Cross-Dialectal Phonetic Variation and Lexical Encoding: Evidence from /s/ Perception in Seville Capital

by

Angela Krak Swain

BA in Spanish, Seton Hill University, 2018

MA in Hispanic Linguistics, University of Pittsburgh, 2020

Submitted to the Graduate Faculty of the

Dietrich School of Arts and Sciences in partial fulfillment

of the requirements for the degree of

Doctor of Philosophy

University of Pittsburgh

2023

UNIVERSITY OF PITTSBURGH

DIETRICH SCHOOL OF ARTS AND SCIENCES

This dissertation was presented

by

Angela Krak Swain

It was defended on

July 12, 2023

and approved by

Matthew Kanwit, Associate Professor, Department of Linguistics

Daniel Villarreal, Assistant Professor, Department of Linguistics

Whitney Chappell, Associate Professor, Department of Modern Languages and Literatures

Dissertation Director: Melinda Fricke, Assistant Professor, Department of Linguistics

Cross-Dialectal Phonetic Variation and Lexical Encoding: Evidence from /s/ Perception in Seville Capital

Angela K. Swain, PhD

University of Pittsburgh, 2023

Compelling evidence on the role of social factors in speech perception has led scholars to posit that linguistic and social information is processed simultaneously and that there is interaction between them (Sumner et al., 2014). The current study addresses multiple gaps in this literature by examining dialect classification, evaluation, and auditory lexical processing of three varieties of Peninsular Spanish (Seville capital, Seville outskirts, and central Spain). Seville capital participants (N = 42) completed a verbal guise, long-term form priming task, brief interview, and language questionnaire. Materials contained words with three variant pronunciations of /s/ associated with different levels of perceived prestige: central Spain (national standard, [s̪]), Seville capital (regional standard, [s̪]), and Seville outskirts (nonstandard, [s̪$^{\theta}$]) (Penny, 2004).

Verbal guise results reflected high dialect classification accuracy overall, but better classification of central talkers and more confusion between Andalusian varieties. Dialect evaluation results supported the attested prestige hierarchy for Peninsular varieties (e.g., Penny, 2004; Santana Marrero, 2022), with central voices evaluated highly for status categories (e.g., formal), while Seville capital voices received high ratings for solidarity categories (e.g., pretty). Results from the auditory word recognition task revealed flexibility in immediate processing of the three variants; /s/ words produced by all talkers were processed similarly. However, there was evidence of greater processing difficulty for the Seville outskirts variant. Changes in recognition of /s/ words after a 20-minute delay provided preliminary evidence for differential encoding, with

processing benefits associated with talkers from Seville capital and central Spain. Additionally, results suggested that different processing strategies affected the /s/ words and control words, with findings interpreted through the lens of bottom-up versus top-down processing.

Different word recognition results for the Seville woman talker compared to the Seville man corresponded with dialect classification and evaluation patterns from the sociophonetic task, reflecting the power of social information in linguistic processing and the importance of combining sociophonetic and psycholinguistic tasks in the same experimental sequence. In summary, this dissertation contributes to the conversation of how sociophonetic variants are evaluated and processed by demonstrating the clear interaction between social and linguistic information in both implicit and explicit task measures.

Key words: spoken word recognition, sociophonetics, Peninsular Spanish, language attitudes

Table of Contents

List of Tables

List of Figures

Acknowledgements

This dissertation is not only the compilation of five years of hard work, but also the result of years of support from the irreplaceable people in my life. I am extremely grateful to my family, friends, mentors, and colleagues for helping me reach the finish line.

To my family:

To Peter, my partner and better half, thank you for brightening each day and making the challenging parts of this process so much more bearable.

To my incredible parents, thank you for your unwavering encouragement and presence during the many different phases of this journey.

Thank you to my three older brothers, Justin, Jonathan, and Andrew, for being the best role models and always giving me the biggest shoes to fill.

A huge thank you to Reina, Diego, and Dante for being the most supportive doggos and providing the best stress relief.

To my mentors:

I would like to start by thanking Melinda Fricke, my phenomenal advisor. I have learned so much from you as a scholar and even more as a person. Thank you for investing countless hours of your time into my development and for being an endless source of wisdom. I am forever grateful for the risk you took in advising me, and without you, this project would not have come to fruition.

To Matt Kanwit, thank you for consistently helping me grow as both a person and a scholar over the past five years. It often felt like I had two academic advisors as a result of your dedication to helping me develop my work beyond the classroom.

To my dissertation committee: Melinda Fricke, Matt Kanwit, Dan Villarreal, and Whitney Chappell. Thank you for making what can be a grueling process for PhD candidates be filled with nothing but constructive and insightful feedback. I know you had absolutely no free time to spare, and I am grateful to have been supported and encouraged by all four of you.

To the professors in Linguistics Department at the University of Pittsburgh, thank you for helping me learn and deepen my knowledge of linguistics in such a positive environment.

