# Perceptions of Dialect Standardness in Puerto Rican Spanish

Jonathan Roig
Advisor: Jason Shaw

*Submitted to the faculty of the Department of Linguistics in partial fulfillment of the requirements for the degree of Bachelor of Arts*

Yale University
May 2018

# Abstract

Dialect perception studies have revealed that speakers tend to have false biases about their own dialect. I tested that claim with Puerto Rican Spanish speakers: do they perceive their dialect as a standard or non-standard one?

To test this question, based on the dialect perception work of Niedzielski (1999), I created a survey in which speakers of Puerto Rican Spanish listen to sentences with a phonological phenomenon specific to their dialect, in this case a syllable-final substitution of [ɾ] with [l]. They then must match the sounds they hear in each sentence to one on a six-point continuum spanning from [ɾ] to [l]. One-third of participants are told that they are listening to a Puerto Rican Spanish speaker, one-third that they are listening to a speaker of Standard Spanish, and one-third are told nothing about the speaker. When asked to identify the sounds they hear, will participants choose sounds that are more similar to Puerto Rican Spanish or more similar to the standard variant? I predicted that Puerto Rican Spanish speakers would identify sounds as less standard when told the speaker was Puerto Rican, and more standard when told that the speaker is a Standard Spanish speaker, despite the fact that the speaker is the same Puerto Rican Spanish speaker in all scenarios. Some effect can be found when looking at differences by age and household income, but the results of the main effect were insignificant ($p = 0.680$) and were therefore inconclusive.

# Acknowledgements

First and foremost, I'd like to thank my advisor, Professor Jason Shaw, for guiding me through this project. Despite many frantic emails and rescheduled meetings, he remained fully supportive throughout the project and led me to a number of great resources, while also imparting a great deal of his own knowledge onto me. I am greatly indebted to him for his time and effort. I would also like to thank my fellow seniors and our Director of Undergraduate Studies, Professor Raffaella Zanuttini, all for being a comforting presence throughout the process, and reminding me that we are all still learning.

Thank you to my friends, who reminded me that my thoughts and ideas were both valid and interesting, and deserved to be heard. An extra thanks to Melina, who proofread my work. Thank you to my mother, who has always supported me in everything I do, and more tangibly in this study supported me by providing recordings for my survey and proofreading of my Spanish. And a special thanks to the people of Puerto Rico, who still responded to this survey despite continuing to struggle with post-Hurricane recovery efforts. For those reading this soon after its completion, if you'd like to support reconstruction efforts, I'd highly recommend donating to the community of Vílla del Río, a neighborhood of Vega Alta that I had the privilege of working with over Spring Break of 2018. Please consider a donation to their community through their GoFundMe page: https://www.gofundme.com/gwpuertoricorelief.

Additional thanks to Benjamin Franklin College for supporting this thesis through a Mellon Undergraduate Research Grant. Without this I would not have been able to distribute the survey for this project so efficiently. My gratitude also goes towards the Yale StatLab, who helped me run tests on my data. Though the results were inconclusive, they were still valuable.

# Contents

# Definitions

- standard: a relatively subjective term used to describe a dialect that is widespread or mainstream within a certain cultural or regional context, and is usually enforced by an elite group within a society; considered to be formal or what would be taught in a classroom
- non-standard: a relatively subjective term used to describe a dialect that is not spoken by the majority of speakers of the same language within a certain cultural regional context; dialects that would not be acceptable for writing in school assignment contexts, and are often stigmatized as "incorrect" ways to speak
- Hispanic: native Spanish speakers from any national backgrounds
- orthography: the writing system of a language
- the states: the 50 states of the United States of America, used to contrast Puerto Rico from the rest of the country, given that other US territories do not factor into this study

# 1 Introduction

Puerto Rican Spanish has been stigmatized among Spanish speakers for having a unique set of sounds compared to many other dialects of Spanish, especially non-Caribbean dialects. I have personally interacted with a number of Spanish speakers who claim that Puerto Rican Spanish is actually hard to understand because of its phonetic differences from their dialects of Spanish. People have referred to Puerto Rican Spanish as hard to understand or "unclean" because it is so different from their dialects, or at least sounds different than expected given Spanish orthography. In my experience, this is not highly discussed within communities of Puerto Rican Spanish speakers, but there is definitely a noticeable difference between Puerto Rican Spanish and, for example, Mexican Spanish, as different dialects of Spanish. On the other hand, such differences are less common between Puerto Rican Spanish and, for example, Dominican Spanish (another type of Caribbean Spanish), which both share similar, though importantly not identical, phonological systems that lead to such different surface phonetic differences from other dialects.

Given that Puerto Rican Spanish receives such criticism, the question of self-perception arises: how do Puerto Rican Spanish speakers perceive their own dialect

of Spanish? Other dialects of Caribbean Spanish do share similar sound systems, though Puerto Rican Spanish is known to exhibit them at higher rates (Figueroa 2017). Puerto Rican Spanish speakers on the island are the particularly interesting case, as they are less often in contact with phonological phenomena from other dialects of Spanish. Have they gotten to the point where they have claimed Puerto Rican Spanish as their own, or do they still see Spanish solely as the colonial language of Spain?

## 1.1 Dialects and Standardness

### 1.1.1 What is a 'standard' dialect?

The line between dialect and language has always been hard to define. Ideally, one might think a language contains many dialects, the same way a language family contains many languages, and within that language, all dialects are mutually intelligible. However, defining when a dialect is no longer part of the same language is a challenge. For example, in a dialect continuum, neighboring areas have mutually intelligible dialects. However, if someone from one end of the continuum were to meet someone from the other end of the continuum, their dialects might not be mutually intelligible despite the chain of dialect similarity that connects them. Despite such challenges, the dialects I refer to here are mutually intelligible dialects contained within the Spanish language. Examples of such varieties of Spanish include Andalusian Spanish, Mexican Spanish, and Puerto Rican Spanish.

### 1.1.2 What is Standard Spanish?

Given the above definition of dialects and languages, what might we consider the "standard" dialect of Spanish? One might assume that it is the Spanish dictated by the Real Academia Española (RAE), the academy in Madrid which seeks to enforce the "correct" way of using Spanish. One might also consider the Academia Puertorriqueña de la Lengua Española, the Puerto Rican equivalent of the RAE that is more sensitive to the dialectical differences in Puerto Rican Spanish. A more general version of this idea is that standard dialects are typically defined by what the elite of a society enforce as standard. These academies are exactly that: organizations run by academics. Such enforced dialects are typically what are taught in a classroom, whereas other dialects are considered to be "incorrect" ways to speak or "low-class". If we consider the RAE to be representative of the elite of Spanish speakers (which is probably fair within the context of Spain), then they indeed are enforcing the use of Standard Spanish, but it is important to note that there is truly

no such thing as correct and incorrect dialects. All intuitive or natural speech in a context can be considered correct, but there is more value in appreciating the differences in the way we communicate the same information, as we try to do in the field of linguistics.

### 1.1.3 Prescriptivism, Descriptivism, and Standardness

Standard is a highly misleading word when referring to dialects. Since it is simply enforced by an elite group of people, it is simply conventional that it is called standard. These dialects simply have the people with the power to make others feel that their dialect is wrong or at least should not be spoken in certain contexts. The implementation of a standard dialect is representative of prescriptivism, an ideology which implies that there are correct and incorrect ways to utilize language. This study (and all of linguistics) does not abide by such persuasions, and rather employs descriptivism: a desire to learn more about language as it is naturally spoken by all people. Rather than using grade-school grammar rules to decide what people should be saying, I will be analyzing data based on the idea that Puerto Rican Spanish speakers, not the RAE or the Academia Puertorriqueña de la Lengua Española, are the best representatives of how their dialect of Spanish is spoken. Therefore, when I use the word "standard", I will not be referring to any one sort of Spanish. Though participants in the study may imagine an aforementioned dialect as standard, I make no such implications. Instead, I use standard to refer to whatever it is that participants think standard implies. I will not be defining standard for them. This might mean that participants think of Puerto Rican Spanish, Madrid Spanish, Mexican Spanish, or some non-existent unilaterally ideal version of Spanish, all of which would lead to interesting discussion of results. The point is to see if they see themselves in or out of this category, however they define it.

## 1.2 Puerto Rican Spanish

Puerto Rican Spanish (PRS) is a dialect of Spanish spoken predominantly on the island of Puerto Rico and in the states of the United States of America. There are currently approximately 3,626,000 L1 speakers of Puerto Rican Spanish on the island and many more on the in the states themselves in the USA (Simons & Fennig 2018). About 5,000,000 more Puerto Ricans are said to live in the mainland (U.S. Census Bureau 2016), but their Spanish-speaking abilities are not accounted for, since information is usually collected only about their English-speaking abilities. In total, there are 39,145,066 (± 94,571) Spanish speakers in the States (U.S. Census Bureau 2012-2016).

