnaire
    a. *The initial screening questionnaire; not conducted live*
2. LEAP-Q Questionnaire
3. North Wind and the Sun Reading Passage
4. Frog, Where Are You? Picture Story Narration Task
    a. *Data and analyses forthcoming in future paper(s)*
5. Carrier Phrase Task
    a. *Data and analyses forthcoming in future paper(s)*
6. Rootedness Metric Survey
    a. *This task was only administered to Puerto Rican speakers from Puerto Rico, namely PRIE or the Puerto Rican Spanish-dominant sequential bilingual control speakers.*

The length of each interview was between 45 to 60 minutes on average, and no significant technical issues arose throughout the process. Upon task completion, all speakers received $20 as compensation.

Summary. The goal of Phase 1 was to identify and gather data from speakers who are members of the target dialect groups. All of the speakers (33) completed the first five tasks, outlined above, while only the Puerto Rican speakers from Puerto Rico (6) were given the RM survey. With the aim of identifying the most canonical in-group members for each simultaneous bilingual speaker group, twice the target final number per group (3) of these speakers were interviewed in preparation for a subsequent screening process in Phase 2. The remaining three per simultaneous bilingual group, alongside the three sequential Spanish-English bilinguals and the three monolingual English control speakers, proceeded from Phase 2 to Phases 3 and 4.

3.4 Phase 2: Canonical Speaker Selection Process

As already noted, the goal of Phase 2 is to determine and select the three most canonical simultaneous Spanish-English bilingual speakers from each speaker group—narrowing the final pool of simultaneous bilinguals in each group from six to three.

3.4.1 Task Design

To that end, a Matched-Guise Box Task was designed. Listeners were provided with ten recordings to group into four Dialect Group boxes through Qualtrics Survey software (see Figure 3.3). Listeners were asked to listen to the clips as many times as necessary and, by dragging and dropping each clip, to group the speakers into each Dialect Group box by those speakers whose 'accents' sounded similar to them (see Appendix F for the full task).

Figure 3.3: Matched-Guise Box Task, as seen by listeners in the Phase 2 study.

The ten recordings included all six speakers from one of the four simultaneous bilingual speaker groups (PRIE, Nuyoricans, Miami Cubans, and Chicanos) as the experimental sample, while the remaining four were control (two monolingual of the English speakers and two of the sequential Mexican Spanish-dominant bilinguals from the control groups identified in Phase 1), for a total of ten clips per item block. The four boxes provided listeners with a space to place speakers from each of the three groups into their own boxes, and the additional fourth box functioned as a distractor.

Because four groups of simultaneous bilingual speakers were tested, four blocks of this task were created. Each listener only completed one of the four blocks that were randomly presented using Qualtrics' survey distribution randomization algorithm. The four Matched-Guise Box item blocks were structured as shown in Table 3.3:

Table 3.3

Item Block Composition by Speaker Group for the Matched-Guise Box Task in Phase 2

Block A		Block B		Block C		Block D	
Speakers	N	Speakers	N	Speakers	N	Speakers	N
PRIE	6	Nuyorican	6	Miami Cuban	6	Chicano	6
English Control	2	English Control	2	English Control	2	English Control	2
Spanish Control	2	Spanish Control	2	Spanish Control	2	Spanish Control	2
Total	*10*	*Total*	*10*	*Total*	*10*	*Total*	*10*

In addition to presenting each block to listeners in a random order, the order of the clips in the task were also randomized to control for any block-internal item order effects.

3.4.2 Listeners

An initial 223 listeners responded to the Phase 2 Qualtrics survey. Of those initial responses, an additional screening was conducted with the following parameters: (1) observing the perceptual judgments of only self-identified native English speakers, and (2) controlling for

response quality by limiting the participant pool to only those who took at least five minutes to respond to the survey (the estimated time to listen to each clip at least once and respond to the demographics questionnaire after the task). After the secondary screening process, a total of 204 responses were recorded for this task—48 in Block A, 49 in Block B, 61 in Block C, and 46 in Block D. The disproportionate final number of listeners in Block C is due to an uneven number of filtered speakers divided in roughly equal amounts across the other three blocks.

3.4.3 Data Analysis and Speaker Selection Process

For this task, listeners grouped 10 speakers per item block (detailed above in Table 3.3) into the Dialect Group boxes to reflect the different dialect groups that they perceived with the given auditory stimuli. The resulting data output, after the screening process, was 204 individual response forms. To tabulate the data, each speaker pairing (e.g., EC1 with EC2) was manually counted as 1 under a corresponding crosstab cell whenever such a pairing occurred until all of the data were accounted for. Those final raw numbers were then double checked for consistency and then converted to a percentage value based on the total number of respondents for that value's item block, illustrated below, resulting in Tables 3.4-3.7 (one for each speaker group):

$$\frac{Cell\ Hits}{N\ Block\ Participants} \times 100$$

The output percentage value is then representative of the relative co-selection frequency between speakers. That percentage value was then used to determine canonicity by speaker group, using a combination of the following weighted criteria, in order of application:

1. Distance from Control Speakers: bilinguals who were less like the Spanish and English controls, indicated by lower control co-selection percentage values,

AND

2. In-Group Selection Frequency: bilinguals who were paired most with speakers from their group, indicated by higher in-group co-selection percentage values.

The resulting percentage values are presented in Tables 3.4-3.7 below, one for each simultaneous Spanish-English bilingual item block. The results in these tables are organized from most perceived like the English Control (EC) speakers to most perceived like the Spanish Control speakers (SC). Finally, the values were gradated using a color-coding scale to demonstrate the results across each respective speaker-to-speaker comparison. The speakers that were ultimately selected, using the criteria outlined above, are underlined and in bold.

Table 3.4

PRIE (PR): Gradated Participant-Speaker Selection Frequency by Speaker Group (N=48)

	EC1	EC2	PR5	PR1	**PR3**	**PR4**	**PR6**	PR2	SC2	SC1
EC1		73%	50%	42%	35%	21%	21%	4%	2%	2%
EC2	73%		63%	44%	33%	29%	23%	4%	0%	0%
PR5	50%	63%		44%	42%	33%	23%	10%	0%	6%
PR1	42%	44%	44%		48%	35%	35%	19%	2%	2%
PR3	35%	33%	42%	48%		35%	25%	27%	15%	19%
PR4	21%	29%	33%	35%	35%		54%	33%	10%	6%
PR6	21%	23%	23%	35%	25%	54%		33%	13%	10%
PR2	4%	4%	10%	19%	27%	33%	33%		46%	42%
SC2	2%	0%	0%	2%	15%	10%	13%	46%		79%
SC1	2%	0%	6%	2%	19%	6%	10%	42%	79%	

Table 3.5

Nuyoricans (NY): Gradated Participant-Speaker Selection Frequency by Speaker Group (N=49)

	EC1	EC2	NY2	NY1	**NY6**	**NY3**	**NY5**	NY4	SC2	SC1
EC1		88%	51%	39%	12%	6%	10%	4%	2%	2%
EC2	88%		51%	43%	6%	8%	6%	0%	0%	0%
NY2	51%	51%		45%	22%	16%	24%	12%	14%	12%
NY1	39%	43%	45%		27%	24%	14%	2%	6%	10%
NY6	12%	6%	22%	27%		51%	31%	31%	14%	12%
NY3	6%	8%	16%	24%	51%		39%	24%	14%	12%
NY5	10%	6%	24%	14%	31%	39%		29%	27%	31%
NY4	4%	0%	12%	2%	31%	24%	29%		53%	51%
SC2	2%	0%	14%	6%	14%	14%	27%	53%		76%
SC1	2%	0%	12%	10%	12%	12%	31%	51%	76%	

Table 3.6

Miami Cubans (CB): Gradated Participant-Speaker Selection Frequency by Speaker Group (N=61)

	EC1	EC2	CB4	**CB5**	CB3	**CB6**	**CB1**	CB2	SC2	SC1
EC1		70%	64%	52%	39%	38%	31%	26%	2%	0%
EC2	70%		44%	43%	48%	30%	33%	13%	0%	0%
CB4	64%	44%		54%	36%	51%	41%	30%	2%	7%
CB5	52%	43%	54%		33%	46%	64%	25%	0%	3%
CB3	39%	48%	36%	33%		26%	34%	21%	11%	13%
CB6	38%	30%	51%	46%	26%		41%	44%	7%	15%
CB1	31%	33%	41%	64%	34%	41%		25%	2%	5%
CB2	26%	13%	30%	25%	21%	44%	25%		20%	21%
SC2	2%	0%	2%	0%	11%	7%	2%	20%		79%
SC1	0%	0%	7%	3%	13%	15%	5%	21%	79%	

Table 3.7

Chicanos (CH): Gradated Participant-Speaker Selection Frequency by Speaker Group (N=46)

	EC1	EC2	CH2	**CH4**	**CH3**	**CH1**	CH5	CH6	SC2	SC1
EC1		70%	52%	48%	46%	26%	9%	9%	0%	4%
EC2	70%		48%	70%	46%	30%	9%	11%	0%	4%
CH2	52%	48%		43%	67%	50%	15%	15%	0%	0%
CH4	48%	70%	43%		57%	39%	11%	15%	4%	4%
CH3	46%	46%	67%	57%		52%	13%	9%	0%	0%
CH1	26%	30%	50%	39%	52%		26%	33%	4%	4%
CH5	9%	9%	15%	11%	13%	26%		50%	37%	39%
CH6	9%	11%	15%	15%	9%	33%	50%		46%	46%
SC2	0%	0%	0%	4%	0%	4%	37%	46%		85%
SC1	4%	4%	0%	4%	0%	4%	39%	46%	85%	

As can be observed in Tables 3.4-3.7, all of the control speakers patterned most strongly with their in-group members, and within each simultaneous bilingual speaker group, each speaker illustrated a relatively strong co-adherence with other in-group speakers. The fact that the simultaneous bilingual groups varied in the color gradients more towards the English controls (PRIE, CB, and CH) or slightly more towards the Spanish controls (only NY) is an effect of speaker-specific language dominance (further detail later in this section). Applying the criteria outlined earlier in this section (distance from control speakers and the strength of in-group selection frequency) yields the resulting speakers that are underlined and in bold in each table, where, for CH, PRIE, and NY speakers, they all adhere closest to one another near the median of the distribution. The averaged co-selection frequencies of the simultaneous bilinguals are provided in Table 3.8.

Table 3.8

Speaker In-Group Co-Selection Frequency Average in Descending Order

PRIE4	38.3%		**NY6**	32.2%
PRIE1	36.3%		**NY3**	31.0%
PRIE3	35.4%		**NY5**	27.3%
PRIE6	34.2%		NY2	24.1%
PRIE5	30.4%		NY1	22.4%
PRIE2	24.6%		NY4	19.6%
CB5	44.3%		**CH1**	40.0%
CB4	42.3%		CH2	38.3%
CB6	41.6%		**CH3**	39.6%
CB1	41.0%		**CH4**	33.0%
CB3	30.2%		CH5	23.0%
CB2	28.9%		CH6	24.3%

The only exception to this median co-adherence observation is CB3, where the speaker appears to be an outlier, not exhibiting strong co-selection values with any of the groups in Table 3.8, with a slightly stronger association to the English controls than their own group (see Table 3.6). This distribution disqualifies CB3 from selection when applying both criteria, in-group selection frequency and distance from control speakers. In the cases of PRIE1, CB4, and CH2, they all exhibited strong co-selection frequency values with their other in-group members (36.3%, 42.3%, and 38.3%, respectively). However, PRIE1, CB4, and CH2 also demonstrated stronger co-selection frequency values with the EC speakers (43%, 54%, and 50%, respectively on average), disqualifying them when applying the first criterion, distance from control speakers.

Based on the results from the Matched-Guise Box Task and the criteria and analyses outlined above, the final selection of the most canonical simultaneous bilingual speakers from each group, which continued on to Phases 3 and 4 , is provided in Table 3.9—each with their respective new speaker reference IDs for the rest of the study.

Table 3.9

Selected Most Canonical Speakers by Simultaneous Bilingual Speaker Groups

Group	Original Speaker IDs	New Selected Speaker IDs, Respectively
PRIE	3, 4, 6	PRIE1, PRIE2, PRIE3
NY	3, 5, 6	NY1, NY2, NY3
CB	1, 5, 6	CB1, CB2, CB3
CH	1, 3, 4	CH1, CH2, CH3

3.4.4 Selected Speaker Language Backgrounds

After the initial screening Phase 1 and the selection process in Phase 2, the speaker pool for the study was narrowed down to 21—comprised of three speakers from each of the seven groups: PRIE, NY, CB, CH, EC, SCPR, and SCMX. For quick reference, the simultaneous bilingual/experimental groups are PRIE, NY, CB, and CH, while the English control group is EC and the Spanish control groups are SCPR and SCMX.

Each speaker group's self-reported language background is reported in Tables 3.10-3.12 below using the initial questions from the LEAP-Q (Blumenfeld & Kaushanskaya, 2007) and BLP questionnaires (Birdsong et al., 2012). To reiterate, the BLP provides an extensive quantification of speakers' sociolinguistic backgrounds by both language domains and overall dominance, while the LEAP-Q provides a more detailed quantification of speakers' language histories in relation to the scores provided by the BLP. This study triangulates the language profiles of all speakers to provide a more concise and thorough background of their domain-specific language proficiencies and overall dominance scores through the BLP, and usage/self-reported proficiency through the LEAP-Q.

Each speaker group's numerical values in this subsection are averaged from the three selected speakers from their respective group. Finally, those scores are rounded and provided down to the second decimal value unless otherwise indicated.

Table 3.10

Speaker Group Average Age and Self-Reported Linguistic History by Language (N=3/Group)

Group	*English*				
	Average Age	Began Acquiring	Became Fluent	Began Reading	Became Fluent Reading
PRIE	28	3	5	5.67	8.67
NY	38	2.33	3.67	5.67	6.67
CB	29	4	4.67	6	7
CH	24	3.33	5	6	8
EC	27	0.33	2	5	6.67
SCPR	29	6	14.67	12.67	18.33
SCMX	25	10.67	16	13	15.33
Group	*Spanish*				
	Average Age	Began Acquiring	Became Fluent	Began Reading	Became Fluent Reading
PRIE	28	1.33	3	5.67	9.67
NY	38	4.67	13.67	12.67	21
CB	29	0.33	6.67	6	12.33
CH	24	0.67	3	4	8.33
EC	27	N/A	N/A	12*	N/A
SCPR	29	0.67	0.67	4	6
SCMX	25	0	3.33	4.33	8

Note: *Only one EC speaker reported an attempt to learn to read Spanish for a brief time.

Table 3.10 illustrates the average self-reported age of exposure to each language for each speaker group. The groups' average ages are also provided in Table 3.10 for reference throughout Tables 3.11-3.12 and are rounded up or down from the nearest decimal score. The four simultaneous bilingual groups all reported a low age of exposure for both Spanish and English, ranging from 0.33 to 4.67, depending on relative dominance. The two Spanish control groups—all sequential bilinguals—reported a much higher age of exposure in English than in Spanish (0 and 0.67 to 6 and 10.67), while the monolingual English control group reported a very low age of exposure for English and no age of exposure for Spanish, as it would be expected. The relative age of exposure for each group patterns well with their relative age of

fluency in both speaking and reading. The results from Table 3.10 pattern well with the perceptual findings in Phase 2, including the inclination of each group's language dominance relative to the perception results discussed in §3.4.3.

Table 3.11

Speaker Group Average Years Spent in Different Environments by Language (N=3/Group)

Group	Environments	Years (*English*)	Years (*Spanish*)
PRIE	Country/Region where (*language*) is spoken	7	24.67
	Family where (*language*) is spoken	11.33	24.67
	School/working environment where (*language*) is spoken	21.67	16.67
NY	Country/Region where (*language*) is spoken	26.67	11
	Family where (*language*) is spoken	26	22.67
	School/working environment where (*language*) is spoken	12.67	14
CB	Country/Region where (*language*) is spoken	28.67	25.33
	Family where (*language*) is spoken	22.67	24
	School/working environment where (*language*) is spoken	25.67	20
CH	Country/Regions where (*language*) is spoken	24	24
	Family where (*language*) is spoken	15	24
	School/working environment where (*language*) is spoken	15.67	12
EC	Country/Regions where (*language*) is spoken	26.67	0
	Family where (*language*) is spoken	26.67	0
	School/working environment where (*language*) is spoken	23.33	0
SCPR	Country/Regions where (*language*) is spoken	6.33	26.67
	Family where (*language*) is spoken	8.67	27.67
	School/working environment where (*language*) is spoken	8.67	25.33
SCMX	Country/Regions where (*language*) is spoken	7.67	17.33
	Family where (*language*) is spoken	5.33	21
	School/working environment where (*language*) is spoken	7	14.33

Table 3.11 shows the average LEAP-Q self-reported years that each group spent in different domains that are crucial to language exposure and acquisition: language of the community through "country/region where (*language*) is spoken," language of the household through "family where (*language*) is spoken," and language in vocational/social networks through "school/working environment where (*language*) is spoken." These categories quantify language exposure by domain.

Because this type of exposure can happen simultaneously through bilingual/mixed monolingual domains, the values for both languages could and should be higher for balanced bilinguals, lopsided for sequential bilinguals, and non-existent for Spanish in EC. The results in Table 3.11 follow those exact expectations for the simultaneous bilingual groups—high numbers in both Spanish and English environments for the PRIE, NY, CB, and CH groups, the Spanish-dominant sequential bilingual groups—more than double the amount of exposure to Spanish than to English in all three domains, and EC, where there has been no significant exposure to Spanish in their language networks. The results from Table 3.11 provide a clearer picture of the language histories and backgrounds for each of the selected speaker groups.

Table 3.12, below, demonstrates the remaining averaged LEAP-Q self-reported results by speaker group for each language, quantifying language exposure (Q1), usage (Q2-6), and perceptions of both the speakers themselves and of those with whom they have interacted (Q7-10). The results from Table 3.12 reinforce the observations made in Tables 3.10-3.11, wherein the simultaneous Spanish-English bilinguals report a balanced affinity for each of their languages, the sequential bilinguals a preference for Spanish, and the English monolinguals a complete functional preference for English.

Table 3.12

Averaged LEAP-Q Results for All Speaker Groups[12] by Question Type (N=3/Group)

#	Question	PRIE	NY	CB	CH	EC	SCPR	SCMX
Q1	% Exposure to **Spanish**/**English** on a typical day	23/77	45/55	35/63	32/67	2/98	37/62	67/27
Q2	% Choice to use **Spanish**/**English** on a typical day	60/38	42/58	52/45	67/33	2/98	67/30	73/27
Q3	Self-rated proficiency in spoken **English** (0: None; 10: Perfect)	9.66	8.66	9.66	9.66	10	9	8.33
Q4	Self-rated proficiency in **English** comprehension (0: None; 10: Perfect)	9.33	9.33	10	10	10	9.33	8.67
Q5	Self-rated proficiency in spoken **Spanish** (0: None; 10: Perfect)	9.33	7.33	8	8.67	1	10	9.33
Q6	Self-rated proficiency in **Spanish** comprehension (0: None; 10: Perfect)	9.33	7.33	8.67	9.33	1.67	10	9.33
Q7	Self-rated foreign accent in **English** (0: None; 10: Pervasive)	4.33	2.67	1	0.67	0	4.33	5.67
Q8	Self-rated foreign accent in **Spanish** (0: None; 10: Pervasive)	0	4.67	4.33	2.67	9	0	0
Q9	Based on your accent, how frequently do others identify you as a non-native speaker of **English**? (0: Never; 10: Always)	3	4.67	0	1.67	0	6.33	7.33
Q10	Based on your accent, how frequently do others identify you as a non-native speaker of **Spanish**? (0: Never; 10: Always)	0	5	1	1.33	10	0	0

Note: Percentages may not add up to 100 due to rounding; decimals are not provided in Q1-2.

Given the results in Tables 3.10-3.12, the expectation is that the averaged BLP dominance scores, in Table 3.13 below, follow the speaker groups' LEAP-Q language attitude and history scores, which were also individually controlled for in the Phase 1 initial recruitment screening process.

Table 3.13 expands on the averaged BLP scores for each speaker group for each language by domain. To reiterate, each domain in each language has a maximum score of score of 54.5 and a minimum score of 0. The domain scores in each language are weighted and added together to generate a total language score, with a maximum of 218 (54.5 times four, one for each domain) and a minimum of 0. Finally, the Spanish language score is subtracted from the English language score to produce the overall dominance score (positive numbers for English dominance and negative numbers for Spanish dominance), which ranges from a final score of 218 (100% English) to -218 (100% Spanish). This scoring system was operationalized to control for bilingual dominance in this study as follows:

i. Simultaneous Spanish-English Bilingual BLP Qualification Score: $\leq$-50 to $\geq$70

ii. Late Spanish-English Sequential Bilingual BLP Qualification Score: >-50

iii. Monolingual English BLP Qualification Score: >200

Additional flexibility is provided towards the positive end of the score thresholds between simultaneous and sequential bilinguals (from -50 to +70 in i. above). This broader range in the selection criteria accounts for the fact that most of the simultaneous bilinguals were raised in a more English-influenced environment (the mainland U.S.). This naturally skews their BLP dominance scores more towards the positive end, despite all their other factors patterning like the simultaneous bilingual PR speakers, who were raised in Puerto Rico.

Table 3.13

Averaged BLP Results from Speaker Groups by Domain with Dominance Scores (N=3/Group)

Domains		PRIE	NY	CB	CH	EC	SCPR	SCMX
Language History	Eng.	40.71	47.52	42.37	42.22	51.45	24.36	23.76
	Spa.	44.49	31.78	40.25	42.37	0	51	45.4
Language Use	Eng.	25.07	42.87	35.97	31.25	54.5	17.44	17.8
	Spa.	29.43	11.63	18.53	23.25	0	37.06	35.97
Language Proficiency	Eng.	54.48	54.48	54.48	54.48	54.48	48.43	44.64
	Spa.	53.72	37.83	44.64	49.94	7.57	54.48	53.72
Language Attitudes	Eng.	50.7	48.43	51.45	37.08	54.48	32.54	32.54
	Spa.	54.48	45.4	46.16	51.45	0.76	54.48	53.72
Total Score	Eng.	170.96	193.3	184.28	165.03	214.91	122.77	118.74
	Spa.	182.13	126.64	149.58	167.02	0	197.02	188.82
Dominance Score		-11.17	66.66	34.69	-1.99	214.91	-74.25	-70.07

The BLP domain and dominance scores provided in Table 3.13 provide a quantified summary of the combined language histories and profiles of each speaker group in this study that can be referenced for the results of the findings in Phases 3 and 4.

3.4.5 Puerto Rican Bilingual Speaker Rootedness Metric Results

With the selected canonical subset of PRIE speakers, a further exploration between their ties to their Puerto Rican identities and their dominance scores can be explored using the Rootedness Metric (RM). While there is prominent sociopolitical impetus for the emergence of a variety of American English in younger Puerto Rican bilinguals, as discussed in Chapter 2, the amelioration of linguistic attitudes towards English functions as a catalyst for its emergence. As such, there is reason to investigate the degree to which the local identities of these Puerto Rican speakers are tied to the island. To that end, this study adopts and adapts RM from Reed (2016 and 2018, among others), a contextually designed in-group measurement tool which aids in quantifying a speaker's attitudinal orientation towards a region.

As the goal in this part of the study is to ascertain to what degree the Puerto Rican speakers' identities are tied to the island, this survey was only given to the PRIE and the PR Spanish-English sequential bilingual control speakers. Much like in Reed (2016), the metric was designed with two principal parts, a question and answer section and an interview section. Item creation for each of those sections was informed by: (a) my own familiarity, as a Puerto Rican, to the aspects that drive local identity in the island, and (b) an adaptation of the same categorical structure used in Reed (2016 and 2018) (see Appendix G for the full survey).

Subsequently, each item or question was quantified through an RM score, derived from two sections of a semi-structured Rootedness interview process (see Table 3.14, below). The closed questions in Section I had a range of scores based on expected answers, depending on the degree of the speakers' positive or negative attitude responses, while the RM scores for the open questions in Section II range from +1 to -1, in line with the same parameters. This RM score structure follows a similar categorical system to that in Reed (2016, p.76), ranging from the importance of Puerto Rico, as a place, in speakers' identities to connections to language and current events. Hence, the minimum and maximum scores for each type of category by section were weighted and adjusted to follow Reed's (2016) RM scoring system parameters. The resulting range of possible scores for the RM in this study extend from -8 to 27, where the higher RM score indicates closer ties to Puerto Rico and/or a locally grounded Puerto Rican identity. Finally, because some of the Puerto Ricans interviewed lived on the island, while others lived in the mainland, two separate sets of items with parallel scoring systems are outlined for this study's RM survey, where the segments of the questions in brackets are separated by slashes. One set of these question segments (PR versus U.S.) was presented to speakers, depending on whether they were living on the island or in the mainland United States, respectively.

Table 3.14

RM Scoring System for Puerto Ricans [Living on the Island / in the Mainland U.S.]

Q ID	Section I: Closed Questions	Answer	RM Score
Q1	Willingness to [Relocate to U.S. / Move back]	Yes No	0 to -1 -2/+2
Q2	Travel Habits to [the U.S. / PR]	Frequent Rare	0 -1/+1
Q3	Origin Self-ID (In-Group)	Puerto Rico (+Variants) U.S. (+Variants)	+2 0
Q4	Origin Self-ID (Out-Group)	Puerto Rico (+Variants) U.S. (+Variants)	+2 0
Q5	Family/Friends Living in PR	5+ Family / 3+ Friends 2-4 Family / 2-3 Friends <2 Family / <2 Friends	+2 +1 0
Q6	Information Medium about Local Events in PR	PR News Networks National/Social Networks Little or None	+2 +1 0
Q7	Areal Identification	[Name of Hometown] La Isla (+Variants) Puerto Rico Latin America The Caribbean The United States	+5 +4 +3 +2 +1 0
Q8	Connection to [PR / U.S. Sports]	PR Teams/Players Others	+1 0
Q9	Connection to [PR / U.S. Popular Culture]	PR Pop Culture Primary US Pop Culture Primary Both or None	+2 -1 0
Q10	Language Self-ID/Opinions	Only/More Spanish Only/More English Both	+1/2 -1/2 0
Q11	ID Tied to Puerto Rico (Averaged with U.S)	Closely Tied Somewhat Tied Not Tied	+3 +2 0
Q12	ID Tied to the United States (Averaged with PR)	Closely Tied Somewhat Tied Not Tied	-3 -1 0
Q ID	**Section II: Open/Interview Questions**		**RM Score**
ID1	Would you say you identify with Puerto Rico? Your hometown? Why?	**+1 to -1 Each**	
ID2	Is there another place that you identify with? Why?		
ID3	What makes it (PR) so special?		
Possible RM Score Range: -8 to 27			

Following the methodological motivations for the RM design in Reed (2016, p.76), each of the categories outlined in Table 3.14 align with listeners' informational and familial connections to Puerto Rico, including the use of local terminologies for regional distinctions. The closed question categories in Section I of the interview cover attachment to or alignment of: place (Q1-Q2), self-identification towards in-group and out-group members (Q3-Q4), geographic centricity of relationships (Q5), informational networks (Q6), self-identification with local and general place names for Puerto Rico (e.g., "La Isla" versus "The United States") (Q7), connection to local sports and popular culture (Q8-Q9), language use (Q10), and identification with place in terms of the island of Puerto Rico versus the mainland United States (Q11-Q12). These scores that quantified answers to the questions in the RM survey were individually tallied for each of the six Puerto Rican bilingual speakers. Before turning to overall scores, Table 3.15 illustrates these individual RM scores by speaker and question ID.

Table 3.15

Individual RM Score Results by Question ID for the PR Speakers Selected in this Study

PR Speaker ID	Rootedness Metric Question ID														
	Q1	Q2	Q3	Q4	Q5	Q6	Q7	Q8	Q9	Q10	Q11	Q12	ID1	ID2	ID3
PR1	2	1	2	0	2	1	5	0	-1	0	3	-1	1	0	1
PR2	2	1	2	2	2	2	5	1	2	0	3	-3	1	1	1
PR3	2	0	2	2	2	2	5	1	1	0	3	-1	1	1	1
SCPR1	-1	0	0	2	2	1	4	0	1	2	2	-3	1	-1	-1
SCPR2	1	-1	1	2	2	1	-1	1	2	0	3	-3	1	0	1
SCPR3	-1	1	2	2	2	2	-1	1	1	0	3	-1	1	0	1

Note: A color-coded gradient is provided to contextualize the results by question ID in each column. Because each question ID has a different range of possible scores (see Table 3.13), the provided color-coded gradient is weighted from column to column, rather than from the overall distribution of scores in the table.

The response patterns to the RM survey questions can be divided into three categories: (1) *Strong PRIE and SCPR Rootedness*, where both Puerto Rican speaker groups demonstrate

strong rootedness; (2) *Strong PRIE Rootedness Only*, where the PRIE group demonstrates stronger rootedness than the SCPR group; and (3) *No Rootedness*, where neither group appears to index rootedness in response to these questions. There were no RM survey questions in which the SCPR group demonstrated more rootedness than the PRIE group.

For the first pattern of responses, both the PRIE and SCPR groups were highly rooted in: identifying as Puerto Ricans to out-group members (Q4), the amount of familiar connections in Puerto Rico over the mainland United States (Q5), sports culture (Q8), popular culture (Q9), identification with Puerto Rico as their home (Q11 and ID1), and in identifying Puerto Rico as a place that is special and unique to them (ID3). These findings are expected, as these questions all identify core elements of Puerto Rican identity, especially to those Puerto Ricans forming part of the new diaspora to the mainland, of which all of these speakers are members.

In the second pattern of responses, PRIE speakers tended to be more rooted than the SCPR speakers in their willingness to move back to Puerto Rico (Q1), frequency of travel back to Puerto Rico (Q2), identify as Puerto Ricans to in-group members (Q3), use local Puerto Rican news sources (Q6), have stronger ties to their hometowns and Puerto Rico overall (Q7), and in underscoring Puerto Rico as the only place with which they identify (ID2). Alongside the fact that the SCPR group did not produce stronger rootedness scores than the PRIE group in any category, this divergence was not expected, as it could be anticipated that the more Spanish-dominant speakers would be more strongly rooted in Puerto Rico. A possible explanation could be stronger social pressure for the SCPR speakers to assimilate to the culture of the mainland. This observation is explored in the overall discussion of the findings further below.

In the third and final pattern of responses, neither group appeared strongly rooted in identifying Spanish as a core element of their identities over English (Q10) or in dissociating

their identities from the mainland United States and choosing to tie themselves only to Puerto Rico (Q12). The results for Q10 and Q11 in the RM survey are not surprising in that they reinforce the discussion in Chapter 2, where recent research indicates that younger generations of Puerto Ricans who are increasingly bilingual and bicultural have ameliorated language attitudes towards English and view cultural ties to the United States as part of their identities. In that sense, these findings reinforce this sociopolitical profile in the Puerto Rican Spanish-English bilingual speakers in this study.

The averaged RM scores in Table 3.16 were determined using the guidelines provided in Table 3.14 and the individual RM results in Table 3.15. For comparative analysis, Table 3.16 provides the BLP dominance scores (from Table 3.13) and the averaged RM scores for all six of the selected Puerto Rican (PRIE and SCPR) speakers. The BLP dominance scores contextualize the speakers' language dominance and usage with their degree of rootedness to a Puerto Rican identity and Puerto Rico as a whole. The speakers' region of residence at the time of the interview is also provided as a reference to their perspectives when answering the RM survey.

Table 3.16

BLP Dominance and Rootedness Scores for Puerto Rican Speakers

Speakers	Residence at Time of Interview	BLP Dominance Score	Rootedness Score
PRIE1	Mainland U.S.	6.72	16
PRIE2	Mainland U.S.	-26.7	22
PRIE3	Puerto Rico	-13.53	22
SCPR1	Mainland U.S.	-92.53	9
SCPR2	Mainland U.S.	-54.76	10
SCPR3	Mainland U.S.	-75.46	13

While the highest possible RM score is 27 and the lowest -8, none of the speakers produced a negative score, with the highest actual score being 22 and the lowest 9. Consistently,

the simultaneous bilinguals (PRIE average RM score: 20; median: 22) demonstrated higher RM scores than the sequential bilinguals (SCPR average RM score: 10.67; median: 10). Furthermore, there was a notably strong positive correlation between the speakers' dominance and rootedness scores (r=0.73). In other words, a Puerto Rican speaker with a higher dominance score (signifying less Spanish dominance to more balanced/English dominance) also had a higher degree of ROOTEDNESS to Puerto Rico; this spread is indicative of a categorical relationship between rootedness and BLP dominance scores. Figure 3.4 illustrates this categorical correlation on a scatterplot, with a set of ovals for each group of data points corresponding to the Puerto Rican speakers' rootedness and BLP dominance scores from Table 3.16. Higher rootedness scores, on the Y-axis, indicate a stronger connection to a Puerto Rican identity.

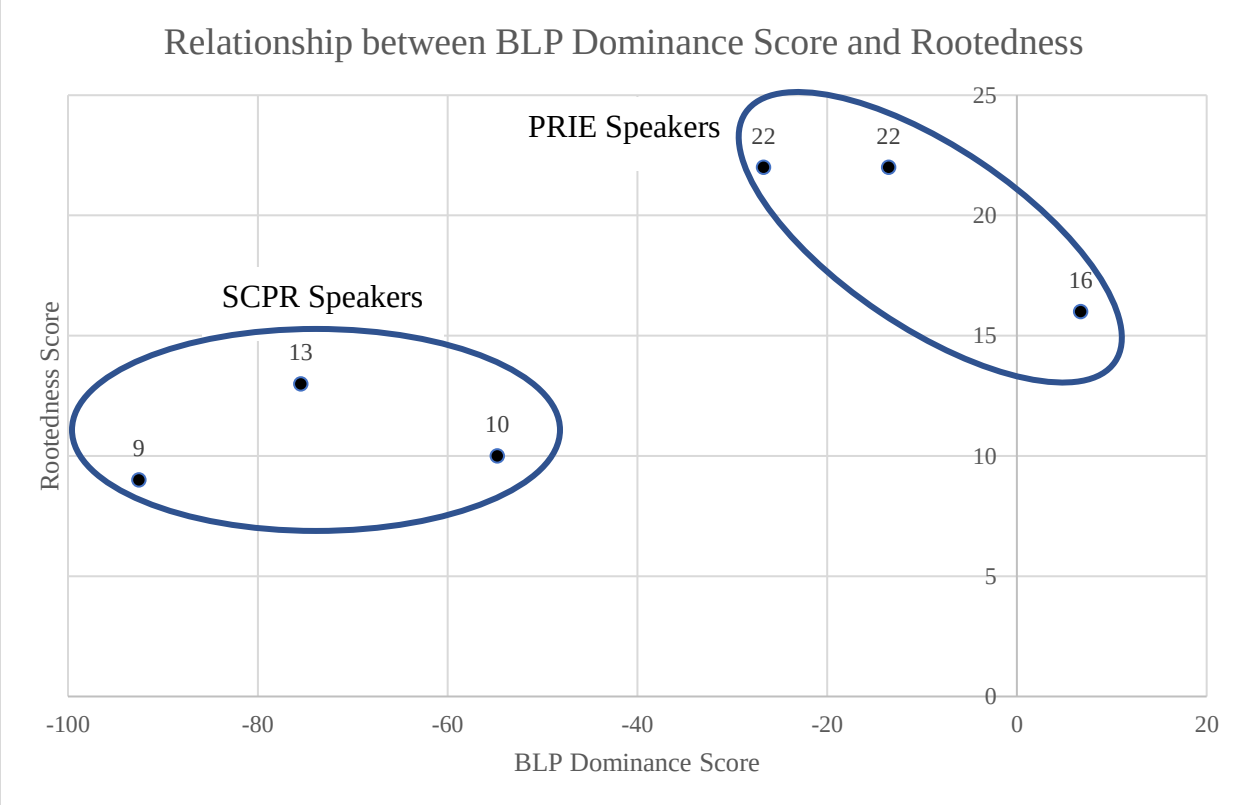


Figure 3.4: Scatterplot demonstrating the relationship between the PRIE and SCPR speakers' BLP dominance and rootedness scores; the ovals underscore the categorical split in the graph.