To my friends:

Thank you to the LingGrads for making it so easy to feel like I belonged in the program from the beginning. I am very grateful for your friendship and mentorship over the past five years.

To my STL friends, thank you for your unwavering companionship and for never asking me when I would be finally done with my PhD.

To my Seville capital friends and participants:

Gracias a Nuria por su ayuda constante y amistad preciosa. Este proyecto no hubiera sido posible realizar sin ella.

Gracias a Mariana por compartir su sabiduría y por haber hecho que me sintiera en casa en Sevilla desde el principio.

Y para concluir: estoy TAN agradecida a todos los sevillanos que participaron en este proyecto. Mil gracias por haber compartido vuestras perspectivas conmigo.

1.0 Introduction

A growing body of psycholinguistic research has investigated the interaction of social and linguistic variables in speech perception. Traditional models of speech perception emphasized abstraction and did not consider social information about the speech signal to play a significant role in processing, instead treating it as noise that needed to be filtered out by the listener (e.g., Jakobson, Fant, & Halle, 1951). This perspective began to shift in the 1990s as a number of studies produced compelling results in support of listeners retaining phonologically non-contrastive acoustic details in memory (e.g., Bradlow, Nygaard, & Pisoni, 1999; Church & Schacter, 1994; Goldinger, 1996; Palmeri, Goldinger, & Pisoni, 1993). These findings led researchers to apply exemplar models of categorization and memory representation (e.g. Hintzman, 1986; Nosofsky, 1988) to speech perception. Contrary to abstract models of speech perception that emphasized reduction of acoustic variability in auditory lexical processing (e.g. models such as TRACE; McClelland & Elman, 1986), *exemplar* theory argues that acoustic detail is stored in memory and that surface variation can be beneficial in processing (Goldinger, 1998; Johnson, 1997; Pierrehumbert, 2001).

Subsequent work transitioned from investigating talker-specific variability to perception of within-dialect surface variants (e.g., McLennan, Luce, & Charles-Luce, 2003; Sumner & Samuel, 2005), a natural step in striving to understand how listeners efficiently accommodate different types of variation associated with the speech signal. Other investigations expanded to consider the treatment of cross-dialectal variants, leveraging the familiarity of the varieties to the listeners (e.g., Clopper, Tamati, & Pierrehumbert, 2016; Floccia, Goslin, Girard, & Konopczynski, 2006; Sumner & Kataoka, 2013; Sumner & Samuel, 2009). This progression of investigation has

led to fruitful conversations about the role of frequency in word recognition, as the most straightforward predictions following from exemplar theory place a heavy emphasis on frequency at all levels of representation (i.e., frequency of words, variants, varieties, etc.). However, a number of results have not been easily accommodated by traditional exemplar models, especially regarding the equivalent processing of multiple surface variants regardless of their frequency of exposure (e.g., Sumner & Samuel, 2005), as well as the apparently preferential treatment of standard varieties in memory (e.g., Clopper et al., 2016; Sumner & Kataoka, 2013).

While early models of exemplar theory primarily focused on the treatment of talker-specific acoustic variability, it was predicted that the models could be extended to account for variability related to social categories such as regional dialects and talker gender (Johnson, 1997; Pierrehumbert, 2001). Additional lines of research have explored the effects of social information on lexical processing, with results demonstrating that the mere activation of social concepts in the experimental environment can affect a participant's perception of who they think they are listening to (e.g., Hay & Drager, 2010; Hay, Nolan, & Drager, 2006a; Niedzielski, 1999). This work suggests that social and linguistic information are both represented in memory and that accessing social categories can affect the perception of linguistic categories (i.e., vowel perception in Hay & Drager, 2010). Updates to traditional exemplar theories have been proposed to account for the role of social information in lexical processing, with Sumner, Kim, King, and McGowan (2014) positing that social and linguistic information are processed simultaneously and in reference to one another.

Cross-dialectal auditory lexical processing has been both a recent and productive topic of investigation, as it provides a lens to explore the interaction of social and linguistic information by testing recognition and encoding of socially meaningful variants. While this area of investigation

has grown rapidly over the last decade, clear gaps persist in the body of literature. First, the vast majority of work focuses on English varieties, with only two recent studies considering perception of Spanish variants (e.g., Gylfadóttir, 2018; López Velarde & Simonet, 2019). Any linguistic theory must be tested across multiple contexts (i.e., variants, varieties, and languages) to ensure that its assumptions are generalizable. Recent findings from Clapp, Vaughn, and Sumner (2023) highlight the importance of this statement, as the inclusion of more diverse talker voices was found to diminish priming effects. Moreover, given that not all talkers associated with the same varieties are perceived equally (e.g., Clapp et al., 2023), it is important to understand how listeners respond to the specific talkers that appear in experimental tasks and how these differences may affect the interpretation of psycholinguistic data. As a result, an additional gap in the current literature is the widespread failure of psycholinguistic experimentation to incorporate dialect classification and evaluation measures.