While PRS speakers in the states are often interacting with Spanish speakers from other dialects, PRS speakers on the island have much less of these interactions, and other dialects certainly stick out among their own, though this says nothing about PRS speakers' opinions of standard dialect. Because of this, I ideally would run this study on PRS speakers who live on the island, not ones in the states. However, due to some limitations that arose throughout the process, this study examines both PRS speakers in Puerto Rico and in the states.

Puerto Rican Spanish is notable for its phonological phenomena related to liquid sounds, such as r's and l's. These phenomena are often what lead Puerto Rican Spanish to be a stigmatized variety of Spanish. Such phenomena can be found in other varieties of Caribbean Spanish, but some are expressed more often or more prominently in Puerto Rican Spanish (Figueroa 2017). In particular, there are three phenomena that are generally perceived as unique to Puerto Rican Spanish. The most prominent is one in which /r/, typically written as "rr" in Spanish orthography, is often realized as [x] or [χ], in contrast to the standard trilled [r]. In fact, many Puerto Ricans do not have a native speaker ability to produce a trilled [r]. This is referred to as the uvularization or velarization of /r/, and is summarized in (1) below. An example of this would be in the word "carro", meaning 'car', which is pronounced [kaxo] instead of [karo].The second phenomenon occurs when /ɾ/, typically written as a single "r" in Spanish orthography, is realized as [l] in syllable-final contexts, as seen in (2) below. This is in contrast to the standard variant, which is simply [ɾ] (Valentín-Márquez 2007).

(1) r → x, χ
(2) ɾ → l / _ $_{\sigma}$

# 2 Motivating a Dialect Perception Study

This study is inspired by a number of previous studies that combine sociolinguistic questions with phonetic studies to uncover information about linguistic biases. The studies described below directly contributed to the design of the methodology for this study and informed the thought process that motivated the research question at hand.

## 2.1 Exposing Bias with Dialect Perception

Niedzielski (1999) led the way with dialect perception studies that seek to uncover information about linguistic biases by studying English speakers in Detroit. Detroit is

located just across the river from Windsor, Ontario in Canada. English speakers from both Detroit and Canada are known to exhibit the phenomenon of Canadian Raising, which is when the diphthong /aw/ is pronounced with a more raised and fronted tongue. This most often shows up in stereotypes of Canadian English involving the word "about", which is exaggerated to "aboot" or "abote" by speakers of American English dialects attempting to reproduce the Canadian pronunciation. Previous work by Niedzielski has shown that Detroit residents associate this phenomenon with Canada but do not notice it in their own speech, despite the fact that both groups have it (1995, 1997). Rather, Detroit speakers associate their speech with Standard American English (SAE), even though they have many features, such as the Canadian Raising, that do not align with SAE. This paper extends the research by attempting to determine what effects social stigma might have on the perception of someone else's speech. Specifically, Niedzielski studied whether Detroit residents used their knowledge of a speaker's "nationality" (in this case, Detroit or Canada) to determine what vowels they were perceiving from that speaker. Niedzielski hypothesizes that listeners do in fact use social biases and stereotypes when perceiving phonological information from speakers, an idea Hay, Nolan, and Drager (2006) expand upon later, and that people can and do make incorrect predictions about their own phonological spaces based on stereotypes they have about their own dialect.

To test her hypotheses, Niedzielski recorded the speech of a Detroit speaker, then created computer-resynthesized vowels on a continuum of six vowels that included the actual vowel spoken by the speaker, the SAE version of the vowel, the resynthesized in-between vowels, and an exaggerated version of each vowel at each end of the continuum. She then asked 41 Detroit area residents to listen to the speaker's recordings, about 50 sentences each, and match the vowel heard in a particular word within each sentence to one of the computer-resynthesized tokens. Niedzielski told half of the participants that the speaker was from Detroit and the other half that the speaker was from Canada, but she gave them all the same set of recordings by the Detroit speaker.

Niedzielski found that for words containing /aw/, nationality labels had a significant effect on most participants' choice of vowels. A majority of participants (53%) who were told the speaker was from Canada chose the actual vowel that corresponded to the speaker. However, very few participants chose the correct vowel when they were told that the speaker was from Detroit (only 15%). This evidence supports the idea that social biases and stereotypes affect the perception of phonological information because participants' perception of their own dialect was entirely changed when told that someone of another nationality said it. It also proves that people can have incorrect stereotypes about their own dialect, since the participants rarely

believed that the speaker had used a vowel that underwent Canadian Raising when told the speaker was from Detroit, and instead chose the more standard variant of the vowel, even though the raised vowel is common for English speakers from Detroit, and Detroit speakers do not usually produce the standard variant.

Niedzielski also explored the Northern Cities Chain Shift (NCCS) in this study to explore whether gender had an effect on stereotyping. In previous studies, she found that women were more likely to have stereotypes about Canadian dialects of English than men, but this study disproved that, since men and women patterned together.

The statistical analyses done in this study are questionable. Niedzielski often looks at the result of a single question rather than comparing the surveys as a whole. She also entirely removes data for some responses (for example, removing the extremes of the spectrum) because they are underutilized overall. Removing this data almost definitely changes the results at least a little, and if it doesn't, there is no need to remove them. I do not find a compelling argument to remove this data.

The study results supported all parts of Niedzielski's hypothesis. Social biases and stereotypes about Detroit and Canadian speakers had direct effects on people's perceptions of the speaker's vowel space, effectively changing the way they heard language as it was spoken to them. This was sufficiently achieved by her study of Canadian Raising. Though she does go on to discuss the NCCS, very little of her NCCS data is used to support her hypothesis. It does, however, correct her previous work's hypothesis that women have stronger stereotypes of Canadian English than men, something that can also be tested for effect in this study.

## 2.2 Exemplar Theory

Hay, Nolan, and Drager (2006) was a direct response to Niedzielski (1999). Hay et al. seek to uncover the biases that New Zealanders hold about Australian English, as well as their own dialect of English. They mostly followed the same methodology as Niedzielski (1999) and made only minor adjustments. The most significant adjustment was that no participants were told explicitly that the speaker was from either New Zealand or Australia. Instead, they simply wrote "New Zealander" or "'Australian" at the top of the answer sheet with no explicit description of how this might be connected to the study. The results were similar to that of Niedzielski (1999). Hay et al. use this information to propose an idea based on exemplar theory. Exemplar theory was an idea in psychology that was introduced to linguistics by Johnson (1996). The theory proposes that phonetic material is stored in the mind is accompanied by social information about the speaker who says it. This means

there would be a lexical entry not only for every word, but for every instance of a word spoken by a different person. Such a memory-heavy operation has undergone criticism and its use in linguistics comes with hesitation, but it is a reasonable explanation for the phenomenon at hand. In the case of these dialect perception studies, learning social information about the speaker or being primed by information that can trigger a social bias can make the perception of sounds related to that social information more salient in the mind, enough to change the perception a listener has of the sounds they hear.

Though they do not explicitly state it, the hypothesis of this paper is effectively the equivalent of Niedzielski (1999), the main idea being that social biases and stereotypes affect perception of phonetic information, and stereotypes about one's own dialect can be and often are inaccurate. The slight difference in the methodology is that the amount of priming necessary to get someone to use such biases for the perception of phonetic information was minimal, only requiring a slight notion (such as writing the word "Australian" on a piece of paper), rather than directly telling someone where a speaker is from. This supports an additional hypothesis that such biases are working with even the slightest association to a social bias, though Hay et al. recommend further research in this area.

Hay et al. base their theory of exemplars both on their own results and Niedzielski's because both point towards the phenomenon of social bias in phonetic perception. Though most speakers in the Hay et al. study acknowledged that they knew the speaker was from New Zealand, they still performed with the same results as Niedzielski (1999), answering with Australian vowels when they were primed to do so. Therefore, Hay et al. believe the exemplars would overpower conscious knowledge of a speaker's nationality, meaning subconscious information either related to the phonetic information itself or notions like writing of a nationality on the top of a page take precedence during speech perception.

A few major issues arise with the methodology of this study. First is that they did not have enough participants, especially because they adjusted their methodology about halfway through their set of participants, with only about 20 participants in each phase, despite having two versions of the survey to distribute. They are unable to make solid claims related to the vowels they used, /æ/ and /ɛ/, because they do not have enough participants to do so. Additionally, as with Niedzielski (1999), there is no control group. Everyone is either primed with "Australian" or "New Zealander" written on the page, but there is no version in which nothing is written on top. Finally, neither study varied the perceived gender of the recorded voice, each only using one speaker. To prevent having gender as a confounding variable, both experiments might have benefited from having multiple genders represented in

the voice recording, including both male, female, and potentially even androgynous voices.