This correlation is unanticipated in that one would expect that those speakers who are more Spanish-dominant (SCPR) would also have stronger ties to the island.[13] However, there is a possible explanation for this correlation in that all of the SCPR speakers were living on the mainland, and their motivation to counteract their strongly Spanish-accented English due to assimilatory social pressures can also lead towards a stronger dissociation to an island-centered identity. On the other side, the balanced simultaneous bilinguals (PRIE) do not exhibit the same identifiably strong accent, which can account for a tendency to be able to compartmentalize their identities from their linguistic performance, allowing for the preservation of stronger ties to the island. This observation is supported by the speakers' self-rated identifiable foreign accent LEAP-Q score discrepancies (PRIE speakers: 3 versus SCPR speakers: 6.33).

While all of the speakers are rooted to Puerto Rico (to varying degrees), there is a marked effect of the ongoing diaspora in their perspectives on the importance of place, language, and identity. Five of the six speakers identified with Puerto Rico the strongest, and only one of them identified with their hometown in the island. Throughout their open-ended responses, there was a consistent metalinguistic theme that combined missing home and accepting movement as a fact of life:

> SCPR1: *I identify myself as Puerto Rican, born and raised, but I don't see myself tied to my hometown anymore because, in PR, I lived in multiple places, so I've learned to move, and I've learned to not create a real connection with a place. Not really. I don't miss my hometown or the places I lived in; I just missed my friends. Hometown is wherever I go. I don't have homesickness. I don't even miss it. I'm okay with that*.

[13] These findings illustrate the complexity of the relationship between rootedness/identity and language. The sequential Spanish bilinguals in this study are less rooted to their Puerto Rican identity than the simultaneous bilinguals. Parallel findings with German Americans in Wisconsin have indicated a similar rootedness/identity to language effect of that observed in this study. The bilingual German Americans were less rooted in a German identity expression than the monolingual English German Americans (Samantha Litty, Personal Communication, January 5, 2021). See Litty (2017) for more information on German Americans in Wisconsin.

> PRIE1*: Puerto Rico was my place of upbringing, where I acquired most of my cultural practices, alongside my languages. My hometown is not quite as important, since it is just a larger part of the whole, and it's not doing as great as it was when I was being raised there. I often think about going back home, but it's just not possible at this time. I do visit my parents at least once a year though.*

Even though there are strong ties to Puerto Rican customs and practices, there is a sense that residing in Puerto Rico is somewhat disconnected from that aspect of their identities due to necessity. The analysis of the RM and dominance scores, in tandem with the background provided by the speakers' qualitative responses and their quantified histories through the BLP and the LEAP-Q questionnaires, contextualize the results in Phases 3 and 4. The Puerto Rican simultaneous bilingual speakers (PRIE) are somewhat more strongly tied to the island, despite the fact that they are either slightly more English-dominant or much less Spanish-dominant than the Puerto Rican sequential bilinguals (SCPR) who are functioning as part of the Spanish controls.

3.5 Summary

This chapter first introduced the pilot study that helped sharpen the research questions for this dissertation project. The pilot study, which focused on surveying perceptions of whether the English of simultaneous Puerto Rican Spanish-English bilinguals was perceived as an L1 English dialect, found that listeners from the Upper Midwest largely perceived those simultaneous bilinguals as native speakers of English. I termed this emerging variety of American English as Puerto Rican Island English (PRIE) in the conclusion of that pilot study. With those findings in mind, the research questions for assess: (1) whether PRIE is emerging as a distinct variety of American English—separate from other varieties of bilingual Englishes in the mainland U.S., (2) whether PRIE is subliminally associated with Puerto Rico—given that it has not yet enregistered,

and (3) whether PRIE presently has any overt regional dialect associations and, if so, to what extent.

The pilot study did present difficulties in terms of assessing both (1) a balance between the language proficiencies and performance of the bilingual speakers, and (2) the effect that prior linguistic knowledge and sensitivity to dialect variation affected the accuracy of the listeners' perception judgments. To that end, the methodology was broken down into four phases. The first two phases, covered extensively in this chapter, addressed the first confound in the pilot study by screening for speakers more extensively in Phase 1 and then determining canonicity in Phase 2 for a more coherent bilingual speaker pool.

Due to the importance of language attitudes in the emergence of PRIE, the Rootedness Metric was also introduced to gauge the strength of the Puerto Rican bilingual speakers' connections to the island in relation to their BLP dominance scores and their results in Phases 3 and 4. In combination with the BLP and LEAP-Q questionnaires, the Rootedness Metric provides additional context for the language histories and attitudes of the Puerto Rican speakers.

The second confound from the pilot study is covered in the methods for Phases 3 and 4 through a series of pre-tests and additional demographic breakdowns in the inferential and descriptive statistical analyses. The methods and results for the remaining two phases are elaborated on further in Chapter 4 for Phase 3 and Chapter 5 for Phase 4.

CHAPTER 4

SIGNAL DETECTION TASKS

This chapter reports on the results of Phase 3, which comprises the Identification (ID) Task and the AX Discrimination Task (AX Task). These tasks are designed to address the first two main research questions through modified approaches in PERCEPTUAL DIALECTOLOGY, discussed earlier in Chapter 3. These modified approaches are grounded in previous synchronic analyses of incipient dialect formation that observed the enregisterment of previously neutral varieties of American English, such as of Alaskan English as a variety influenced by migration from the Upper Midwest (Purnell, Raimy & Salmons, 2009) and of Wisconsin English by way of college-aged speakers (Schuld et al., 2017). The framework of PERCEPTUAL DIALECTOLOGY explains listeners' motivations to interpret linguistic features from a sociolinguistic perspective—both in relation to their own social groups and in understanding a speaker's dialect as a factor in determining social and/or regional group membership.

In this dissertation, PERCEPTUAL DIALECTOLOGY is taken a step further and employed as a means to explore NEW DIALECT FORMATION as a byproduct of a dialect's perceptual saliency to listeners and the dialectic sociolinguistic awareness of the speakers of that dialect. This kind of approach was first exemplified in Schuld et al. (2017), where selected recordings were presented to listeners around the country and employed to survey the status of Wisconsin English as a developing regional dialect. Building on that approach, this chapter explores ongoing dialect formation by incorporating SIGNAL DETECTION THEORY, listeners' language histories and linguistic awareness, and inferential statistical modeling to extant tools and methods in PERCEPTUAL DIALECTOLOGY.

To that end, the first section of this chapter introduces the inferential statistical models used to analyze the results of the Identification and the AX Tasks in Phase 3. The second section summarizes the methodological motivations for each task in Phase 3. The third and fourth sections examine the methodology and results for each of the signal detection tasks in Phase 3, respectively. Finally, the fifth and last section discusses the overall findings.

4.1 Statistical Analysis in Phase 3

As initially reviewed in Chapter 2, SIGNAL DETECTION THEORY measures the discriminability or salience of a set of signals, for machines, or stimuli, in the case of listeners. These stimuli or targets are paired against a noise category, often referred to as a distractor, and listeners are tasked with distinguishing between the two in a set of trials.

Table 4.1

2x2 Stimulus-Response Combinations for Signal Detection Test

Stimulus	Participant Response - Present	Participant Response - Absent
Present (Target)	Hit	Miss
Absent (Distractor)	False Alarm	Correct Rejection

Table 4.1 illustrates the possible responses in a set of signal detection trials. The goal of this model is to capture participants' discriminability between predefined stimuli and distractors. As summarized earlier in the introduction of the pilot study results in Chapter 3, an expected positive response is identified as Hit, an unexpected negative response as a Miss, an unexpected positive response as a False Alarm, and an expected negative response as a Correct Rejection. Those results are then used to calculate the: Hit Rate (H), which tells us how accurate listeners were at correctly identifying a signal; False Alarm Rate (FA), which tells us how often participants incorrectly identified a distractor as a stimulus. These measures can then be converted to calculate dependability and accuracy—to be expanded on the discussion of A-Prime (A′) in §4.1.1.

In psychology, signal detection has been widely employed to measure decision making, perceptual discriminability, and response bias (Pallier, 2002). Given the appropriate adaptational methodological designs, signal detection tasks can function as a powerful tool in PERCEPTUAL DIALECTOLOGY to correct for the participant bias that is particularly implicit in these data and to improve on the generalizability of the output of the results.

The chief difference between the application of SIGNAL DETECTION THEORY to PERCEPTUAL DIALECTOLOGY in this study and the standard methods is that the stimuli are not immediately discrete, such as a machine generated vowel sound or a set of pictures. Instead, the targets and distractors are real speakers reading a text, where the speakers' voices are a complex amalgamation of their language backgrounds and linguistic performances. The targets and distractors are complexified in comparison to the aforementioned examples. The results from a signal detection analysis in this context are intended to be revelatory, rather than examinational. In other words, the signal detection analysis explores the strength of the perceptual relationship that listeners draw between selected speaker groups, instead of quantifying performance on a strict set of baseline expectations.

In this dissertation, with the qualification outlined above, signal detection is employed to test both participants' discriminability/response biases of predefined stimuli (through the Non-Parametric A′ measure) in the results of the AX Task and overall signal strength of the target stimuli in participants' perception judgments (through what I term a Differential A′ measure) in the results of the ID Task. Both approaches are detailed in §§4.1.1 and 4.1.2.

4.1.1 Non-Parametric Analysis: A-Prime

In statistics, methodologies can vary between parametric or non-parametric approaches based on whether the data are normally distributed. Although a parametric analysis of a data set is more powerful, it requires both normally distributed results and the same standard deviation between the distributions (Stanislaw & Todorov, 1999, p.140). Because the linguistic data from these results represent an unknown distribution of the population and an unknown number of variables, which is typical in studies involving human subject responses, a nonparametric approach (A′) was chosen instead of its parametric equivalent, D-Prime.

In A′, five values are produced with the measures outlined in the discussion of Table 4.1 above: Hit Rate (H), False Alarm Rate (FA), Percent Correct (% C), the A-Prime score (A′), and the Response Bias score (B''_D). H represents the percentage of the target trials in which listeners correctly identified a stimulus, while the FA Rate indicates the percentage of the distractor trials in which listeners incorrectly identified a distractor as a target. The Percent Correct score denotes the percentage of all trials in which listeners correctly identified the targets and distractors as intended. The H and FA Rates are then used to calculate the A′ and B''_D scores.[14]

A′ can range from 0 to 1, where 0 is no discriminability, 0.5 is chance discriminability between targets and distractors, and 1 is perfect performance. Most A′ scores typically range from 0.5 to 1 (Stanislaw & Todorov, 1999, p.140), and that range of performance can be used to quantify overall discriminability across the categories or groups forming part of the trials for listeners; answers can then be derived from that comparative performance. Finally, B''_D measures the response bias from the sample and can range from -1 to 1, where -1 is an extreme liberal bias (tendency to identify a stimulus as a target), 0 is no response bias, and 1 is an extreme

[14] See Stanislaw & Todorov (1999) or Pallier (2002) for a more detailed explanation of the mathematical principles behind A′ and B''_D score calculations.

conservative bias (tendency to identify a stimulus as a distractor). More information on the A′ procedures specific to the AX Task is provided in §4.4.3.

4.1.2 Differential A-Prime Approach

While, as detailed above, A′ is typically used to control for bias and accuracy in discriminability tasks, this dissertation repurposes the tool to identify listeners' salient perceptions of the selected dialect groups. This analysis is controlled to address the specific research questions by: (1) a task design that only uses the audio clips of these dialect speakers as the stimuli, (2) question prompts that guide listeners' responses to the stimuli, and (3) audio clips of all the speakers reading the same text out loud.

Therefore, because A′ can be used to determine whether listeners can correctly discriminate between fixed stimuli, A′ can also be repurposed to observe the strength of the signals from variable stimuli (speakers), and how these signals pattern in comparison to one another. With that premise in mind, the aim of the Differential A-Prime (dA′) approach is to analyze the A′ results categorically. This categorical analysis can be employed to determine the dialect group(s) that generate the strongest response by manipulating which trial stimuli groups are defined as targets or distractors in the A′ calculation process accordingly. The resulting highest dA′ scores are compared against one another to determine the maximal signal. The maximal signal is reanalyzed as the 'target.' Based on the results from listeners' judgments in the task, the stimuli group that listeners identify as the maximal signal is derived to be the de facto target of that signal detection task. The results from that dA′ analysis can then find answers to the relevant research question.

This dA′ approach addresses the research question relevant to the ID Task in Phase 3. More information on the categorical selection process and analysis procedures that are specific to that task is provided in §4.3.3.

4.2 Phase 3: Signal Discrimination Tasks

Chapter 3 introduced Phases 1 and 2, which covered the initial concerns that emerged from the pilot study through a much more rigorous speaker selection and screening process. A second confound, identified in the pilot, was to control for listeners' preexisting language experiences and their overall sensitivity in identifying dialectal differences among the speakers. This confound is accounted for in the methods for Phases 3 and 4 through a series of pre-tests and additional demographic breakdowns in the inferential and descriptive statistical analyses for the remaining tasks in Phases 3 and 4.

With these methodological adjustments in mind, the tasks in Phase 3 (the ID Task and the AX Discrimination Task) provide evidence on the first two specific research questions, respectively:

1. Based on the speakers' audio recordings and no other background information, can listeners identify PRIE speakers distinctly from other similar varieties of mainland American Englishes (Nuyorican, Chicano English, and Miami Cuban English)?
2. How do speakers of Puerto Rican Englishes (PRIE and Nuyorican) compare to other selected recognized varieties of mainland American Spanish-English bilingual varieties (Chicano English and Miami Cuban English)?

Both of the signal discrimination tasks in Phase 3 (ID and AX Tasks) were provided in the same Qualtrics online survey and, subsequently, share the same listeners. Listeners completed the same demographics questionnaire used for Phase 2 (see Appendix A for the

demographics questionnaire). In this subsection, the methodology for each of the tasks is first introduced, followed by their results and an analysis of the findings.

4.3 The Identification Task

The Yes-No Forced Choice Identification Task (ID Task) aimed to answer the first research question of the study, to observe how listeners grouped the PRIE speaker group when listening to all seven of the speaker groups included in this study (PRIE, CB, CH, NY, EC, SCPR, and SCMX). The ID Task provides insight on listeners' initial categorical perceptions of the PRIE, NY, and SCPR speakers in this study in comparison to the other non-Puerto Rican speaker groups. This task examines whether listeners can identify the Puerto Rican speakers as a separate group, and particularly, whether perceptions of PRIE speakers pattern with the NY and/or the SCPR speakers, or not at all.

4.3.1 Methodology

Listeners responded to the one-sentence clips of speakers reading the excerpt from "The North Wind and the Sun" described in Chapter 3; the clips from all 21 speakers were used for this task. Listeners were given the following prompt question:

- *Main Prompt Question: Does this speaker sound like they have Puerto Rican ancestry? That is, either from Puerto Rican communities in the mainland United States or from the territory of Puerto Rico*.

Listeners then heard an audio clip corresponding to a speaker and then answered 'yes' or 'no' to the prompt question (see Appendix H for a sample item of the task). Since there are 21 speakers in this task, listeners responded to 21 experimental trials and five distractor trials, for a total of 26 items. The five distractor trials included five randomly selected unique speakers from the initial 21-speaker pool. Those five speaker clips were repeated in the task once more, with the exception that the same prompt question above was modified to the following:

- *Distractor Prompt Question: Does this speaker sound like they have **Chicano** ancestry? That is, from **Chicano** communities in the mainland United States.*

All 26 items for the ID Task were presented in sets of three (and one of two) per survey page, and each set was randomly presented, following the Qualtrics page presentation randomization algorithm. Listeners were not allowed to change their answers once they were submitted.

4.3.2 Listeners

An initial 130 listeners responded to the Phase 3 Qualtrics survey. Of those initial responses, an additional screening was conducted with the following parameters: (1) observing the perceptual judgments of only self-identified native English speakers, (2) controlling for listeners who reported that they had not had sustained contact with a Spanish-speaking community, and (3) removing the only two listeners who identified as a member of a Spanish-speaking community, as the sample size for that subgroup is too small to consider for analyses. After the secondary screening process, a total of 91 responses were recorded for this task. Of the remaining 91 listeners, 72 (79%) were living in the Upper Midwest or the Northeast, and the remaining 19 (21%) listeners were living in other regions in the mainland U.S. All listeners reported having lived in the United States throughout their childhoods.

4.3.3 Data Analysis

In terms of expectations for the ID Task, consider Table 4.2 below:

Table 4.2

Analysis by Grouping Types in Response to the Yes-No Question in the ID Task

Grouping Type	Speaker Group(s) Identified as Puerto Rican	Interpretation
PRIE Only	PRIE	PRIE Awareness
Sequential PR Only	SCPR	Awareness of L2 English PR PRIE speakers undetected
Nuyoricans Only	NY	Association of enregistered to island PRIE speakers undetected
Enregistered L1 Community	PRIE + NY	Recognition of PR accent substrate
Enregistered + Accented	NY + SCPR	PR accent and ancestry connection PRIE speakers undetected
All Puerto Rican Speakers	PRIE + NY + SCPR	Perfect discrimination of PR speakers
Only Sequential Speakers	SCPR + SCMX	Spanish accent association to PR PRIE speakers undetected
Low Familiarity	CH + Any	Low familiarity of these dialects
Hispanic Speakers	All - EC	Hispanic speakers in general = PR
No Differentiation (Y Bias)	All Groups	No differentiation between speakers
No Differentiation	Random/NONE	No differentiation between speakers

Depending on how listeners responded, the results were analyzed according to the categorical breakdown provided in Table 4.2. Given the sociohistorical status of Puerto Rico, the current suspected status of PRIE, and the enregisterment of Nuyorican English, the following results are not expected: SCPR, All-EC, All Groups, CH + Any, Random/NONE, among other combinations not included in Table 4.2. The listeners' respective response patterns depend on the degree of enregisterment that they may have about the selected dialects. With the expectation that PRIE has not been overtly enregistered, NY+SCPR is the expected predominant result. Otherwise, should PRIE be enregistered as a distinct dialect of Puerto Rican English, it would pattern together with NY. Nevertheless, any of the plausible results where PRIE is not part of the

group composition of overtly identified dialects would suggest that PRIE has not been enregistered. In this scenario, it is plausible that PRIE is perceived as its own category, neither following the perceptual patterns of any of the other speaker groups nor being overtly identified as part of a Puerto Rican community.

Differential A-Prime. To reiterate, the aim of the dA′ approach is to analyze the A′ results categorically. This categorical analysis can be employed to determine the dialect group(s) that generate the strongest response by manipulating which trial stimuli groups are defined as targets or distractors in the A′ calculation process accordingly. The process is essentially an iterative A′ analysis, determining different speaker groups as the target stimuli to identify strong signals according to listeners' judgments. This categorical analysis follows the combinations and interpretations listed in Table 4.2 above.

Since the results from each A′ test are compared against each other, the data result in two tables, one illustrating the dA′ results by speaker group combination, and the other the A′ results by individual speaker groups. The latter provides a baseline for signal strength analyses of the former. The output of the dA′ tables is provided in the five traditional metrics for A′: H, FA, %C, A′, and B''_D. All A′ tests were run with extreme rates correction, where rates of zero are calculated at (0 ÷ N Signal or Noise Trials) to correct for errors produced in the A′ calculation when dividing by zero; this correction measure is used in most computational models to deal with extreme score errors (see Stanislaw & Todorov, 1999, pp.143-144 for more details).

4.3.4 Identification Task Results

Table 4.3 demonstrates the dA′ results for the ID Task. These results represent the maximal group composite A′ score and compares them against each other to identify the strongest signal for listeners. The Hit Rates (H), False Alarm Rates (FA), Percent Correct Rates (%C), A-Prime Scores (A′), and Non-Parametric Bias Scores (B''_D) are provided.

Table 4.3

Maximal Multiple Speaker Group dA′ Results for the ID Task

Differential Target Comparison	**H**	**FA**	**%C**	**A′**	**B''_D**
NY SCPR SCMX	0.69	0.18	0.77	0.83	0.22
SCPR SCMX	0.72	0.27	0.73	0.80	-0.07
NY SCMX	0.68	0.28	0.71	0.77	-0.03
NY SCPR	0.66	0.29	0.70	0.76	-0.02

The maximal signal maximizes the resulting A′ score, while minimizing the FA. The composite speaker group that meets these criteria is NY+SCPR+SCMX (FA=0.18; A′=0.83), while subset combinations of this group follow in the next three highest scores. There is also minimal bias in listeners' responses, ranging from a minimum -0.02 liberal bias to a maximum 0.22 conservative bias, with minimal variance in the %C scores. The results from Table 4.3 indicate that, when listeners were asked to identify Puerto Rican ancestry in these speakers' voices, the listeners' perceptual judgments were guided by a combination of two factors: (1) enregisterment (NY) and (2) marked accent (SCPR & SCMX). In other words, listeners' perceptual judgments focused on the saliency of the speakers in each group. This analysis is further strengthened by the fact that only the combination of the speaker groups in the maximal

dA′ signal generate stronger A′ scores and lower FA rates than any individual speaker group's score. The individual speaker group scores are provided in Table 4.4 for consideration.

Table 4.4

Comparative A′ Results by Individual Speaker Group for the ID Task

Individual Target Comparison	**H**	**FA**	**%C**	**A′**	$\mathbf{B''_D}$
SCMX	0.73	0.34	0.67	0.76	-0.31
SCPR	0.70	0.35	0.66	0.76	-0.20
NY	0.62	0.36	0.64	0.67	-0.12
PRIE	0.35	0.40	0.56	0.44	0.38
CB	0.15	0.44	0.50	0.26	0.71
CH	0.15	0.44	0.50	0.26	0.67
EC	0.07	0.45	0.48	0.20	0.86

Each composite speaker group in Table 4.3 exceeds the highest individual group A′ scores in Table 4.4, while also exhibiting lower FAs, which is indicative of a stronger signal for listeners. The scores in Table 4.4 also reveal that marked usage was the primary factor that guided listeners' judgments, where the most Spanish-dominant groups, SCMX and SCPR, were given the highest individual A′ scores. Upon being asked to identify Puerto Rican speakers, listeners focused on identifying traces of Spanish in each speaker, ultimately concentrating on the common association that Puerto Rico is a primarily Spanish-speaking region. This outcome is expected in naïve speakers, who do not have a shared language history and background with Spanish speakers from which to draw in making their perceptual judgments for this task.

With these observations in mind, the individual group A′ score breakdown in Table 4.4, in combination from the composite dA′ results in Table 4.3, suggests a categorical result, with three perceptual thresholds: maximal/top (SCMX, SCPR, NY) tier, an ambivalent middle tier (PRIE), and a lowest/bottom (CH, CH, and EC) tier. The speaker groups in the maximal tier were highly associated to Puerto Rico, suggesting a stronger degree of associations to a marked

accent, while the speaker groups in the lowest tier were the least associated at relatively uniform response rates (A′=0.26, 0.26, and 0.20, respectively, with nearly identical FAs).

PRIE speakers, however, occupied a middle tier, where their results were distinctly ambivalent in comparison to the other two speaker group clusters. This categorical differentiation is important in that it shows that listeners perceived PRIE distinctively from the rest of the dialect groups. These results are discussed in context with the results from the AX Task, introduced below.

4.4 The AX Discrimination Task

The AX Discrimination Task (AX Task), also known as a Same-Different Task, aimed to answer the second research question, to observe whether listeners perceived differences between the simultaneous bilingual speaker groups (PRIE, CH, CB, and NY). The results from this task aid in observing how listeners group the members of each of the target bilingual speech groups with respect to each other. This comparison is important to understand whether these bilingual speaker groups contrast from each other, despite their shared language background history. This task was provided to the same group of listeners as the ID Task, and all listeners completed this task first. As such, the listeners for Phase 3 were not made aware that they were comparing different bilingual groups until they reached the ID Task.

4.4.1 Methodology

Before listeners were presented with the main task, they completed a practice task with four trials—two same and two different. The practice trials included four clips extracted from publicly available videos. Each clip for the practice task was 4-7 seconds in length and provided a framework for listeners to understand the activity. The four speakers in the practice clips were canonical White male speakers from Wisconsin and Georgia, two from each state. The dialects of these four speakers all had heavily marked regional variants, purposefully chosen to highlight phonetic differences and prime listeners for the main task (more on these speakers in the discussion of the Mental Map Task's pre- and post-tests in Chapter 5).

In the main task, listeners responded to the same one-sentence clips of speakers reading the excerpt from "The North Wind and the Sun." Listeners were presented with a pair of speaker clips for each trial. In each trial, the listeners classified as "same" or "different" based on the following prompt question:

- *Prompt Question: Do you think that these two speakers have the same accent?*

These perception judgments are useful in comparing perceptions of different bilingual accents by similar Spanish-English speaker groups. Therefore, unlike the ID Task, only the simultaneous bilingual speaker groups were included in the AX Task, which resulted in a total of 12 speaker clips (see Appendix I for a sample item of the task).

With the aim of establishing how perceptions of the PRIE speakers pattern in comparison to those of the other speaker groups, the target group is PRIE, while the other groups for comparison are the CB and CH speakers (both native English varieties of other Spanish-speaking and/or enregistered communities). The NY group serves as a foil, as it is both a mainland variety of English with shared Puerto Rican ancestry and the most marked in comparison to the other

bilingual groups in this task. The objective is to have listeners compare the simultaneous bilingual speakers and observe whether listeners perceive PRIE speakers to be closer to one of the other bilingual groups, or whether they pattern as a standalone dialect group, which would reinforce the notion of its current development as an incipient dialect of American English. Combinations between PRIE and NY can vary depending on whether listeners can identify or associate PRIE and Nuyorican speakers to the same community. Following McGuire (2010), this task is designed with roughly an equal number of expected same and different trials, with comparisons between CB & CH as foils and PRIE & NY as unknown outcomes; these combinations are illustrated in Table 4.5:

Table 4.5

Expected Combination of Responses by Speaker Groups for the AX Task

	PRIE	CB	CH	NY
PRIE	Y	N	N	Y/N
CB	N	Y	N	N
CH	N	N	Y	N
NY	Y/N	N	N	Y

With a total of 12 individual speaker clips and 66 one-way combinations, three item blocks were created for this task. Each listener only completed one of the three blocks that were randomly presented using Qualtrics' survey distribution randomization algorithm. Listeners were presented with three trials on each survey page. The order of each trial and the trial pages were also randomized using the respective randomization algorithms in Qualtrics to control for any block-internal item order effects. Each block was comprised of a total of 30 items, which included all of the 1:1 same pair trials in Table 4.6 (12) and an assorted 18 of the different pair trials in Table 4.7; each letter from A-D stands for one of the speaker groups (the selection of each set of trials per block in Table 4.7 is color-coded as green, yellow, or gray).

Table 4.6

Same Trials Pair Combinations for the AX Task (N=12)

A to A	B to B	C to C	D to D
A1A2	B1B2	C1C2	D1D2
A2A3	B2B3	C2C3	D2D3
A3A1	B3B1	C3C1	D3D1

Table 4.7

Different Trials Pair Combinations for the AX Task, Color Coded by Block (N=54)

A to B	A to C	A to D	B to C	B to D	C to D
A1B1	A1C1	A1D1	B1C1	B1D1	C1D1
A2B1	A2C1	A2D1	B2C1	B2D1	C2D1
A3B1	A3C1	A3D1	B3C1	B3D1	C3D1
A1B2	A1C2	A1D2	B1C2	B1D2	C1D2
A2B2	A2C2	A2D2	B2C2	B2D2	C2D2
A3B2	A3C2	A3D2	B3C2	B3D2	C3D2
A1B3	A1C3	A1D3	B1C3	B1D3	C1D3
A2B3	A2C3	A2D3	B2C3	B2D3	C2D3
A3B3	A3C3	A3D3	B3C3	B3D3	C3D3

The set of same trials is repeated in all three blocks because of the inherent difference that comes with pairing four groups against one another in trials of two stimuli per group. To compensate, the different trials in Table 4.7 function as distractors and are shuffled throughout the three blocks. The combination of the set of 12 same trials and one of the sets of 18 different trials results in the final block design provided in Table 4.8.

Table 4.8

Block Design for the AX Task

Block A		Block B		Block C	
Item Code	Trial	Item Code	Trial	Item Code	Trial
same_01	A1A2	same_01	A1A2	same_01	A1A2
same_02	A2A3	same_02	A2A3	same_02	A2A3
same_03	A3A1	same_03	A3A1	same_03	A3A1
same_04	B1B2	same_04	B1B2	same_04	B1B2
same_05	B2B3	same_05	B2B3	same_05	B2B3
same_06	B3B1	same_06	B3B1	same_06	B3B1
same_07	C1C2	same_07	C1C2	same_07	C1C2
same_08	C2C3	same_08	C2C3	same_08	C2C3
same_09	C3C1	same_09	C3C1	same_09	C3C1
same_10	D1D2	same_10	D1D2	same_10	D1D2
same_11	D2D3	same_11	D2D3	same_11	D2D3
same_12	D3D1	same_12	D3D1	same_12	D3D1
diff_01	A1B1	diff_04	A1B2	diff_07	A1B3
diff_02	A2B1	diff_05	A2B2	diff_08	A2B3
diff_03	A3B1	diff_06	A3B2	diff_09	A3B3
diff_13	A1C2	diff_16	A1C3	diff_10	A1C1
diff_14	A2C2	diff_17	A2C3	diff_11	A2C1
diff_15	A3C2	diff_18	A3C3	diff_12	A3C1
diff_25	A1D3	diff_19	A1D1	diff_22	A1D2
diff_26	A2D3	diff_20	A2D1	diff_23	A2D2
diff_27	A3D3	diff_21	A3D1	diff_24	A3D2
diff_28	B1C1	diff_31	B1C2	diff_34	B1C3
diff_29	B2C1	diff_32	B2C2	diff_35	B2C3
diff_30	B3C1	diff_33	B3C2	diff_36	B3C3
diff_40	B1D2	diff_43	B1D3	diff_37	B1D1
diff_41	B2D2	diff_44	B2D3	diff_38	B2D1
diff_42	B3D2	diff_45	B3D3	diff_39	B3D1
diff_52	C1D3	diff_46	C1D1	diff_49	C1D2
diff_53	C2D3	diff_47	C2D1	diff_50	C2D2
diff_54	C3D3	diff_48	C3D1	diff_51	C3D2

4.4.2 Listeners

The listeners who responded to this task are the same ones who responded to the ID task, since both tasks were included in the same Qualtrics online survey. To reiterate, an initial 130 listeners responded to the Phase 3 Qualtrics survey. After the secondary screening process, a total of 91 responses were recorded for this task—34 in Block A, 27 in Block B, and 30 in Block C. 72 (79%) of the remaining 91 listeners were from the Upper Midwest or the Northeast, and the remaining 19 (21%) listeners were from other regions in the mainland U.S. The slight disproportions in the final number of listeners per block are due to an uneven number of filtered speakers in Blocks B and C once the screening criteria were applied.

4.4.3 Data Analysis

Different kinds of grouping can provide varying insights on perceptions of PRIE relative to the other selected speech community groups. The potential combinations of results of interest for this task are summarized in Table 4.9 below:

Table 4.9

Expected Potential Same-Different Group Combination Result Interpretations

Combination	Interpretation(s)
PRIE ≠ All Else	PRIE speakers only pattern together; possible perception of PRIE as a separate and enregistered dialect.
PRIE + CH + CB ≠ NY	PRIE speakers grouped more with most mainland varieties, but Nuyoricans pattern more with the Spanish-influenced speakers; PRIE is less marked in context and shows a non-enregistered status, similar to the remaining mainland bilingual American Englishes
PRIE + NY ≠ All Else	PRIE speakers associated with Nuyoricans; underlying perceptions of an L1 English Puerto Rican accent.
PRIE + CB + NY ≠ CH	Speakers with a Caribbean Spanish substrate; potential subliminal perceptions of a regionally constrained Spanish influence.
All Same/Different	No recognition OR all Spanish-influenced speakers grouped together; the results for PRIE are inconclusive.

The analysis of the results in the AX Task is conducted comparatively through A′ to observe how listeners' perceptions of PRIE pattern in relation to those of the other simultaneous Spanish-English bilingual speaker groups. Out of the possible combinations outlined above, there are two possibilities that are considered most probable, given the hypotheses and expectations detailed in Chapters 2 and 3: (1) PRIE ≠ All Else, wherein PRIE is perceived as a variety that is unlike any of the others to which it is being compared, in which case there is a stronger argument to be made in favor of its status as an emerging dialect; and (2) PRIE + CH + CB ≠ All Else, wherein the most enregistered variety of the four in this task (NY) is more prominent and patterns away from the remaining bilingual varieties, in which case PRIE is either a more incipient variety than initially hypothesized, or it is more similar to the remaining bilingual American Englishes. In this latter interpretation, it is also possible that only one of NY and CH appears to be the most prominent. Ultimately, the results are dependent on listeners' interpretations of the overt and covert features of these Englishes in relation to one another.

A-Prime. Following the procedures detailed in Pallier (2002), a non-parametric A′ signal detection measure was conducted, as described in §4.1. All A′ measures were conducted with extreme rates correction, as detailed in Stanislaw & Todorov (1999, pp.143-144). Additional corrections for extreme B''_D biases were also conducted by filtering out all listeners who had a B''_D score that was <-0.75 or more than >0.75. However, after applying that correction, the results remained unchanged down to the hundredths.

Overall A′ results are conducted on comparisons with PRIE across all groups—focusing generally on listeners' performance in detecting PRIE versus the other three simultaneous Spanish-English bilingual groups and specifically on listeners' performance in distinguishing PRIE from each group individually. In other words, the A′ measures are calculated by comparing

PRIE as the target and defining the speakers from every other group as distractors and by comparing PRIE as the target and defining the speakers from each group as distractors one at a time. This A′ analysis provides insight on the strength of PRIE as a signal in different comparative contexts, which allows us to extrapolate results for comparison to the interpretations and expectations provided in Table 4.9.

4.4.4 AX Task Results

The A′ results for the AX Task are explored as follows: (1) by observing the results of the PRIE group when compared to all of the other simultaneous Spanish-English bilingual speaker groups (CB, CH, and NY) as distractors, (2) by further examining those results with one-on-one comparisons between PRIE and each of the subgroups, and (3) by contrasting the results in (1) and (2) with the overall performance of each subgroup as individual targets. These results are provided in Tables 4.10, 4.11, and 4.12, respectively; the A′ results are formatted in the same way as in the discussion for the ID Task above.