Finally, while evidence from immediate word recognition tasks suggests that multiple variants can be accommodated by experienced listeners, as measured by similar response times and/or lexical acceptability (Sumner & Samuel, 2009; López Velarde & Simonet, 2019), this pattern in immediate recognition does not necessarily capture the structure of listeners' long-term representations. Though long-term priming paradigms have not been as heavily researched in cross-dialectal speech perception, findings to this point suggest that the canonical variant, or one that is most associated with a perceived standard, is prioritized in memory (Clopper et al., 2016; Sumner & Samuel, 2009; Sumner et al., 2014). An important remaining question therefore concerns the long-term representation of socially salient dialectal variants and how this representation may be affected by factors such as standardness, the experience of the listener, and the broader sociopolitical context in which the listeners, talkers, variants, and varieties coexist.

My dissertation addresses these gaps by implementing a combination of sociophonetic (i.e., Verbal Guise) and psycholinguistic (i.e., Long-Term Form Priming) tasks to explore dialect classification, evaluation, word recognition, and encoding of three varieties of Peninsular Spanish: Seville capital, Seville outskirts, and central Spain (i.e., Madrid and Castilla-La Mancha). Items in these tasks feature a novel operationalization of three word-initial /s/ variants, each associated with phonetic, geographic, and social differences (i.e., [s̪], central Spain, national standard; [s̺], Seville capital, regional standard; [s̺$^{\theta}$], Seville outskirts, nonstandard). While /s/ variation is a prolific topic of investigation in the field of Hispanic Linguistics (e.g., Colantoni, 2011), with Peninsular variants receiving ample attention in production and perception literature (e.g., coda /s/ weakening; *seseo*, *ceceo*, *distinción* 'distinction'; e.g., Hernandez-Campoy & Villena-Ponsoda, 2009; Iglesias, 2003; Penny, 2004; Regan, 2017a, 2019, 2022; Samper Padilla, 2011; Santana Marrero, 2016-2017, 2018b, 2022), the word-initial /s/ variants in the current study provide a unique opportunity to explore the interaction of social and linguistic information in processing, as they are more reliably associated with the geographic locations and varieties of interest. Moreover, the relationship between Andalusian and north-central varieties from the perspective of media representation is complex; north-central varieties are most frequently featured in national media, while many phonetic features of Andalusian varieties are neutralized by southern news anchors and politicians (e.g., Cruz Ortiz, 2020; León-Castro, 2016; Santana Marrero, 2022; Penny, 2004).

The dissertation will overview key literature in chapter 2, ending with an introduction of the research questions addressed in the current study. Chapter 3 includes an overview of the task methodology and descriptions of talker and listener demographic information. Next, analysis chapter 4 focuses on the Verbal Guise, outlining findings from dialect classification and evaluation portions of the task. Analysis chapter 5 focuses on the Long-Term Form Priming task results, with

separate analyses of block one and block two word recognition performance. Discussion chapter 6 includes a summary of noteworthy results from both tasks, extending the implications of the findings from the sociophonetic and psycholinguistic tasks to consider the role of the experimental sequence. Finally, conclusions, limitations, and future directions are provided in chapter 7.

2.0 Literature Review

In this chapter, I begin by outlining the concepts of word recognition, encoding, and representation in auditory lexical processing. In Section 2.1, I situate the study of spoken word recognition within exemplar models of speech perception (subsection 2.1.1). I also introduce the differences between immediate, short-term, and long-term priming paradigms, outlining key constructs and establishing the theoretical landscape of the literature (2.1.2). Section 2.2 focuses on the processing of within-dialect variants, emphasizing previous work that has considered the question of frequency of exposure. Section 2.3 then surveys recent work on auditory word recognition in cross-dialectal contexts, paying special attention to the differences between immediate and long-term processing. Section 2.4 highlights trends in production for salient Peninsular /s/ phenomena, with the goal of motivating the dialectal variants that are the focus of the present study. Returning again to speech perception, Section 2.5 outlines perception of sociophonetic variants in the Spanish-speaking world. Focusing on Seville, Spain, Section 2.6 details dialect classification, evaluation, and lexical processing work that has focused on the capital city of Andalusia. Finally, Section 2.7 summarizes the gaps in the literature that this study has been designed to address and presents the research questions for this dissertation.

2.1 Perceptual Processes within Exemplar Models of Speech Perception

2.1.1 Spoken Word Recognition and Exemplar Theory

The perception of spoken language can be discussed in terms of three main components: word recognition, encoding, and representation. In line with the general theoretical framework of this dissertation, these components will be outlined from the perspective of exemplar theory. In exemplar models of speech perception, the cognitive representation of language is comprised of detailed memories of linguistic episodes, where a linguistic episode is generally conceptualized as corresponding to a listener's experience of hearing a single word (Goldinger, 1998; Johnson, 1997, Pierrehumbert, 2001). When a listener hears a spoken word, the token activates other instances (i.e., exemplars) that were previously stored in memory, and word recognition occurs when the acoustic signal is accurately matched with the most similar tokens in memory (Goldinger, 1998; Johnson, 1997, Pierrehumbert, 2001). Exemplar theory assumes that a given word's representation in the lexicon consists of a cloud of all exemplars that have been previously recognized. As a result, frequency of exposure is predicted to play a critical role in word recognition, leading to more frequently experienced episodes being robustly represented in memory. An example of how a frequency effect could manifest during word recognition is via processing speed, with faster recognition times occurring when a speaker has repeatedly encountered a given word (Pierrehumbert, 2001; Radeau, Morais, & Segui, 1995; Sumner & Samuel, 2005). For example, a Pittsburgh native listener would likely have a large number of exemplars for the word ‘yinz’ as a result of its frequent use in the city (Johnstone et al., 2006).