# 3 The Puzzle

## 3.1 Motivating the Study

The Niedzielski study inspired this study for a number of reasons. First and foremost, it introduced the idea of using nationality as a potential bias that could affect speech perception. Given the way that Detroit speakers responded to their own dialect given different initial information, I wanted to do a similar test on a contextually different population. This serves as a test of the reproducibility of previous experiments, but also attempts to uncover results for a unique population of speakers who do not have a parallel to Detroit's relationship with Windsor. Given that Puerto Rico is an island, Puerto Rican Spanish speakers are likely not in nearly as much contact with other dialects of Spanish as Detroit English and Windsor/Canadian English. Additionally, the use of a vowel continuum inspired the use of a continuum to test dialect perception. However, given that vowel-related phenomena are not what distinguish Puerto Rican Spanish from most other dialects of Spanish, an [ɾ] to [l] continuum is used instead, targeting the aforementioned phonological rule relating these two sounds. Finally, the idea of collecting information about the speaker, especially gender, comes from this study. Collecting such data can lead to the discovery of correlations between social groups and certain responses, which will be explored in the analysis and discussion sections.

The Hay et al. study was important for deciding how to deliver information about the speaker. As in the Niedzielski, I found it important to introduce the speaker's supposed origin straight from the beginning in hopes of getting strong results. However, Hay et al. introduced the idea of putting a more subtle notice throughout the study. As a result, throughout this study, I inserted a phrase at the top of each page that says what type of Spanish the speaker is supposedly speaking, where applicable. They are still told about the speaker's origin as in the Niedzielski study, but get the additional reminder in the fashion of the Hay et al. study. In the end, I do not make it subtle at all, but the goal of this study is not to test subtlety of notions, but rather implicit bias. I believe that the more times the social group of the speaker is emphasized, the greater chance we have at seeing significant results.

Some inspirations from these studies also come from aspects that felt missing. Solving the issue of a single-gender voice was outside of the scope of this study, but to address it, the speaker of the practice sentences is a male voice and the speaker

of the actual task sentences is a female voice. Additionally, since neither study had a control group, I decided to add one to test if it is indeed functioning as a control. The control version has no information about the speaker anywhere in the survey. I also decided to distribute my survey online in an attempt to eliminate effects due to bias about the experimenter.

Both experiments set the groundwork for the number of participants. Both had a goal of having about 20 participants in each condition, though the Hay et al. (2006) did not achieve this in the end. Therefore, there is a goal of 60 participants for this study (20 each in 3 groups), with the understanding that some data may not be usable for various reasons, but enough data should remain for analysis.

## 3.2 Hypothesis

Many Spanish speakers might assume that the "standard" or "prestige" dialect of Spanish would be one spoken in Spain, like in Madrid. However, Puerto Rican Spanish speakers rarely have interactions with Spanish speakers from Spain, especially those who have lived their whole lives on the island. This means that Puerto Rican Spanish speakers are a special case when it comes to perceptions of a standard dialect. They might perceive PRS as a standard dialect, or they might assume they speak a non-standard dialect, whether that is due to their relatively small population or the active marginalization and stigmatization they experience. Do they perceive their own dialect of Spanish as standard, since it is the dominant dialect in their lives, or non-standard, especially with the knowledge that the language came from Spain?

My prediction is that Puerto Rican Spanish speakers do not see their own dialect as a standard form of Spanish. Therefore, using a six-step [ɾ] to [l] continuum similar to that of Niedzielski, where 1 is on the [ɾ] side and 6 is on the [l] side, when told that they are listening to a Puerto Rican Spanish speaker, they will hear the /ɾ/ => [l] phenomenon, whereas when told that they are listening to a Standard Spanish speaker, the phenomenon will be masked by social bias and they will hear [ɾ]. For the control group, they must pattern with one and only one group to be considered a control. Otherwise, if it patterns totally differently, they must be studied as a separate group with its own effect with its own potential reasons for performing differently. If it patterns the same as both, no conclusion can be made since no results would be significant or conclusive. I predict the control will pattern with the Standard group because they are likely to perceive the survey as a formal affair, and as a result they will either assume that the speaker is speaking formally, therefore accessing parts of the lexicon marked as formal speech, or they will simply not trust

| Survey Group | Predicted Side of the Continuum | Significantly Different From |
|---|---|---|
| No Info | [ɾ] | Either, but not both |
| PRS | [l] | Standard |
| Standard | [ɾ] | PRS |

Table 1: The hypothesis as it is predicted to appear in each survey result, and which surveys each survey should be significantly different from

their own instincts about what they are hearing.

I chose the phonological phenomenon described in (2) in Section 1.2 because it is neither unnoticed nor blatantly obvious. This allowed me to make a reasonable attempt at masking the identity of the PRS speaker I recorded in the surveys that do not give accurate information about the speaker. Had I chosen the phenomenon described in (1) in Section 1.2, it would be too obvious that the speaker was Puerto Rican. Ideally, this phenomenon will not be actively noticed outside of the version of the survey where participants are told that the speaker is Puerto Rican.

# 4 Methodology

## 4.1 The Survey

The hypothesis was tested through a survey designed using Qualtrics and distributed on Amazon Mechanical Turk (mTurk). All text in the survey was translated into Spanish. The survey started with an introduction that had them confirm that they were a Puerto Rican Spanish speaker. For the purposes of this study, this meant they had spoken PRS for the first 14 years of their lives and still used PRS at least once per week. For each trial, the sentence is first displayed on the screen with the target word in bold. It remains on the screen throughout all steps of the given trial. The recording of the sentence is then played as the participant continues to see the sentence on the screen. The recording of the sentence is immediately followed by playing each of the 6 sounds of an [ɾ] to [l] continuum, each clearly numbered by the speaker. The participant then must choose the sound from the continuum that they think best matches the sound they heard by choosing that number as their response to the question.

The survey contains 20 test sentences with sentence-medial, word-final [ɾ], each of which is produced on the surface with an [l]. The environment is controlled for the preceding vowel, meaning the /ɾ/ is always in the final syllable /eɾ/. The word

never repeats across the sentences. The word also always has its primary stress on the last syllable. An additional 10 filler sentences contain sentence-medial, word-final /l/, produced [l], in order to both keep the attention of the participants and mask the target phenomenon. These are not controlled for vowel as they were not designed to be used for any data analysis. 4 practice sentences were also used to train each participant for the task. 2 of these sentences contain sentence-medial /ɾ/ produced as [ɾ]; the other 2 contain sentence-medial, syllable-final /l/ produced as [l] so that both ends of the spectrum are employed in the practice. Again, they were not controlled for vowel since they were not designed to be used for any data analysis. The practice sentences were read by a different speaker to prevent creating expectations about the speaker from the test sentences during the practice portion. The total number of sentences is 34, as opposed to the 50 the previous studies used. This is because I tried to make the survey short enough to avoid losing the attention span of participants and to pay participants appropriately for their time while keeping the cost reasonable. Test sentences were created with special attention to other phonological phenomena that may give away the dialect of the speaker. For example, words that contain an /r/ appear on the surface as [x] (or a similar production), and therefore were not included in the test sentences, as this is a phenomenon that tends to be considered a quintessential sound of Puerto Rican Spanish. Not all phenomena could be removed, however, due to their prevalence in the language. For example, syllable-final /s/-deletion is common in Puerto Rican Spanish, but eliminating all instances of word-final /s/ is difficult if not impossible in Spanish sentences.

There will be three editions of the survey: one in which participants are told that the speaker is from Puerto Rico (PRS group), one in which participants are told that the speaker is a speaker of Standard Spanish (Standard group), and one where participants are given no information about the speaker (No Info group). For PRS group and Standard group, the phrase "Puerto Rican Spanish" and "Standard Spanish" will show up at the top of each screen to remind participants of what they were told at the beginning, hopefully leading to a more noticeable result. Standard is not defined for the participants. The No Info group simply has question numbers on their screens.

The speaker was a 48-year-old female Puerto Rican Spanish speaker living in New Jersey. Spanish was her first language, but was learned alongside English starting at around age 5. Therefore, she is a native speaker of both. She speaks Puerto Rican Spanish daily, but also teaches Spanish to K-8 students. She is able to perform a controlled switch between the formal Spanish she uses for the classroom and Puerto Rican Spanish. The recordings were made with her conscious knowledge that she should be speaking in her Puerto Rican Spanish mode.

## 4.2 The Continuum

To make the continuum, the speaker was asked to produce the traditional names of the Spanish letters 'l' and 'r' (i.e. [ele] and [eɾe]). Using Praat, I compared the waveforms and spectrograms to compare differences between the segments that could be manipulated to create the continuum. The main observable difference between the speaker's production of [l] and [ɾ] was the duration of the consonant: approximately 92ms (not including vowels) for [l] (see Figure 1) and 21ms for [ɾ] (see Figure 2). Given that this was the most apparent difference between the two consonants, it played a vital role in the manipulation for the continuum. Additionally, there was a sharp drop in F2 for [l] upon the transition into the consonant closure, whereas [ɾ] has a gradual drop in F2 during its closure.