Table 4.10 shows the general A′ results with PRIE defined as the target stimuli for the Same-Different AX Task analysis. In this task, listeners were asked to listen to two clips at a time and determine whether they believed that the speakers were members of the same or a different dialect group, without additional context.

Table 4.10

General A′ Results for the AX Task with PRIE as Target Stimuli

Target Speaker Group	**H**	**FA**	**%C**	**A′**	$\mathbf{B''_D}$
PRIE	0.54	0.43	0.56	0.58	0.04

The general results in Table 4.10 are slightly above chance for the H rate, %C, and the A′ scores—alongside a relatively high FA rate (0.43). These general results suggest that listeners could not reliably differentiate between the members of each speaker group. This observation is

a significant finding in that the PRIE speakers were not perceived as notably different from the other simultaneous bilinguals from the mainland (CB, CH, and NY). This initial outcome supports the status of PRIE as an incipient L1 variety of American English. However, the results from Table 4.10 warrant further inspection.

Table 4.11 contextualizes the general findings in Table 4.10 by comparing the performance of PRIE to each individual simultaneous Spanish-English bilingual group.

Table 4.11

Subset A′ Results for the AX Task with PRIE as Target Stimuli in Comparison to Individual Subgroups

Subset Comparison	**H**	**FA**	**%C**	**A′**	$\mathbf{B''_D}$
PRIE to CB	0.52	0.42	0.55	0.59	0.09
PRIE to CH	0.52	0.45	0.54	0.55	0.04
PRIE to NY	0.52	0.37	0.58	0.63	0.18

Note: H remains the same in all instances because the target stimuli, PRIE, is the same one in all comparisons.

The subset comparison in Table 4.11 reveals slight differences in the results between PRIE and the other subgroups, with similar FA and %C Scores. Nonetheless, although the differences are less pronounced, a similar pattern to that of the results of the ID Task emerges from these results. Listeners' perceptual judgments of CB and CH (FA=0.42 and 0.45, and A′=0.59 and 0.55, respectively) closely follow each other, with PRIE patterning closer with them than with the NY speaker group, where there is some divergence (FA=0.37 and A′=0.63). Table 4.12 explores these findings further by considering the relative performance of the speaker groups other than PRIE when they are observed as the target stimulus.

Table 4.12

A′ Results for the AX Task with Other Subgroup as Target Stimuli

Target Speaker Group	H	FA	%C	A′	B''_D
CB	0.64	0.44	0.59	0.65	-0.19
CH	0.66	0.43	0.60	0.68	-0.17
NY	0.45	0.21	0.66	0.63	0.49

These results follow the patterns outlined above, with CB and CH illustrating similar results, while NY notably diverges from them. While the A′ scores appear similar on the surface, NY shows much lower H and FA scores than CB or CH, which indicates that, although listeners were less accurate in identifying the NY speakers, they were also much less likely to misidentify members of other groups as NY speakers.

The findings from the AX Task reveal that, when asked to compare the simultaneous Spanish-English bilinguals from the mainland (CB, CH, and NY) with PRIE, listeners drew starker contrasts between NY and the other three groups, instead of PRIE. When comparing each speaker group individually, listeners same-different perceptual judgments of PRIE pattern slightly closer to CB and CH than with NY, which is the more relatively marked variety. Contrastively, there were fewer contrasts when analyzing listeners' same-different judgments of PRIE with the other three bilingual speaker groups from the mainland, confirming the status of PRIE as an incipient L1 variety of American English.

4.5 Discussion

The previous sections explored the results of the signal detection tasks using A′ as the statistical measurement tool, with promising results on the status of PRIE as an emerging dialect of American English. The ID Task aimed to investigate whether listeners would identify the Puerto Rican speakers as a separate group, and whether the PRIE speakers would pattern with the speaker groups included in this study or as an individual speaker group, potentially indicating a process of enregisterment if the latter were the case. With the aim of establishing how perceptions of PRIE pattern in comparison to the other speaker groups, the results from the AX Task aided in observing how listeners group the members of each of the target bilingual speech groups with respect to each other. Both tasks were provided to the same set of naïve listeners, who were mostly from the Upper Midwest and Northeast regions.

A prevailing theme throughout the signal detection tasks was that listeners' perception judgments of PRIE patterned closer to those of CH and CB, while NY acted more as a control, following the sequential speakers in the ID Task and diverging from CH, CB, and PRIE in the AX Task, where only the simultaneous bilinguals were included as speakers. As observed in Phase 2 (Table 3.6 of Chapter 3), although the language profiles of these simultaneous bilingual groups are balanced, the results of NY suggested that listeners associate these speakers more with the Spanish control groups, while CH and CB were more associated with the English control group. While it is the case that PRIE was more closely associated with CH and CB in the AX Task, it occupied a distinctive space in the ID Task. To contextualize this discussion with the initial expectations for the results in the ID Task, consider Table 4.2 from earlier on in this chapter, provided again below.

Table 4.2

Analysis by Grouping Types in Response to the Yes-No Question in the ID Task

Grouping Type	Speaker Group(s) Identified as Puerto Rican	Interpretation
PRIE Only	PRIE	PRIE Awareness
Sequential PR Only	SCPR	Awareness of L2 English PR PRIE speakers undetected
Nuyoricans Only	NY	Association of enregistered to island PRIE speakers undetected
Enregistered L1 Community	PRIE + NY	Recognition of PR accent substrate
Enregistered + Accented	NY + SCPR	PR accent and ancestry connection PRIE speakers undetected
All Puerto Rican Speakers	PRIE + NY + SCPR	Perfect discrimination of PR speakers
Only Sequential Speakers	SCPR + SCMX	Spanish accent association to PR PRIE speakers undetected
Low Familiarity	CH + Any	Low familiarity of these dialects
Hispanic Speakers	All - EC	Hispanic speakers in general = PR
No Differentiation (Y Bias)	All Groups	No differentiation between speakers
No Differentiation	Random/NONE	No differentiation between speakers

In the ID Task, PRIE did not pattern with either CB and CH, which were more associated with EC, or NY, which was most associated with SCPR and SCMX, the Spanish-dominant sequential bilingual groups. This categorical differentiation in the ID Task is important in that it shows that listeners perceived PRIE distinctively from the rest of the dialect groups (PRIE Only in Table 4.2). It is from this distinctive position that the argument for PRIE as an emerging dialect gains strength. The defining characteristic of PRIE as an L1 English variety in these results is in its disconnection from listeners' associations to the more Spanish-like dialect groups in the ID Task. This observation is further reinforced by the results in the AX Task, where, once the Spanish and English controls were removed and listeners judged PRIE in relation to the other bilingual varieties, PRIE was also ultimately associated with the groups that listeners judged to pattern more closely with the monolingual English, rather than the Spanish-dominant controls.

The ID Task was designed to examine whether listeners would identify PRIE as a distinct variety of American English in comparison to two control groups and several simultaneous Spanish-English bilingual mainland dialects. The findings show that listeners did, in fact, identify PRIE as distinct from the other speaker groups in the task. PRIE categorically occupied a middle point between the more recognized mainland varieties and the Spanish controls. Despite this differentiation, PRIE tended to pattern more towards the English-dominant (EC) side of the gradient of results, starkly different from SCPR and NY, the more perceptually salient varieties of Puerto Rican-affiliated varieties in the speaker group samples.

The AX Task exclusively investigated the perceptual relationship between PRIE and the other selected simultaneous bilingual dialect groups in this study, CB, CH, and NY. Considering Table 4.9, with the initial expectations for the AX Task from earlier in the chapter (provided again below for reference), the findings track the results from the ID Task, where listeners' perceptions of PRIE aligned closer to those of other mainland bilingual American Englishes.

Table 4.9

Expected Potential Same-Different Group Combination Result Interpretations

Combination	Interpretation(s)
PRIE ≠ All Else	PRIE speakers only pattern together; possible perception of PRIE as a separate and enregistered dialect.
PRIE + CH + CB ≠ NY	PRIE speakers grouped more with most mainland varieties, but Nuyoricans pattern more with the Spanish-influenced speakers; PRIE is less marked in context and shows a non-enregistered status, similar to the remaining mainland bilingual American Englishes
PRIE + NY ≠ All Else	PRIE speakers associated with Nuyoricans; underlying perceptions of an L1 English Puerto Rican accent.
PRIE + CB + NY ≠ CH	Speakers with a Caribbean Spanish substrate; potential subliminal perceptions of a regionally constrained Spanish influence.
All Same/Different	No recognition OR all Spanish-influenced speakers grouped together; the results for PRIE are inconclusive.

In the general speaker group comparison results, listeners did not identify notable differences across the speaker groups, confirming that PRIE holds a similar perceptual status as the other simultaneous bilingual varieties of mainland American English. Upon closer look, the results of the one-to-one speaker group comparisons revealed that listeners associated PRIE slightly more with the CB and CH speaker groups, while NY played a similar role as in the ID Task, acting as a diametrically opposed anchor. Following the expectations outlined in Table 4.9, the results from both signal detection tasks indicate that naïve listeners largely perceived PRIE as a variety of American English that is perceptually on par with the Miami Cuban and Chicano English varieties.

As initially predicted in §2.3 using Schneider's model of post-colonial Englishes (2003), the results in this chapter confirm that PRIE is at least in a parallel status with Phase IV in that enough local standards and norms have been informally established for PRIE to approximate mainland varieties of American English. However, with these findings, it is not particularly clear whether PRIE appears to be at least in the first stages of enregisterment (Phase V). The results in this chapter suggest that that it is not the case. One method to further investigate the potential enregisterment for PRIE is to observe the regional associations that listeners ascribe to PRIE speakers in comparison to the other selected dialects for this study. This perspective is explored in the Mental Map Task in Chapter 5.

CHAPTER 5

REGIONAL DIALECT PERCEPTIONS

The Mental Map Task aimed to answer the third and final research question of the study, to observe the dialect region(s) that listeners most associated to the PRIE speaker group in contrast to the other six speaker groups included in this study (CB, CH, NY, EC, SCPR, and SCMX). For reference, the third research question is provided below.

3. From a perceptual perspective, to what dialect region(s) of the United States are the selected American Spanish-English varieties associated vis-à-vis the Spanish and monolingual English control speakers?

The Mental Map Task provides insight on listeners' categorical regional associations for each of the speaker groups. These associations provide two notable data points: (1) the type of regiolect association for each dialect depending on currently known dialect distributions across North America, and (2) the relationship that these regiolect associations share across the different selected dialects for this study. The former allows us to observe connections between listeners' perceptions of PRIE and their preexisting knowledge of American English dialects, while the latter details the connections that listeners make between their perceptions of PRIE and those of the dialects that are better known and are, therefore, more recognizable, such as CH, NY, or EC.

It is reasonable to expect that listeners' accuracy in their regional judgments (see the discussion on the findings in Preston (1996) in §5.1.4) to carry over to their regional mapping of PRIE, which provides information on the features of PRIE to which listeners are paying the most attention and the associations that come from those observations. To that end, the Mental Map Task complements the results from the other perception tasks and provides information on listeners' understanding of the regiolect of the speaker groups based on their prior judgments.

The first section of this chapter presents the methodology for the Mental Map Task, including a detailed discussion on the pre- and post-tests, how dialect regions were defined, and the data analysis approach for the results. The second section presents the findings of the main task, which are then discussed in the final section.

5.1 Methodology

The Mental Map Task in Phase 4 was administered through an online Qualtrics survey, separate from the Phase 3 tasks (see Appendix J for a sample item from the task). The survey included three sections: a set of pre- and post-tests, the main task, and the same demographics questionnaire that was administered to every listener since Phase 2 (see Appendix A for the demographics questionnaire). This subsection first explains the general map selection and design progress, and then reports on the pre- and post-tests as well the main task methods.

5.1.1 Defining Dialect Regions

Beyond what we know of NEW DIALECT FORMATION and the political distinctions between what constitutes a language or a dialect, the distribution of dialects is understood to be fluid. Dialects are more generally distributed over continua rather than contained within discrete regions, often better observed as isogloss scatterplots than highlighted geopolitical maps. With this in mind, careful consideration must be taken when designing a map based on dialect regions, rather than continua, as there will always be gray areas, such as dialect transition zones, that are left out in comparison to more nuanced visualization approaches.

To that end, the multiple-choice regional maps designed for the Mental Map Task are largely based a combination of the dialect regions provided in the Atlas of North American English (ANAE) maps (Labov, Ash & Boberg, 2006, p.148). The selected map, in Figure 5.1 below, breaks down the United States into 18 dialect regions, some of which are macro-regions (e.g., the South, the West, etc.) that are representative of broader swaths of land and others that

are micro-regions (e.g., Texas South, St. Louis Corridor, etc.) that capture pockets of dialects that form part of those broader regions. Each of these dialect regions are representative of the isoglosses and the history that represents them, where areas that have experienced more sound change, towards the east, are more diversified than those areas in the western half of the country that experienced migration and colonial settlement more recently.

Figure 5.1: A map providing an overall view of the dialect regions of North America, as defined in the ANAE by isogloss (Labov, Ash & Boberg, 2006, p.148).

From the dialect regions in Figure 5.1, the most important for present purposes are the following seven: the West, the Midland, the South, Texas South, Florida, the North, NYC, and the combined East Northeast (ENE) & West Northeast (WNE). The region definition process for this task considers the historical connections of the southwestern crescent region of the United States, which would be comprised of a combination of the West and the Texas South in Figure

5.1. As overviewed in Chapter 3, the southwestern crescent of the United States, ranging from eastern Texas to northern or central California, has been seen as a Spanish-English language contact region, one that is considered the locus of Chicano English (Fought, 2006, pp.79-80) and other Spanish-influenced L1 varieties of American English. Consider Figures 5.2a and 5.2b below for demographic visualizations of this contact region by raw numbers and share of the population, respectively.

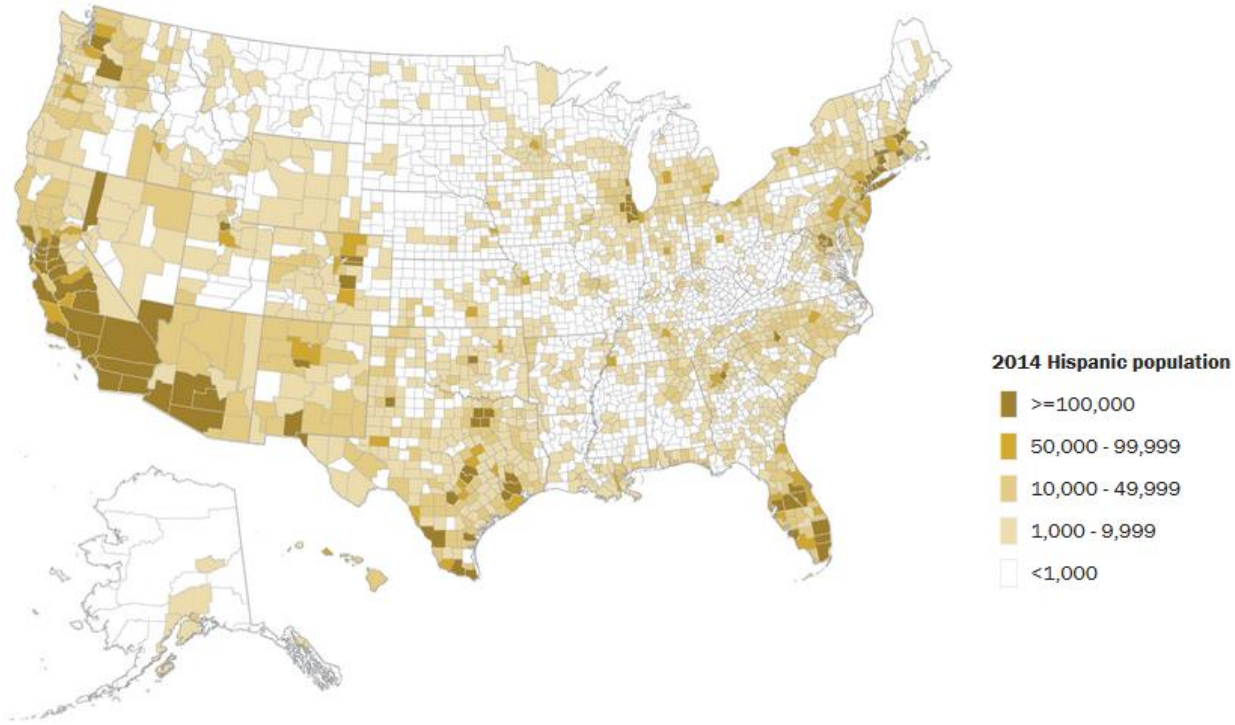


Figure 5.2a: A 2014 map of the southwestern crescent, a region of the United States known for its historical roots in Spanish and Spanish-influenced English varieties. Colors highlight the number of the population in each county that identified as Hispanic (Pew Research Center, 2016).

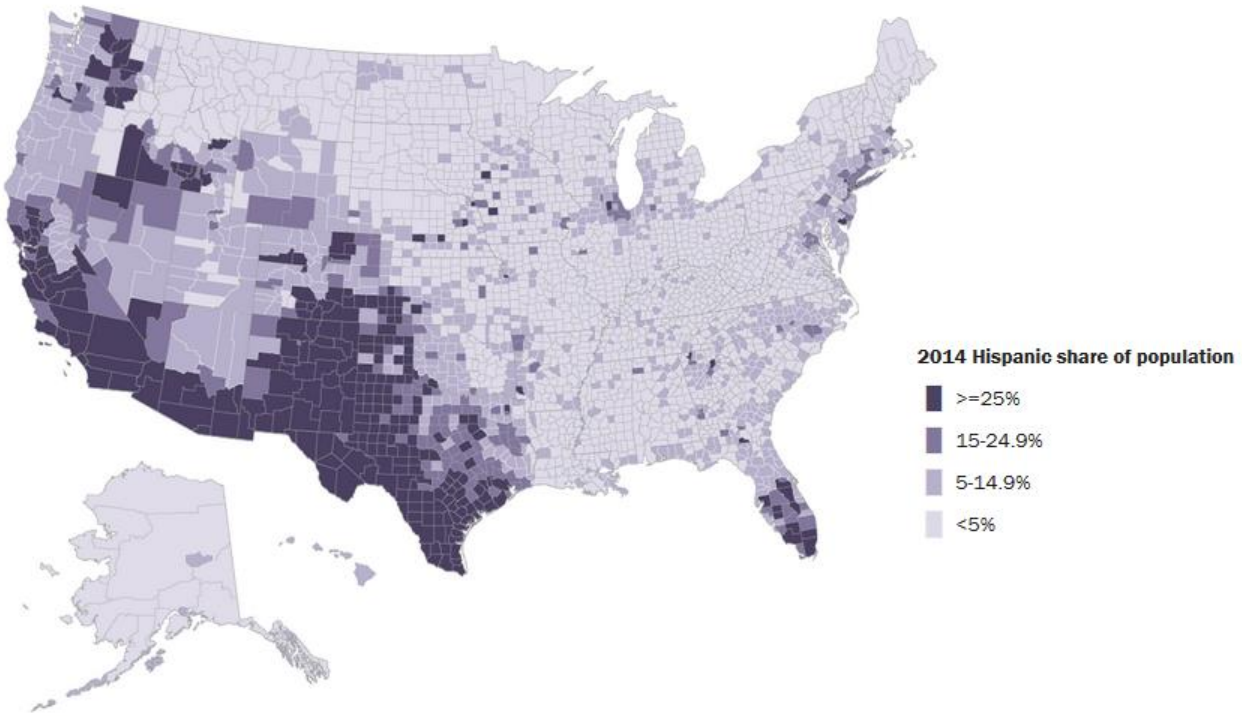


Figure 5.2b: A 2014 map of the southwestern crescent. Colors highlight the percentage share of the population in each county that identified as Hispanic (Pew Research Center, 2016).

Defining the southwestern crescent is important for the purposes of this study because results from the pilot study indicated that listeners identified that region for the speakers that they perceived to be more Spanish-dominant. This suggests that this region exists in listeners' mental representations of American English dialects as a more Spanish-influenced area. Should listeners perceive a speaker or group of speakers of a dialect to be more Spanish influenced, it is expected that those listeners generally identify them as speakers of a dialect from the southwestern crescent. Similar considerations can be given to enregistered varieties, such as Nuyoricans to NYC and Chicano English to the southwestern crescent as well.

A final consideration in designing the maps for this task is to maintain a regional parallel across the speaker groups. In other words, each of these Spanish-English bilingual groups should also include analogue regions from which their Spanish dialects originate—Cuba for CB, Mexico for CH, and Puerto Rico for SCPR and PRIE. The selection of dialect regions for the

multiple-choice main task must, then, consider this factor alongside the information from both Figures 5.1, 5.2a, and 5.2b. The resulting map for the main task is provided in Figure 5.3.

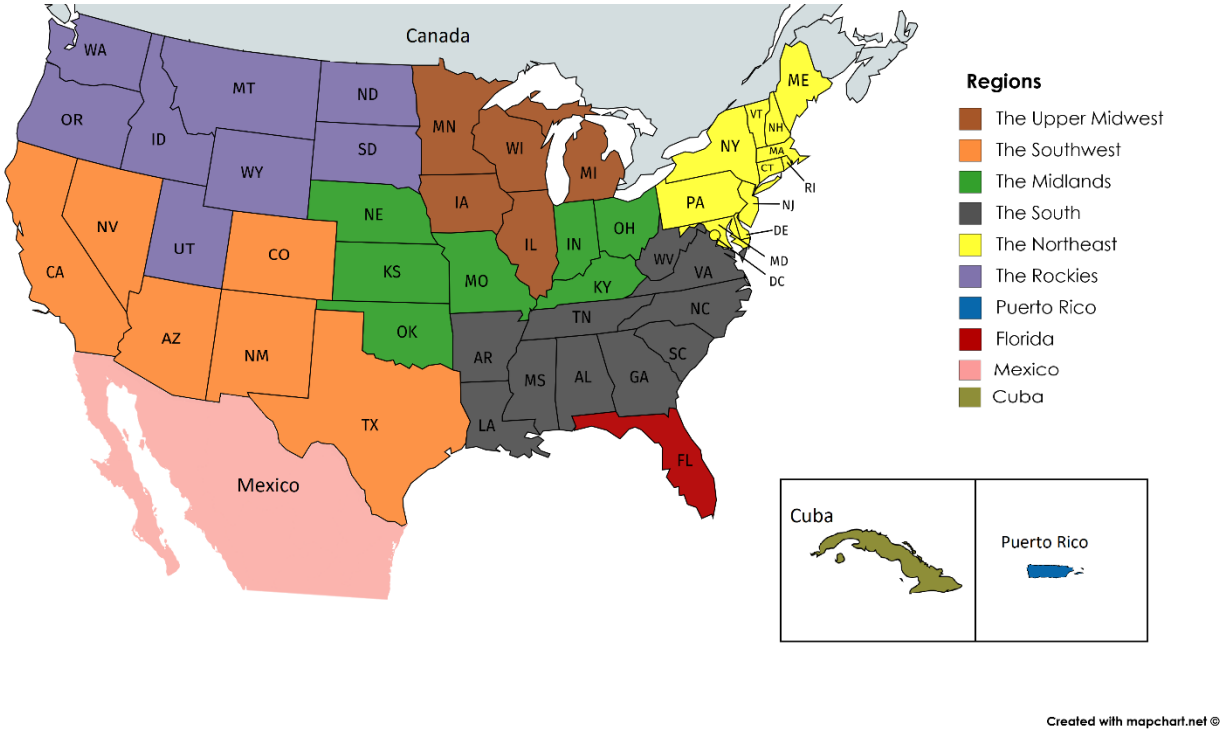


Figure 5.3: Map for the main Mental Map Task with a legend by each test region with the information that was shown to listeners.[15]

The map in Figure 5.3 largely follows the macro-regions from the ANAE in Figure 5.1. Because this map follows discrete political state boundaries, it does not exhibit the same level of detail in the dialect boundaries as the ANAE map. However, it does provide a reliable baseline for non-linguist listeners. Aside from boundary limitations, the biggest difference between Figures 5.1 and 5.3 is in the West; that region was divided between the southwestern crescent (defined as "The Southwest" in Figure 5.3) and the remaining northern half (defined as "The Rockies"[16] in Figure 5.3). This dialect boundary division is motivated by the discussion of that

[15] The base map was generated using Map Chart: https://mapchart.net.
[16] "The Rockies" here refers to the dialect boundary region as defined above, rather than a name based on the geographical location of the Rocky Mountains, which starts west of the Dakotas.

region in Figures 5.2a and 5.2b. In the initial instructions for the survey, listeners were informed that this map was to function as the reference for their multiple-choice selections.

5.1.2 Pre/Post-Test Design

Before listeners were presented with the main task, they completed a pre-test; this same pre-test was re-administered as the post-test after listeners completed the main task. In these pre- and post-tests, listeners were provided with six audio clips, the pre- and post-test map, and a drag-and-drop box area corresponding to six locations in the United States: The Upper Midwest, The South, The Southwest, Texas, New York City, and Boston. Listeners were given the following prompt:

- *<u>Instructions</u>: The goal of this task is for you to listen to the clips below and place the one speaker into each of the regional dialect groups, based on how you believe that they pair up. To do so, you can drag and drop a speaker into a dialect group box.*

  If you had to guess, where do you think that these speakers are from? There is only one speaker per dialect region, such that each of the six boxes should have only one of the six speakers by the end of the activity.

  Please feel free to use the map below as a reference for each of the regions. The provided regional shading serves as a guideline and does not include the entirety of each region.

See Figure 5.4, below, for an example of the drag-and-drop task.

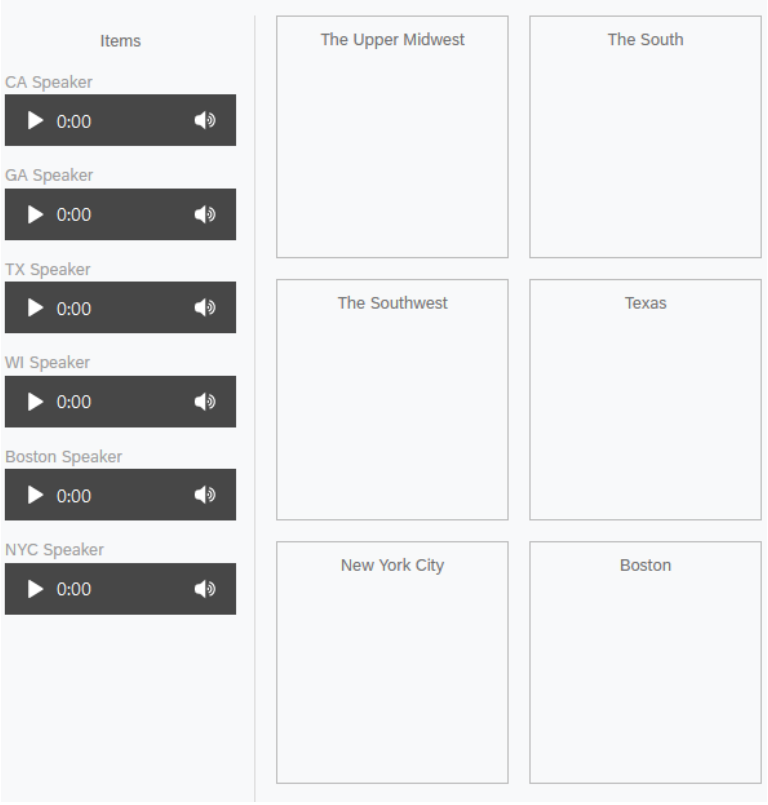


Figure 5.4: Drag-and-drop boxes with draggable audio clips for the pre- and post-tests.

Each clip was 4-7 seconds in length, and the audio was extracted from publicly available interview videos. The six speakers in these clips were canonical White males from Boston (Kenneth Wormald), Brooklyn/NYC (Steve Schirripa), California (Tyler Posey), Georgia (Travis Denning), Texas (Matthew McConaughey), and Wisconsin (Chris Bangle), corresponding to one of the six areas in the drag-and-drop boxes. Each speaker was representative of a dialect that has been historically associated to one of those areas, with heavily marked regional features, purposefully chosen to highlight phonetic differences and prime listeners for the main task. The order in which the clips were presented to listeners was randomized using Qualtrics' item randomization algorithm. Figure 5.5 illustrates the map used for the pre- and post-tests, which is largely based on the map for the main task (Figure 5.3), with a few exceptions.

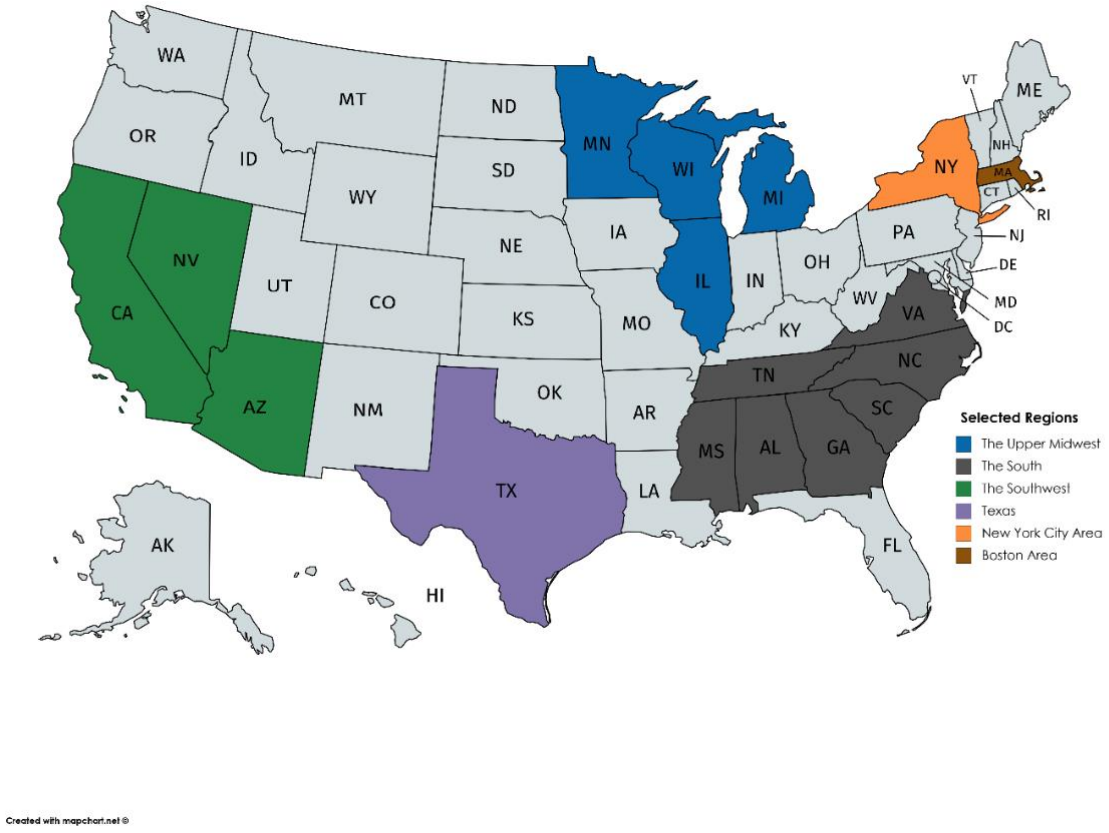


Figure 5.5: Pre- and post-test map for the Mental Map Task with a legend by each test region, as shown to listeners—the Southwest in green, Texas in purple, the Upper Midwest in blue, the South in black, New York City in orange, and Boston in brown.[17]

Since only one speaker from each of these regions is being tested, Cuba, Mexico, Puerto Rico, among others were not included in these tests. Furthermore, the scope of each area was reduced to specific locations, such as NYC and Boston over "the Northeast," to provide listeners with a clearer idea of the relationship between the chosen speaker and the dialect region to which they correspond. The grayed-out states also serve as buffer zones to underscore distinctions between each location on the map. The results from these pre- and post-tests are discussed at length and contextualized with the of the listeners who responded to this task further below.

[17] The base map was generated using Map Chart: https://mapchart.net/.

The pre- and post-tests fulfilled two purposes. The first was to have listeners practice and become familiar with the activity. The second was to use those results to better gauge each listener's knowledge of and sensitivity to regional dialects of American English. These results, combined with the demographic screening process, allows for the data to be observed by listeners' accuracy in identifying regional American English dialects.

5.1.3 Main Task Design

In the main task, listeners responded to the same one-sentence clips of speakers reading the excerpt from "The North Wind and the Sun." Listeners were presented with one clip, the map from Figure 5.3 (above), and a single select multiple-choice question for each of the 21 speakers from all the speaker groups (three of each from the PRIE, CB, CH, NY, EC, SCPR, and SCMX speaker groups). Listeners were given the following prompt question alongside the audio clip:

- *<u>Prompt Question</u>: If you had to guess based on the dialect, what region do you think that this speaker is from? Please feel free to use the map below as a reference.*

After listening to the audio clip of a speaker, and based on the regions on the map in Figure 5.3, listeners then assigned the speaker to one of the following ten regions: the Upper Midwest, the Southwest, the Midlands, the South, the Northeast, the Rockies, Puerto Rico, Florida, Mexico, and Cuba. Both the order in which each speaker page was presented and the order in which the regions were presented in the single select multiple-choice question were randomized using Qualtrics' respective randomization algorithms. After answering the Mental Map Task questions for all 21 speakers, listeners completed the post-test.

5.1.4 Listeners

An initial 59 listeners responded to the Mental Map Task. Of those initial responses, and as in other phases in this study, an additional screening was conducted with the following parameters: (1) observing the perceptual judgments of only self-identified native English speakers, (2) controlling for listeners who reported that they had not had sustained contact with a Spanish-speaking community, and (3) removing the only listener who identified as a member of a Spanish-speaking community. After the secondary screening process, a total of 43 responses were recorded. 26 (61%) of the remaining 43 naïve listeners were from the Upper Midwest, 10 (23%) from the Northeast, and the remaining 7 (16%) listeners were from other regions in the mainland U.S. These listeners also reported having lived in the United States throughout their childhoods.

The results from listeners' responses to the pre- and post-tests were used to determine three control factors: (1) listeners' sensitivity to dialect variation, (2) listeners' ability to identify those dialects accurately, and (3) response consistency. Figure 5.6 shows the hierarchy of regional dialects of present-day American English in relation to each other's shared features and relative distance. This hierarchy expands on the map in Figure 5.1 and forms the basis for the definition of the Macro Regions in Table 5.1, which sets the constraints for the data analysis of the results from the pre- and post-tests. Accuracy in the pre- and post-tests was defined semi-rigidly, where some degree of flexibility in the listeners' responses was afforded accordingly to dialect similarity and featural distance. Each node represents a common set of shared features anchored by two major phonemic phenomena in American Englishes, the fronting of back vowels and the low-back merger.