After successful word recognition, the token is *encoded*, or stored, in the cloud of exemplars associated with the corresponding lexical category. Continuing with the Pittsburgh

example, each time a listener hears the word ‘yinz,’ the token would activate the numerous prior exemplars that were previously recognized and encoded as ‘yinz.’ An additional assumption inherent to exemplar models is that the encoding of a given token is highly phonetically detailed, and it can also include a variety of social and contextual information. Thus, while the phonetic information that comprises the word would be available and activated during word recognition, research has also demonstrated the powerful influence of social information in lexical processing, especially relating to regional associations (e.g., Hay & Drager, 2010). In this way, social categories related to ‘yinz’ would also be activated, such as dialect (e.g., Pittsburgh), information about the talker and the talker’s gender (e.g., Dad), and other social and contextual information that is potentially unique to the listener’s cognitive space (Johnson, 1997; Pierrehumbert, 2001).

Finally, *representation* in this framework corresponds to the set of all exemplars that have been previously recognized and encoded (Clopper et al., 2016). It is this component of word recognition that relates to long-term memory, in that the long-term memory representation of a word is comprised of the linguistic episodes that a listener has identified and stored. With the assumption that the listener is from Pittsburgh, the lexical representation of ‘yinz’ in memory would consist of a highly populated exemplar cloud that includes a dense network of links to corresponding related social categories as a result of having often heard the term (Johnson, 1997; Pierrehumbert, 2001).

2.1.2 Priming Paradigms in Spoken Word Recognition

A large body of research has sought to investigate how variation in the speech signal affects lexical processing, especially given that in the past, variation was considered to be a hindrance to accurate speech perception (Goldinger, 1998). Recall that abstract models of speech perception assume that prior to word recognition, the speech signal is filtered to only include the most essential information, thus discarding many factors that constitute acoustic and social variability (e.g., Norris, McQueen, & Cutler, 2003). In exemplar models, however, social information such as the voice of the speaker is assumed to be stored in great detail (Goldinger, 1998, Johnson, 1997).

Priming paradigms are commonly implemented to examine implicit memory, which is the subconscious recollection of previously presented items (Church & Schacter, 1994). In the context of exemplar models of speech perception, the ability of a given word token to 'prime' other tokens of the same word can help reveal the extent to which representations contain overlapping information. Priming effects from repetition-based paradigms manifest as quicker reaction times (RTs) and better accuracy for previously encountered items when compared to newly presented items, a result of the repeated activation of form-based representations leading to faster processing (Church & Schacter, 1994; McLennan et al., 2003). Priming paradigms examining implicit memory can further be divided into two categories: immediate and long-term priming. Immediate priming refers to a small amount of time passing between the presentation of a prime and target word, such as under one second (e.g. López Velarde & Simonet, 2019). On the contrary, long-term paradigms include a much larger gap between the primes and targets (e.g., 20 minutes). A crucial difference between them, then, is that immediate priming tasks are thought to tap into immediate word recognition processes, while long-term priming tasks provide information about whether and

how previously perceived tokens were encoded into long-term memory (Sumner & Samuel, 2009), as only items that were encoded into memory after their first presentation can possibly have an effect on items presented several minutes later.

Prior research has supported the notion that auditory priming is sensitive to specific acoustic representations (e.g., vocal similarity, specific talker voices; Church & Schacter, 1994; Palmeri et al., 1993). These early findings inspired subsequent work exploring the effects of phonetic variability in immediate and long-term auditory lexical processing. Within priming paradigms, there is also a distinction between form priming and semantic priming. Form priming is repetition-based and is used by researchers to examine the dimensions of similarity between variable surface structures, as only representations that overlap in form can prime one another. This is useful in exploring whether listeners respond in similar ways to multiple variants of the same lexical item (Sumner & Samuel, 2009). In auditory form-priming paradigms, listeners are presented with a prime, immediately followed by a phonologically related target. Priming is typically measured as a function of reaction time and occurs when listeners respond faster to either identical or similar targets to the prime when compared to control (i.e., phonologically different) items (Radeau et al., 1995).