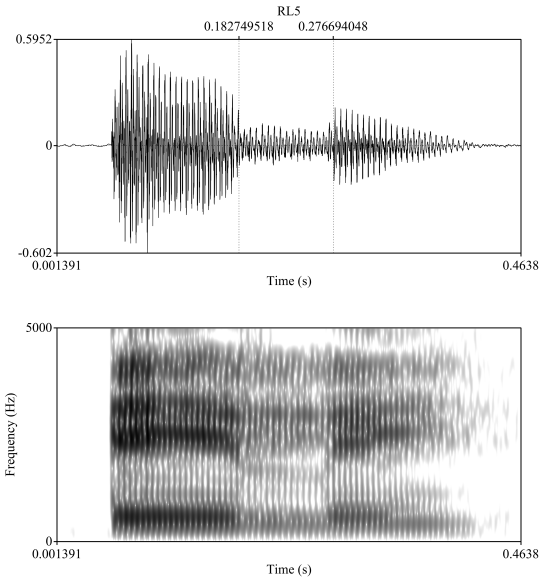


Figure 1: The waveform and spectrogram for Token 5. The two vertical lines mark the beginning and end of the consonant [l].

The continuum is made up of 6 tokens based on the previous experiments described above. They are labeled 1 to 6 from most [ɾ]-like to most [l]-like. Token 2 is the actual [ɾ] token recorded by the speaker, and Token 5 is the actual [l] token recorded by the speaker. Tokens 1 and 6 are the "extreme" tokens, meaning they are manipulated to be further on the spectrum of duration than the actual consonants. These extreme tokens are based on similar tokens in the Niedzielski study's continuum. Token 1 is cut shorter than the actual [ɾ] token, to approximately 7ms,

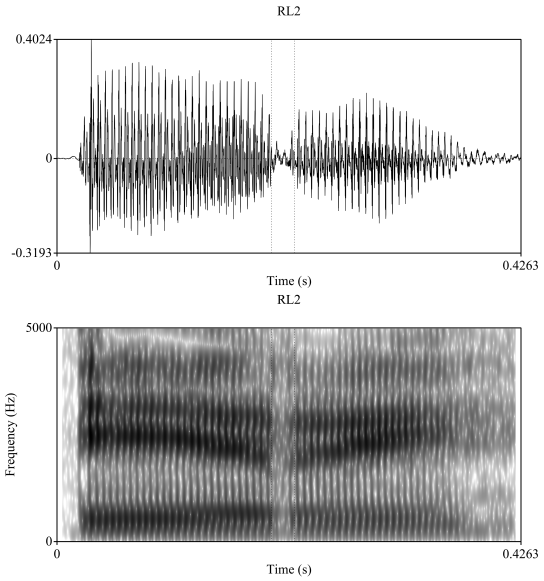


Figure 2: The waveform and spectrogram for Token 2. The two vertical lines mark the beginning and end of the consonant [ɾ].

whereas Token 6 is extended to be longer than the actual [l] token, to approximately 104ms. Token 6 was extended by copying a part of the [l] consonant audio and repasting it into the sound file within the preexisting consonant sound. The audio file was cut at a point when the waveform hit an amplitude of 0, going from positive to negative. (see Figure 3). This was done to maintain realistic audio quality and prevent any "skipping" sounds during playback. As a result, the audio file sounds like an elongated [l] as opposed to two adjacent [l] tokens stitched together.

To keep the durations evenly spaced across tokens 2-5, the duration of the consonant must increase by 20ms for each consecutive token. Therefore, with adjustments due to the aforementioned smoothing, the consonant in Token 3 has a duration of 72ms, and the consonant in Token 4 has a duration of 65ms.

Given that F1 is approximately the same between [ɾ] and [l], and up until the point of the drop in F2, their F2 values are close (within 100Hz, typically), the difference between the transitions into the consonant closures was not manipulated. Instead, these transitions for Tokens 3-6 all come from the [l] recording. The transitions for leaving the consonant comes from the [ɾ] in Tokens 1-4. The difference between Tokens 3 and 4 is simply that slightly more than half of Token 3 is made from the [ɾ], whereas slightly more than half of Token 4 is made from the [l] recording.

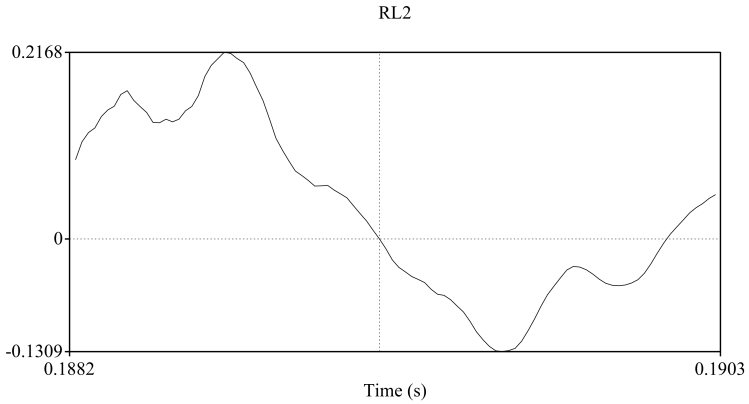


Figure 3: An example of a point in which the audio file was cut in order to provide a smooth audio file despite the synthetic manipulation.

# 5 Results

## 5.1 Demographic Disparities

| Survey Group | Participants in PR | Participants not in PR | Total |
|---|---|---|---|
| No Info | 4 | 13 | 17 |
| PRS | 6 | 17 | 23 |
| Standard | 6 | 13 | 19 |
| Total | 16 | 43 | 59 |

Table 2: The breakdown of participant location between Puerto Rico and the states

A total of 63 participants completed the survey. However, four responses were removed due to concern over faked responses that were simply completed for the payment, which is a known concern of posting surveys through mTurk. 2 of these surveys consisted of solely '1' responses on the continuum for each and every question. The other two were removed because all the written answers were made up of incoherent strings of letters that were clearly not answers to the questions asked of them, and rather random text they typed in to satisfy the requirement that each question has an input answer before moving on. This leaves the total number of participants at 59. There were 17 participants in the No Info group, 23 in the PRS group, and 19 in the Standard group (see Table 2). These numbers are not balanced both because of removed data and because the survey that each participant received

| Survey Group | Male Participants | Female Participants | Total |
|---|---|---|---|
| No Info | 14 | 3 | 17 |
| PRS | 12 | 11 | 23 |
| Standard | 12 | 7 | 19 |
| Total | 38 | 21 | 59 |

Table 3: The breakdown of self-reported gender by survey

was chosen randomly, not based on any demographic information, to keep a true random sample.

The survey was distributed across the United States, including Puerto Rico. The survey did not seek to balance the location (or any other demographic information) of the participants. It simply limited responses to Puerto Rican Spanish speakers living in the United States. 43 participants were living in the mainland, while 16 lived in Puerto Rico itself. Ideally, the survey would have been entirely distributed to only Puerto Ricans living in Puerto Rico, but the limitations both due to the functionality of mTurk and the likelihood of a lack of data due to Hurricane Maria made this unreasonable and impossible. Additionally, 38 respondents identified as male, while 21 identified as female. Male includes people who identified as "masculino" or "hombre". Female includes people who identified as "femenino", "mujer", or "señora" (see Table 3). The question was left open-ended to allow participants to identify as comfortable, but no one identified outside of these two groups as described. The difference in sizes of these groups is considerable, and causes some challenge in the later statistical analysis. They should be considered when discussing the meaning data might hold. For example, the No Info group only included 3 female participants out of its 17 total participants. Further research with larger, more diverse sample sizes would be required to confirm the claims that will be made in relation to this demographic data. mTurk did not bring in as wide a variety of people as hoped.

## 5.2 Simplifying Data

The data was initially organized by taking the means of each set of survey responses. This was done by first taking the mean of each user's survey responses, then taking the mean of all the participant means within a survey. The result was checked by doing the same calculation, but instead taking the mean of all participant responses to each question, then taking the mean of all question means within a survey. The results are depicted in Table 4.

| Survey Group | Mean |
|---|---|
| No Info | 3.356 |
| PRS | 3.415 |
| Standard | 3.592 |

Table 4: The mean responses to the continuum by survey

Given that the continuum spanned numbers 1 to 6, a true average response would hypothetically be centered at 3.5. On a cursory look of these means, none of them seem far from 3.5, and all are relatively close to one other. Despite this, analyses were done to test the significance or lack thereof not only between surveys, but also within surveys. Self-reported gender, age, income, and more were analyzed for potential effects, as well.

A concern about a survey with a continuum such as this one is that some participants may have treated the continuum differently than others. For example, one person may have distinguished [ɾ] from [l] using only 5 and 6, while someone else may have done the same using only 1 and 2. Therefore, data was standardized by participant and means between questions were compared using this data. Surveys could not be compared to one another in this way because the definition of standardization causes all survey means to be 0 when results are standardized. This data can be found in Appendix 9.2, Table 11.