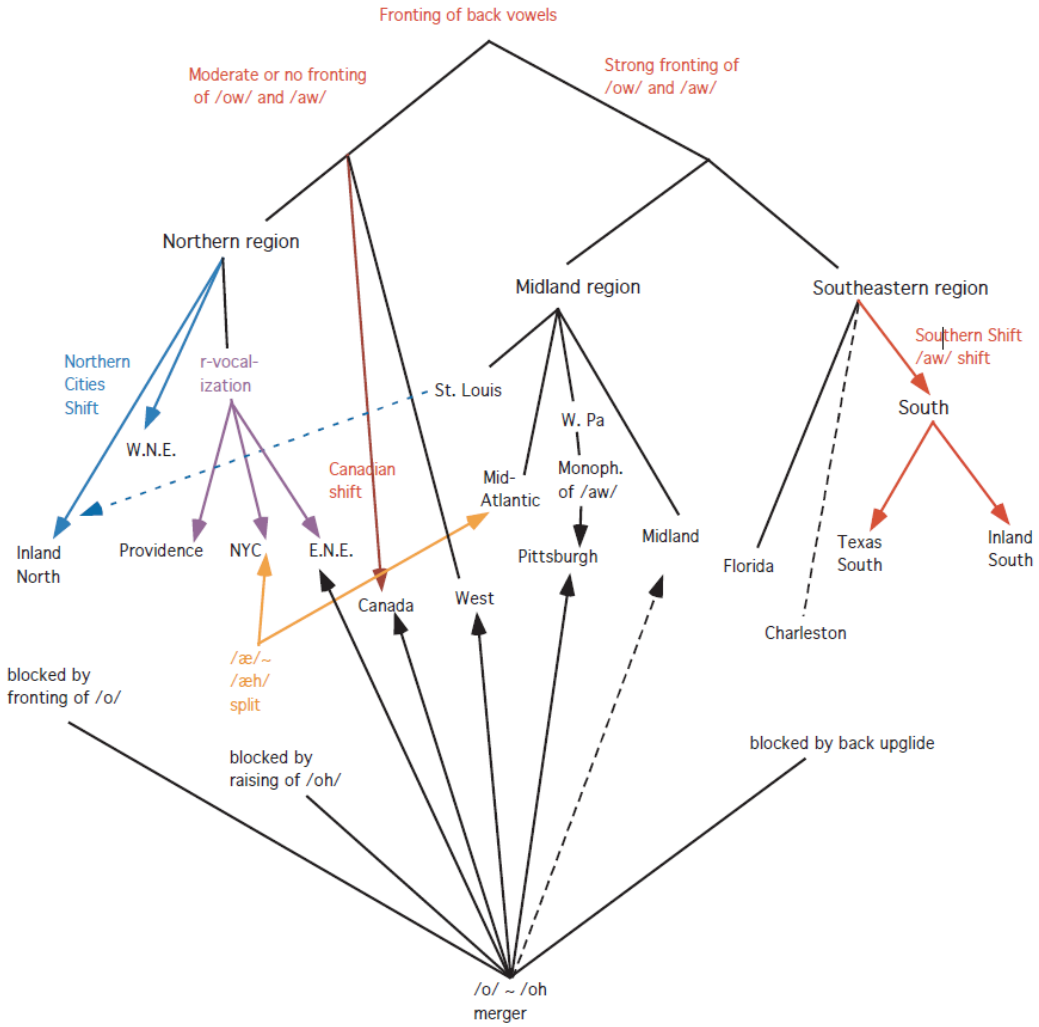


Figure 5.6: Hierarchical structure of North American dialects, as illustrated in Labov, Ash & Boberg (2006, p.147).

In combination with the information in Figure 5.6, this methodological decision is based on the results from Preston's north to south Mississippi River Mental Map Task (1996), which revealed that perceptual distinctiveness broke along the lines of North versus South, rather than state by state, revealing three macro dialects, the North, the Midlands, and the South. In other words, the degree of dialect distinctiveness a listener has depends on both their preexisting knowledge and their relative orientation to a speaker's dialect. For instance, listeners living in a southern state (e.g., Alabama) would be able to better distinguish between their own dialect and that of a speaker from Louisiana. Conversely, a speaker living in Wisconsin would be more

sensitive to dialect variation in Upper Midwestern dialects than a southern speaker would. This perceptual gap is a result of a listener's dialect knowledge framework, largely built upon their own experiences as both speakers and listeners.

This task accounts for that perceptual gap by allowing some degree of flexibility in defining accuracy by the relevant nodes in Figure 5.6. Thus, to identify whether a listener is accurate in detecting dialect variation, a listener does not have to strictly identify each canonical speaker to their state, but rather, at the very least, closely approximate the target region. The only exception to this flexibility in assessing listeners' accuracy is in identifying the Wisconsin speaker, as most of these listeners are either from the Upper Midwest or have experience with Upper Midwestern Englishes. Under this observation, listeners who are identified as MORE ACCURATE are expected to be able to strictly categorize the Wisconsin speaker correctly. Table 5.1 outlines the accuracy evaluation process for the results from the pre- and post-tests. The Macro Region column describes the general acceptable categorical target under the outlined semi-rigid guidelines. The Target Region column provides the categories provided to listeners in the pre- and post-tests, and the Expected Response Spread column includes the acceptable range of speaker selections for each category should a listener qualify for the MORE ACCURATE subset.

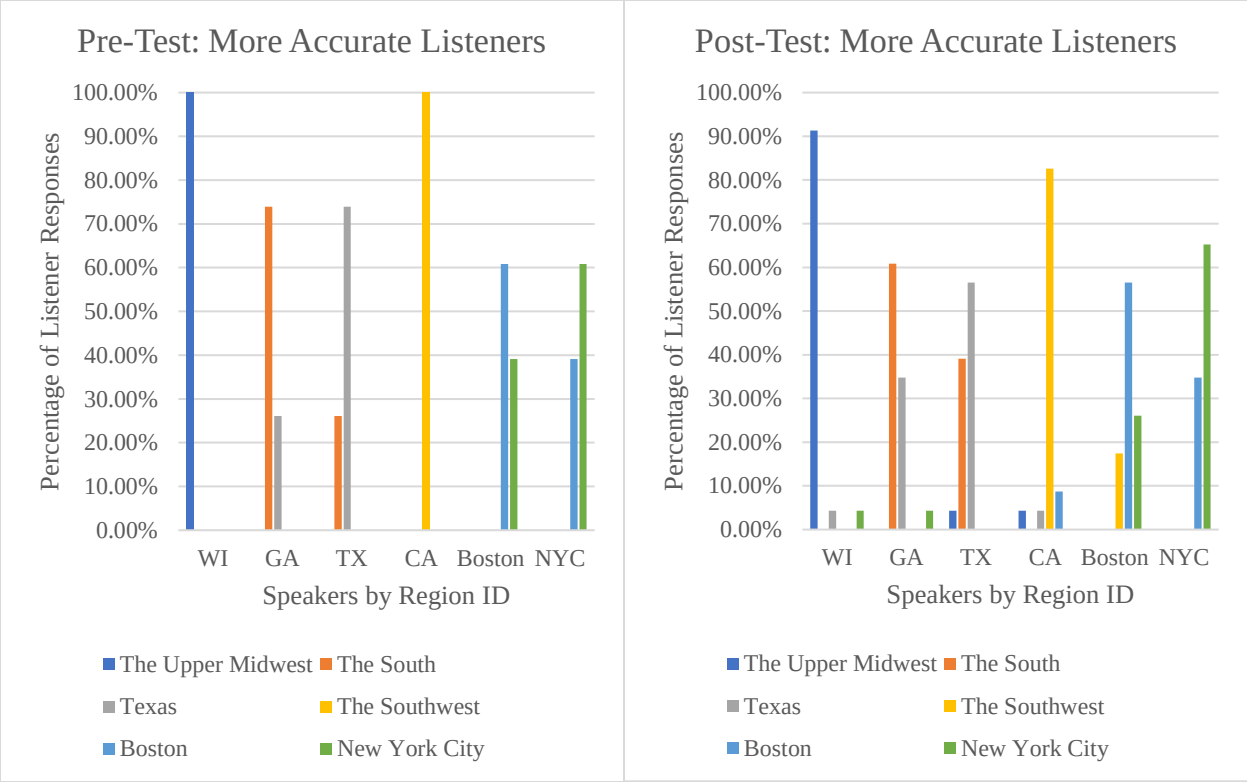
Table 5.1

Accuracy Constraints by Macro Region, Target Region in Pre/Post Tests, and Expected Responses

Macro Region	Target Region in Pre/Post-Test	Expected Response Spread
The Midwest	Upper Midwest	Only WI
The Greater South	South	GA, TX
	Texas	TX, GA
The Southern Crescent	Southwest	CA, TX
	Texas	TX
The Northeast	NYC	NYC, Boston
	Boston	Boston, NYC

Listeners who followed the expected response spread in the pre-test qualified as the more accurate listeners, while those who did not were classified as less accurate listeners for the purposes of analyzing their performances in the pre- and post-tests as well as the Mental Map Task. The same selection process was not applied to the post-test results in order to observe their consistency or potential variance in listeners' responses after they completed the Mental Map Task. Once the constraints from Table 5.1 were applied, the 43 listeners were divided into the two categories for analysis, MORE ACCURATE (N=23) and LESS ACCURATE (N=20).

Consider Figures 5.7-5.9 (more accurate listeners) and 5.10-5.12 (less accurate listeners), below. The first two figures of each set provide those listeners' pre- and post-test results, and the last figure of each set provides the different in their results between the pre- and post-tests. For the more accurate listener results, the pre-test results reveal their overall perception judgment spreads under the categories from Table 5.1, while the post-test results underscore variance in their responses to the same speakers. For the less accurate listener results, the pre-test results highlight the contrast in their initial performance versus that observed in the more accurate listeners, while the post-test results emphasize any changes in their perception judgments after the main task. The and pre- and post-test differential graphs quantify each subgroup's overall accuracy and consistency by comparing their results between the pre- and post-tests. The differential graphs highlight the variation in each listener group's performance by visualizing response rate discrepancies to the same stimuli after the main task.



Figures 5.7 and 5.8: Pre-test and post-test results (respectively) for the more accurate listeners by speaker to region allocation. Colors in the legend correspond to listeners' allocation responses.

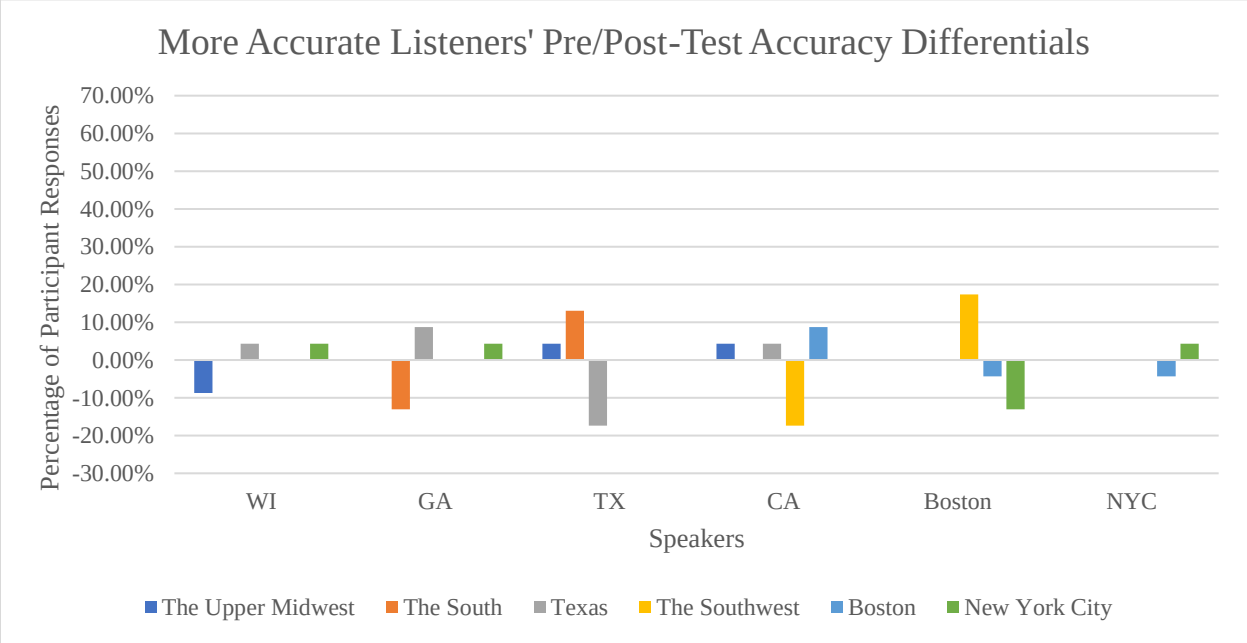
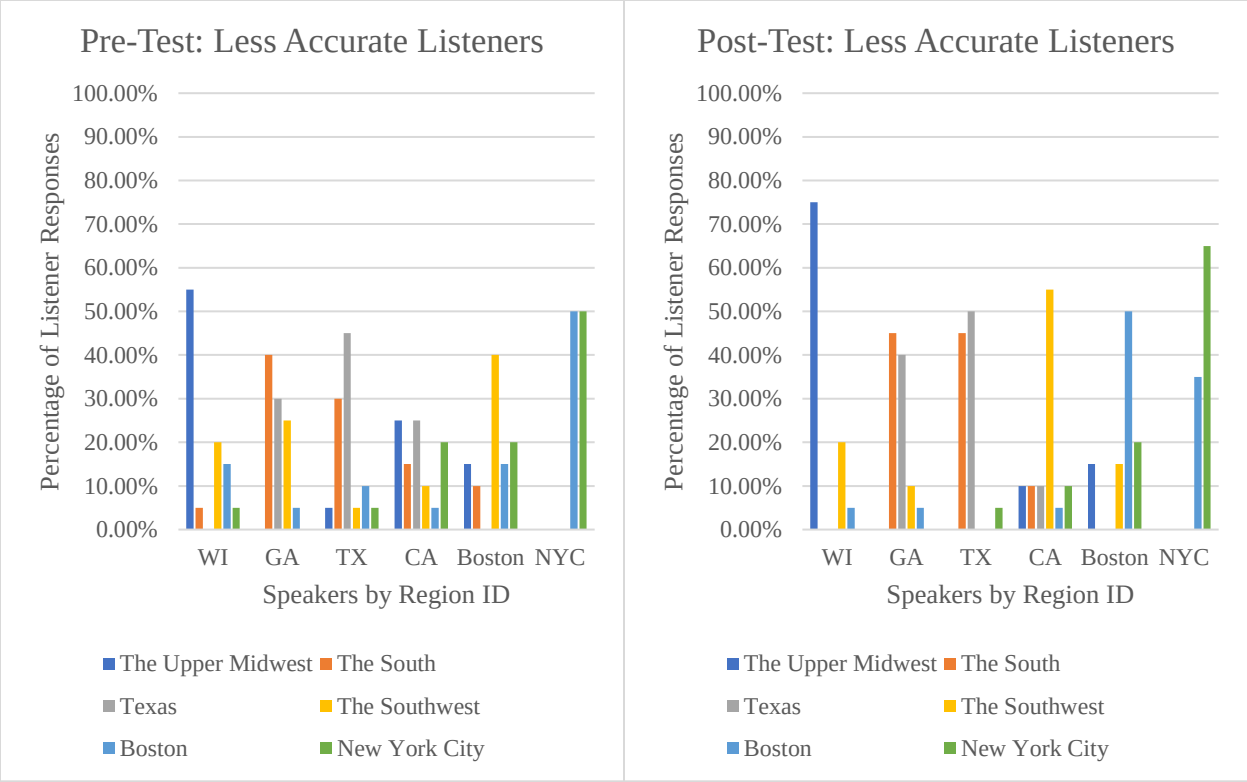


Figure 5.9: Difference between the pre-test and post-test results for the more accurate listeners.

Figures 5.7 and 5.8 illustrate the performance of the more accurate group of listeners in the pre- and post-tests for the Mental Map Task. Observing the pre-test results in Figure 5.7, the

more accurate listeners correctly placed the WI and CA test speakers with 100% accuracy. Furthermore, they correctly identified the GA and TX test speakers with over 70% accuracy, with the remaining percentage of the trials split between the two state categories. Finally, the more accurate listeners were least accurate in identifying the differences between the Boston and NYC test speakers, with a 60:40 split between the two. The post-test results in Figure 5.8 indicate some degree of variance in this group's second attempt at the task. However, the percentage split across the categories was similar, and the group was generally consistent in accurately placing a test speaker to their target region across the two tests.

Figure 5.9 shows the difference in response rates between the pre- and post-test results for the more accurate listeners to scale. While these listeners were nearly 100% accurate for WI and CA, their perceptual sensitivity for the remaining four targets was categorically less accurate for two of the macro regions identified earlier in Table 5.1, the Greater South (GA + TX) and the Northeast (Boston + NYC), in descending order of accuracy. Given that these regions are both quite distant from the Upper Midwest, the performance differential between these two categories could be attributed to the degree of difference in perceptual saliency. In other words, at a phonetic level, these listeners likely had more phonetic cues to distinguish between the two speakers from the states in each region to varying degrees of reliability. Nevertheless, the lowest differential was at 4% (Boston) and the highest two were tied at 17% (TX and CA)—representing a mean accuracy decline of 9%, where only responses for NY nominally improved. The overall categorical difference in their responses was minimal, despite this variance. Ultimately, using the evaluation criteria for these tests that was outlined earlier, the more accurate group of listeners was both notably more precise and consistent, particularly when compared to the less accurate listeners.



Figures 5.10 and 5.11: Pre-test and post-test results (respectively) for the less accurate listeners by speaker to region allocation. Colors in the legend correspond to listeners' allocation responses.

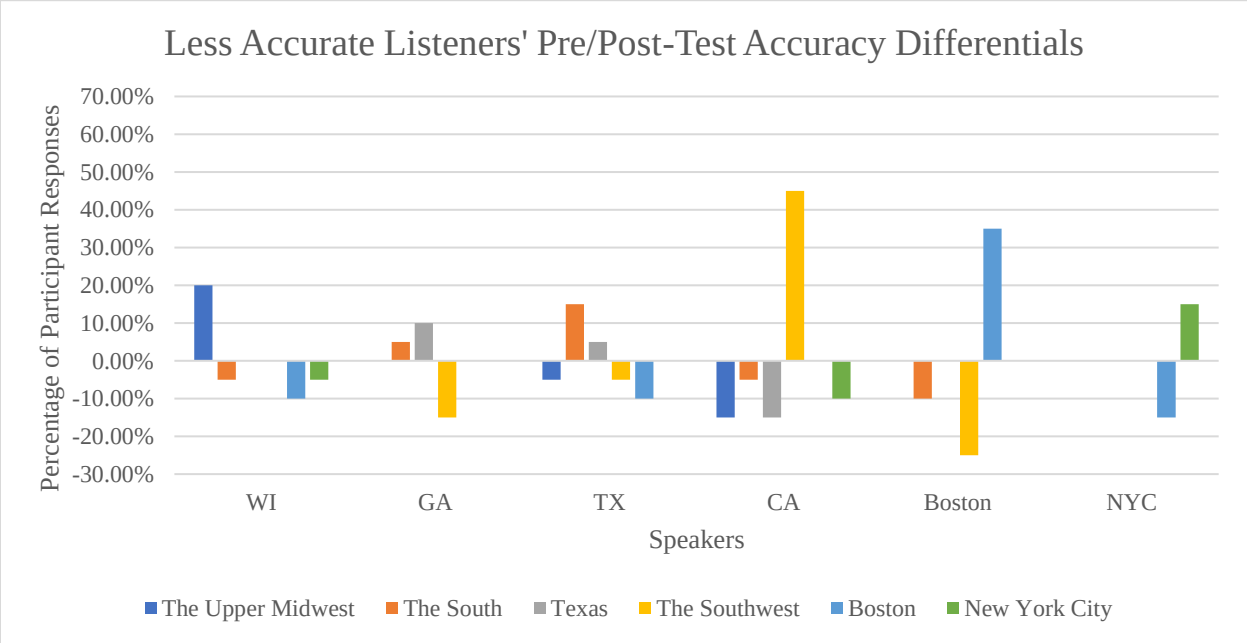


Figure 5.12: Difference between the pre-test and post-test results for the less accurate listeners.

Figures 5.10 and 5.11 illustrate the performance of the less accurate group of listeners in the pre- and post-tests for the Mental Map Task. The pre-test results for the less accurate listeners in Figure 5.10 are markedly different from those of the more accurate listeners. There is a considerably broad spread across all of the categories, such that nearly all of the test speakers were placed into all of the categories to some degree. Additionally, no speaker was accurately placed into a region in more than 50% of the cases, except for the WI speaker. In some cases, the speaker allocated the most into a region was incorrectly placed, such as the Boston speaker being mostly identified as a speaker of the Southwest.

However, observing the post-test results in Figure 5.11, the less accurate speakers demonstrated marked improvement in the accuracy of their perception judgments after completing the main Mental Map Task (Figure 5.12). Their post-test results for each category trace much closer to the overall performance of the more accurate speakers, even though there is still a broader degree of selection variance and, thus, less consistency. Figure 5.12 underscores the difference in response rates between the pre- and post-test results for the less accurate listeners to scale. In comparison to the more accurate listeners, the results for the less accurate listeners were both inconsistent and imprecise in general, even when there was notable improvement in the post-test results when compared to their pre-test performance, ranging between the lowest at 5% (TX and GA) and highest at 45% (CA)—representing a mean improvement of 21% from their pre-test results.

The improvement in the less accurate listeners' pre- and post-test results suggests that the less accurate listeners became more sensitive to differences in dialect features as they gained familiarity with the speakers and the task over the course of the survey. Although, to the best of my knowledge, there is no research on the methodological performance impact on less accurate

listeners in tasks analogous to the Mental Map Task, previous parallel studies have been conducted in foreign accent rating tasks, finding comparable patterns indicating that increased familiarity can lead to variable performance (McDermott, 1986; Munro, Derwing & Morton, 2006 in Schmid & Hopp, 2014, p.384). However, it appears that the methodological impact on the effects of familiarity and training on accuracy and performance has not been explored in foreign accent rating tasks either (Jesney, 2004, p.8). The potential of future work to explore this line of inquiry is detailed in Chapter 6.

Despite the less accurate listeners' improvement across the board, it is important to note that similar patterns emerged from the results from both groups of listeners: WI was the most accurately placed in all tests;[18] GA and TX as well as Boston and NYC patterned inversely, where listeners from both groups co-associated the test speakers from each respective region the most, while still largely preferring the target speaker for each region. The results from these pre- and post-tests provide the basis for comparative analyses in the main Mental Map Task.

5.1.5 Data Analysis

The results from the Mental Map Task are output in two methods: (1) as perception selection gradience tables, providing a purely quantitative overview of listeners' dialect region perception allocations, and (2) as gradient dialect region maps, which offer a clearer picture on listeners' regional dialect allocations for each speaker group. Both the quantitative tables and the speaker group-specific dialect region results maps are observed comparatively between the results from the less accurate and more accurate listener groups.

These differentiated results allow us to observe how listeners' varying sensitivity to dialect variation and the discrepancy in their abilities to identify those dialects accurately can

[18] The listeners in this part of the study, who were mostly from the Upper Midwest, were generally more accurate in identifying other Upper Midwesterners, which confirms one of the core findings in Schuld et al. (2017).

affect the ultimate outcome of their perceptual judgments of the same dialect stimuli. From a methodological standpoint, the resulting data provide deeper insights on the gaps between accuracy and performance in perceptual tasks. From the standpoint of the main Mental Map Task, the findings highlight the intrinsic associations that listeners make between PRIE and the provided dialect regions vis-à-vis the other speaker groups in this study.

5.2 Mental Map Task Results[19]

The results from the main Mental Map Task are presented by speaker group, with individual data points and figures for the less accurate and more accurate listeners. In Figures 5.14-5.27, the results from listeners' perceptual judgments by dialect regions are illustrated using a color gradient to represent regional response rates. For ease of reference, Figure 5.13 shows a larger version of the color key.

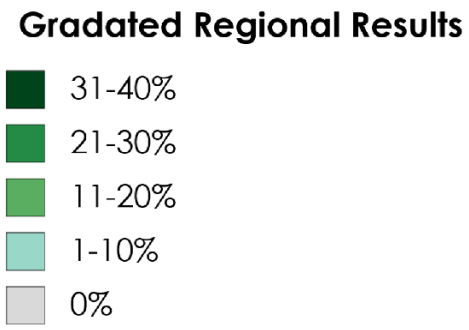


Figure 5.13: Enlarged color key for regional dialect results in Figures 5.14-5.27.

As a starting point, the results for the control groups, EC, SCPR, and SCMX, are discussed to establish a baseline for the results from the simultaneous Spanish-English bilingual groups, PRIE, CB, CH, and NY. In all cases, the results from the less accurate listeners for a speaker group are presented first, followed by results from the more accurate listeners.

[19] The base maps for this section were generated using Map Chart: https://mapchart.net.

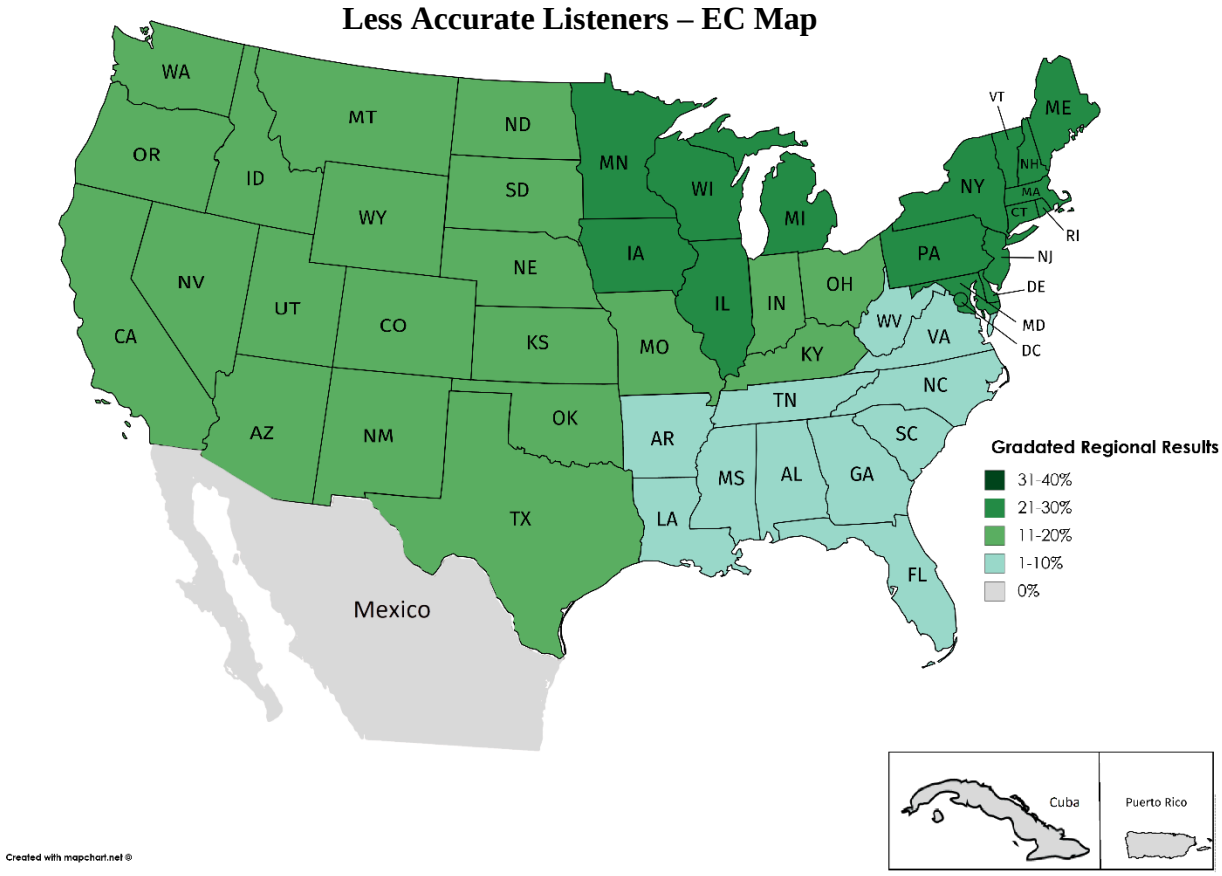


Figure 5.14: Less accurate listeners' composite regional map of the EC speaker group.

In Figure 5.14, less accurate listeners primarily associated the monolingual English control group with the Upper Midwest and the Northeast, with secondary associations to the Midlands, Southwest, and Rockies regions. There were minor tertiary associations to the South and Florida, and no associations to the primarily Spanish-speaking regions of Cuba, Puerto Rico, and Mexico. Although these perception judgments approximate the EC speakers who are from a same city in the Upper Midwest, no regional results exceeded 30%, suggesting lower overall precision for this listener group.

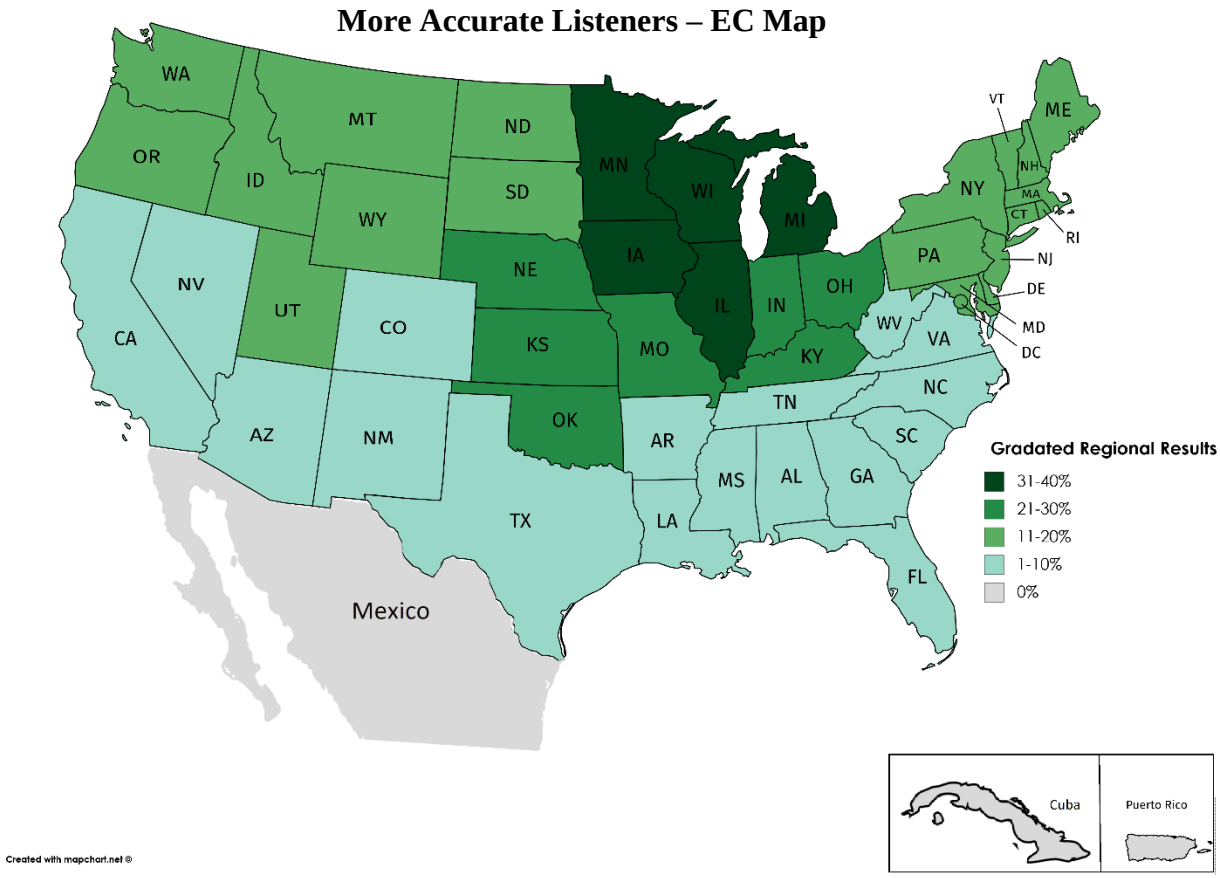


Figure 5.15: More accurate listeners' composite regional map of the EC speaker group.

In contrast, the more accurate listeners' results for EC in Figure 5.15, follow a similar pattern as the less accurate listeners in Figure 5.14, with the distinction that the more accurate listeners' results were more focused. More accurate listeners' responses followed a wave-like pattern, with the most responses in the expected region, the Upper Midwest, and a secondary association to the nearest region, the Midlands. Likewise, there were minimal associations to the South and the more Spanish-influenced regions of the United States, and no associations to the Spanish-dominant dialect regions. One pattern is evident from this initial contrast: the more distant the dialect region from the target region for EC, the fewer responses from this group. In this contrast, the more accurate listeners demonstrate higher accuracy and precision in their perceptual judgments of the EC speaker group.

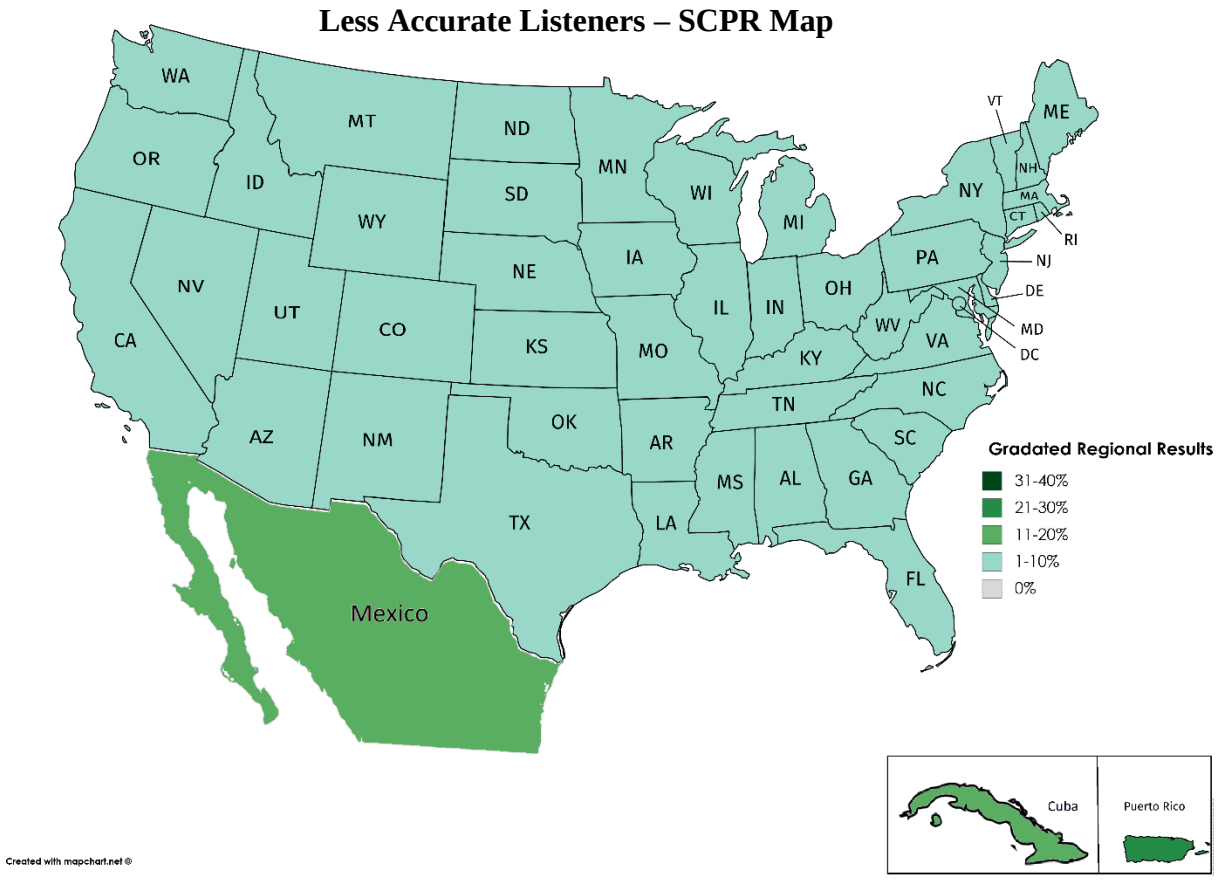


Figure 5.16: Less accurate listeners' composite regional map of the SCPR speaker group.