Semantic priming, on the other hand, tests whether different variants can activate the same lexical representation, thus probing a deeper (i.e., meaning-based) level of processing. In this paradigm, priming occurs when participants respond more quickly to semantically related sets of prime-target pairs (e.g., king-queen) than unrelated pairs (e.g., bell-queen). For example, Andruski Blumstein, and Burton (1994) manipulated voice onset time (VOT) of word-initial stops to examine the effects of acoustic variability on lexical processing. They found slower lexical decision times when primes contained a voiced counterpart (e.g., pet-bet), leading to the hypothesis

that phonologically similar competitors are also activated during the initial stages of lexical processing.

Now having established the relevant theoretical and methodological background, the following sections will provide an overview of how these constructs and paradigms have been used in the context of the study of within-dialect phonetic variation.

2.2 Within-Dialect Variation and Word Recognition Processes

A fundamental question that spans both early and current investigations of spoken word recognition relates to how listeners are able to cope with the vast amount of variation present in the acoustic signals we hear on a daily basis. Spoken words are highly dynamic in nature, yet despite the multitude of segmental (i.e., use of multiple surface variants) or suprasegmental (i.e., changes in prosody or speech style) adjustments that can accompany production, comprehension still ensues with seemingly little processing cost. To better understand the nature of spoken word recognition, early research primarily explored the question of whether talker-specific information was represented in memory for listeners (e.g., Bradlow et al., 1999; Church & Schacter, 1994; Goldinger, 1996; Palmeri et al., 1993).

In a study that greatly influenced the trajectory of the field, Palmeri et al. (1993) examined whether word recognition was affected by the presentation of stimuli in the same vs. different voices (i.e., 'talker specificity'). The authors implemented a continuous repetition memory paradigm, in which participants identified whether each lexical item was old (i.e., repeated) or

new. Palmeri et al. (1993) reasoned that if any acoustic information about the talker was encoded, then participants should demonstrate faster and more efficient recognition of old items when the talker who produced the initial token remained constant. The results indicated that indeed, a processing benefit (i.e., faster RTs and higher accuracy scores) was associated with the presentation of a repeated item in the same voice when compared to a new voice. Moreover, listeners were slower to recognize old items when they were presented in a new voice. These findings led Palmeri et al. (1993) to conclude that acoustic details were encoded in memory.

Subsequent research in the speech perception literature replicated talker specificity effects across different paradigms and contexts (e.g., Bradlow et al., 1999; Church & Schacter, 1994; Goldinger, 1996). For instance, Church and Schacter (1994) found that auditory priming effects were sensitive to acoustic representation, demonstrating faster recognition of repeated items that contained subtle f0 adjustments. In another study, Goldinger (1996) found talker specificity effects, in that words repeated by the same person were recognized more efficiently than those produced by a new talker, also demonstrating a repetition effect for acoustic similarity across voices. Lastly, Bradlow et al. (1999) replicated the processing benefit for recognition of old items presented by the same talker but also found repetition effects according to speech rate. In other words, recognition of old items was more efficient when the speech rate matched that of the first token presentation (i.e., slow-slow, fast-fast) as opposed to a mismatched condition (e.g., slow-fast).

After results from numerous early studies on spoken word recognition corroborated the hypothesis that listeners encode acoustic details such as talker shifts, speech rates, and f0 changes, the question of how phonetic and phonological alternations are recognized and encoded became a topic of interest. Research on perception of linguistic variants has included both free variation

(e.g., Sumner & Samuel, 2005) and allophonic variation (e.g., McLennan et al., 2003). Compared to allophonic variants that exist in complementary distribution, free variation is considered to be less constrained by the linguistic context, though other factors can affect the use of one variant over another (e.g., perceived formality; Sumner & Samuel, 2005).

Speech perception research on free variation in English has examined word-final /t/ production (e.g., Sumner & Samuel, 2005) and nasal assimilation (e.g., Gow, 2001), among other variable phonetic contexts. In Sumner and Samuel (2005), the authors investigated the perception of word-final /t/ to determine if all variants were equally effective in priming. The three /t/ variants under investigation were canonical (e.g., *flute* [flut]), coarticulated (e.g., [fluʔt]), and glottal (e.g., [fluʔ]). Critically, these variants were associated with different frequencies of exposure, with the coarticulated variant most commonly occurring in casual speech. Through the implementation of a semantic priming task, Sumner and Samuel (2005) found that all three variants facilitated priming (e.g., participants were equally fast in recognizing the semantically related target word 'music' when it was preceded by [flut], [flu?t], and [flu?]). The authors concluded that variation did not hinder semantic activation in spoken word recognition and that no processing benefits emerged between a canonical variant and the most frequent form (i.e., the RTs for targets preceded by [t] did not significantly differ from each other), given that all three variants were acceptable in the linguistic context.