## 5.3 Conscious Perception and Other Raw Data

Each participant answered additional questions about their backgrounds and thoughts related to the survey. For example, information about the number participants' organized by primary household income (Table 12) and age (Table 13) within each survey and across all surveys can be found in Appendix 9.2. Given the aforementioned issue with the mTurk population not being diverse enough, only a few potentially meaningful groupings could be made for income and age to test for significance. The range of these groups was designed to avoid having just 1 or 0 participants fall into a given group in any of the surveys. However, this led to creating divisions that were not preferable, such as a >34 age group. Though having groups with few participants was avoided, a wider range and more precise set of groups could be included in the data analysis if a larger, more diverse population was surveyed.

One additional question that was considered aside from demographic information asked participants directly if they personally thought Puerto Rican Spanish is a

standard dialect (Table 14). This, as well as the demographic questions, was asked at the end to avoid influencing the results of the survey. This table can also be found in Appendix 9.2. This question was asked in contrast to the rest of the survey, which seeks to uncover subconscious perceptions about the status of standardness of Puerto Rican Spanish, with the potential to find contrast between conscious and subconscious perceptions of PRS.

# 6 Statistical Analysis of Data

## 6.1 Main Effect

In order to find any statistical significance in the data, I ran a series of ANOVAs and two-sample t-tests. ANOVAs were used when a variable (where "variable" refers to gender, age, etc.) with multiple subgroups was considered in the calculation. All statistical tests were run using the program Minitab unless otherwise specified. Data were checked to make sure the assumptions of each test were met. Before running an ANOVA, I ran a Levene's test to confirm that the variances for that data could be considered equal. For both ANOVAs and t-tests, I checked that distributions were approximately normal by observing normal quantile plots. While there was slight skewing in some situations when looking within the surveys, there was nothing of particular concern, especially since such skewing did not happen when looking at data across all three surveys. Given that there were around 20 participants per survey, some deviance is to be expected due to sparse data, particularly when dividing those groups further by demographic groupings. P-values for ANOVAs and two-sample t-tests were compared to significance level of $\alpha = 0.05$, meaning p-values needed to be less than 0.05 to be considered significant.

The main test was to see if there was a statistically significant difference in the means across the three surveys. An ANOVA was run over the three surveys, using the means of each participant as the data points within each survey. The results were not significant at $p = 0.680$. This is not surprising given how close the means of each survey were. However, I did not stop the analysis here since there is a possibility for patterns based on the demographic information provided by the participants.

## 6.2 Effects by Group

To find potential effects based on the additional collected information, I tested each variable along with each survey condition for their effects on the means in a two-way ANOVA. This was to ensure that there was no significant interaction between

the condition and the variables looked at in each analysis, that way the data across all three surveys could be combined for the analyses. For example, gender and condition were tested for their effect on mean in a two-way ANOVA. If their effect was insignificant ($p \geq 0.05$), there was basis for collapsing the data of all three surveys into one dataset to test for effects based on gender. All of the variables passed this test. Data was combined and one-way ANOVAs were run on each variable to find effects on the mean. I also ran ANOVAs on the same demographic groups within each survey to see if there were any additional effects within each condition. P-values for all ANOVAS and t-tests done by group can be found in Table 10 in Appendix 9.2.

### 6.2.1 Test Sentence Effects

As mentioned previously, participants may have interacted with the survey differently, meaning that standardization of the results could be interesting. Results were standardized using Excel before being analyzed further in Minitab. They were standardized using the mean and standard deviation of each participant. This means that each participant's answers averaged to 0, and therefore all survey averages were 0. However, this gave me a chance to see if the results of any questions stood out among the rest. An ANOVA was run over the averages of all the questions, both with combined survey data, and one survey at a time. Unfortunately, none of the results were significant.

### 6.2.2 Demographic Effects

Effects from gender were first analyzed using the combined data from all surveys. I ran a two-sample t-test to look for effects of gender on mean response. No significant difference was found between those who identified as male and those who identified as female. I also ran a two-sample t-test on gender and mean within each survey. As mentioned before, though the test can be run, it may not mean much in the case of the No Info survey, which had 3 female respondents and 14 male respondents. The difference is less noticeable in the other two surveys but present (see Table 3). Regardless, none of the results were statistically significant, meaning that the means cannot be said to be different between males and females in any of the three surveys, just as in the combined data.

Given the aforementioned issues with the age range of the sample population tested being relatively young (see Table 13 in Appendix tables), the grouping of ages is relatively arbitrary and likely doesn't say much since there was not a significant older population. Nevertheless, I made an attempt to find significant differences

| Age Group | Mean |
|---|---|
| <25 | 3.512 |
| 26-29 | 3.314 |
| 30-33 | 3.375 |
| >34 | 5.375 |

Table 5: The breakdown of means by age group in the Standard group

between the age groups I had to work with. Analysis using an ANOVA on the combined data from all three surveys did not return a significant result. However, running an ANOVA within the surveys returned the first significant result. For the Standard group, those aged 34 and above had a statistically significant difference from the rest of the age groups at $p = 0.018$. This may be because only 2 participants landed in this group for this particular survey, but it is still worth considering. Table 5 contains the means for each age group within the Standard group. The mean for the >34 group is visibly quite higher than the rest, which was confirmed by a Tukey Pairwise Comparison. It would be unfair to claim that this means older groups are more likely to choose the [l] side of the continuum than younger folks, given that >34 covers quite a range. This could be the case, but it must be tested further in order to make a more supported claim. Additionally, such an idea is only supported within the Standard condition. The other two versions of the survey do not have a significant result here, and the means of these groups are usually not the highest, which does not support the idea that older groups are more likely to choose from the [l] side of the continuum. The tension between the results of these surveys emphasizes the need for further study with participants that have a greater diversity in age.

The analyses based on income had the most exciting results. Income was tested using the aforementioned income divisions (see Table 12 in Appendix 9.2). When tested for significance in the combined data from all surveys, the results of the ANOVA were technically still insignificant at $p = 0.083$. However, it is reasonably close to the generally accepted $\alpha$ value of 0.05 to be considered and discussed. Since the p-value was not below 0.05, none of the groups were shown as significantly different from one another in a Tukey Pairwise Comparison; however, Figure 4 makes clear the unusual pattern present in the data. The lowest and highest income groups pattern together, and are noticeably lower than the middle four income groups (especially so for the highest income group). It seems unusual that the lowest and highest groups would pattern together, and we cannot say for sure that that is what is happening here given the lack of significant results, but this is possibly the case

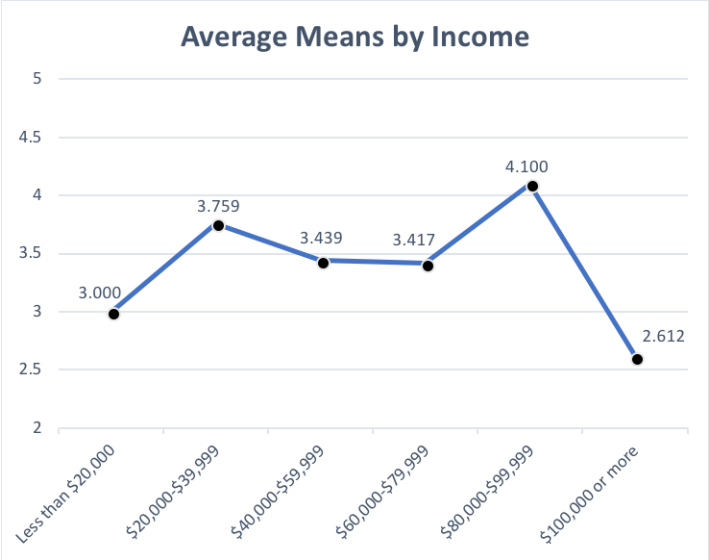


Figure 4: The average means of participant responses across all surveys plotted by income group

and will be discussed further in Section 7.

As before, I also ran an ANOVA within each survey by income group. A significant p-value was again obtained only in the Standard group, coming in at $p = 0.024$. However, a Tukey Pairwise Comparison does not group any of the income categories separately; none are shown as significantly different. Given the means of each group (available in Table 6), the most likely to be significantly different are the less than \$20,000 category and the \$20,000-\$39,999. This reflects the pattern present in the combined data (see Figure 5 for comparison, but note that there is no data point for \$80,000 - \$99,999 in the Standard Survey). The significance may be a result of the fact that there is only 1 participant in the less than \$20,000 group for this survey, which is not a large enough sample to make judgments or claims about. However, since it does have support from the analysis across all surveys, it is worth considering further in the discussion.