For the SCPR control group, less accurate listeners primarily associated these speakers to the Spanish-dominant regions in the map, Mexico, Cuba, and Puerto Rico, in Figure 5.16. These listeners produced a low response density throughout all of the remaining mainland dialect regions. Interestingly, Puerto Rico was the area that represented the most responses from this group (over 20%), with Cuba and Mexico functioning as secondary regions (at over 10%, but under 20% each). Despite their performance with EC, the less accurate group appropriately identified the target region for SCPR and followed a similar wave-like response pattern, with target dialect region distance functioning inversely over response rate.

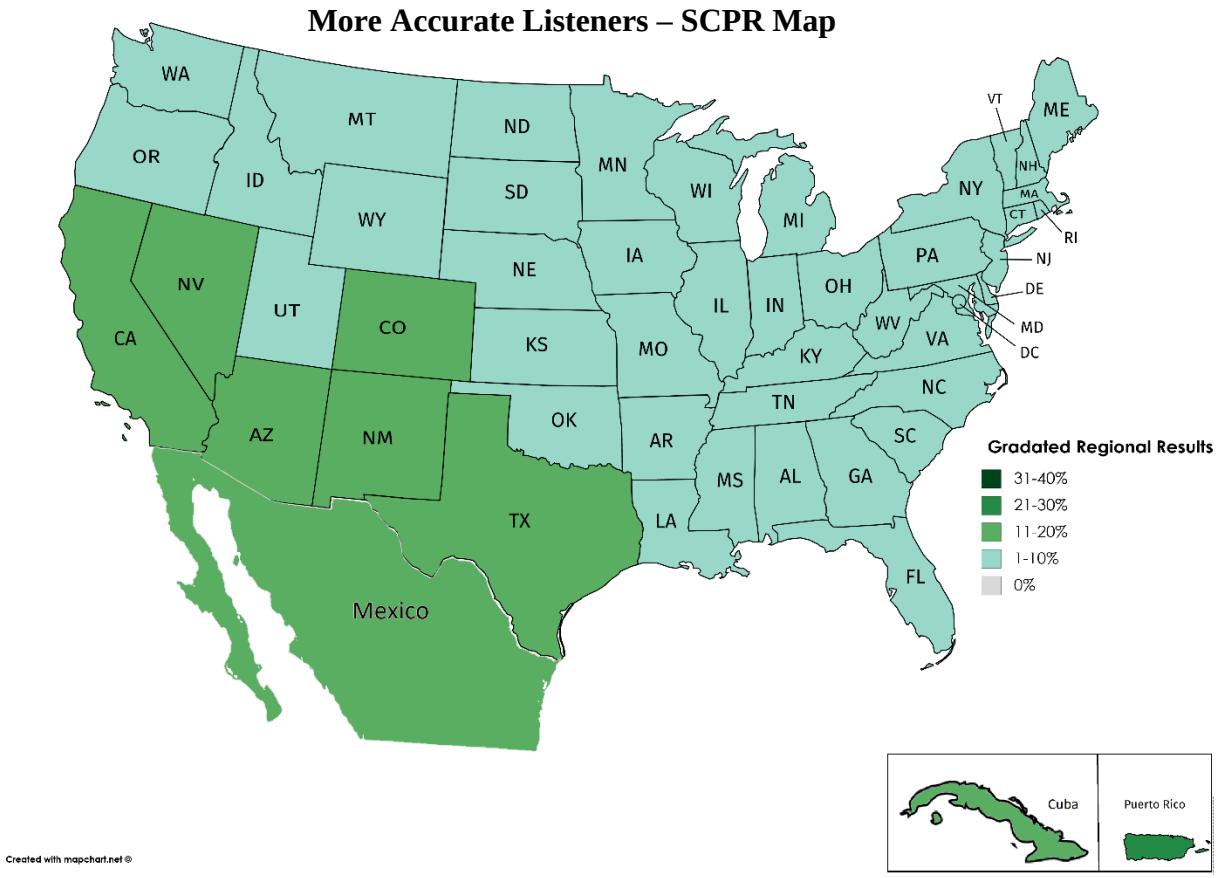


Figure 5.17: More accurate listeners' composite regional map of the SCPR speaker group.

In comparison, more accurate listeners in Figure 5.17 performed similarly to the less accurate listeners for the SCPR speaker group. More accurate listeners illustrated the same relative response rate as the less accurate listeners for each of the dialect regions, except for the Southwest, where more accurate listeners diverged by placing the SCPR speakers there at a similar rate to the Spanish-dominant regions (Puerto Rico at over 20%, and the remaining regions in the lightest shade of green at over 10% and under 20%). Unlike the less accurate listeners, the more accurate listeners connected SCPR to both the Spanish-dominant regions and the Southwestern Crescent, referenced earlier as a region influenced by Spanish, both historically and at present.

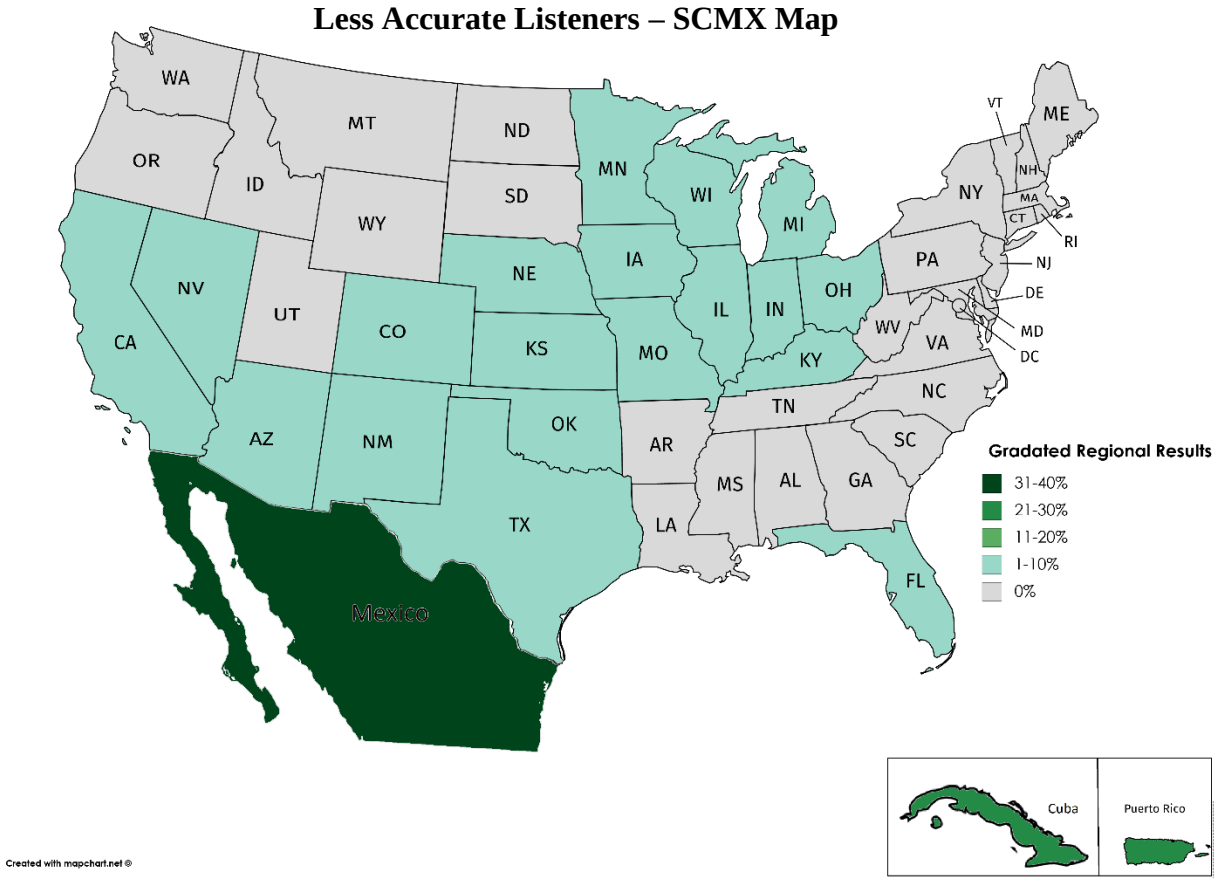


Figure 5.18: Less accurate listeners' composite regional map of the SCMX speaker group.

Contrasting with overall responses to the SCPR speaker group, less accurate listeners demonstrated both more accurate and more precise perceptual judgments of SCMX in Figure 5.18. Less accurate listeners reported strong associations of the Mexican Spanish control group to Mexico, followed by Cuba and Puerto Rico as secondary regions. Much weaker tertiary responses rates were scattered throughout the mainland United States, with no responses reported in the South, the Northeast, and the Rockies. In comparison to regional association to SCPR in Figure 5.16, less accurate listeners followed a similar pattern; the difference in this map is that the responses are more densely packed into the target region.

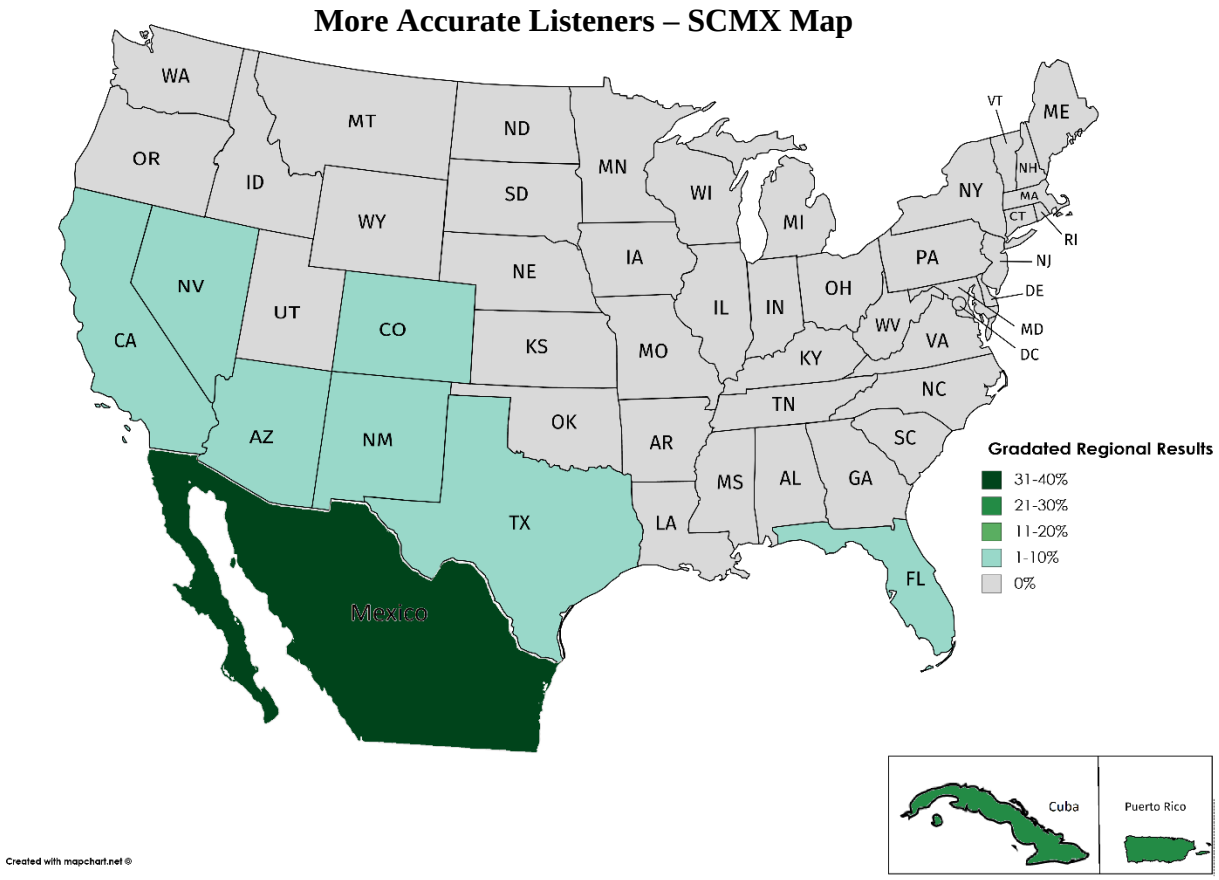


Figure 5.19: More accurate listeners' composite regional map of the SCMX speaker group.

The more accurate listeners' dialect region associations for SCMX in Figure 5.19 draw a starker contrast. Their regional associations of the SCMX speakers were even more concentrated than those of the less accurate listeners. Much like the less accurate listeners, the more accurate listeners correctly placed the SCMX speakers within the target region, but their responses were particularly focused only on the Spanish-dominant and Spanish-influenced dialect regions—Mexico/Puerto Rico/Cuba and the Southwestern Crescent/Florida, respectively. No responses were reported for any of the other dialect regions in the map. For SCMX, a similar pattern can be observed between the less accurate and more accurate listeners, where both groups identify the target regions at higher rates, but the more accurate listeners do so more precisely and consistently, following the wave-like response pattern.

With the results of the pre- and post-tests separating the less accurate and more accurate listeners and a control baseline established with EC, SCPR, and SCMX to confirm their response patterns, we can now expand on their mental map judgments of the simultaneous bilinguals.

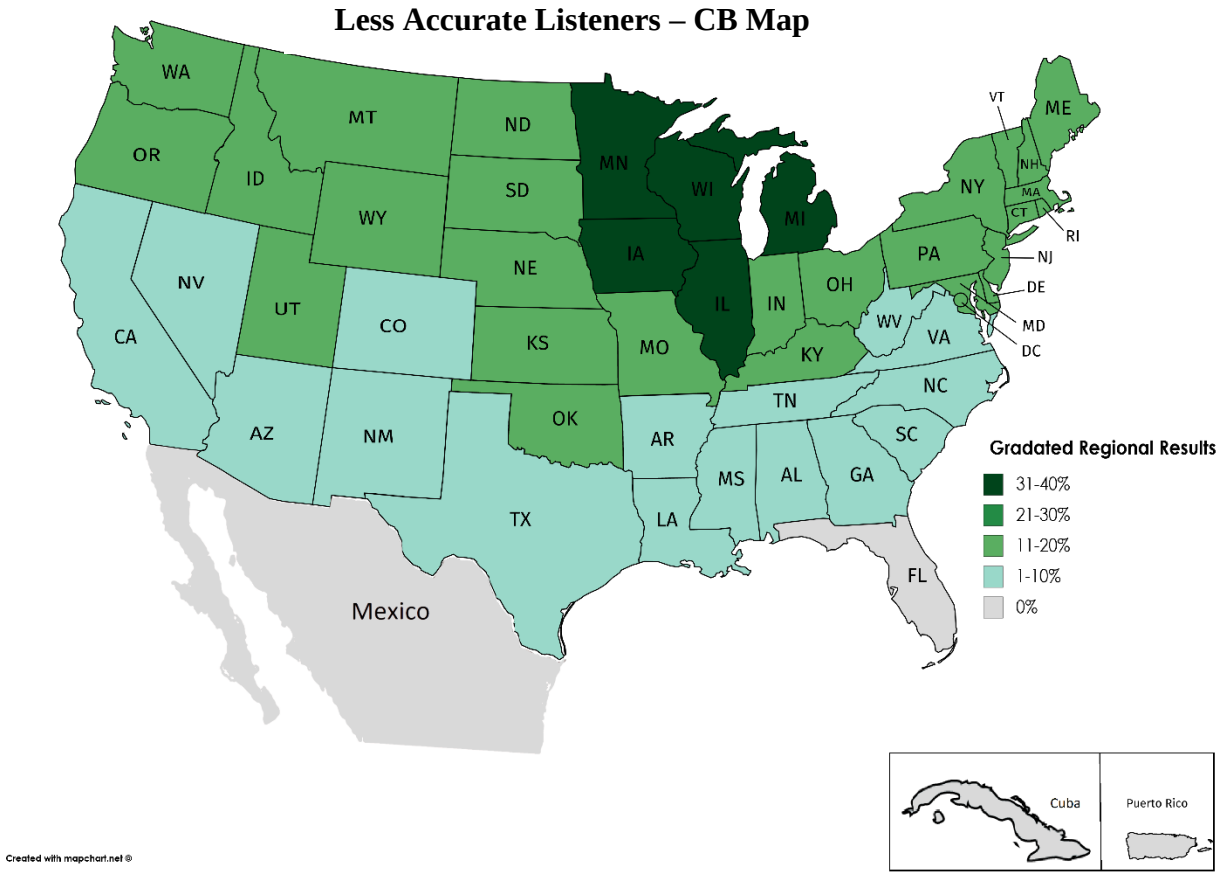


Figure 5.20: Less accurate listeners' composite regional map of the CB speaker group.

Figure 5.20 illustrates the less accurate listeners' mental map for CB. These listeners associated the CB speakers most with the Upper Midwest, followed by the Midlands, the Northeast, and the Rockies, and finally the South and the Southwest at a much-reduced response rate. This map once again shows two consistent patterns: (1) the wave-like response pattern, and (2) the reliable patterning between CB and EC, which was readily apparent in the signal detection tasks as well. As it has been observed for speakers that have been strongly associated to the Upper Midwest, there were no responses in the Spanish-dominated or influenced regions.

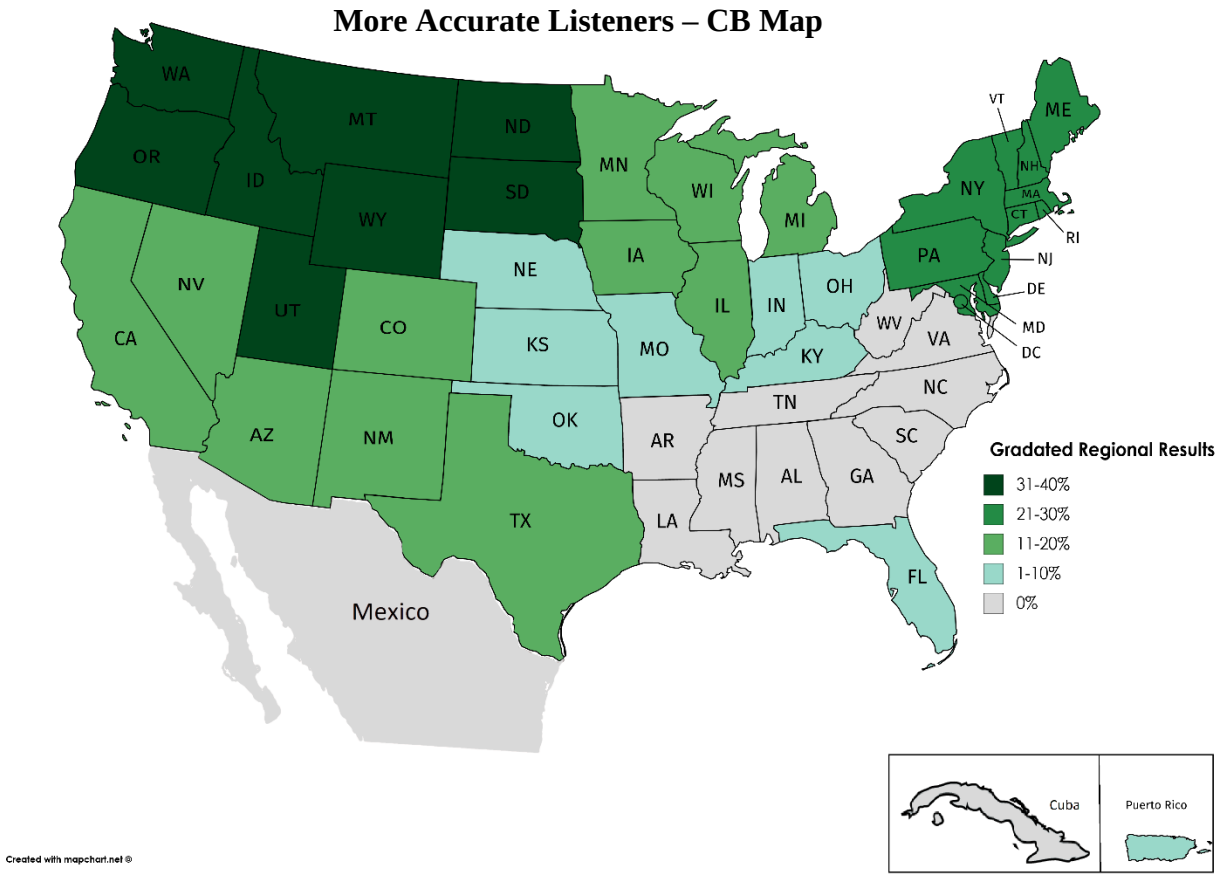


Figure 5.21: More accurate listeners' composite regional map of the CB speaker group.

The more accurate listeners diverge from the less accurate listeners in their perceptual map of CB in Figure 5.21. These listeners strongly associated the CB speakers more with the Rockies area, a region that is largely perceived as unmarked due to more recent migration and settlement in comparison to the eastern United States. Unlike previous maps, this map also breaks away slightly from the wave-like pattern, with the Northeast receiving the second-most number of responses, the Upper Midwest and the Southwestern Crescent third, the Midlands, Florida, and Puerto Rico last, and no responses in the South. Strong associations to the Northwest indicate that the more accurate listeners likely perceived the CB speaker group to be unmarked. This mental map is the only one that strongly links any of the speaker groups to the Rockies region.

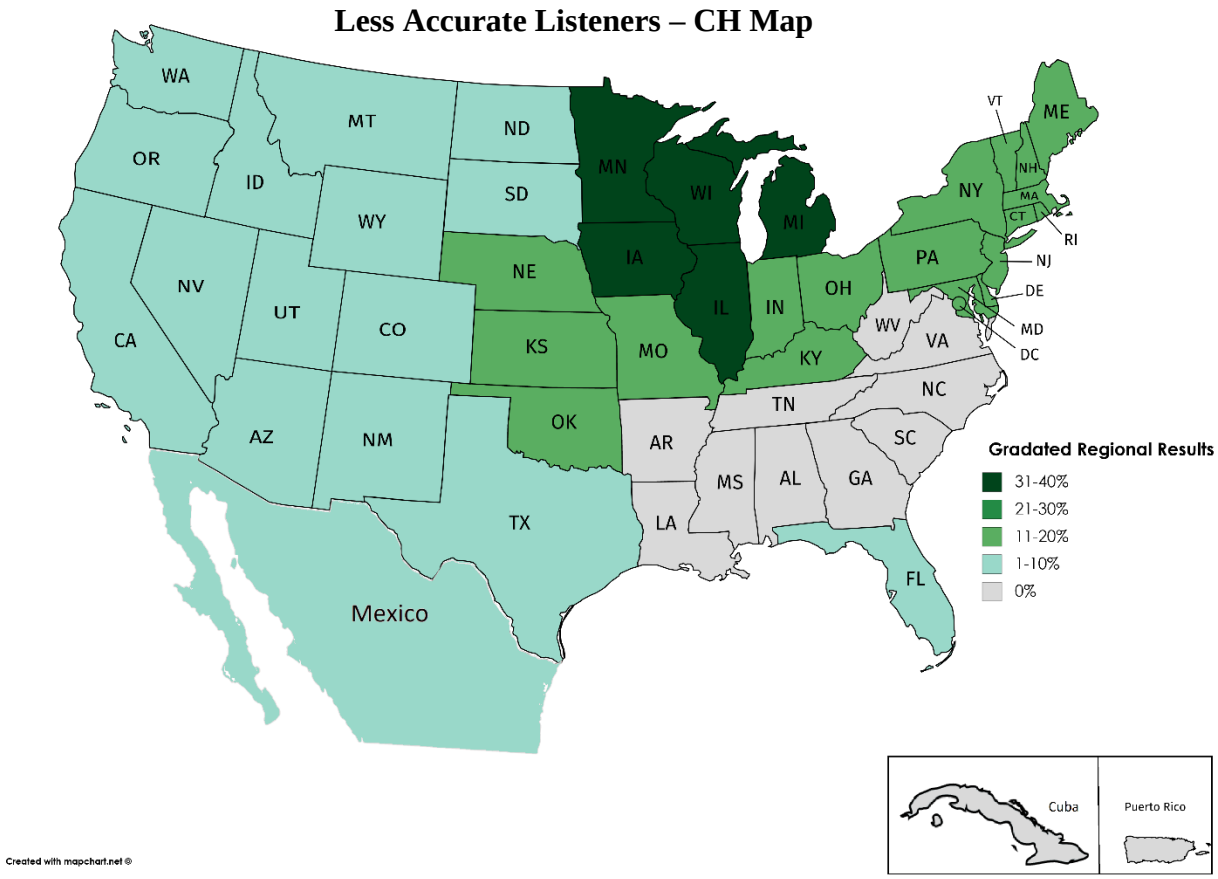


Figure 5.22: Less accurate listeners' composite regional map of the CH speaker group.

The less accurate listeners' mental map for CH in Figure 5.22 follows a similar pattern as their judgments for CB in Figure 5.20. The CH speakers are strongly associated with the Upper Midwest and secondarily with its surrounding regions, with little to no responses in the more distant regions on the map. As discussed for CB in Figure 5.20, the less accurate listeners' perception judgments for CH have also been observed to pattern with CB and EC in the previous tasks, and this mental map in Figure 5.22 further reinforces that pattern. It is possible that the less accurate listeners are not identifying any enregistered markers in CB and CH and are, thus, defaulting to their own dialect region to signify perceived lack of markedness. This pattern was observed to a larger extent in the pilot study results, namely because the listeners were not controlled for preexisting knowledge and sensitivity to regional dialect variation.

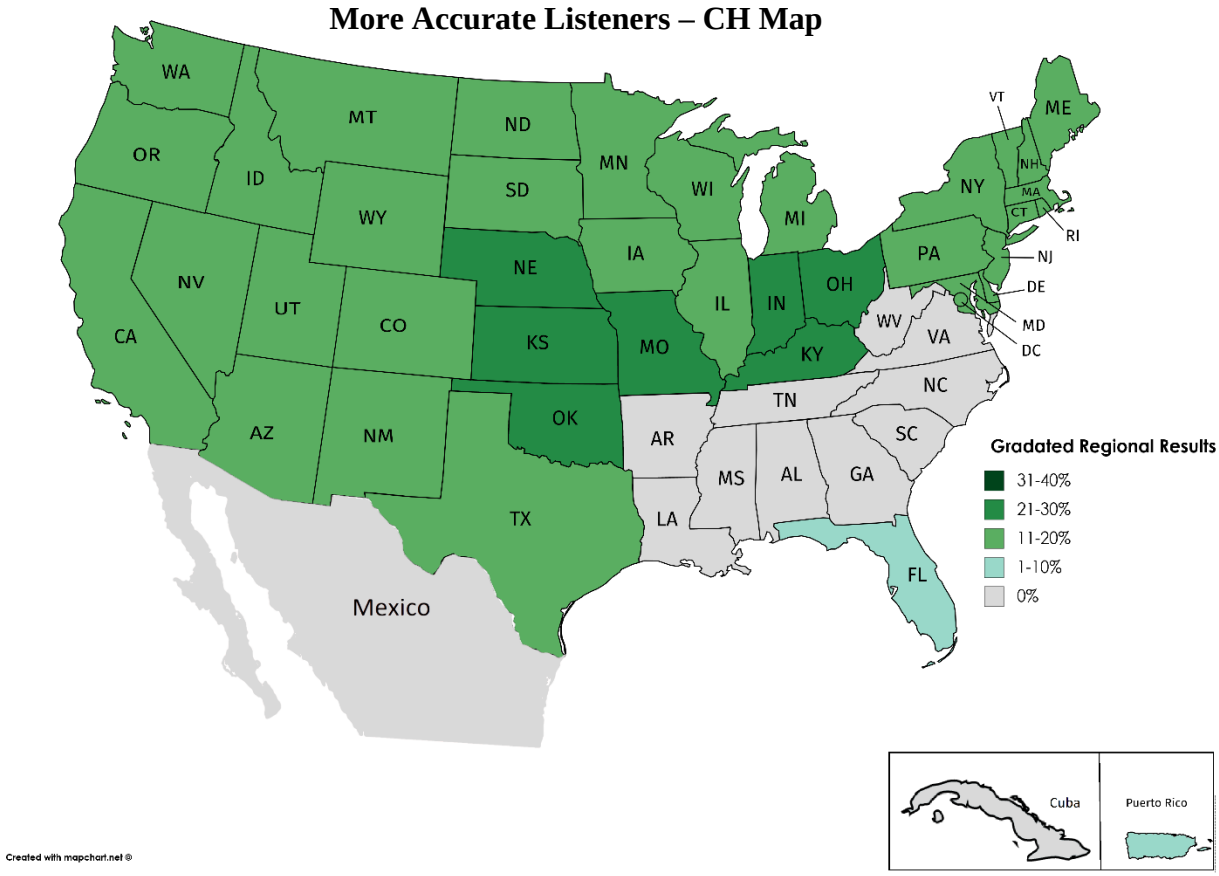


Figure 5.23: More accurate listeners' composite regional map of the CH speaker group.

The more accurate listeners once again diverged from the less accurate listeners in their mental map judgments of CH in Figure 5.23. More accurate listeners placed CH primarily in the Midlands, although not strongly in comparison (<30%), and relatively equally across the Upper Midwest, the Northeast, the Rockies, and the Southwestern Crescent. Florida and Puerto Rico drew a minimal number of responses, while no responses were recorded in the South, Cuba, or Mexico. The Midlands were once associated to General American English, an informally recognized standard that was perceived to be devoid of marked speech. This region, since then, has differentiated thanks to urban-oriented sound change (Labov, Ash & Boberg, 2006, pp.133-137). Associations to this region can largely be connected to more recent varieties of Urban Englishes and are explored further in the discussion section for this task, §5.3.

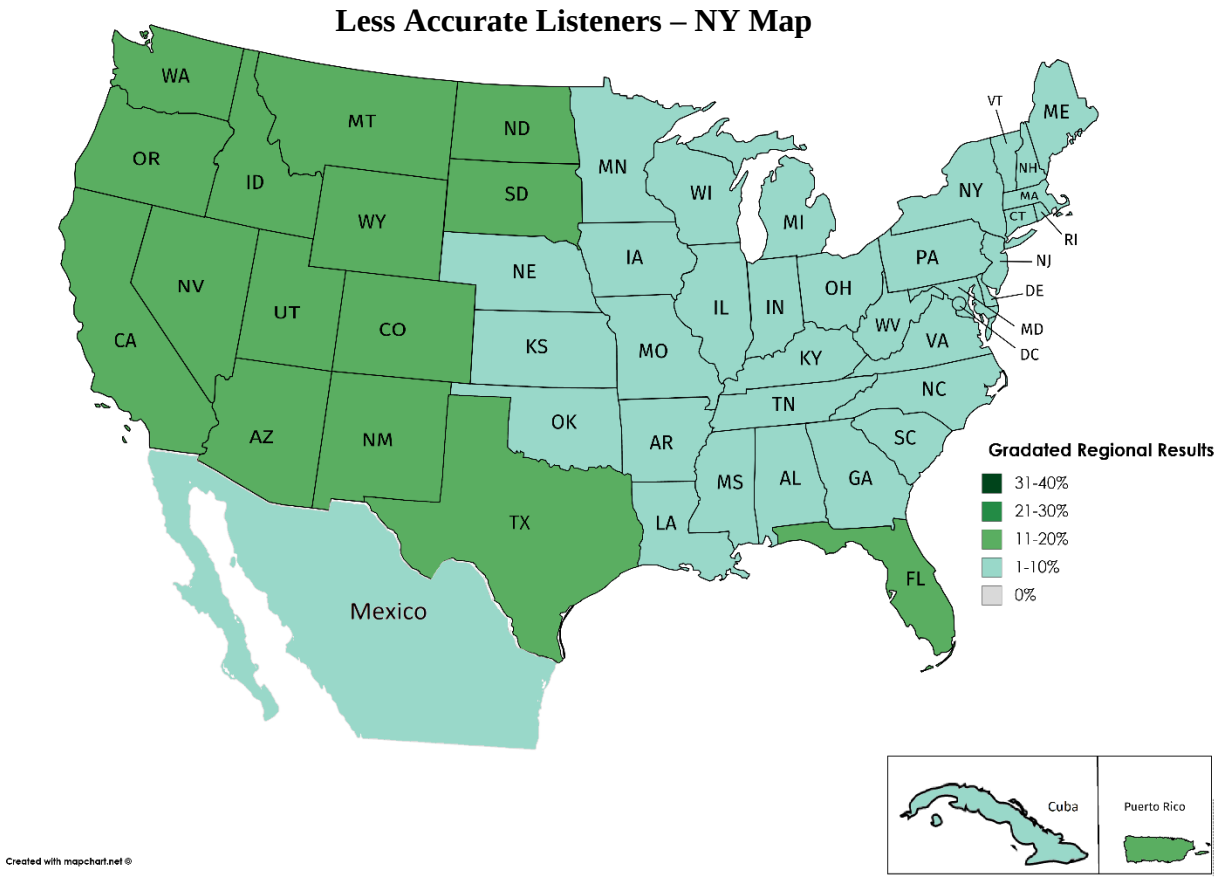


Figure 5.24: Less accurate listeners' composite regional map of the NY speaker group.

Figure 5.24 illustrates the less accurate listeners' mental map for NY, with surprising findings to a certain extent. These listeners associated the NY speakers equally across the Southwestern Crescent, the Rockies, Florida, and Puerto Rico, while there were minimal responses across the board in all of the remaining dialect regions. On the one hand, this map is surprising in the sense that one would expect Nuyoricans, speakers of an enregistered variety to be more associated to the Northeast, which includes New York. On the other hand, the map does show another recurring pattern in the findings thus far, a reliable patterning between listeners' perceptions of NY and the Spanish Control groups, where there is consistent divergence between EC/CB/CH and NY/SCPR/SCMX, with results of PRIE either in between them or more towards the EC speaker group cluster.