In a series of short-term priming experiments, McLennan et al. (2003) examined the processing of flaps in American English (e.g., the use of [ɾ] instead of intervocalic [t] and [d]) to test whether casual phonetic variants activate their underlying phonemes. In other words, would an ambiguous stimulus such as [æɾəm] activate/prime both 'atom' and 'Adam'? Importantly, the use of flapping in American English is primarily associated with casual speech, while production

of /t/ and /d/ stops most commonly occurs in careful speech. Findings from a series of repetition priming tasks in McLennan et al. (2003) indicate that flapped variants were effective in priming carefully produced variants, while carefully produced variants also primed flaps. One interesting outcome when task difficulty was manipulated was that only the difficult lexical decision task facilitated equal priming for flaps and stops. More specifically, repetition effects were only found when the nonwords featured in the experiment were phonologically plausible (e.g., '*bacov')* when compared to pseudowords with phonemically impossible patterns in English (e.g., '*thushthudge').* The authors reasoned that the enhanced task difficulty required a deep level of processing as a result of the nonwords activating similarly stored lexical items, ultimately leading to priming effects. These results are supported by Andruski et al. (1994) in which reaction times were slowed for items that contained phonological lexical competitors (e.g., pet-bet). It should be noted that while McLennan et al. (2003) referred to their experiments as 'long-term repetition priming,' blocks of primes and targets only included 24 items, which according to some (e.g., López Velarde & Simonet, 2019) would not introduce a sufficiently long lag to be truly considered as 'long-term.'

While research on immediate and short-term processing has demonstrated that acceptable variants can effectively prime each other (though it should be noted that this effect is not uniform; see Pitt (2009) for differing findings in post-nasal [t] deletion), a discrepancy has emerged in the long-term. For example, Sumner and Samuel (2005) conducted a long-term repetition task to investigate priming of the same three word-final /t/ variants with the implementation of longer blocks of primes and targets. In the design, each block contained 360 items comprised of real and pseudowords, and participants made a lexical decision for each token. The authors wanted to examine whether all forms of /t/ facilitated long-term priming, meaning that previously presented words would be recognized more quickly in the target block regardless of the phonetic variant

presented in the prime block. The results surprisingly indicated that only the canonical [t] variant served as a long-term prime, even though this is not the most frequent variant in production. The authors reasoned that a traditional exemplar model would not be sufficient to explain the results, given that the coarticulated variant should have been most represented in memory as a result of having the highest frequency in speech.

The inconsistency in results for immediate and long-term processing will now be examined from a cross-dialectal perspective, which offers parallels to within-dialect variation regarding the mapping of multiple forms to one lexical item. One crucial way in which cross-dialectal variation differs from within-dialect phonetic variation pertains to listener experience. For instance, it is difficult to control for the distribution and frequency of phonetic variants in the lexicon of each individual listener due to interspeaker variation. However, investigating native and non-native varieties acts as a control for frequency, as non-native and unfamiliar variants will undoubtedly be less represented in the lexicon due to smaller amounts of exposure.

2.3 Cross-Dialectal Word Recognition

Cross-dialectal speech perception has been a recent topic of interest, with research exploring lexical processing in variable contexts including rhoticity (Sumner & Kataoka, 2013; Sumner & Samuel, 2009), vowel quality (Clopper et al., 2016; Dossey, Jones, & Clopper, 2023;

Floccia et al., 2006) and *shesheo* in *norteño* Spanish[1] (López Velarde & Simonet, 2019). Other factors under consideration have been the effects of listener experience on spoken word recognition (Clapp et al., 2023; Clopper et al., 2016; Floccia et al., 2006; Sumner & Samuel, 2009), listener expectation (Hay & Drager, 2010, Hay et al., 2006a; Niedzielski, 1999), the standardness of the varieties (Clopper et al., 2016; Floccia et al., 2006; Sumner & Samuel, 2009; Sumner & Kataoka, 2013), and the testing location (Clapp et al., 2023; Clopper, 2017; Dossey et al., 2023).

One of the first studies on cross-dialectal word recognition posed the question of how processing was affected by an accent's level of familiarity to the listener (Floccia et al., 2006). In a series of lexical decision tasks, the authors explored the perception of five regional French accents, also including multiple talkers per variety to test variable recognition of voices associated with the same regions. In one of the experiments, participants from the Franche-Comté region in northeastern France were significantly slower in recognizing words presented in an unfamiliar accent (i.e., the southern Aix-en-Provence variety) when compared to a familiar variety associated with a media standard (i.e., Parisian) or their native variety. Listeners demonstrated equivalence in processing the familiar and native varieties, as there were no significant differences in response times between the two.

Subsequent research has explored cross-dialectal word recognition in different varieties of English. In a study exploring the role of listener familiarity in auditory word recognition, Sumner and Samuel (2009) implemented a series of immediate and long-term priming tasks to examine the perception of nonrhoticity in New York City English (e.g., [beɪkə], 'bak-uh,' as an r-less coda

[1] *Shesheo* refers to the use of [ʃ] for [t͡ʃ]; the nonstandard [ʃ] variant is a feature of *norteño* Spanish (i.e., the Spanish spoken in northwestern Mexico; López Velarde & Simonet, 2019).

instead of [beɪkɚ], ‘bak-er’). In their experiments, three groups of participants were tested: individuals from NYC who frequently produced r-dropping in the exit interviews (i.e., Overt-NYC), individuals from NYC who did not produce r-dropping in their own speech, but had native experience with the variant from growing up in NYC (Covert-NYC), and General American (GA) listeners with no previous experience with r-less variants in perception or production who had recently relocated to NYC for college. All participants (N = 48) were students attending Stony Brook University in Long Island who were monolingual speakers of English.