Since the survey was distributed both in and out of Puerto Rico, one might be concerned that those in Puerto Rico might respond differently than those in the states. This is especially important given that the ideal version of this study would have only been run in Puerto Rico and not the states at all. However, the t-tests on the combined data, as well as the data within each survey, all had insignificant results, determining that there is no significant difference in the results between those

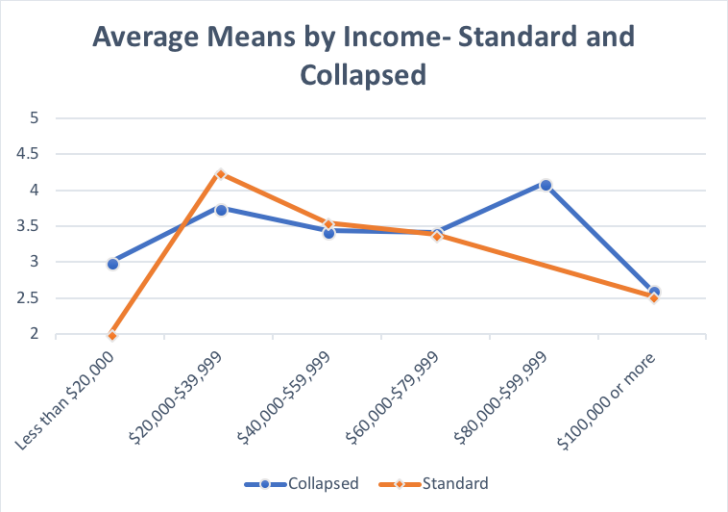


Figure 5: The average means of participant responses across all surveys plotted by income group, overlaid on the average means of participant responses within only the Standard group

| Income | Mean |
|---|---|
| Less than $20,000 | 2.000 |
| $20,000 - $39,999 | 4.250 |
| $40,000 - $59,999 | 3.550 |
| $60,000 - $79,999 | 3.383 |
| $80,000 - $99,999 | no participants |
| $100,000 or more | 2.525 |

Table 6: The breakdown of means by income group in the Standard group

in Puerto Rico and those elsewhere in the country.

Participants were asked directly at the end of the survey if they thought Puerto Rican Spanish was a standard dialect to compare their conscious thoughts to their subconscious actions in the survey (see Table 14 in Appendix 9.2 for this data). No significant results were found in either the t-tests across the combined data or within each survey. However, it is interesting to simply observe the results to this question as raw data. 41 respondents answered that Puerto Rican Spanish is in fact standard while 18 respondents answered that it was not. That is, about two-thirds of respondents thought Puerto Rican Spanish was standard. Interestingly, 7 of the

18 "no" responses came from the participants who were in Puerto Rico, almost half of that group, which had 16 total. This is in contrast to the approximately one-quarter of the group of participants from the states who also did not see Puerto Rican Spanish as standard.

# 7 Discussion

## 7.1 Implications of the Data

The lack of significance in the ANOVA between surveys means that the main hypothesis that Puerto Rican Spanish speakers do not see their own dialect as standard cannot be supported. Ideally, this would mean that Puerto Rican Spanish speakers do see their own dialect as standard. Reflecting on Niedzielski (1999) and Hay, Nolan, and Drager (2006), one might in fact expect the opposite hypothesis from the one I predicted, since the participants in those studies treated their own dialect as standard even when it technically was not. However, that would require statistical significance in the opposite direction, and is therefore also not supported by this data. This question must be tested further in order to come to conclusive results, and this study indeed calls for future studies in dialect perception for Puerto Rican Spanish speakers both in and out of the island. There are many reasons that may have caused such results, ranging from oversights in survey design to the specific population tested, and a plethora of other reasons I could not even begin to imagine. Recommendations for changes to be made for future studies can be found in Section 8.

As seen in Section 6, many of the other tests also did not support the hypothesis for their specific groups, with a few exceptions. However, such results were not always a bad thing. The lack of significance in the ANOVA done over the standardized data means that none of the questions had a particularly unusual pattern compared to the rest, which could be a sign that questions were designed well enough to avoid triggering a particular response on their own. Though having certain questions stick out from the rest would have been interesting, it is better for the integrity of the survey that no question means were significantly different from the rest. It could have been a stroke of luck related to the fact that the surveys all had means relatively close to the middle of the continuum, and the individual questions simply managed to reflect this, but at least for this study, no questions needed to be removed from the data and none had an effect on the outcome of the data.

The lack of significance in the two-sample t-test run on gender on the combined data and within each survey simply means that the way males and females are inter-

acting with the survey, and hopefully the way they are interacting with perceptions of dialect standardness, are virtually the same. This is what should be expected according to Hay et al. (2006). While an effect might've been interesting, it is reassuring that there is not one in the hopes that that means neither represented gender is experiencing more extreme internalization of stigma against their dialect than the other.

Since it's within the Standard group that the >34 age group leaned more towards the [l] side of the continuum, it is possible that this supports the idea that Puerto Rican Spanish speakers do in fact treat Puerto Rican Spanish as Standard. Since they were receptive in hearing the 'l's that were present in the recording despite being told that the speaker spoke Standard Spanish, it may be that Standard Spanish did not trigger the idea of a non-Puerto Rican Spanish speaker the way I expected. However, since this is not supported by the results of the ANOVAs across the surveys, this speculation must be confirmed by further testing. As mentioned before, the age of the sample population was not representative of the whole population, and so these results have minimal meaning for the population at large.

Even with the significant result in the Standard group and the similar pattern in the combined data, there is not much to be made of the ANOVAs analyzing the effects of household income on the survey results. This was included to seek out potential differences in class, since there are previous works that highlight prejudice against the phenomena that signify Puerto Rican Spanish even within Puerto Rico. Some phonological phenomena of Puerto Rican Spanish are referred to as "low-class" by Puerto Rican Spanish speakers who perceive themselves as part of the upper-class (Valentín Márquez 2007). However, since the largest and only significant difference in means is between adjacent income groups, these prejudices cannot be considered to be the reason for the significance, or at least not on their own. I would expect there to be a linear trend, with lower income groups (especially within the PRS group) leaning more towards [l] because of their active use of the [l] and awareness of such use due to experiences with marginalization and stigmatization of their dialect. I would expect higher income groups to lean more towards [ɾ] because they are less likely to experience stigmatization against their dialect in their context as upper-class citizens. However, it is the middle groups that lean more towards [l]. This pattern is not unreasonable, though unexpected. Given that higher income groups might be more likely to stigmatize others, they may perceive their specific version of Puerto Rican Spanish as standard, but may not have the [l] phenomenon because of their active rejection of it. Therefore, they would choose the [ɾ] end of the spectrum more often no matter what, since it would be both the standard and Puerto Rican Spanish variant in their perception. The only potential difference would be in the

No Info group. Lower income groups may choose [ɾ] more often because they more likely have the phenomenon and may experience more stigmatization for it, following the expected path of the main hypothesis. I would still expect them to lean more towards [ɾ] in the standard case, but this does not relate to the significant result of the Standard group, and there is in fact no significant data to argue for or against this.

The lack of significant difference in the tests based on location may be a good thing for the validity of the data collected. This means that the fact that the survey was distributed to both those in the states and on the island may not have affected the results in the way I thought it might. I thought the results might be different if only tested on the island as opposed to testing in the states as well, but this result shows that it is possible that living in Puerto Rico or in the states makes no difference on the perception Puerto Rican Spanish speakers have of their dialect. This may be because the marginalization of the dialect is felt in equal parts both on the island and the states. It would be worth running further tests with a larger, more balanced population from the island to see if this is the case, given that they were only about one-quarter of participants in this study.

## 7.2 Defining Standardness

I asked participants if they thought Puerto Rican Spanish was a standard dialect to see if there might be a difference between the way they answered the survey questions and their conscious thoughts about the question. There may not have been a conclusive answer based on the survey, but there were interesting results when they were asked the direct question. I was surprised that there were so many "yes" answers, since I had assumed more Puerto Rican Spanish speakers would answer "no" due to the stigmatization of Puerto Rican Spanish, especially in the states. It likely still has some effect, given that one-quarter still answered "no", which is not a negligible amount by any means. Chances are that stigmatization of their dialect is the reason many of these people answered no, but I would have expected a greater percentage to answer this way.

Even more surprising was the fact that nearly half of the participants in Puerto Rico answered "no" to the question about standardness, in contrast to about one-quarter of the participants in the states. I find this surprising because I would expect those in the states, who likely have much more contact with other dialects of Spanish, to experience stigmatization about their dialects from speakers of other dialects. I cannot explain why the answer is different within Puerto Rico, and the fact that I do not know the answer is part of why I ran this study to begin with. I know there is

widespread stigmatization about Puerto Rican Spanish in the states, but it is unclear whether the dynamic is different on the island. Is it more prevalent in Puerto Rico? Or at least more discussed there? These questions emphasize the need for further study on Puerto Rican Spanish, as well as Latin American dialects of Spanish in general.