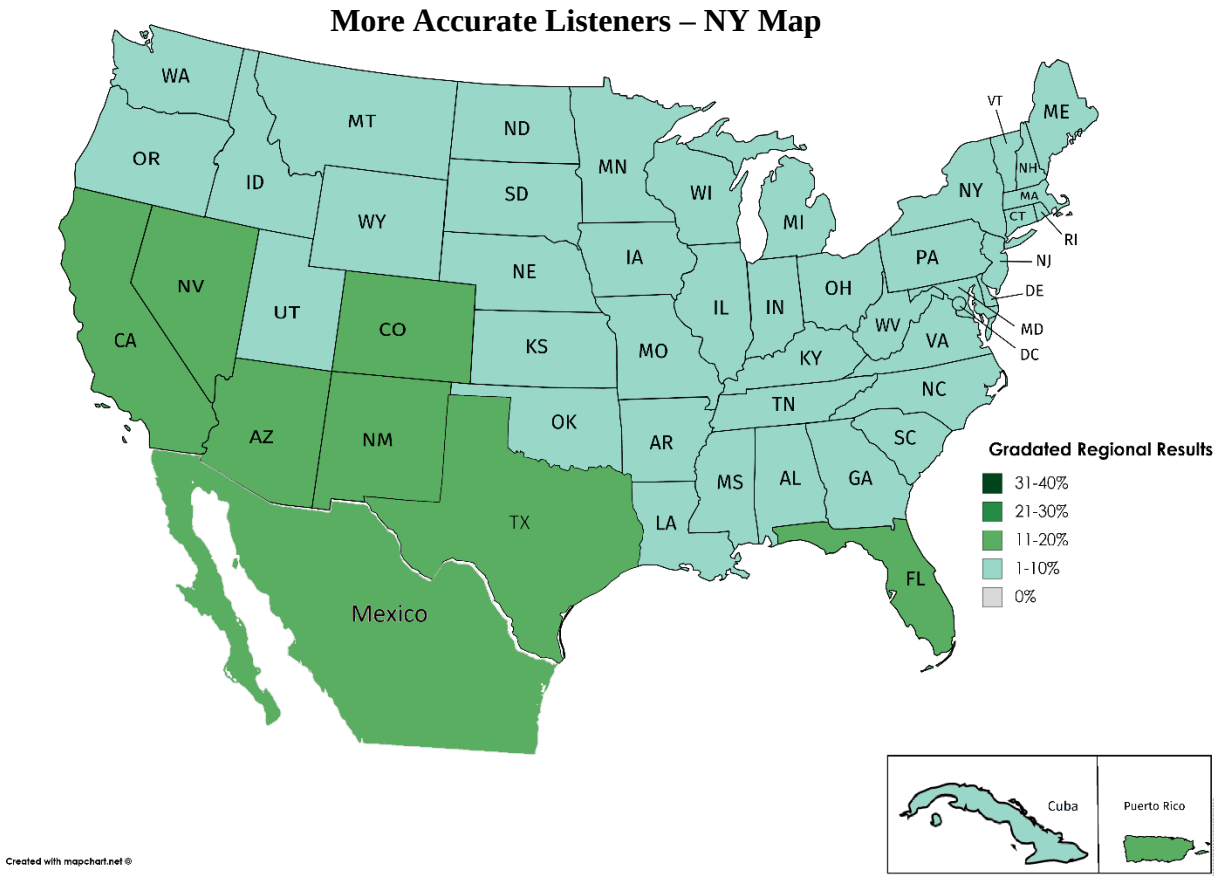


Figure 5.25: More accurate listeners' composite regional map of the NY speaker group.

Examining the more accurate listeners' mental map of NY in Figure 5.25, the pattern observed Figure 5.24 with the results from the less accurate listeners remains largely the same. The more accurate listeners' perceptions of NY were roughly equally divided between the Mexico, the Southwestern Crescent, Florida, and Puerto Rico, with minimal responses for each of the remaining mainland dialect regions. For these listeners, perceptions of NY are less informed by the variety's enregisterment and are more shaped by their associations to Spanish-influenced or Spanish-dominant dialect regions. Building on previous observations on the differences between the less and more accurate listeners, the more accurate listeners' responses were more focused. Even though these listeners did not indicate the initially expected dialect region, they were, nevertheless, more consistent than the less accurate listeners.

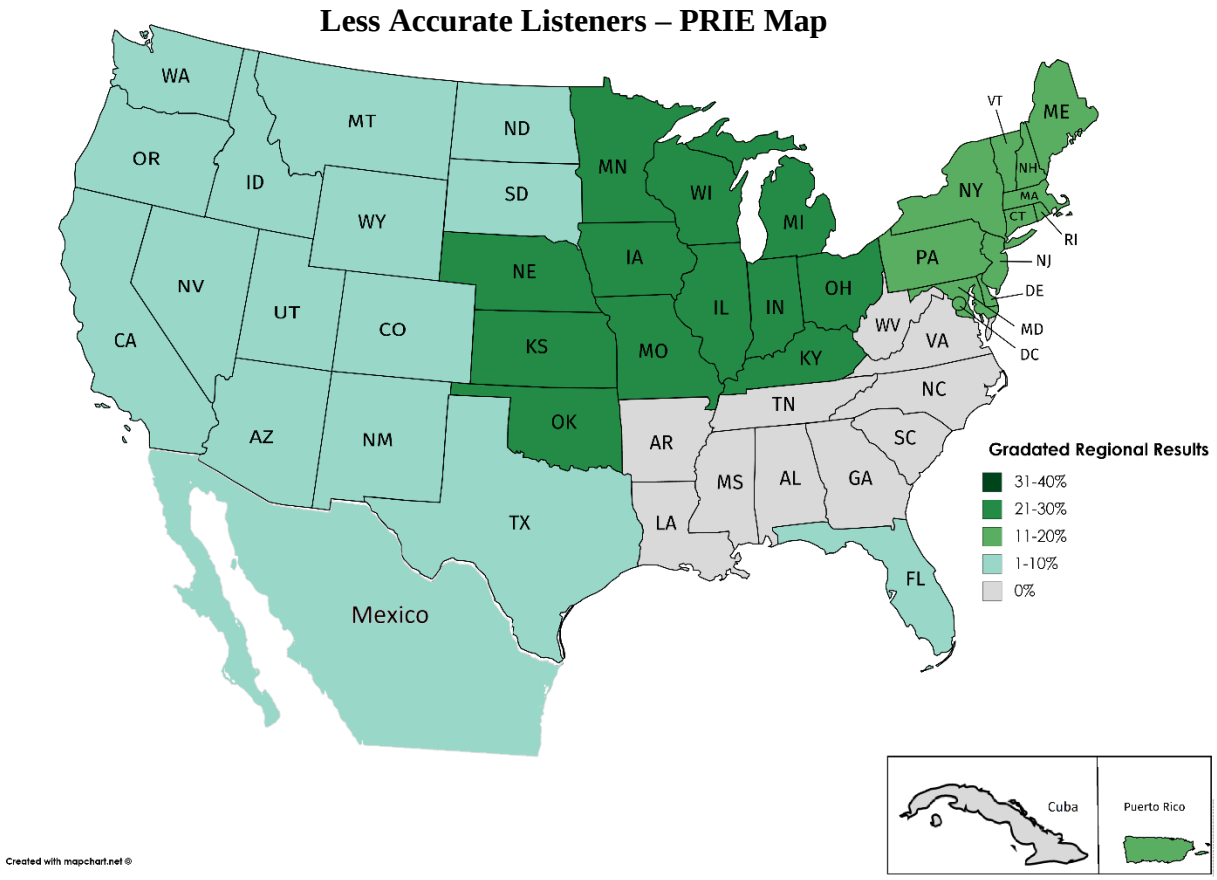


Figure 5.26: Less accurate listeners' composite regional map of the PRIE speaker group.

The less accurate listeners' mental map for PRIE, in Figure 5.26, shows a relatively strong primary association of the PRIE speakers to the Upper Midwest and Midlands regions, a secondary association to the Northeast and Puerto Rico, and minimal associations to every other region except Cuba and the South. This map represents a similar pattern to that of the CH group in Figure 5.21 for the same group of listeners, with the exception that instead of a strong association only to the Upper Midwest, there is a more generalized association across the Upper Midwest and the Midlands. Much like CB and CH, the results are reminiscent of the perceptions of the English control group, where the loci of responses are in the heartland. However, even though this PRIE map is a mixture of the results across the CB, CH, and EC, the PRIE speakers are the only ones among this northern pattern set that were also identified with Puerto Rico.

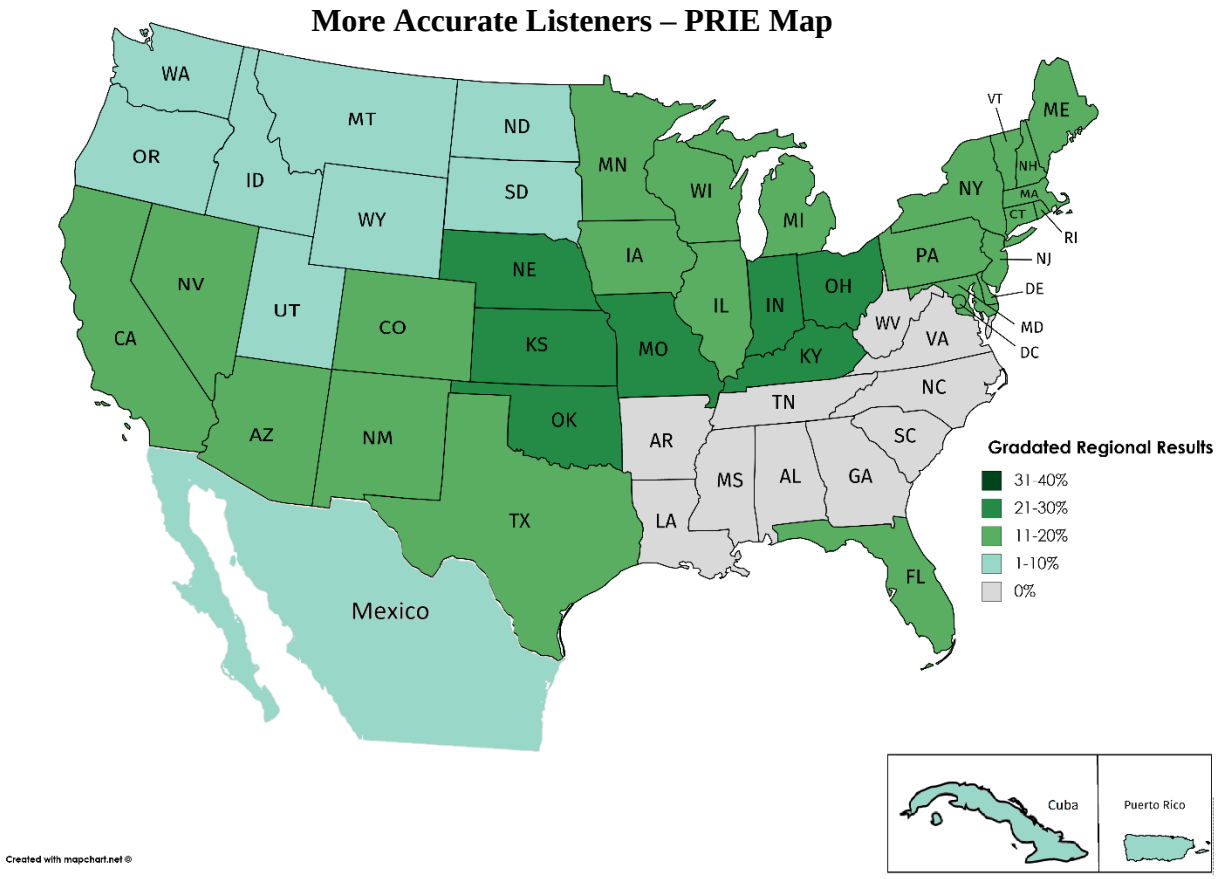


Figure 5.27: More accurate listeners' composite regional map of the PRIE speaker group.

The more accurate listeners' mental map for PRIE, in Figure 5.27, presents a slightly different perspective. In contrast with the less accurate listeners, these listeners identified the PRIE speakers more with the Midlands than any other region, demonstrating a similar result to what the more accurate listeners provided for the CH group. More so, PRIE is only minimally associated with the more Spanish-dominant regions, Cuba, Mexico, and Puerto Rico, while there is a secondary association to the Upper Midwest, the Southwestern Crescent, and Florida—the latter two of which are Spanish-influenced regions, and the last of which has one of the largest populations of Puerto Ricans in the mainland, Florida. For both the less and more accurate participants, the only region with which PRIE speakers were not identified to some extent was the South.

In order to more closely observe the main Mental Map Task results and compare them across the two listener groups, the numerical findings that were used to generate the maps in Figures 5.14-5.27 are provided below in Tables 5.2 and 5.3. The values in both tables were gradated using a color-coding scale and organized from most to least like EC to demonstrate the results across the speaker-to-region comparisons.

Table 5.2 summarizes the less accurate listeners' overall mental map results for each of the speaker groups by the dialect region to which the groups were associated.

Table 5.2

Less Accurate Listeners' Speaker Group Identification Allocation Percentages by Region (N=20)

	Identified Dialect Region									
Speaker Group	UM	Midlands	Rockies	NE	South	FL	SW	PR	Cuba	MX
EC	25%	12%	17%	23%	8%	2%	13%	0%	0%	0%
CH	38%	20%	10%	18%	0%	2%	8%	0%	0%	3%
CB	38%	18%	17%	15%	2%	0%	10%	0%	0%	0%
PRIE	22%	22%	8%	12%	0%	8%	10%	15%	0%	3%
NY	5%	10%	15%	7%	3%	15%	12%	18%	7%	8%
SCPR	3%	5%	7%	7%	5%	5%	8%	27%	17%	17%
SCMX	5%	2%	0%	0%	0%	7%	5%	22%	28%	32%

Note: UM: Upper Midwest. NE: Northeast. FL: Florida. SW: Southwest. MX: Mexico.

As discussed in the map results for Figures 5.14-5.27 above, the results from the main Mental Map Task resemble those of the ID Task in Chapter 4, where listeners grouped CH and CB with EC and NY with SCPR and SCMX. The PRIE row, enclosed in a bold rectangle, represents the breakoff point between these two groupings, which shows that PRIE, once more, occupies a middle point between the two groups for the less accurate listeners. For these listeners, PRIE is equally split between the Upper Midwest and the Midlands, with secondary associations to the Northeast and Puerto Rico. Comparatively, the results for PRIE appear as a combination between the results for CB/CH and NY, where there is a heavier preference towards

the northern regions, with a notable number of responses for some of the regions associated with Spanish, particularly Puerto Rico.

To compare, Table 5.3 shows the more accurate listeners' overall mental map results for each of the speaker groups by the dialect region to which the groups were associated.

Table 5.3

More Accurate Listeners' Speaker Group Identification Allocation Percentages by Region (N=23)

	Identified Dialect Region									
Speaker Group	UM	Midlands	Rockies	NE	South	FL	SW	PR	Cuba	MX
EC	38%	25%	12%	12%	7%	3%	4%	0%	0%	0%
CH	17%	23%	16%	20%	0%	4%	14%	4%	0%	0%
CB	16%	10%	32%	25%	0%	1%	14%	1%	0%	0%
PRIE	14%	22%	9%	12%	0%	14%	20%	6%	1%	1%
NY	3%	10%	6%	6%	6%	16%	17%	14%	7%	14%
SCPR	1%	1%	4%	3%	1%	10%	19%	26%	14%	19%
SCMX	0%	0%	0%	0%	0%	1%	6%	30%	30%	32%

Note: UM: Upper Midwest. NE: Northeast. FL: Florida. SW: Southwest. MX: Mexico.

The overall distribution of responses for each speaker group is different for the more accurate listeners in Table 5.3. These listeners were generally more consistent and precise than the less accurate listeners, which can be evidenced in the higher concentration of responses for the target dialect regions in the control groups. The more accurate listeners' responses for CH and CB still pattern the most with EC, and the same is evident with NY and SCPR/SCMX. A similar pattern with regards to PRIE appears for the more accurate listeners, also enclosed in a bold rectangle to underscore the same three-tier split pattern. However, PRIE is slightly more divided between the Midlands and the Southwestern crescent than the Midlands and the Upper Midwest when comparing across the listener groups.

A final overall observation is that the lack of responses for the South for both listener groups is a positive result in that no speakers demonstrated features that would be overtly associated with the South. Similarly, all the simultaneous bilingual groups received non-minimal

responses in the Southwestern Crescent, indicating listeners' implicit awareness of the language background in Spanish that the speakers share. In combination with the pre- and post-test results, the results for the main task strengthen the consistency of these listeners' mental map intuitions.

5.3 Discussion

The results for the Mental Map Task yield three important observations: (1) there is a notable difference in performance across the different listener groups that were divided based on their pre- and post-test results; (2) all listeners' response patterns generally follow wave-like patterns, as predicted in the design process when observing mental map results from previous studies; and (3) the associations that listeners applied to each speaker group follow the established patterns from the results of the signal detection tasks, where the seven speaker groups were roughly divided into three tiers—largely like EC, largely like SC, and mostly like neither.

To the best of my knowledge, no previous study in perceptual dialectology has attempted to divide listeners by level of sensitivity and preexisting relevant background knowledge as done in this Mental Map Task. The results from this task show that there are substantial underlying differences once listeners are sorted out into subgroups according to those criteria. In general, the less accurate listeners were both less consistent and less likely to associate a stimulus to the target. Contrastively, the more accurate listeners were generally more consistent and closer to the expected target, when applicable.

Regardless of the listener group, however, listeners' responses followed a wave-like pattern, where the regions adjacent to those regions that received the highest response rates were secondary or tertiary response foci. For instance, no group that was primarily associated to the Upper Midwest or the Midlands was secondarily associated to the South or one of the Spanish-dominant dialect regions on the map. As discussed in the methods design process for this task,

Preston's north-to-south Mississippi river mental map task (1996) demonstrated similar recognition patterns. This distribution can likely be attributed to regionally specific dialect knowledge in that listeners have more knowledge about their immediately relevant language contexts than those that are more distant. Even for the less accurate listeners, these results show that listeners are sensitive to dialect variation and the relationship between distance and marked changes in regional dialect features.

The Southwestern Crescent was purposefully selected as a dialect region for this task because of its historical and contemporary coherence as a Spanish-influenced region and the results from the pilot, where listeners associated Spanish-dominant speakers with that area. Listeners' options to that end were expanded on in this task by including Spanish-dominant regions (Puerto Rico, Mexico, and Cuba) that were also ancestrally or directly parallel to the simultaneous bilingual speaker groups. This modification was intended to not only make the mental map more analogically relevant to the speaker groups, but to also test how the strength of Spanish-influenced associations are in comparison to the Southwestern Crescent. The findings for the Spanish control groups' results reveal that listeners overwhelmingly associated a marked Spanish accent to the Spanish-speaking regions. In contrast, the roughly equally dominant simultaneous Spanish-English bilingual groups were not generally associated to the Spanish-speaking regions, but they were secondarily connected to the Southwestern Crescent, despite patterning closer to EC. These results suggest that some of the listeners were implicitly aware of some degree of influence from Spanish in their Englishes.

The dialect region termed 'the Rockies' in this study contrasts with the Southwestern Crescent. Descriptively, the Rockies comprises most of the dialect region defined as "The West" in Figure 5.1. It is a region that is defined by its lack of homogeneous dialect features that would

differentiate it from the dialect regions in the eastern United States (Labov, Ash & Boberg, 2006, p.284). The eastern United States was colonized by English speakers much earlier in history and has had more time to differentiate and develop distinguishing regional features. Because of that paucity of enregistered regional dialect features, perceptions of the Rockies are similar. Of the speaker groups in this task, CH and CB were strongly associated to the Rockies by the less and more accurate listeners, respectively, which suggests that these listener groups did not identify any overt dialect features in these speakers.

The Midlands are especially relevant to this discussion, given that both listener groups associated the PRIE speakers to that dialect region with some degree of prominence. The Midlands was a region (Figure 5.1) that was once associated with General American English, the widely perceived informal standard of American English. As noted in the presentation of the mental map results for CH, sound changes in cities like Cincinnati and partial influence of the Northern Cities Shift in cities such as St. Louis have differentiated the region away from its once-perceived unmarked informal standard.

The Midlands region is distinguished by features such as the low-back merger as well as /ʌ/ and /u/ fronting[20] (Labov, Ash & Boberg, 2006, pp.133-135). In fact, Labov, Ash & Boberg speculate that "there is reason to believe that the Midland is becoming the default system of North American English" as a function of migratory patterns into cities outside of the dialect region, such as Norfolk, Charleston, and Atlanta (2006, p.135). Preston (2003) observes that the Midlands region is not as broad reaching as initially believed, but that it is a rather tiny "skinny Midland" found between the North and the South, one that is perceptually real (p.240) for listeners and representationally real in terms of its shared lexical, phonetic, and phonemic

[20] See the cited source for more information on these sound changes within the context of the dialect region and North American Englishes.

systems (Labov, Ash & Boberg, 2006). The Midlands, in effect, is a small, yet cohesive, boundary region between the northern and southern dialect regions, with observable features that distinguish it. This is no longer the story of the Midlands as the locus for General American English, but rather as a categorically distinct region that is perceptually salient. As Labov, Ash & Boberg (2006) observe, it shares features that are predominant in newer urban Englishes in the United States.

Regardless of the Midland's standardness status, it is reasonable to observe that it carries widespread influence in contemporary American English, given its expanding reach outside of the region. What follows that increased scope is a potential rise in the spread of the Midland's dialect features,[21] which would be perceptually indistinguishable for outside listeners. As such, associations of an emerging dialect carrying those features to the Midlands would be what one would expect out of an emerging dialect of American English that has yet to enregister, as is the case for PRIE in these findings. Although a phonological analysis of PRIE has not been conducted in this study, given the results of the Mental Map Task, it is expected that the findings of such an analysis would reveal similar sound patterns as those that are usually ascribed to the Midlands, such as those observed above.[22]

A final factor to consider in looking at the results for PRIE is the overall percentage spread of responses for both listener groups. A higher concentration of responses, particularly at least over 20%, given the abundance of categories functioning as noise, indicates some degree of consistency in allocating a speaker group to a region. However, a broader response spread across multiple or all regions should be interpreted as a lack of a regional association. This observation applies generally to the NY speaker group, whose response rates patterned slightly more with the

[21] Preston (2003) terms these features that are distinctive of dialect boundaries, DIALECTEMES.
[22] See Chapter 6 for more on future studies following up on the findings in this dissertation.

Spanish controls yet were, nevertheless, spread out roughly evenly across the board. It is, then, important to study both pattern and spread when analyzing the results. Patterns indicate a consistent association of a speaker group to a region or set of regions; these are useful to examine comparatively to determine the relative performance of a group. Spreads are indicative of a listener group's overall confidence in a pattern, where a tighter spread is indicative of higher confidence and vice-versa.

In terms of spread, the PRIE group, much like the NY group, is more spread out than the other five groups, and particularly, the control groups, which shows less overall confidence in the pattern for PRIE. Listeners' responses for PRIE behave similarly to those for NY in that sense, even though there are slightly stronger preferences and tendencies with PRIE than NY. However, in terms of pattern, a consistent observation is that the Midlands is among the highest response regions for PRIE for both listeners groups, behaving similarly to the two simultaneous mainland Spanish-English bilingual speaker groups that are most associated with the English control group, CB and CH. These results are indicative of a broader theme observed in the results for the signal detection tasks in Chapter 4, where the results for PRIE do not indicate that it is a marked variety, particularly when compared to NY and the SC groups.

To summarize, the perception task results largely indicate that PRIE behaves uniquely from each of the two group tiers described above. Thematically, however, it tends to pattern most with the mainland simultaneous bilinguals that most align with the English control speakers. The results from the signal detection tasks in Chapter 4 found that PRIE held a similar perceptual status as those of the other simultaneous bilingual varieties of mainland American English—that is, one that is categorically recognized as a variety of American English whose enregisterment status was unclear. Given the discussion in this section, the results for the Mental Map Task

suggest that PRIE is not enregistered to a specific region, but rather it behaves as an emerging dialect whose features default to a generalized Midlands status, following similar developing trends in urban areas throughout the mainland United States.

CHAPTER 6

CONCLUSION

This dissertation employed approaches and concepts from PERCEPTUAL DIALECTOLOGY, SIGNAL DETECTION THEORY, and ROOTEDNESS to explore the process of ENREGISTERMENT and NEW DIALECT FORMATION in PRIE as an emerging variety of American English. This study developed a novel methodology to approach the examination of emerging dialects as a product of contrastive perceptual salience, juxtaposing broadly enregistered dialects with emerging dialects, while accounting for confounds in the experiential framework of listeners making these perceptual judgments. The research framework employed throughout this dissertation partitioned the perceptual study into four phases to best capture listeners' representations of each dialect in comparison to the target dialect, PRIE: speaker identification, speaker selection, signal detection for inferential statistical analysis, and the identification of regional dialect allocation for a descriptive representation of listeners' mental dialect maps.

From these analyses, I conclude that PRIE is largely perceived as a variety of American English that is perceptually on par with similar bilingual English dialects in the mainland, such as Miami Cuban and Chicano English varieties. Furthermore, I also found that PRIE is not enregistered to a specific region, but rather it behaves as an emerging dialect whose features default to general Midlands status, following similar developing trends in urban areas throughout the mainland United States.

In this chapter, I summarize my findings, with reference to the initial hypotheses, research questions, and overall results to answer those research questions. I then provide an overview of the broader implication of these findings to contextualize those implications in terms of contributions to the field. Finally, I observe the limitations of this study and explore future

research directions for both this methodological framework and the study of PRIE as an emerging dialect of American English.

6.1 Summary of Findings

This dissertation investigated the emergence of PRIE as a historical circumstance of the complex sociolinguistic, historical, and ecopolitical relationship between Puerto Rico and the United States within the broader scope of World Englishes. This context shows that language attitudes towards English in Puerto Rico have ameliorated in younger Puerto Rican generations, resulting in a shifting linguistic landscape, where both L1 Spanish and English coexist in an increasingly bilingual population. As discussed throughout, this context is a mirror image of similar phenomena witnessed in bilingual Spanish-English communities in the mainland United States, such as Miami Cuban English. Given the literature discussed in Chapter 2, a set of hypotheses were formulated to further explore the existence of PRIE:

H0: There is *no* PRIE.
H1: PRIE is emerging *in Puerto Rico*.
H2: PRIE is *not* emerging *as a unique form, but rather, it is a subset of existing mainland American English varieties*.

To investigate these hypotheses, the present study examined the status of PRIE in relation to other varieties of American English based on the following research questions:

1. Based on speakers' audio recordings and no other background information, can listeners identify PRIE speakers distinctly from other similar varieties of mainland American Englishes (Nuyorican, Chicano English, and Miami Cuban English)?
2. How do speakers of Puerto Rican Englishes (PRIE and Nuyorican) compare to other selected recognized varieties of mainland American Spanish-English bilingual varieties (Chicano English and Miami Cuban English)?

3. From a perceptual perspective, to what dialect region(s) of the United States are the selected American Spanish-English varieties associated vis-à-vis the Spanish and monolingual English control speakers?

To answer these questions, the study was divided into four phases. The first two phases of the study delimited the speaker selection process, while Phases 3 and 4 addressed the research questions with the speaker data gathered and screened for in Phases 1 and 2, respectively. In accordance with the results discussed in Chapters 4 and 5, the answers to each of these questions are reviewed and contextualized in the subsections below.

6.1.1 Signal Detection Tasks Summary

To the best of my knowledge, no prior studies have been conducted to identify and observe the gradual shift of the role of English in Puerto Rico from an L2 to an L1 variety among younger bilingual Puerto Ricans. Analyses in this study departed from the observation that, unless accidentally discovered, an object of study must first rise to a level of awareness at which we can recognize it as sufficiently different to investigate. Aspects of approaches in PERCEPTUAL DIALECTOLOGY also began from the premise that perception judgments guide representations of linguistic realities in listeners from all backgrounds—realities that affect how we interact with the world around us and vice-versa. Following quantitative methods, I aimed to unify the perspectives in PERCEPTUAL DIALECTOLOGY with inferential approaches, which ultimately strengthen the generalizability, verifiability, and control of the resulting data.

Combined, the ID and AX signal detection tasks were concerned with the first two research questions of this study, each of them respectively tasked with identifying whether listeners heard a signal that would indicate what they believed to be a Puerto Rican dialect and whether the PRIE speakers were perceived differently from speakers of other Spanish-English

bilingual dialects from the mainland United States. The findings from the signal detection tasks confirmed that listeners did, in fact, identify PRIE as distinct from the speaker groups with more dominance toward one language and that PRIE holds a similar perceptual status as the other simultaneous bilingual varieties of mainland American English.

In the ID Task, PRIE categorically occupied a middle point between the more recognized mainland varieties and the Spanish controls. Despite this differentiation, PRIE tended to pattern more towards the English-dominant (EC) side of the gradient of results, drawing a stark difference from SCPR and NY, the more perceptually salient varieties of Puerto Rican-affiliated varieties in the speaker group samples. In the AX Task, PRIE illustrated a similar perceptual status as the other simultaneous Spanish-English bilingual varieties of mainland American English. These findings, where naïve listeners largely identified a signal that categorically distinguished PRIE from the other speaker groups, confirm H1 in that PRIE is undergoing NEW DIALECT FORMATION in Puerto Rico. Listeners consistently indicate the perceptual reality of PRIE in juxtaposition with other simultaneous and sequential bilingual Englishes that run parallel to it to varying extents.

It is important to note, however, that, although the signal detection tasks underscored the categorically distinctive position that PRIE presently occupies, listeners were ambivalent in uniquely identifying it as a variety of Puerto Rican English. In fact, the NY and Spanish Control groups were the most salient in that regard. These results suggest that the perceptual status of PRIE is not straightforward, but rather a product of a more complex metalinguistic awareness, a status that is not all that different from that of Miami Cuban English (as attested in Carter, López Valdez & Sims, 2020, pp.126-127). The salience of PRIE implicates some degree of phonetic

differentiation of which listeners are implicitly aware. The phonological status of PRIE remains unexplored and is a subject of future research.

These findings are indicative of a linguistic landscape in Puerto Rico that is in flux, guided by a long and, at times, difficult history with the United States. The world of English continues to expand in unpredictable ways, as the sociopolitical and economic influence of the United States continues to proliferate, and the language and sociocultural attitudes of younger generations continue to change. The recent emergence of PRIE as a variety of American English in the culturally distinct and geographically isolated island that is Puerto Rico is an example of this phenomenon.

6.1.2 Mental Map Task Summary

The Mental Map Task, following previous designs (such as Preston & Robinson, 2005) complemented the findings from the signal detection tasks by providing listeners' perspectives on the regional associations that they had for each of these dialects in relation to PRIE. These geographically anchored perspectives paint a picture for not only PRIE, but also the existing reference framework that listeners generally had in making perception judgments. This task addressed the third research question by observing enregisterment via the dialect region(s) that listeners most associated to the PRIE speaker group in contrast to the other six speaker groups included in this study. In other words, it captured degrees of regional associations or enregisterment of PRIE; these judgments were counterbalanced with the speakers from the other six dialect groups.

The analyses of the Mental Map Task largely concluded that PRIE did not appear as a form that was uniquely enregistered to Puerto Rico. This finding was expected, given that little attention has been given to this recent emerging dialect. The interesting aspect of this task,

however, was to survey the connections that listeners did draw, despite an absent framework for PRIE. The results in this task also detailed a categorical separation of PRIE from the other dialect groups, where the more English- and Spanish-associated varieties clustered together at opposites ends of a regional spectrum.

Perceptions of PRIE, nevertheless, patterned closer with that of the mainland simultaneous bilingual speaker groups (CB and CH) that were most associated to the English control speakers, suggesting that PRIE held a similar perceptual status to those of these other mainland Spanish-English bilingual dialects. Despite the overall associations from a categorical viewpoint, PRIE was the only speaker group that was principally identified with the Midlands dialect region. It is important to note, however, that this association was not as strong as the strength of the regional associations made for any of the other groups. While the ultimate result for PRIE in the Mental Map Task was consistent, it was also less confident in that the response distribution was more spread out.

The primary goal of this task was met by identifying whether PRIE was enregistered to its region of origin, Puerto Rico, and in comparing those results to the associations made for the other speaker groups. Listeners responded as expected, demonstrating no strong associations of PRIE to Puerto Rico, but instead responding to PRIE in a pattern that was closer to the English control group and further from the Spanish-dominant groups. These expectations were drawn from initial results from the pilot study and were observed to a stronger degree in a more detailed Mental Map Task in this study. A secondary goal of this task was to build on the present findings to explore potential avenues of future research.

The principal association of PRIE to the Midlands sets the stage for a broader range of research inquiry on the nature of this association. Are listeners inclined to associate PRIE to the

Midlands dialect because they hear a set of positive cues that connect PRIE to the Midlands dialect or because they hear no positive northern or southern features? In other words, there are two ways in which listeners may be interpreting the Midlands, either: (1) as a dialect that they can positively identify and associate to PRIE, or (2) as a default ascription of PRIE due to PRIE's perceived lack of regionally identifiable features (the Midlands as a negatively defined dialect). Both factors are not mutually exclusive for the broader set of listeners, and a combination of them could be guiding the results that have been observed.

It is imperative to observe which of these factors are at play in listeners' initial associations of PRIE to the Midlands and to what extent. On the one hand, it may be that there is a set of positively identified features in PRIE that are linked to the Midlands, an explanation which follows the observations that Labov, Ash & Boberg made in the ANAE (2006, p.135), where the Midlands appears to be playing an influential role in the development of newer dialects of American Englishes. In this case, a bigger question is to identify exactly how perceptually salient features and aspects from the Midlands dialect found their way to younger generations of simultaneous bilinguals in Puerto Rico. On the other hand, PRIE may be identified as a Midlands dialect as a product of listeners constructing perceptions of the Midlands as a negatively defined dialect, one that is neither northern nor southern, but something that is in between or lacks the distinctive traits of either. This train of thought follows Preston (2003, p.249), where the Midlands essentially functions as a small transitional boundary between the northern and southern dialect regions.

An explanation for the perceptual status of PRIE could lie in its potential development as a KOINE dialect, a result of the nature of PRIE's long-distance relationship with dialects from all over the mainland United States through its media and cultural markets. PRIE, then, is a dialect

that builds from different regional input from an assortment of dialects from the mainland, which is then reinforced through language use in the broader community of simultaneous bilinguals in Puerto Rico. The resulting dialect in development could be largely perceived as regionally neutral, at least until the point that it solidifies its national identity and becomes enregistered. These findings on PRIE, independently and jointly rooted in the longer discussion on what defines dialects in the minds of its speakers and listeners, warrant further investigation.

6.2 Implications

This study purposefully integrates approaches in PERCEPTUAL DIALECTOLOGY and the use of SIGNAL DETECTION THEORY, focusing on an experiential framework for speakers and listeners. In considering participants' histories, priority was given to specific a priori group classifications, particularly age and gender for speakers—namely to control for generational boundaries and potential gender effects in perceptions. The impetus in this theoretical paradigm is to observe how the effects of background knowledge, sensitivity (for listeners), and performance (for speakers) interact in the outcome of perceptual experiments.

The resulting framework begins to address a multitude of concerns regarding the effects of idiosyncratic background knowledge that guides listeners' perception judgments. This factor is usually unaccounted for in most traditional approaches to perceptual experiments in linguistics, which affects replicability and verifiability. To this end, the methods in this study provide a robust amount of operationalized and quantified background information for each set of participants with measurements for speakers, such as the ROOTEDNESS METRIC, and the triangulation between traditional demographic information with results from the BLP and the LEAP-Q surveys in the screening process.

In addition to the careful screening process for speakers, listeners were controlled for by region of origin and residence, previous relevant language experience, and response bias in the signal detection tasks. In the Mental Map Task, an additional process focusing on dialect sensitivity (more versus less accurate) was included in the analyses. This methodological consideration was implemented in acknowledgment that the individual and their experiences ultimately inform and shape their perception judgments—a phenomenon that is important to consider in the context of listener-centric analyses. In observing the nuance in the results between the listener categories, the methodological significance of this decision cannot be understated.