In an immediate form-priming task, Sumner and Samuel (2009) found that both NYC groups of listeners experienced equal priming for both r-less and r-ful variants. In other words, [beɪkə] (‘bak-uh’) primed [beɪkɚ] (‘baker’) and vice versa for NYC listeners. However, no priming effects were found for GA listeners when [beɪkə] was a target (i.e., inexperienced listeners were not faster to recognize the r-less variant, despite having been previously primed with an r-ful form). In a semantic priming experiment, Sumner and Samuel (2009) found similar results according to the experience of the listeners with the r-less variant. Both NYC groups showed effective priming regardless of the variant of the prime (i.e., ‘slender’ and ‘slend-uh’ both led to faster reaction times when hearing a semantically related target such as ‘thin’). However, r-less forms were not effective primes for GA listeners (i.e., ‘slend-uh’ did not lead to faster reaction times for the related target ‘thin’). Sumner and Samuel (2009) reasoned that tokens containing the r-less variant were treated more like nonwords for inexperienced listeners, as they are less likely to activate semantic associations than their r-ful counterparts.

In a third experiment, Sumner and Samuel (2009) implemented a long-term repetition priming task to analyze how the variants were represented in memory. As outlined in subsection 2.1.1, one critical difference between immediate and long-term processing pertains to encoding,

the mechanism by which variants are stored in the lexicon. For this reason, a long-term paradigm is necessary to understand the representational nature of the variants, as encoding proceeds word recognition. The prime and target blocks were each comprised of 560 tokens to ensure that 20-30 minutes would separate the presentation of critical primes and targets. Results from the long-term priming task indicated a difference between the two NYC groups, in which only Overt-NYC listeners experienced priming after the long lag. In other words, Overt-NYC listeners were faster to recognize target items that had been previously presented in the prime block, and both variants facilitated recognition. The Covert-NYC participants patterned similarly to the GA group, in which only the r-ful variant was strongly encoded and represented in memory. This lack of encoding manifested as no observable priming effects when primes were produced in the r-less variant (e.g., participants were not faster in recognizing 'baker' when [beɪkə] was the prime).

Taken together, findings from the three experiments in Sumner and Samuel (2009) highlighted the striking difference between immediate and long-term processing for experienced listeners and also called into question what it means to have a dialect. The authors proposed that individuals may have a dialect not only in production, but also in perception and representation; the dialect (i.e., as operationalized by rhoticity) of Overt-NYC participants was reflected in production, perception, and representation, while it was predominantly present in perception for Covert-NYC participants, and not at all for GA participants. These findings imply that dialectal differences can affect word recognition among monolingual speakers of English, but exposure to the dialect can allow listeners to become more flexible and better able to process variation.

In a subsequent investigation, Sumner and Kataoka (2013) expanded the scope of r-dropping on speech perception to include three varieties of English: GA, NYC, and a standard variety of British English (BE). The authors selected a BE variety because, although it employs

the same trait of r-dropping, it is typically regarded as a prestigious variety of English, while the r-dropping associated with NYC speech is socially stigmatized. In a semantic priming task with GA listeners, Sumner and Kataoka (2013) found that targets were recognized more quickly when preceded by either GA or BE primes but that no priming occurred for NYC primes. Moreover, priming was equivalent for both GA and BE varieties, despite none of the listeners being BE speakers. The findings offer an intriguing discrepancy in the perception of non-native dialectal variants, as the same linguistic trait of r-dropping affected processing differently in two varieties by inexperienced listeners. Sumner and Kataoka (2013) conclude that the prestige of BE allowed it to become more socially salient, thus granting it recognition equivalence to the native GE variety of the listeners. The results provide another example of findings unsupported by a model of speech perception that only accounts for frequency, as the participants were inexperienced listeners of the BE variety.

Other research has also added nuance to the attested familiarity effects in cross-dialectal speech perception through investigating variable vowel productions in Northern and Midland varieties of U.S. English (e.g., Clopper et al., 2016; Dossey et al., 2023; Jones & Clopper, 2019). In one study, Clopper et al. (2016) examined word recognition of two varieties of GA English: Midland (standard) and Northern (nonstandard). All participants were living in the Midland region at the time of experimentation and were familiar with the variety, while half of them were born in the Northern region. The authors conducted a lexical decision task and a memory task to test both recognition and encoding of these dialectal variants, including a noise condition to increase task difficulty and assess variable encoding under adverse listening conditions. The authors found a processing benefit for the Midland variety, represented by faster RTs and better representation in memory when noise was introduced, for both groups of listeners. These findings were surprising,

given that the Northern participants had lived in their region until the age of 18. Clopper et al. (2016) hypothesized that local varieties are both recognized and encoded better, regardless of the native variety of listeners, especially if the local variety is also associated with standardness. The results also highlight the difference between word recognition and encoding in lexical processing, as Midland targets demonstrated stronger encoding in the noise condition, despite having a lower word-recognition accuracy in the study phase. Clopper et al. (2016) concluded with evidence of differential encoding and representation for dialectal variants as a function of both dialect (and within dialect, standardness and local context) and task difficulty.