This data also sparks the question: what does standardness really mean, in or out of linguistics? Outside of linguistics, people might claim it is the correct way to speak. In the survey, I did not define it for them. In linguistics, it can be defined as the dialect spoken by an elite group that can enforce its dialect as the correct way to speak. Whether or not this definition is enough, how far can a standard dialect travel? Maybe Standard American English is widespread enough that the whole country can be included in the area that experiences elitism based on that standard dialect. I'd argue that the same is almost certainly not true of Spanish as it is spoken in Spain, at least not for Latin America. Many speakers of Latin American dialects are interacting with either predominantly their own dialect, or a series of other, mostly Latin American dialects. I imagine that with time, the idea that the Spanish that is spoken in Spain is the correct way to speak has faded, and standardness can be redefined more locally. Maybe Puerto Rican Spanish speakers are more likely to consider Puerto Rican Spanish a standard dialect because it is truly the standard dialect of Puerto Rico, regardless of what the rest of the Hispanic world says. Though not everyone even just in Puerto Rico would agree with that concept, I think that there is room for standardness to have varying scopes over societies. There is certainly room for further discussion on this topic within the realm of linguistics, and further studies on the validity of standard dialects as a concept the way it currently exists.

# 8 Conclusion and Further Research

The most important point to be taken from this project is that there is a need for a series of further studies. Generally, further study needs to be done on the idea of dialect standardness. More specifically, there needs to be more testing done to find an answer to the original question posed: do Puerto Rican Spanish speakers perceive their dialect as standard or non-standard? This study was not able to find a definitive answer to the overall question. Relatively few parts of the data returned significant results for finding differences in answers among groups. Further dialect perception studies on Puerto Rican Spanish could lead to more insight on the topic and potentially a more conclusive answer.

Future versions of this study should first and foremost be tested on a larger

population. The assumption that the 20 people per survey that previous studies used would be enough was incorrect, potentially due to the addition of a third survey. Given that the averages were in the middle of the range and the standard deviations were around 1, more data might result in a clearer pattern.

A number of smaller changes might also improve the accuracy of results. The continuum, for example, was synthesized from recordings of the speaker saying [ele] and [eɾe], which are also the names of those letters in the Spanish alphabet. This is in contrast to previous studies, which simply made the sound of the letter, since those letters represented vowel sounds. That was not a possibility given the use of consonants, but it might be better practice not to use the name of the letter and use some other vocal material around the letter sounds of the continuum to avoid any effects based on the use of the letter name. Additionally, little information is known about how participants interacted with the continuum. Though I attempted to account for this by standardizing the data, there is a limited amount of tests that can be done with such information. Further efforts into the design of a continuum and study about the interactions participants have with it could be helpful for a dialect perception study such as this one.

Another important test that was unsuccessful within this study was that the control group was not significantly different from either of the groups. This means that there is still no proof that the assumption made by previous studies that such a group is not needed is a valid one. Future dialect perception studies, no matter what languages or dialect they study, should include a control group until we have valid reason to believe it is not necessary. If it is not significantly different from one of the groups, then it may be that we do not need a control. However, in the case that it is different from both, we may need to reconsider what the results of previous dialect perception studies really mean, depending on what that difference is. There is no harm in adding a third group for control. There is concern for the added cost and time, of course, but it is necessary to test if dialect perception studies are to continue this way.

Part of the benefit of distributing the survey through an online service, rather than going to Puerto Rico and collecting data in person, was supposed to be that bias effects based on perceptions about the experimenter would be eliminated. However, due to the consent form at the beginning which states that I am an investigator from Yale (though not necessarily that I am a student), and the fact that my name is listed might introduce bias based on any conscious or subconscious assumptions they may make. Any errors in the Spanish text that may have flagged that I am not a fully native speaker of Spanish could have also triggered some bias effect. It is hard to go beyond this scenario to eliminate bias due to the presence of one or many

researchers, but further attempts should be made to control for this.

Other limitations arose from the functionality of the online systems. In Qualtrics, making a system in which the audio autoplays and then moves to the next page was impossible. Limiting the amount of times someone can play an audio file also did not exist. Therefore, I could not control the number of times participants listened to each sentence. This mattered less for the continuum, which they heard plenty of times, but they would ideally hear each sentence the same number of times as the rest of the participants. It is hard to tell if this had an effect on the results, but if such an effect were to exist, it could be eliminated in a system that only plays each test sentence once and plays the continuum once per question. mTurk was also an imperfect system for distributing the survey the way I intended. I initially distributed the survey to only mTurk users in Puerto Rico. I received no responses to this survey. However, when I redistributed the survey across the United States, I received 16 responses from participants in Puerto Rico. Because of this, I was not only dissatisfied with the way mTurk was working, but also became more concerned with how future studies might be affected, and confused at how they decided who was in Puerto Rico and who wasn't. I did not miss out on testing exclusively in Puerto Rico because of external issues such as electricity loss from the hurricane like I thought, but rather because mTurk does not have a functioning system for deciding who is in Puerto Rico. Such a roadblock certainly has the ability to stifle further research on Puerto Rico in many fields. Based on these struggles with online survey distribution, I would recommend that future studies do such testing in person if possible.

Given that I never defined standard for the participants, but asked them if they thought Puerto Rican Spanish was a standard dialect, it might've been interesting to hear how they define standardness. I also could have defined it given the linguistic definition and asked if their answers changed. These were missed opportunities that could be remedied in a future study.

The results of this study are ultimately inconclusive, but there is a promising amount of further study to be done before the topic can be abandoned. Puerto Rico is a special case for study, both because it is a relatively small population with one major dialect in its dominant language, and also because it is a place that is understudied across many and most fields. I hope that future studies continue to examine the uniqueness of the island, as well as the complicated nature of standardness and dialect perception.

# 9 Appendices

## 9.1 Test Sentences

The following tables include all the sentences used in the survey, along with translations and IPA transcriptions of the target word in each sentence as they would be said by a Puerto Rican Spanish speaker. Guidance for transcriptions was provided by the phonological description of Spanish available in Whitney (2002).

### 9.1.1 Practice Sentences

| Sentence | Gloss | IPA Transcription |
|---|---|---|
| Elena necesita **papel** para hacer su tarea. | Elena needs paper to do her homework. | [papel] |
| A Sebi no le gusta usar loción en su **piel** porque se siente extraña. | Sebi doesn't like to use lotion on his skin because it feels strange. | [pjel] |
| Emma no sabe como **pronunciar** palabras en alemán. | Emma doesn't know how to say words in German. | [pr̃õnũnsjal] |
| Alejandro necesita **ganar** este juego de béisbol. | Alejandro needs to win this game of baseball. | [be͡isbol] |

### 9.1.2 Task Sentences

| Sentence | Gloss | IPA Transcription |
|---|---|---|
| Diego quiere **ser** un ingeniero mecánico. | Diego wants to be a mechanical engineer. | [sel] |
| Miguel siempre necesita **tener** el juego nuevo. | Miguel always needs to have the new game. | [tẽnel] |
| Paola preparó la comida para **comer** con su familia. | Paola prepared dinner to eat with her family. | [kõmel] |
| Ana está intentando **aprender** alemán. | Ana is trying to learn German. | [apr̃ẽndel] |
| Jesús nunca quiere **ver** películas románticas. | Jesús never wants to watch romantic movies. | [βel] |
| José abrió una tienda para **vender** obras de ficción. | Jose opened a store to sell fiction books. | [βẽndel] |

| | | |
|---|---|---|
| Mari no puede **leer** mucho antes de acostarse. | Mari can't read much before bed. | [leːl] |
| Cami no quiere **romper** su televisión. | Cami doesn't want to break her television. | [rõmpel] |
| Carlos no pudo **entender** su libro de química. | Carlos couldn't understand his chemistry book. | [ẽntẽndel] |
| Esperanza va a **obtener** un nuevo pasaporte. | Esperanza is going to get a new passport. | [oβtẽnel] |
| Juan necesita **escoger** un aspirante para el puesto de director. | Juan needs to choose an applicant for the position of director. | [eskoxel] |
| César no sabe cuando **hacer** su tarea. | César doesn't know when to do his homework. | [asel] |
| Ricardo escribió una carta para **agradecer** a su abuela por el regalo. | Ricardo wrote a letter to thank his grandmother for the gift. | [aɤɾeðesel] |
| Pati no puede esperar a **volver** a su casa. | Pati can't wait to return to her house. | [βolbel] |
| Milo paró de **responder** a mis mensajes. | Milo stopped responding to my messages. | [ɾespõndel] |
| Antonio trata de **conocer** a todos sus estudiantes. | Antonio tries to get to know all of the students. | [kõnosel] |
| Victoria no tiene nada que **perder** porque ya lo perdió todo. | Victoria has nothing to lose because she already lost it all. | [peldel] |
| Sara tiene que **beber** agua con su cena. | Sara needs to drink water with her dinner. | [βeβel] |
| Sofía no prefiere **depender** de sus padres. | Sofia doesn't prefer to depend on her parents. | [ðepẽndel] |
| Gabi hizo un flan para **traer** a la fiesta. | Gabi made a flan to bring to the party. | [tra(j)el] |