The results of this study also have implications for future research on the role of English in Puerto Rico and, by extension, PRIE. Puerto Rico has been mostly seen as an L1 Spanish and L2 English-speaking region. While this continues to be the case for most older generations of Puerto Ricans and a considerable segment of younger Puerto Ricans, there is more complexity to the linguistic context. As a product of the amelioration of language attitudes towards English (Vicente Vélez, 2000) and the subsequent disentanglement of the monolingual view of a Puerto Rican identity (Domínguez-Rosado, 2015), there is a growing number of younger Puerto Ricans who are being raised as simultaneous Spanish-English bilinguals, as exemplified throughout this dissertation. In many households, such as my own, English is seen as a gateway of opportunity on an island that is besieged by economic turmoil, exacerbated by a combination of ineffective local governance and the disadvantages of Puerto Rico's now longstanding colonial relationship with the United States.

Although almost every Puerto Rican knows English to some extent, the status of English is not monolithic. In this sense, Puerto Rico finds itself at a junction where both English and

Spanish coexist to different extents in communities of speakers that vary depending on education, age, and identity, at the very least. As an increasing number of younger Puerto Ricans come into the fold, the composition of the speakers that fall into the L1 Spanish and L2 English category may continue to dwindle as the number of simultaneous bilingual PRIE speakers increases. The directionality of this shift in the future remains unclear; it may continue down the direction described in Chapter 2, towards increasing Spanish-English bilingualism, or any number of future sociopolitical events may reverse this trend. As it stands, the number of PRIE speakers continues to grow, which carries significant implications for future language-based studies in Puerto Rico—even beyond the field of linguistics. An L1 Spanish-only Puerto Rico is quickly becoming a relic of the past, and the importance of the emergence of native English (PRIE) speakers throughout the island cannot be overstated.

A final set of implications to consider regards the role in operationalizing approaches in PERCEPTUAL DIALECTOLOGY to identify NEW DIALECT FORMATION and ENREGISTERMENT. Schuld et al. (2017) pioneered the use of signal detection tasks to identify enregisterment by comparing the strength of regionally associated signals that listeners heard in Wisconsin speakers from different generations, despite the absence of salient regional dialect features in their recordings. Modeling on and adapting from that methodology, this dissertation combined approaches from SIGNAL DETECTION THEORY and PERCEPTUAL DIALECTOLOGY to explore NEW DIALECT FORMATION and potential ENREGISTERMENT in simultaneous bilinguals in Puerto Rico. The principal challenge was to operationalize this methodological adaptation to accommodate for the additional complexities that bilingual speakers exhibit.

In the context of PRIE, this process translated to drawing speech samples from parallel bilingual varieties in the mainland United States. Furthermore, unlike Wisconsin English, which

has features that have been enregistered to some extent,[23] PRIE is practically unknown to listeners from the mainland. This methodology departed from Schuld et al. (2017) in that it observed PRIE in the context of other parallel bilingual dialects—as opposed to diachronic speech samples from the same target region—and in that the results were further complemented by a Mental Map Task to contextualize listeners' regional dialect associations for the included varieties.

The resulting theoretical and methodological framework provides a broader view of a target dialect that has not yet been enregistered, which creates a foundation for future similar exploratory work in the relationship between NEW DIALECT FORMATION and ENREGISTERMENT. The findings show that listeners perceived PRIE as an incipient variety that is both distinct and different from the selected parallel dialects. It occupied a distinct categorical slot onto itself, where it patterned more towards the English control cluster of dialects than the Spanish control cluster. Unlike the other dialects in these tasks, the results for PRIE suggest that listeners did not perceive a strong signal for PRIE, but rather a categorical difference where PRIE was not like either of the other two signal clusters, suggesting a disconnect between recognition and confidence. These results provide a baseline through which NEW DIALECT FORMATION and ENREGISTERMENT can be explored, where observing a new dialect that is perceptually salient yet categorically ambivalent suggests a lack of ENREGISTERMENT.

The findings of this study draw from the combination of two approaches in the exploration of the emergence of PRIE as a new dialect of American English. Analyses throughout this dissertation underscore the importance of language attitudes, a complex bilingual

[23] The enregisterment of Wisconsin and Upper Midwestern English at the national level is underway through media and films. Some examples of these include the film *Fargo* and, more recently, *The Manitowoc Minute*, the latter of which focuses on Wisconsin English.

context, identity, and a shared sociopolitical history as driving forces in this shift in a subset of communities in Puerto Rico. This shift represents a newly developing coexistence between Spanish and English as L1 varieties in Puerto Rico. Future studies aiming to explore NEW DIALECT FORMATION and ENREGISTERMENT should consider adopting, adapting, and building on the approaches employed in this dissertation.

6.3 Contributions

This dissertation has been innovative in at least five ways. First and foremost, it has proposed, explored, and demonstrated the existence of PRIE, a new dialect of American English in Puerto Rico spoken by younger bilingual Puerto Ricans that mirrors the context of other bilingual Englishes in the mainland United States. This finding represents the start of a fruitful and robust research agenda for PRIE, ranging from codifying its segmental and suprasegmental features from a synchronic perspective to identifying its historical origins and the dialect's potential developmental directionality from a diachronic perspective (more on this in Future Directions, §6.6). Recognizing the existence of PRIE also changes the narrative in terms of the role of English in Puerto Rico and the factors that catalyzed the formation of a new dialect in a region that had historically rejected English since its re-colonization by the United States at the end of the Spanish-American War in 1898.

Second, it has expanded on methodological approaches in PERCEPTUAL DIALECTOLOGY by combining a descriptive Mental Map Task (in the line of Preston & Robinson, 2005) and a modified set of inferential signal detection tasks (drawing inspiration and adapted from Schuld et al., 2017) to explore NEW DIALECT FORMATION as a product of categorically distinct perceptual saliency. This retooled approach is useful in expanding on perceptual (signal) recognition studies in linguistics. Observing degree of contrast between results provides a foundation in analyzing

more complex, rather than discrete, signals in context—much of which is applicable to most language use contexts. Future studies in perceptual dialectology and sociophonetics need to consider incorporating and/or adapting this approach with a focus on contrastive analyses to determine salience, recognizability, and perceptual status in the minds of listeners. These factors are critical in understanding the role of language use in context and in society.

Third, it has accounted and controlled for experiential knowledge that informs listeners performing perceptual tasks through what I term the LISTENER ACCURACY FRAMEWORK. In this framework, pre-/post-test performance and practice tasks that establish baselines, such as the pre-/post-tests in the Mental Map Task, can be used to discern listener accuracy and sensitivity for analyses. This methodology addresses individual response variance as a product of sensitivity to and expected performance in the task, complex factors that cannot be assumed to be equivalent across listeners. As the diverging results between types of listeners in Chapter 5 demonstrate, this additional step provides for a greater degree of nuance in understanding response patterns and relating individual performance to the whole. The LISTENER ACCURACY FRAMEWORK represents a new way to account for listener response variance in perceptual tasks and begins to account for performance variance. It is, likewise, my hope that other researchers conducting studies in PERCEPTUAL DIALECTOLOGY will consider adopting and building on this framework.

Fourth, it has adapted the Rootedness Metric to observe the roles that ROOTEDNESS may play in the phenomena observed and described throughout this study. This adaptation is a first pass at building on Reed's (2016 and 2018) framework to examine a bilingual and bicultural community of speakers experiencing rapid change. Unfortunately, this first venture did not meet the initial expectations for a number of possible reasons: (1) working on speakers of an incipient dialect form, (2) adapting to a multilingual setting where the target language does not carry as

much historicity as Appalachian English, (3) the limited amount of respondents in comparison to other rootedness studies, and (4) the overall structure of this RM that could have been formulated differently to better capture attachment to place as a function of language. The results do underscore the need to further explore ROOTEDNESS in PRIE and challenge the expectations that one may draw based on ROOTEDNESS and language usage in multilingual settings.

Finally, this dissertation employed an interdisciplinary approach, drawing from a broad set of methodologies and theoretical frameworks—NEW DIALECT FORMATION, PERCEPTUAL DIALECTOLOGY, SIGNAL DETECTION THEORY, BILINGUALISM, ROOTEDNESS, WORLD ENGLISHES, and LANGUAGE ATTITUDES. These methods came together to explore the emergence of a new dialect of American English in Puerto Rico that is driven by a confluence of the dynamic sociolinguistic, ecopolitical, and historical circumstances that made this phenomenon possible. While there is still much work to be done to further discuss and elaborate on PRIE, the framework in this study can prove valuable for other researchers to expand on current approaches to explore language change, and perceptions thereof, in context.

6.4 Limitations

In this study, I can identify two limitations: the caveats of only selecting a more focused listener sample with regards to their language experiences and the limited composition of the listener regional sample.

Even though this study successfully represented the target listener population—those with little to no knowledge of Spanish and with extremely limited experiences with Spanish speaking communities—the general impact of this methodological decision could have been made clearer if a relatively equal sample size of listeners with extensive experiences in the same contexts would have also been gathered. This approach would have allowed for a clearer view of

the extent to which PRIE has enregistered by highlighting the contrast between the current group of listeners and those who have had more exposure to Spanish communities in general. At the very least, and in the same train of thought, I suspect that a listener sample gathered exclusively from Puerto Rico would draw an even starker contrast with the latter, largely due to their in-group status.

It would have also been valuable to have recruited equally weighted samples from all of the dialect regions identified in the Mental Map Task to account for potential regional dialect in-group versus out-group effects. As discussed in Chapter 5, listeners are better at identifying variation within their own regions and circles. An equally weighted sample by dialect region would have also allowed to observe the degrees to which regional familiarity would have affected listeners' allocation judgments of the speaker groups in the Mental Map Task.

Some of these limitations are attributable to maintaining a more focused scope on the subject matter in this study, and despite these limitations, the study was successful in achieving its aims. Future research will, nevertheless, address these concerns directly.

6.5 Final Remarks

This study's goal was to observe the emergence of PRIE as a dialect of American English through a combination of methods in PERCEPTUAL DIALECTOLOGY and statistical analyses. The findings throughout this research confirm the existence of PRIE as a new dialect of American English that does not yet appear to have enregistered for naïve listeners in the mainland.

The emergence of PRIE is unique in its position with regards to World Englishes and dialect formation, in general. While Puerto Rico is part of the United States, it remains to be a region with an extensive colonial history, having been a colony once to Spain and then to the United States. Although it is true that the findings here point to PRIE following the patterns of a

dialect of American English, PRIE is also emerging from a complicated historical amalgamation of sociopolitical circumstances. Puerto Rico retains a long-held identity that is uniquely its own, and the emergence of PRIE is made possible by an ongoing redefinition of that identity—of what it means to be Puerto Rican—and not despite it. In this sense, PRIE is both a new dialect of American English and a product of a colonial political status at the same time. This challenging combination of sociopolitical factors makes it so that PRIE fits neither here nor there when it comes to World Englishes from the traditional post-colonial perspective.

Puerto Rico has been almost entirely viewed as an L1 Spanish/L2 English-speaking region. Even with the increasing sociocultural and ecopolitical influence of the mainland United States and the amelioration of language attitudes toward English, the findings in this dissertation begin to reshape the conversation on the role of English on the island and dispel some of the myths that come associated with a lack of knowledge about the linguistic diversity and changing linguistic landscape of Puerto Rico.

6.6 Future Directions

Given the findings in this dissertation and how quickly the linguistic context in Puerto Rico continues to change, at least in terms of the use of English, much work remains to be done to account for the large gap in research.

An avenue of research can focus on qualitatively understanding the origins and development of PRIE, particularly within the present-day context. This fieldwork would elucidate what PRIE and its usage represents to its speakers on the island. While this discussion could and should be informed by the phonological work on PRIE that remains to be done, it is another important aspect that needs further investigation in order to understand the current status of PRIE in the minds of its speakers.

Further down this avenue is also an elaboration of the findings on ROOTEDNESS in this study, where the PRIE speakers were more rooted to Puerto Rico than the sequential bilinguals living in the mainland United States. These findings were surprising in that one would expect those speakers who have a stronger and more dominant connection to Spanish would also be more rooted to the island, where Spanish is notably more prevalent. I suspect that this divergence is related to pressures to assimilate to the mainland, where the late sequential speakers have to necessarily invest more resources to improve their English, which may, in turn, sever their primary previous connections to an identity that is rooted in Spanish.

An alternative possibility to the findings on ROOTEDNESS is that the redefinition of the Puerto Rican identity away from the "language one" culture (i.e., Spanish only, as Domínguez-Rosado, 2015 observes) has a broader reaching effect in both speaker groups. From this perspective, the sequential speakers do not see a stronger embrace of English as a rejection of their identities as Puerto Ricans, and the PRIE speakers do not see an equal embrace of both languages as the same, both of which would have skewed the RM results. To address these concerns, two additional steps should be taken. First, more open-ended qualitative data on Puerto Rican speakers' views on English-Spanish language use and identity should be gathered to then observe their effects in greater detail. Second, adjustments to the questions asked in the RM for this study must be made to better reflect the former observations, focusing away from directly interpreting language usage preferences as an indexer of ROOTEDNESS, specifically in this bilingual and bicultural community of speakers where English has less historicity. Certainly, further investigation is warranted.

Regarding the speaker selection process, the methodological limitations on speaker selection were discussed in Chapter 3. Following Poplack's (1978, pp.97-102) findings on

degrees of stylistic variance in production based on gender differences, this study only included male speakers to control for these factors on variant performance as much as possible. Now, with this baseline established, future research should include female speakers as well to better understand the range of variation in production, if any, across all PRIE speakers.

The stark contrast in the use of English in younger versus older generations of Puerto Ricans also presents ideal circumstances to replicate an adapted version of Schuld et al.'s (2017) study. Younger Puerto Ricans, PRIE speakers, could be recorded and compared against the English of older Puerto Ricans and a set of distractor speakers from the mainland United States. Although, considering the findings in this study, I would expect an outcome in the contrast that would further confirm the generational rift of English use in Puerto Rico.

Finally, in terms of understanding the enregisterment of PRIE, future research can also focus on its phonology, identifying the phonological features that may be developing as future regionally identifiable characteristics. Its status as an English dialect that coexists with Spanish creates a language contact scenario with sizable potential for movement of features, such as Voice Onset Time, vowel length, vowel trajectory directions, intrinsic allophonic variation, increased use of loanwords and borrowings, the effect and use of Puerto Rican Spanglish, among many other aspects. This line of inquiry can parallel Carter, López Valdez & Sims (2020), who have begun documenting similar kinds of featural variation in Miami Cuban English; it can also help us begin to understand why listeners from the mainland United States mostly perceived the PRIE speakers as Midlands dialect speakers in this study. A robust research agenda can be built on detailing PRIE's segmental and suprasegmental features alone, which I plan to begin to investigate in a near future.

Whichever way one decides to pursue further research on PRIE, there are a myriad of possibilities to further explore, describe, and understand about PRIE as a new dialect of American English. I, for one, look forward to it.

References

Antonio Barreto, Amílcar. 2000. Speaking English in Puerto Rico: The Impact of Affluence, Education, and Return Migration. *CENTRO Journal* 12(1). 5-17.

Argyri, Efrosyni & Antonella Sorace. 2007. Crosslinguistic Influence and Language Dominance in Older Bilingual Children. *Bilingualism: Language and Cognition* 10(1). 79-99.

Audacity Team. 2015. Audacity (Version 2.1.0) [Computer program]. Retrieved from: http://audacityteam.org/

Baird, Brandon O., Marcos Rohena-Madrazo & Caroline Cating. 2018. Perceptions of Lexically Specific Phonology Switches on Spanish-Origin Loanwords in American English. *American Speech* 93(1). 79-107.

Baker, Wendy & William G. Eggington. 1999. Bilingual Creativity, Multidimensional Analysis, and World Englishes. *World Englishes* 18(3). 343-358.

Benmamoun, Elabbas, Silvina Montrul & Maria Polinski. 2013. Defining an "Ideal" Heritage Speaker: Theoretical and Methodological Challenges. *Theoretical Linguistics* 39(3-4). 259-294.

Birdsong, David, Libby M. Gertken & Matthew Amengual. 2012. Bilingual Language Profile: An Easy-to-Use Instrument to Assess Bilingualism. *Center for Open Educational Resources and Language Learning* (COERLL), University of Texas at Austin.

Bischoff, Shannon. 2017. Is English a Language Barrier to Public Higher Education in Puerto Rico? *Multilingua* 36(3). 281-311.

Blumenfeld, Marian & Margarita Kaushanskaya. 2007. The Language Experience and Proficiency Questionnaire (LEAP-Q): Assessing Language Profiles in Bilinguals and Multilinguals. *Journal of Speech and Hearing Research* 50(4). 940-967.

Calamai, Silvia. 2015. Between Linguistics and Social Psychology of Language: The Perception of Non-Native Accents. *Studi e saggi linguistici* 53(2). 289-305.

Carstens, Adelia. 2016. Translanguaging as a Vehicle for L2 Acquisition and L1 Development: Students' Perceptions. *Language Matters* 47(2). 203-222.

Carter, Phillip M., Lydda López Valdez, and Nandi Sims. 2020. New Dialect Formation Through Language Contact: Vocalic and Prosodic Development in Miami English. *American Speech* 95(2). 119-148.

Casillas, Joseph V., and Miquel Simonet. 2018. Perceptual Categorization and Bilingual Language Modes: Assessing the Double Phonemic Boundary in Early and Late Bilinguals. *Journal of Phonetics* 71. 51-64.

Chand, Vineeta. 2009. *Who Owns English? Political, Social, and Linguistic Dimensions of Urban Indian English Language Practices*. University of California, Davis dissertation.

Ciriza, Maria. 2010. *Dialect Divergence and Identity in Basque Spanish*. University of Illinois at Urbana-Champaign dissertation.

Cockcroft, Kate, Mandy Wigdorowitz & Luzanne Liversage. 2017. A Multilingual Advantage in the Components of Working Memory. *Cognition* 22(1). 15-29. Retrieved from: https://doi.org/10.1017/S1366728917000475

Cornwell, Sarah & Yasaman Rafat. 2017. English Interdental Fricative Production in Dutch Heritage Speakers Living in Canada. *Ilha do Desterro* 70(3). 95-115.

Domínguez-Rosado, Brenda. 2015. *The Unlinking of Language and Puerto Rican Identity: New Trends in Sight*. Cambridge, UK: Cambridge Scholars Publishing.

Duany, Jorge. 2001. Irse pa' fuera': Los modos de vida móviles de los migrantes circulares entre Puerto Rico y los Estados Unidos. *Temas: Cultura, Ideología, Sociedad* 26. 39-49.

Fabiano-Smith, Leah, Rebecca Shuriff, Jessica A. Barlow & Brian A. Goldstein. 2014. Dialect Density in Bilingual Puerto Rican Spanish-English Speaking Children. *Linguistic Approaches to Bilingualism* 4(1). 34-60.

Falcón, Sandra María. 2004. *Attitudes toward Languages in Contact: Spanish, English, and Intellectuals in Puerto Rico*. Georgetown University dissertation.

Fought, Carmen. 2006. *Language and Ethnicity: Key Topics in Sociolinguistics*. New York: Cambridge University Press.

Gonzales, Kalim, Krista Byers-Heinlein & Andrew J. Lotto. 2019. How Bilinguals Perceive Speech Depends on Which Language They Think They're Hearing. *Cognition* 182. 318-330. Retrieved from: https://doi.org/10.1016/j.cognition.2018.08.021

González-Martínez, Juan. 2008. Metodología para el estudio de las actitudes lingüísticas. Conference Presentation at the *Simposio Internacional de la Sociedad Española de Lingüística*, University of Navarra.

González Rivera, Melvin & Luis A. Ortiz López. 2018. El español y el inglés en Puerto Rico: una polémica de más de un siglo. *Centro Journal* 30(1). 106-131.

Haswell, Christopher. 2013. A Global Model of English: How New Modeling Can Improve the Appreciation of English Usage in the Asia Pacific Region. *Asia Pacific World* 4(2). 122-137.

Henriksen, Nicholas. 2015. Acoustic Analysis of the Rhotic Contrast in Chicagoland Spanish: An Intergenerational Study. *Linguistic Approaches to Bilingualism* 5(3). 285-321.

Hickling, Robert. 2000. Simultaneous Versus Early Sequential Bilingualism. *Zassaku* 8. 39-48. Retrieved from: http://ebookcentral.proquest.com

Ingle, Jennifer K., Richard Wright & Alicia Wassink. 2005. Pacific Northwest Vowels: A Seattle Neighborhood Dialect Study. *The Journal of the Acoustic Society of America* 117 (4). 2459.

International Phonetic Association. 1999. *Handbook of the International Phonetic Association: A Guide to the Use of the International Phonetic Alphabet*. Cambridge, UK: Cambridge University Press.

Janson, Tore. 2004. *Natural History of Latin: The Story of the World's Most Successful Language*. New York: Oxford University Press. Retrieved from: http://ebookcentral.proquest.com

Jesney, Karen. 2004. *The Use of Global Foreign Accent Rating in Studies of L2 Acquisition*. Calgary, AB: Language Research Centre, University of Calgary report.

Kachru, Braj B. 1990. World Englishes and Applied Linguistics. *World Englishes* 9(1). 3-20.

Kerswill, Paul. 2005. Migration and Language. In Klaus Mattheier, Ulrich Ammon & Peter Trudgill (eds.), *Sociolinguistics/Soziolinguistik: An International Handbook of the Science of Language and Society*, 2(2). 1-27. Berlin: De Gruyter.

Knight, Rachael-Anne. 2011. Assessing the Temporal Reliability of Rhythm Metrics. *Journal of the International Phonetic Association* 41(3). 271-281.

Kondo, Mariko, Hajime Tsubaki & Yoshinori Sagisaka. 2015. Segmental Variation of Japanese Speakers' English: Analysis of "the North Wind and the Sun" in AESOP Corpus. *Journal of the Phonetic Society of Japan* 19(1). 3-17.

Labov, William. 1972. *Sociolinguistic Patterns*. Philadelphia, Pennsylvania: University of Pennsylvania Press, Inc.

Labov, William, Sharon Ash & Charles Boberg. 2006. *The Atlas of North American English*. New York: Mouton de Gruyter.

Litty, Samantha M. 2017. *We Talk German Now Yet: The Sociolinguistic Development of Voice Onset Time and Final Obstruent Neutralization in Wisconsin German and English Varieties, 1863-2013*. University of Wisconsin – Madison dissertation.

Mack, Sara. 2010. A Sociophonetic Analysis of Perception of Sexual Orientation in Puerto Rican Spanish. *Journal of Language and Politics* 1(1). 41–63.

Mayer, Mercer. 1969. *Frog, Where Are You?* New York: Dial Books for Young Readers.

McGuire, Grant. 2010. A Brief Primer on Experimental Designs for Speech Perception. UCSC Linguistics Research Center laboratory report. Retrieved from: https://people.ucsc.edu/~gmcguir1/LabReport/McGuire1LRC.pdf

Montalvo, Ana J. & Lynda Laughlin. 2017. An Island in Crisis? A Statistical Portrait of Recent Puerto Rican Migration and Socioeconomic Outcomes between 2005 and 2015. Population Association of America. U.S. Census Bureau. Retrieved from: https://www.census.gov/library/working-papers/2017/demo/SEHSD-WP2017-21.html

Montgomery, Christopher & Philipp Stoeckle. 2013. Geographic Information Systems and Perceptual Dialectology: A Method for Processing Draw-A-Map Data. *Journal of Linguistic Geography* 1(1). 52-85.

Montrul, Silvina. 2014. Ultimate Attainment in Spanish L2 Acquisition. In Kimberly L. Geeslin, (ed.), *The Handbook of Spanish Second Language Acquisition*, 353-368. Malden, MA: John Wiley & Sons, Inc. Retrieved from: http://ebookcentral.proquest.com

Moreno Fernández, Francisco. 2009. Actitudes lingüísticas. Capítulo IX en *Principios de sociolingüística y sociología del lenguaje*. Barcelona, Spain: Book Print Digital.

Mukherjee, Joybrato. 2010. The Development of the English Language in India. In Andy Kirkpatrick (ed.), *The Routledge Handbook of World Englishes*. 167-180. New York: Routledge.

Nash, Manning. 1996. The Core Elements of Ethnicity. In John Hutchinson & Anthony D. Smith (eds.), *Ethnicity*, 24-28. New York: Oxford University Press.

Nash, Rose. 1968. Intonational Interference in the Speech of Puerto Rican Bilinguals, an Instrumental Study Based on Oral Readings of a Juan Bobo Story. San Juan, Puerto Rico: Inter American University.

Niedzielski, Nancy. 1999. The Effect of Information on the Perception of Sociolinguistic Variables. *Journal of Language and Social Psychology* 18(1). 62–85.

Pallier, Cristophe. 2002. Computing Discriminability and Bias with the R Software. Retrieved from: http://www.pallier.org/

Pérez Casas, Marisol. 2008. *Codeswitching and Identity among Island Puerto Rican Bilinguals*. Washington, D.C.: Georgetown University dissertation.

Pew Research Center. 2014. Puerto Rican Population Declines on Island, Grows on U.S. Mainland. Retrieved from: https://www.pewhispanic.org/2014/08/11/puerto-rican-population-declines-on-island-grows-on-u-s-mainland/ (Accessed 6 April, 2019)

Pew Research Center. 2016. Hispanic Population Growth and Dispersion Across U.S. Counties, 1980-2014. Retrieved from: https://www.pewresearch.org/hispanic/interactives/hispanic-population-by-county/ (Accessed 24 September, 2020)

Poplack, Shana. 1978. Dialect Acquisition among Puerto Rican Bilinguals. *Language in Society* 7(1). 89-103.

Pousada, Alicia. 2010. English-Speaking Enclaves in Puerto Rico. Proceedings of the *College English Association Conference*. Rio Piedras, Puerto Rico: University of Puerto Rico.

Preston, Dennis R. 1996. Where the Worst English is Spoken. In Edgar W. Schneider (ed.), *Focus on the USA*. Philadelphia, PA: John Benjamins Publishing Company.

Preston, Dennis R. 2003. Where Are the Dialects of American English at Anyhow? *American Speech* 78(3). 235–254.

Preston, Dennis R. 2013. Language with an Attitude. In Jack K. Chambers & Natalie Schilling-Estes (eds.), *The Handbook of Language Variation and Change* 2, 157-182. Malden, MA: John Wiley & Sons, Inc.

Preston, Dennis R. 2018. Perceptual Dialectology. In Charles Boberg, John Nerbonne & Dominic Watt (eds.), *The Handbook of Dialectology* 1, 177-203. Malden, MA: John Wiley & Sons, Inc.

Preston, Dennis R. & Gregory C. Robinson. 2005. Dialect Perception and Attitudes to Variation. In Martin J. Ball (ed.), *Clinical Sociolinguistics*,133-149. Malden, MA: Blackwell Publishing.

Purnell, Thomas, Eric Raimy & Joseph Salmons. 2009. Defining Dialect, Perceiving Dialect, and New Dialect Formation: Sarah Palin's Speech. *Journal of English Linguistics* 37(4). 331-355.

Purnell, Thomas, William Idsardi & John Baugh. 1999. Perceptual Phonetic Experiments on American Dialect Identification. *Journal of Language and Social Psychology* 18(1). 10-30.

Ramírez, Marta & Miquel Simonet. 2018. Language Dominance and the Perception of the Majorcan Catalan /ʎ/−/ʒ/ Contrast: Asymmetrical Phonological Representations. *International Journal of Bilingualism* 22(6). 638-352.

Reed, Paul E. 2016. *Sounding Appalachian: /aɪ/ Monophthongization, Rising Pitch Accents, and Rootedness*. University of South Carolina dissertation.

Reed, Paul E. 2018. The Importance of Appalachian Identity: A Case Study in Rootedness. *American Speech* 93(3-4). 409-424.

Rivera, Magaly. 2019. Puerto Rico's History. Retrieved from: http://www.topuertorico.org/history5.shtml (Accessed 1 April, 2019)

Rodríguez, Clara E. 2000. Puerto Ricans: Immigrants and Migrants: A Historical Perspective. Retrieved from: https://americansall.org/sites/default/files/879-PUERTO%20RICAN%20AM%20C_T%20MedLoRes_0.pdf (Accessed 11 July, 2019)

Schmid, Monika S. & Holger Hopp. 2014. Comparing Foreign Accent in L1 Attrition and L2 Acquisition: Range and Rater Effects. *Language Testing* 31(3). 367-388.

Schneider, Edgar W. 2003. The Dynamics of New Englishes: From Identity Construction to Dialect Birth. *Language* 79(2). 233-281.

Schuld, Danielle, Joseph Salmons, Thomas Purnell & Eric Raimy. 2017. "Subliminal Accent": Reactions to the Rise of Wisconsin English. *Journal of Linguistic Geography* 4. 15-30.

Smallwood Ramos, Jacelyn. 2020. *The Current State of Specialized Public Bilingual Schools in Puerto Rico: Teacher and Administrator Perspectives*. University of Puerto Rico, Rio Piedras dissertation.

Stanislaw, Harold & Natasha Todorov. 1999. Calculation of Signal Detection Theory Measures. *Behavior Research Methods, Instruments, & Computers* 31(1). 137-149.

Trudgill, Peter. 2008. Colonial Dialect Contact in the History of European Languages: On the Irrelevance of Identity to New-Dialect Formation. *Language in Society* 37(2). 241-254.

Trudgill, Peter, Elizabeth Gordon, Gillian Lewis & Margaret MacLagan. 2000. Dialect Formation and the Genesis of New Zealand English. *Journal of Linguistics* 36(2). 299-318. UK: Cambridge University Press.

U.S. Census Bureau. 2010. 2010 Census: Puerto Rico Profile. Retrieved from: https://www2.census.gov/geo/maps/dc10_thematic/2010_Profile/2010_Profile_Map_Puerto_Rico.pdf (Accessed 6 April, 2019)

U.S. Census Bureau. 2019a. Language Spoken at Home: 2013-2017 (Report No. S1601). Retrieved from: https://factfinder.census.gov/faces/tableservices/jsf/pages/productview.xhtml?src=CF (Accessed 6 April, 2019)

U.S. Census Bureau. 2019b. Quick Facts about Puerto Rico: 2013-2017 (V2018). Retrieved from: https://www.census.gov/quickfacts/fact/map/pr/PST120218 (Accessed 6 April, 2019)

Vicente Vélez, Nancy Vanessa. 2000. *The Effect of Political Affiliation on Motivation to Learn English in Puerto Rico*. University of Puerto Rico – Mayagüez Campus thesis.

Whelan, Richard. 2002. The American Spelling Reform Movement. *The Language Quarterly* 27(4). 2-7.

Wolfram, Walt & Natalie Schilling-Estes. 2006. *American English*. Malden, MA: Blackwell Publishing.

Wong, Galston & Bee Chin Ng. 2018. Moral Judgement in Early Bilinguals: Language Dominance Influences Responses to Moral Dilemmas. *Frontiers in Psychology* 9. 1-10. Retrieved from: https://doi.org/10.3389/fpsyg.2018.01070

Zentella, Ana Celia. 1997. *Growing Up Bilingual*. Malden, MA: Blackwell Publishers.

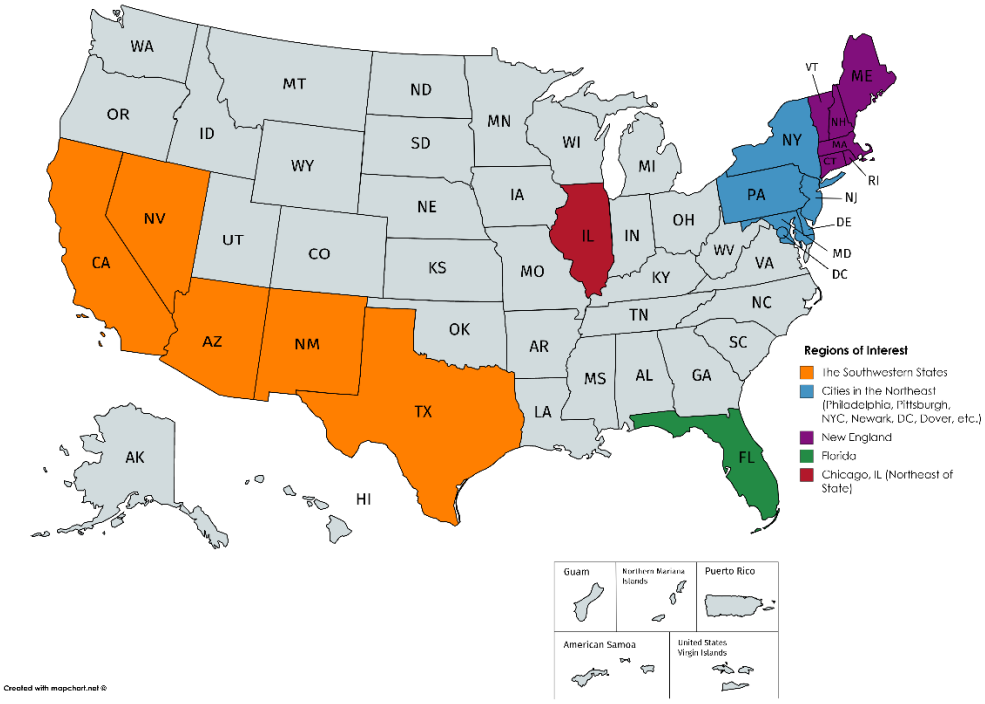
Zentella, Ana Celia. 2003. Returned Migration, Language, and Identity: Puerto Rican Bilinguals in Dos Worlds/Two Mundos. In Matthew C. Gutmann, Félix V. Matos Rodríguez, Lynn Stephen & Patricia Zavella (eds.), *Perspectives on Las Americas: A Reader in Culture, History & Representation*, 245-258. Malden, MA: Blackwell Publishers.

Appendix A: Demographics Questionnaire

In this part of the activity, you will be asked to respond to demographic questions about yourself. Please recall that your participation is confidential, and any personal information that may emerge from responses will be omitted from the data. The answers to these questions assist in contextualizing your responses, but they do not identify you directly.

1. What is your age?
    a. [Drop Down Tab Response]
2. What is your highest level of education attained?
    a. High School or Less
    b. Some College
    c. College Degree
    d. Professional or Graduate School
3. Are you a native speaker of? (Click all that apply)
    a. English
    b. Spanish
    c. Other (Please Specify): _____
4. [IF 3=A] Do you consider yourself a speaker of American English?
    a. Yes
    b. No
5. In the last ten years, how much time have you spent interacting in English with native Spanish speakers?
    a. None or next to none of the time
    b. Less than half of the time
    c. About half of the time
    d. Most, if not all, of the time
6. Did you live in the U.S. or its territories for most of the time between the ages of 4 and 12?
    a. Yes
    b. No
    c. Other (Please Explain): _____

7. [IF 6=A] Do you or have you lived in the following areas of the United States? [The Southwest, Cities in the Northeast, New England, Florida, Chicago]



   a. Yes
   b. No
8. [IF 6=A AND 7=B] In what state or territory do you currently live?
   a. [Drop Down Tab Response]
9. [IF 6=A AND 7=A] Which one(s)? (Click all that apply)
   a. [Multiple-Select Options Listing the Previous Areas]
10. Do you consider yourself of Hispanic, Latino, or Spanish origin?
   a. No
   b. Yes, Mexican, Mexican-American, Chicano
   c. Yes, Puerto Rican
   d. Yes, Cuban
   e. Yes, another Hispanic, Latino, or Spanish origin–for example, Argentinian, Colombian, Dominican, Nicaraguan, Salvadorian, so on.
11. Have you lived in Puerto Rico?
   a. Yes
   b. No
12. Have you had sustained contact with Puerto Rican communities?
   a. Yes
   b. No
13. [IF 12=Yes] How long (approximately)?
   [Year/Month Drop-down Tab Format]

Appendix B: English-Spanish BLP Questionnaire (Birdsong, Gertken & Amengual, 2012)

II. Language history

In this section, we would like you to answer some factual questions about your language history by placing a check in the appropriate box.