### 9.1.3 Filler Sentences

| Sentence | Gloss | IPA Transcription |
|---|---|---|
| Los prisioneros intentaron escapar de la **cárcel**, pero no tuvieron éxito. | The prisoners tried to escape from jail, but were not successful. | [kalsel] |

| | | |
|---|---|---|
| Doctores no recomiendan dar **miel** a los bebés. | Doctors do not recommend giving honey to babies. | [mjel] |
| Santiago va a un **hotel** en Ponce para el fin de semana. | Santiago is going to a hotel in Ponce for the weekend. | [otel] |
| Algunas ciudades tienen un **nivel** alto de educación. | Some cities have a high level of education. | [niβel] |
| Marta está lista para plantar un **árbol** enfrente de la casa. | Marta is ready to plant a tree in front of the house. | [albol] |
| La girafa es el **animal** favorito de Adriana. | The giraffe is Adriana's favorite animal. | [ãnĩmal] |
| Isabel va al **hospital** para tener una cirugía en su pierna. | Isabel is going to the hospital to have foot surgery. | [ospital] |
| Jorge, pásame la **sal** y la pimienta, por favor. | Jorge, please pass me the salt and pepper. | [sal] |
| Alán usó el **metal** para crear una máquina nueva. | Alán used the metal to create a new machine. | [makĩna] |
| Eduardo está aprendiendo como ser más **profesional** para su empresa. | Eduardo is learning how to be more professional for his business. | [profesjõnal] |

## 9.2 Additional Data

| Variable | No Info | PRS | Standard | Combined |
|---|---|---|---|---|
| Main Effect | N/A | N/A | N/A | 0.680 |
| Question | 0.835 | 0.758 | 0.654 | 0.651 |
| Gender | 0.944 | 0.176 | 0.096 | 0.774 |
| Age | 0.273 | 0.815 | 0.018 | 0.266 |
| Income | 0.135 | 0.281 | 0.024 | 0.083 |
| Location | 0.715 | 0.605 | 0.575 | 0.346 |
| Standardness | 0.341 | 0.716 | 0.670 | 0.881 |

Table 10: The p-values of the ANOVAs and t-tests done on the combined data across surveys and within each survey

| Question Number | No Info | PRS | Standard |
|---|---|---|---|
| 1 | 0.067 | -0.061 | -0.076 |
| 2 | -0.167 | -0.247 | 0.048 |
| 3 | -0.139 | 0.260 | -0.135 |
| 4 | -0.126 | -0.068 | -0.232 |
| 5 | 0.232 | -0.022 | -0.288 |
| 6 | 0.061 | 0.178 | -0.088 |
| 7 | 0.176 | 0.197 | 0.326 |
| 8 | 0.127 | -0.245 | 0.165 |
| 9 | -0.153 | -0.139 | -0.106 |
| 10 | -0.148 | 0.169 | 0.452 |
| 11 | 0.157 | -0.013 | 0.028 |
| 12 | 0.401 | -0.183 | -0.124 |
| 13 | -0.115 | -0.234 | -0.203 |
| 14 | -0.036 | -0.035 | -0.021 |
| 15 | -0.453 | -0.157 | 0.388 |
| 16 | 0.041 | 0.193 | -0.018 |
| 17 | 0.091 | 0.269 | -0.224 |
| 18 | 0.011 | -0.140 | -0.148 |
| 19 | 0.229 | 0.093 | 0.044 |
| 20 | -0.235 | 0.184 | -0.08496 |

Table 11: Standardized means of each question by survey

| Income | No Info | PRS | Standard | Total |
|---|---|---|---|---|
| Less than $20,000 | 3 | 2 | 1 | 6 |
| $20,000 - $39,999 | 5 | 4 | 7 | 16 |
| $40,000 - $59,999 | 5 | 7 | 6 | 18 |
| $60,000 - $79,999 | 4 | 5 | 3 | 12 |
| $80,000 - $99,999 | 0 | 3 | 0 | 3 |
| $100,000 or more | 0 | 2 | 2 | 4 |
| Total | 17 | 23 | 19 | 59 |

Table 12: The breakdown of primary household income by survey

| Age Group | No Info | PRS | Standard | Total |
|---|---|---|---|---|
| <25 | 6 | 7 | 4 | 17 |
| 26-29 | 3 | 6 | 7 | 16 |
| 30-33 | 6 | 3 | 6 | 15 |
| >34 | 2 | 7 | 2 | 13 |
| Total | 17 | 23 | 19 | 59 |

Table 13: The breakdown of age ranges by survey

| Answer | No Info | PRS | Standard | Total |
|---|---|---|---|---|
| Yes | 11 | 15 | 15 | 41 |
| No | 6 | 8 | 4 | 18 |
| Total | 17 | 23 | 19 | 59 |

Table 14: The breakdown of answers to the question "In your opinion, is Puerto Rican Spanish a standard dialect?" by survey

## 9.3 Sample Question

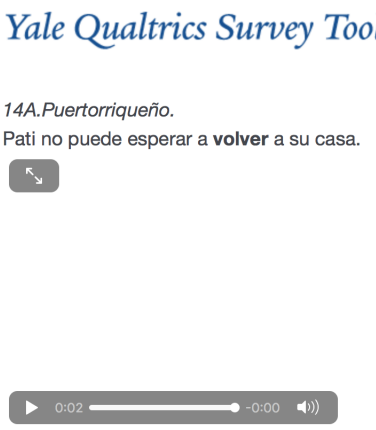


Figure 6: An example of the first page a participant would see in a trial. The sentence is displayed and an audio recording of the sentence is played.

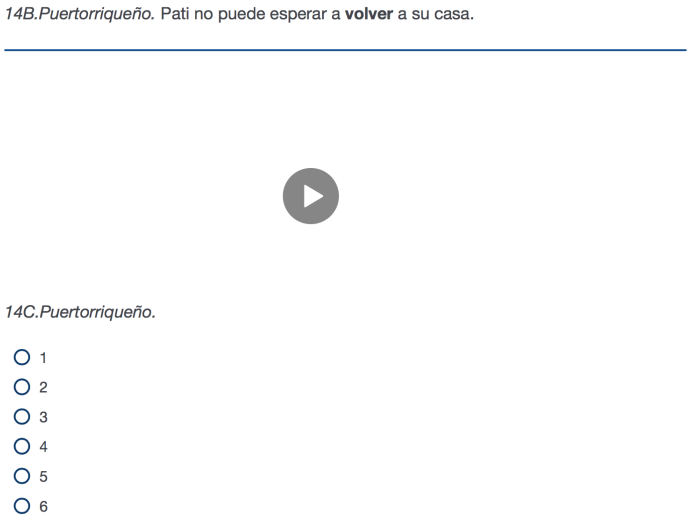


Figure 7: An example of the second page a participant would see in a trial. The sentence is displayed and an audio recording of the continuum is played. The participant selects the number that corresponds to the sound they have chosen.

# References


[1] Figueroa, Nicholas J. 2017. *Variation in /ɾ/ Outcomes in Speech of U.S.-Born Puerto Rican and Dominican Heritage Spanish Speakers*. Albany, NY: State University of New York dissertation.

[2] Hay, Jennifer, Aaron Nolan & Katie Drager. 2006. From *fush* to *feesh*: Exemplar priming in speech perception. *The Linguistic Review* 23. 351-379.

[3] Niedzielski, Nancy. 1995. Acoustic analysis and language attitudes in Detroit. In M. Meyerhoff (ed.), *(N)Waves and means: University of Pennsylvania working papers in lingusitics*, vol. 3, 73-86. Philadelphia: University of Pennsylvania Press.

[4] Niedzielski, Nancy. 1997. *The effect of social information on the phonetic perception of sociolinguistic variables*. Santa Barbara, CA: University of California, Santa Barbara dissertation.

[5] Niedzielski, Nancy. 1999. The effect of social information on the perception of sociolinguistic variables. *Journal of Language and Social Psychology* 18(1). 62-85.

[6] Johnson, K. 1996. Speech perception without speaker normalization. in Johnson, K. and Mullennix (eds.) *Talker Variability in Speech Processing*. San Diego. Academic Press.

[7] Simons, Gary F. and Charles D. Fennig (eds.). 2018. *Ethnologue: Languages of the World, Twenty-first edition*. Dallas, Texas: SIL International. Online version: http://www.ethnologue.com.

[8] U.S. Census Bureau. 2016. *2012-2016 American Community Survey 5-Year Estimates, Language Spoken at Home*. https://factfinder.census.gov/

[9] U.S. Census Bureau. 2016. *2016 American Community Survey 1-Year Estimates, Hispanic or Latino Origin by Specific Origin*. https://factfinder.census.gov/

[10] Valentín-Márquez, Wilfredo. 2007. *Doing being boricua: Perceptions of national identity and the sociolinguistic distribution of liquid variables in Puerto Rican Spanish*. Ann Arbor, MI: University of Michigan dissertation.

[11] Whitney, M. Stanley. 2002. *Spanish/English Contrasts: A Course in Spanish Linguistics, Second Edition*. Washington, DC: Georgetown University Press.