1. At what age did you **start learning** the following languages?

English

☐ Since birth ☐ 1 ☐ 2 ☐ 3 ☐ 4 ☐ 5 ☐ 6 ☐ 7 ☐ 8 ☐ 9 ☐ 10 ☐ 11 ☐ 12 ☐ 13 ☐ 14 ☐ 15 ☐ 16 ☐ 17 ☐ 18 ☐ 19 ☐ 20+

Spanish

☐ Since birth ☐ 1 ☐ 2 ☐ 3 ☐ 4 ☐ 5 ☐ 6 ☐ 7 ☐ 8 ☐ 9 ☐ 10 ☐ 11 ☐ 12 ☐ 13 ☐ 14 ☐ 15 ☐ 16 ☐ 17 ☐ 18 ☐ 19 ☐ 20+

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

English

☐ As early as I can remember ☐ 1 ☐ 2 ☐ 3 ☐ 4 ☐ 5 ☐ 6 ☐ 7 ☐ 8 ☐ 9 ☐ 10 ☐ 11 ☐ 12 ☐ 13 ☐ 14 ☐ 15 ☐ 16 ☐ 17 ☐ 18 ☐ 19 ☐ 20+ ☐ not yet

Spanish

☐ As early as I can remember ☐ 1 ☐ 2 ☐ 3 ☐ 4 ☐ 5 ☐ 6 ☐ 7 ☐ 8 ☐ 9 ☐ 10 ☐ 11 ☐ 12 ☐ 13 ☐ 14 ☐ 15 ☐ 16 ☐ 17 ☐ 18 ☐ 19 ☐ 20+ ☐ not yet

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

English

☐ 0 ☐ 1 ☐ 2 ☐ 3 ☐ 4 ☐ 5 ☐ 6 ☐ 7 ☐ 8 ☐ 9 ☐ 10 ☐ 11 ☐ 12 ☐ 13 ☐ 14 ☐ 15 ☐ 16 ☐ 17 ☐ 18 ☐ 19 ☐ 20+

Spanish

☐ 0 ☐ 1 ☐ 2 ☐ 3 ☐ 4 ☐ 5 ☐ 6 ☐ 7 ☐ 8 ☐ 9 ☐ 10 ☐ 11 ☐ 12 ☐ 13 ☐ 14 ☐ 15 ☐ 16 ☐ 17 ☐ 18 ☐ 19 ☐ 20+

4. How many years have you spent in a **country/region** where the following languages are spoken?

English

☐ 0 ☐ 1 ☐ 2 ☐ 3 ☐ 4 ☐ 5 ☐ 6 ☐ 7 ☐ 8 ☐ 9 ☐ 10 ☐ 11 ☐ 12 ☐ 13 ☐ 14 ☐ 15 ☐ 16 ☐ 17 ☐ 18 ☐ 19 ☐ 20+

Spanish

☐ 0 ☐ 1 ☐ 2 ☐ 3 ☐ 4 ☐ 5 ☐ 6 ☐ 7 ☐ 8 ☐ 9 ☐ 10 ☐ 11 ☐ 12 ☐ 13 ☐ 14 ☐ 15 ☐ 16 ☐ 17 ☐ 18 ☐ 19 ☐ 20+

5. How many years have you spent in a **family** where the following languages are spoken?

English

☐ 0 ☐ 1 ☐ 2 ☐ 3 ☐ 4 ☐ 5 ☐ 6 ☐ 7 ☐ 8 ☐ 9 ☐ 10 ☐ 11 ☐ 12 ☐ 13 ☐ 14 ☐ 15 ☐ 16 ☐ 17 ☐ 18 ☐ 19 ☐ 20+

Spanish

☐ 0 ☐ 1 ☐ 2 ☐ 3 ☐ 4 ☐ 5 ☐ 6 ☐ 7 ☐ 8 ☐ 9 ☐ 10 ☐ 11 ☐ 12 ☐ 13 ☐ 14 ☐ 15 ☐ 16 ☐ 17 ☐ 18 ☐ 19 ☐ 20+

6. How many years have you spent in a **work environment** where the following languages are spoken?

English

☐ 0 ☐ 1 ☐ 2 ☐ 3 ☐ 4 ☐ 5 ☐ 6 ☐ 7 ☐ 8 ☐ 9 ☐ 10 ☐ 11 ☐ 12 ☐ 13 ☐ 14 ☐ 15 ☐ 16 ☐ 17 ☐ 18 ☐ 19 ☐ 20+

Spanish

☐ 0 ☐ 1 ☐ 2 ☐ 3 ☐ 4 ☐ 5 ☐ 6 ☐ 7 ☐ 8 ☐ 9 ☐ 10 ☐ 11 ☐ 12 ☐ 13 ☐ 14 ☐ 15 ☐ 16 ☐ 17 ☐ 18 ☐ 19 ☐ 20+

III. Language use

In this section, we would like you to answer some questions about your language use by placing a check in the appropriate box. Total use fo
all languages in a given question should equal 100%.

7. In an average week, what percentage of the time do you use the following languages **with friends**?

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

8. In an average week, what percentage of the time do you use the following languages **with family**?

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

9. In an average week, what percentage of the time do you use the following languages **at school/work**?

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

10. When you talk to yourself, how often do you **talk to yourself** in the following languages?

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

11. When you count, how often do you **count** in the following languages?

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

IV. Language proficiency

In this section, we would like you to rate your language proficiency by giving marks from 0 to 6.

	0=not well at all						*6=very well*
12. a. How well do you speak **English**?	☐0	☐1	☐2	☐3	☐4	☐5	☐6
b. How well do you speak **Spanish**?	☐0	☐1	☐2	☐3	☐4	☐5	☐6
13. a. How well do you understand **English**?	☐0	☐1	☐2	☐3	☐4	☐5	☐6
b. How well do you understand **Spanish**?	☐0	☐1	☐2	☐3	☐4	☐5	☐6
14. a. How well do you read **English**?	☐0	☐1	☐2	☐3	☐4	☐5	☐6
b. How well do you read **Spanish**?	☐0	☐1	☐2	☐3	☐4	☐5	☐6
15. a. How well do you write **English**?	☐0	☐1	☐2	☐3	☐4	☐5	☐6
b. How well do you write **Spanish**?	☐0	☐1	☐2	☐3	☐4	☐5	☐6

V. Language attitudes

In this section, we would like you to respond to statements about language attitudes by giving marks from 0-6.

	0=disagree						*6=agree*
16. a. I feel like myself when I speak **English**.	☐0	☐1	☐2	☐3	☐4	☐5	☐6
b. I feel like myself when I speak **Spanish**.	☐0	☐1	☐2	☐3	☐4	☐5	☐6
17. a. I identify with an **English-speaking** culture.	☐0	☐1	☐2	☐3	☐4	☐5	☐6
b. I identify with a **Spanish-speaking** culture.	☐0	☐1	☐2	☐3	☐4	☐5	☐6
18. a. It is important to me to use (or eventually use) **English** like a native speaker.	☐0	☐1	☐2	☐3	☐4	☐5	☐6
b. It is important to me to use (or eventually use) **Spanish** like a native speaker.	☐0	☐1	☐2	☐3	☐4	☐5	☐6
19. a. I want others to think I am a native speaker of **English**.	☐0	☐1	☐2	☐3	☐4	☐5	☐6
b. I want others to think I am a native speaker of **Spanish**.	☐0	☐1	☐2	☐3	☐4	☐5	☐6

Appendix C: LEAP-Q Survey for Two Languages (Blumenfeld & Kaushanskaya, 2007)

Northwestern Bilingualism & Psycholinguistics Research Laboratory

Please cite Marian, Blumenfeld, & Kaushanskaya (2007). The Language Experience and Proficiency Questionnaire (LEAP-Q): Assessing language profiles in bilinguals and multilinguals. *Journal of Speech Language and Hearing Research, 50* (4), 940-967.

Language Experience and Proficiency Questionnaire (LEAP-Q)

Last Name	N/A	First Name	N/A	Today's Date	
Age		Date of Birth	N/A	Male ☐	Female ☐

(1) Please list all the languages you know **in order of dominance**:

1	2	3	4	5

(2) Please list all the languages you know **in order of acquisition** (your native language first):

1	2	3	4	5

(3) Please list what percentage of the time you are *currently* and *on average* exposed to each language.
(*Your percentages should add up to 100%*):

List language here:					
List percentage here:					

(4) When choosing to read a text available in all your languages, in what percentage of cases would you choose to read it in each of your languages? Assume that the original was written in another language, which is unknown to you.
(*Your percentages should add up to 100%*):

List language here					
List percentage here:					

(5) When choosing a language to speak with a person who is equally fluent in all your languages, what percentage of time would you choose to speak each language? Please report percent of total time.
(*Your percentages should add up to 100%*):

List language here					
List percentage here:					

(6) Please name the cultures with which you identify. On a scale from zero to ten, please rate the extent to which you identify with each culture. (Examples of possible cultures include US-American, Chinese, Jewish-Orthodox, etc):

List cultures here					
	(click here for scale	(click here for scale	(click here for scale	(click here for scale	(click here for scale

(7) How many years of formal education do you have?
Please check your highest education level (or the approximate US equivalent to a degree obtained in another country):

☐ Less than High School ☐ Some College ☐ Masters
☐ High School ☐ College ☐ Ph.D./M.D./J.D.
☐ Professional Training ☐ Some Graduate School ☐ Other:

(8) Date of immigration to the USA, if applicable _____
If you have ever immigrated to another country, please provide name of country and date of immigration here.

(9) Have you ever had a vision problem ☐, hearing impairment ☐, language disability ☐, or learning disability ☐ ? (Check all applicable). If yes, please explain (including any corrections):

Language:

This is my (please select from pull-down menu) language.

All questions below refer to your knowledge of .

(1) Age when you…:

began acquiring :	*became fluent* in :	*began reading* in :	*became fluent reading* in :

(2) Please list the number of years and months you spent in each language environment:

	Years	Months
A country where is spoken		
A family where is spoken		
A school and/or working environment where is spoken		

(3) On a scale from zero to ten, please select your *level of **proficiency*** in speaking, understanding, and reading from the scroll-down menus:

Speaking	(click here for scale	Understanding spoken language	(click here for scale)	Reading	(click here for scale

(4) On a scale from zero to ten, please select how much the following factors contributed to you learning :

Interacting with friends	(click here for pull-down scale)	Language tapes/self instruction	(click here for pull-down scale
Interacting with family	(click here for pull-down scale)	Watching TV	(click here for pull-down scale
Reading	(click here for pull-down scale)	Listening to the radio	(click here for pull-down scale

(5) Please rate to what extent you are currently exposed to in the following contexts:

Interacting with friends	(click here for pull-down scale)	Listening to radio/music	(click here for pull-down scale
Interacting with family	(click here for pull-down scale)	Reading	(click here for pull-down scale
Watching TV	(click here for pull-down scale)	Language-lab/self-instruction	(click here for pull-down scale

(6) In your perception, how much of a foreign accent do you have in ?

(click here for pull-down scale)

(7) Please rate how frequently others identify you as a non-native speaker based on your accent in :

(click here for pull-down scale)

Language:

This is my (please select from pull-down menu) language.

All questions below refer to your knowledge of .

(1) Age when you…:

began acquiring :	*became fluent* in :	*began reading* in :	*became fluent reading* in :

(2) Please list the number of years and months you spent in each language environment:

	Years	Months
A country where is spoken		
A family where is spoken		
A school and/or working environment where is spoken		

(3) On a scale from zero to ten, please select your *level of **proficiency*** in speaking, understanding, and reading from the scroll-down menus:

Speaking	(click here for scale	Understanding spoken language	(click here for scale)	Reading	(click here for scale

(4) On a scale from zero to ten, please select how much the following factors contributed to you learning :

Interacting with friends	(click here for pull-down scale)	Language tapes/self instruction	(click here for pull-down scale
Interacting with family	(click here for pull-down scale)	Watching TV	(click here for pull-down scale
Reading	(click here for pull-down scale)	Listening to the radio	(click here for pull-down scale

(5) Please rate to what extent you are currently exposed to in the following contexts:

Interacting with friends	(click here for pull-down scale)	Listening to radio/music	(click here for pull-down scale
Interacting with family	(click here for pull-down scale)	Reading	(click here for pull-down scale
Watching TV	(click here for pull-down scale)	Language-lab/self-instruction	(click here for pull-down scale

(6) In your perception, how much of a foreign accent do you have in ?

(click here for pull-down scale)

(7) Please rate how frequently others identify you as a non-native speaker based on your accent in :

(click here for pull-down scale)

Appendix D: BLP Results for All Speakers

ID	II. History (weight= .454)		III. Use (weight=1.09)		IV. Proficiency (weight=2.27)		V. Attitudes (weight= 2.27)		English	Spanish	Dominance
	English	Spanish	English	Spanish	English	Spanish	English	Spanish			
CB1	51.302	46.762	41.42	13.08	54.48	34.05	54.48	38.59	201.68	132.482	69.2
CB2	44.492	45.4	31.61	22.89	52.21	49.94	54.48	54.48	182.79	172.71	10.082
CB3	49.94	49.486	39.24	15.26	54.48	43.13	54.48	54.48	198.14	162.356	35.784
CB4	46.762	40.86	47.96	6.54	54.48	40.86	54.48	47.67	203.68	135.93	67.752
CB5	32.688	34.504	35.97	18.53	54.48	45.4	47.67	52.21	170.81	150.644	20.164
CB6	43.13	39.498	30.52	23.98	54.48	54.48	52.21	47.67	180.34	165.628	14.712
CH1	42.222	42.222	27.25	27.25	54.48	54.48	27.24	54.48	151.19	178.432	-27.24
CH2	42.676	28.602	38.15	16.35	52.21	45.4	47.67	29.51	180.71	119.862	60.844
CH3	39.498	39.952	30.52	23.98	54.48	52.21	31.78	54.48	156.28	170.622	-14.344
CH4	44.946	44.946	35.97	18.53	54.48	43.13	52.21	45.4	187.61	152.006	35.6
CH5	49.486	49.486	32.7	21.8	52.21	47.67	43.13	52.21	177.53	171.166	6.36
CH6	35.866	38.136	41.42	13.08	54.48	34.05	54.48	49.94	186.25	135.206	51.04
EC1	51.756	0	54.5	0	54.48	0	54.48	0	215.22	0	215.216
EC2	51.302	0	54.5	0	54.48	22.7	54.48	2.27	214.76	0	214.762
EC3	51.302	0	54.5	0	54.48	0	54.48	0	214.76	0	214.762
NY1	52.664	39.498	35.97	17.44	54.48	36.32	27.24	27.24	170.35	120.498	49.856
NY2	46.762	37.682	49.05	5.45	54.48	36.32	47.67	52.21	197.96	131.662	66.3
NY3	47.67	30.872	51.23	3.27	54.48	31.78	40.86	31.78	194.24	97.702	96.538
NY4	42.222	43.13	21.8	32.7	47.67	36.32	43.13	43.13	154.82	155.28	-0.458
NY5	49.032	19.522	35.97	18.53	54.48	38.59	54.48	49.94	193.96	126.582	67.38
NY6	45.854	44.946	41.42	13.08	54.48	43.13	49.94	54.48	191.69	155.636	36.058
PR1	28.148	42.676	27.25	27.25	49.94	52.21	49.94	54.48	155.28	176.616	-21.338
PR2	32.688	44.492	35.97	18.53	49.94	49.94	40.86	54.48	159.46	167.442	-7.984
PR3	45.854	49.94	33.79	20.71	54.48	54.48	52.21	54.48	186.33	179.61	6.724
PR4	35.866	42.676	20.71	33.79	54.48	54.48	47.67	54.48	158.73	185.426	-26.7
PR5	27.694	41.768	37.06	17.44	52.21	34.05	43.13	52.21	160.09	145.468	14.626
PR6	40.406	40.86	20.71	33.79	54.48	52.21	52.21	54.48	167.81	181.34	-13.534
SCMX1	25.424	46.308	21.8	30.52	54.48	54.48	38.59	54.48	140.29	185.788	-45.494
SCMX2	25.424	35.412	15.26	39.24	38.59	52.21	40.86	52.21	120.13	179.072	-58.938
SCMX3	20.43	54.48	16.35	38.15	40.86	54.48	18.16	54.48	95.8	201.59	-105.79
SCPR1	16.798	49.032	15.26	39.24	52.21	54.48	20.43	54.48	104.7	197.232	-92.534
SCPR2	33.596	52.664	19.62	34.88	47.67	54.48	40.86	54.48	141.75	196.504	-54.758
SCPR3	22.7	51.302	17.44	37.06	45.4	54.48	36.32	54.48	121.86	197.322	-75.462

Appendix E: Full "The North Wind and the Sun" Reading

The North Wind and the Sun were disputing which was the stronger, when a traveler came along wrapped in a warm cloak.

They agreed that the one who first succeeded in making the traveler take his cloak off should be considered stronger than the other.

Then the North Wind blew as hard as he could, but the more he blew the more closely did the traveler fold his cloak around him; and at last the North Wind gave up the attempt.

Then the Sun shined out warmly, and immediately the traveler took off his cloak. And so the North Wind was obliged to confess that the Sun was the stronger of the two.

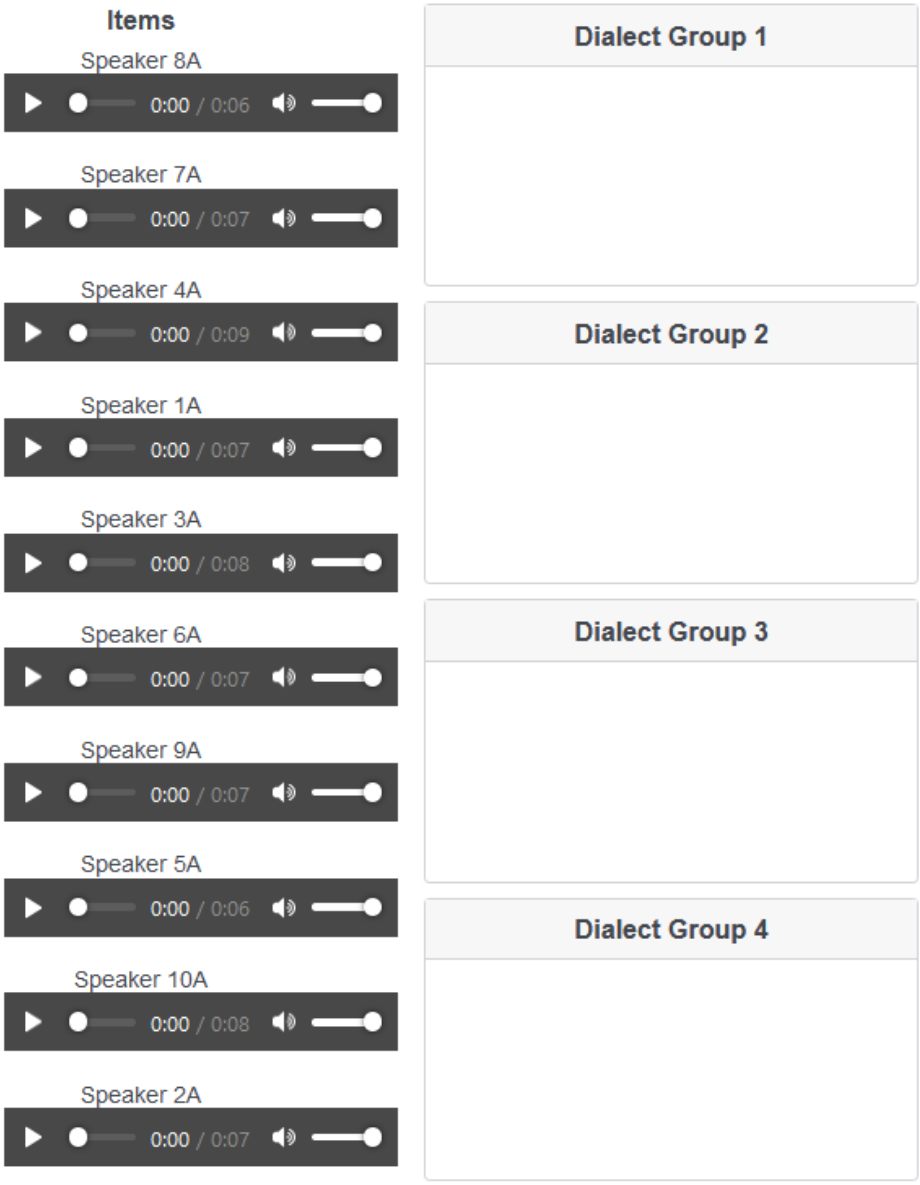
Appendix F: Phase 2 – Matched-Guise Box Task

The goal of this task is for you to listen to the clips below and group the speakers into dialect groups based on how you believe that they pair up. To group them together, you can drag and drop any number of speakers into a dialect group box. There are 10 clips in total, each corresponding to one unique speaker.

First, listen to all the clips, keeping note of speakers whose accents sound similar (who you would like to group together into the same box). You can listen to the clips as many times as necessary.

Once you think a speaker belongs in the same *Dialect Group* box (i.e., it seems that they speak the same dialect), you can drag and drop those speakers into the same *Dialect Group* box, repeating this process until all 10 of the speakers have been put into boxes.

You may use **any number** of the provided boxes; **you may not need all of them**.



Appendix G: Rootedness Metric Questionnaires

For Puerto Ricans in the Mainland United States

1. Are there any circumstances in which you might see yourself moving back to Puerto Rico? ……. Yes No
If yes, what kinds of circumstances might lead you to that decision? _____
_____
If you would be willing to live somewhere else, could you see yourself living in Puerto Rican communities in the States? ……..…… Yes No
Why or why not? _____
_____

2. How often do you visit Puerto Rico? _____

3. When you speak to another native Spanish speaker, where do you say you're from? ____

4. If you traveled far away to some other place in the U.S. and met someone who asked where you were from, what would you tell them? _____

For Puerto Ricans Living in Puerto Rico

1. Are there any circumstances in which you might see yourself moving away from Puerto Rico? ……. Yes No
If yes, what kinds of circumstances might lead you to that decision? _____
_____
If you would be willing to live somewhere else, could you see yourself living in Puerto Rican communities in the States? ……..…… Yes No
Why or why not? _____
How often would you want to visit if you left? _____

2. How often do you visit the mainland United States? _____

3. If you visit a state with a large Spanish-speaking population, where do you say you're from?
_____

4. If you traveled to some other place in the U.S. and met someone who asked where you were from, what would you tell them? _____

For all Puerto Ricans

5. How many friends/family members do you have living in Puerto Rico? _____
How about in the states? _____

6. Do you follow news from Puerto Rico? ……..……Yes No
If yes, from what sites or sources? _____

7. Rank the following (1-7) in the order that you most identify with:

Puerto Rico ___	The United States ___	La Metro ___
Your State ___ (___)	Latin America ___	
The Caribbean ___	La Isla ___	

8. To what degree do you follow sports? . . . . . A lot | Somewhat | Not at all
If you do, do you root for Puerto Rican or U.S. teams? Why? _____

9. What kind of popular music do you follow? _____
How about from *Puerto Rico/the U.S.* (depending on initial answer)? _____

10. Do you identify more with English, Spanish, or Both? . . . . . English | Spanish | Both
Why? _____
What is your opinion on each of these languages? _____

11. On a scale of 1 to 5, to what degree you would say your identity is tied to Puerto Rico? (1= Not at all tied, 3= Somewhat tied, 5= Closely tied) _____

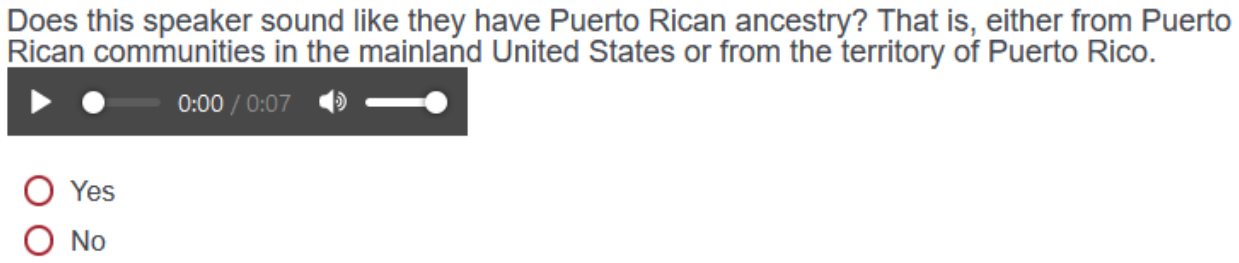
12. On a scale of 1 to 5, to what degree you would say your identity is tied to the United States? (1= Not at all tied, 3= Somewhat tied 5= Closely tied) _____

Open-Ended Identity Questions

1. Would you say you identify with Puerto Rico? Your hometown? Why?
2. Is there another place that you identify with? Why?
3. What makes it so special?

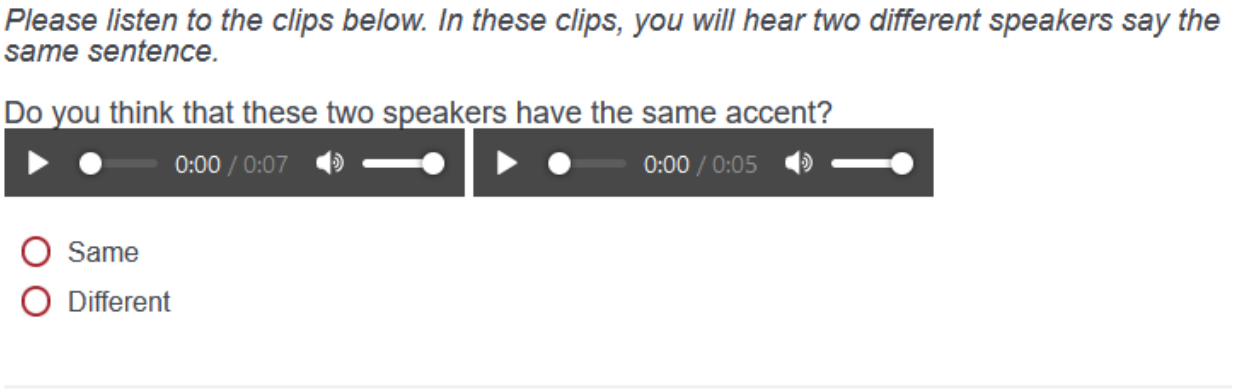
Appendix H: Phase 3 – ID Task Item Sample

In this second task, you will be asked to listen to one speaker at a time. After listening to a speaker, you will be prompted to answer one yes-no question based on your perception of that speaker's accent.



Appendix I: Phase 3 – AX Task Item Sample

In this activity, you will be asked to listen to two speakers at a time. After listening to both clips, you will indicate whether you believe that they are speakers of a same dialect or of a different dialect.



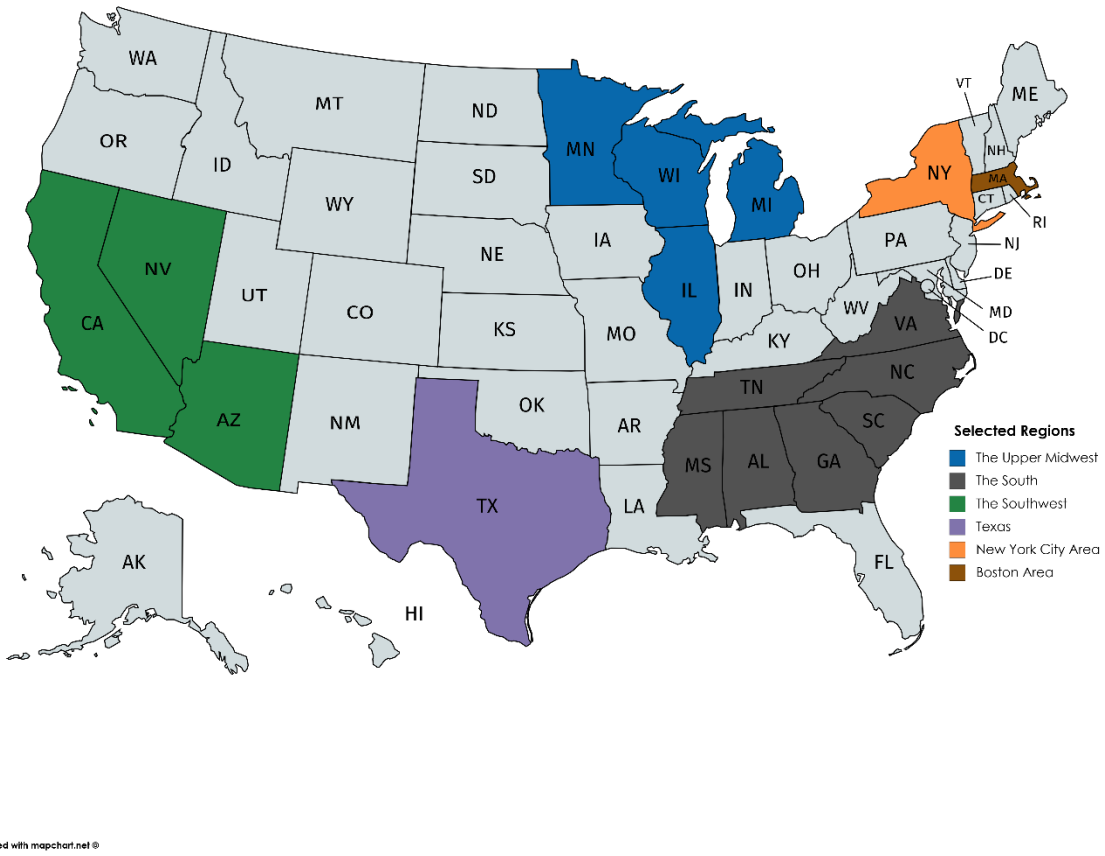
Appendix J: Phase 4 – Mental Map Task

Pre- and Post-Test Prompt

The goal of this task is for you to listen to the clips below and place the one speaker into each of the regional dialect groups, based on how you believe that they pair up. To do so, you can drag and drop a speaker into a dialect group box.

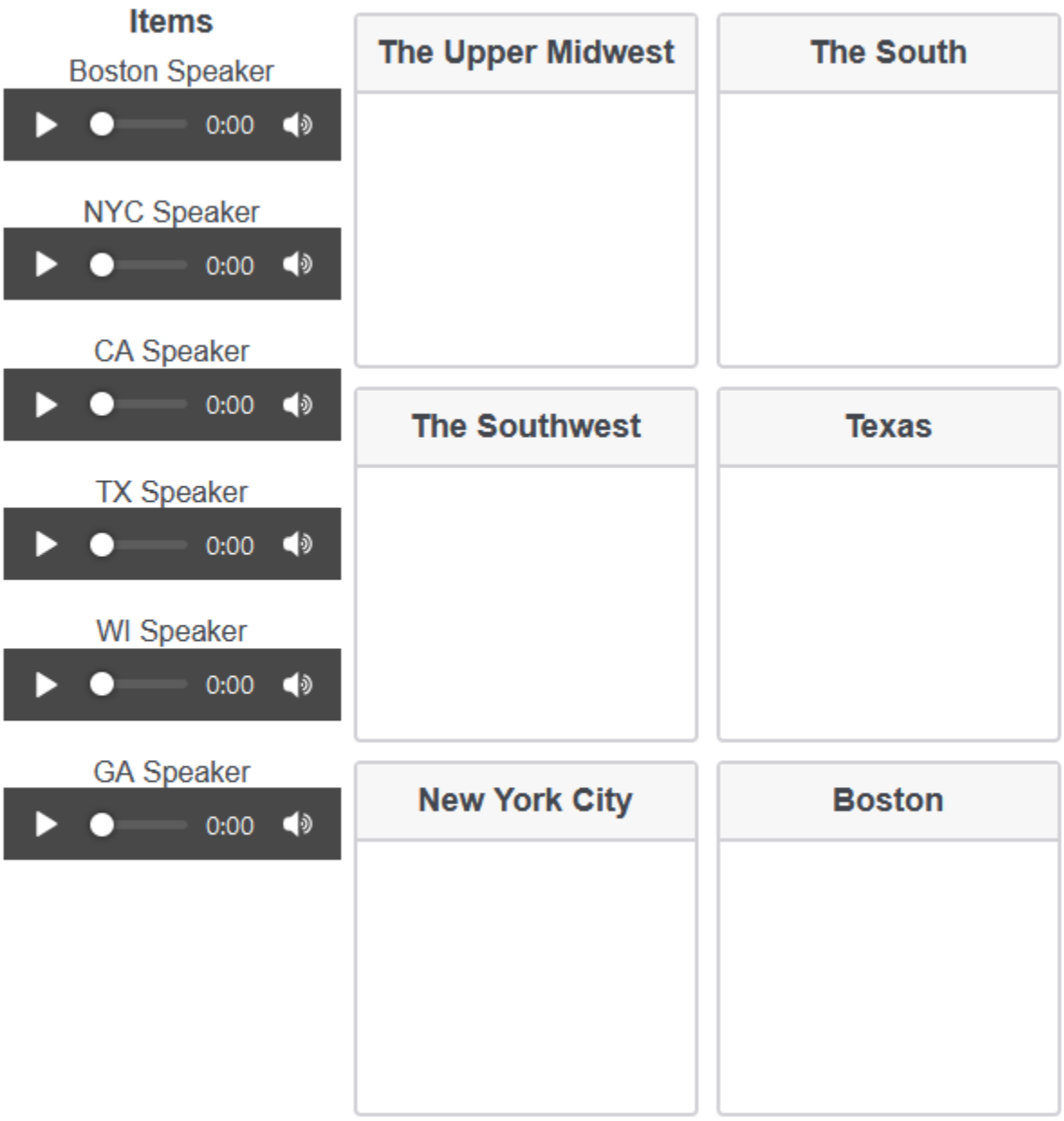
If you had to guess, where do you think that these speakers are from? There is only **one** speaker per dialect region, such that each of the six boxes should have only one of the six speakers by the end of the activity.

Please feel free to use the map below as a reference for each of the regions. The provided regional shading serves as a guideline and does not include the entirety of each region.



Created with mapchart.net ©

<u>Pre- and Post-Test Task</u> (Speaker Regional IDs Added for Clarity)



Main Mental Map Task Item Sample

For the main activity, you will be asked to listen to one speaker at a time. Based on how they speak, you will guess the region that you believe their dialect belongs to on the map.

If you had to guess based on the dialect, from what region do you think that this speaker is from?

0:00 / 0:07

Please feel free to use the map below as a reference.

- ○ The Midlands
- ○ Mexico
- ○ Cuba
- ○ Florida
- ○ The Northeast
- ○ Puerto Rico
- ○ The Southwest
- ○ The Rockies
- ○ The Upper Midwest
- ○ The South