Moving to languages other than English, the investigation of lexical processing in Spanish cross-dialectal speech perception is rare. In one recent study, López Velarde and Simonet (2019) analyzed the perception of *shesheo* in the *norteño* variety of Mexican Spanish.[2] In this variety, [t͡ʃ] and [ʃ] are both possible variants for orthographic <ch> (e.g., for the word *charco*, 'puddle,' [t͡ʃ]arco and [ʃ]arco). López Velarde and Simonet (2019) found that speakers from northwestern Mexico equally accepted both variants as words, and that both variants primed each other in an immediate form-priming task. Despite dual acceptability and priming, participant RTs were faster when associated with the canonical [t͡ʃ], even though [ʃ] is the regional variant (López Velarde & Simonet, 2019).

[2] The difference between within-dialect and cross-dialectal variation is less defined in research that features experienced listeners; while both variants occur in Hermosillo, the study is considered cross-dialectal because of the regional association of each variant, as well as the possibility of 'having a dialect' in perception and representation along with production (Sumner & Samuel, 2009). However, connections can be made to within-dialect research, as the variants are in free variation for the Hermosillo participants.

The findings from *norteño* Spanish have parallels to Sumner and Samuel (2009) regarding a less frequent variant having a processing benefit, which is partially theorized to be a result of orthographic transparency (e.g., all listeners in Sumner and Samuel [2009] had fewer lexical decision errors when primes were presented in GA, the r-ful variety). Similarly, though [ʃ] is the regional variant, all lexical items are orthographically written as <ch> and not <sh>. However, one challenge in interpreting results from López Velarde and Simonet (2019) pertains to disparate findings in the literature for the frequency of the variants: it is unclear whether [t͡ʃ] or [ʃ] is more commonly produced in the *norteño* speech community. Regarding the representation of the variants, the authors hypothesize that more than one phonetic variant may be represented in the lexicon of speakers who live in a region with sociophonetic variation (López Velarde & Simonet, 2019). Thus, familiarity and exposure to dialectal variants can facilitate speech perception, though processing costs may be associated with the nonstandard variant.

2.3.1 Listener Expectations and Increased Talker/Listener Diversity

A particularly compelling line of research has investigated how expectations about who individuals may be listening to can affect speech perception. Hay and Drager (2010) explored how the placement of regionally bound stuffed animals (i.e., kangaroos/koalas and kiwis; which are associated with Australia and New Zealand, respectively) could affect listeners' vowel perception. This work was an extension of Niedzielski (1999) and Hay et al.'s (2006) research that demonstrated effects for the placement of regional labels on participants' test sheets. Results from Hay and Drager (2010) revealed that women listeners' vowel perception on a continuum of Australian-New Zealand-like productions mirrored the association of the stuffed animals in the room. The authors propose that the experimental environment can deeply influence results, and

that activation of a social concept can influence perception, regardless of whether the talker actually belongs to the particular region.

Recent research in cross-dialectal lexical processing has worked to target more diverse talker and listener populations in multiple ways, expanding beyond homogenous groups who are often overrepresented in this area of research (i.e., undergraduate students). For example, Dossey et al. (2023) and Clopper (2017) recruited participants who were visiting a local science museum, thus representing a larger age range among individuals. Gylfadóttir (2018) explored categorization and semantic priming of *seseo* vs. *distinción* processes in Seville, Spain, testing native Spanish-speaking participants outside of a lab setting. These contexts allow for the generalizability of attested trends in the literature to be explored.

Recent findings from Clapp et al. (2023) support the necessity of replication studies on different talker and listener groups. The authors replicated the continuous recognition memory paradigm from Palmeri et al. (1993), with a critical modification: the use of a socially diverse set of talkers and listeners (i.e., individuals spanning different races, dialects, and genders). In line with the results of Palmeri et al. (1993), findings from Clapp et al. (2023) demonstrated a robust repetition effect for words repeated by the same talker. However, where the original study did not yield a significant processing cost on different trials in which the number of talker voices was variable, Clapp et al. (2023) demonstrated increasing task difficulty when the talker set was more diverse. As the number of diverse talkers increased, listeners experienced processing costs for accuracy and response times for both old and new trials. As repetition effects were not uniformly found, Clapp et al. (2023) concluded that listeners have better memory for some talkers than others, even when they are associated with the same dialect region.

2.3.2 Updates to Existing Speech Perception Models

The aforementioned findings in cross-dialectal speech perception present several challenges for an exemplar model of speech perception that only accounts for frequency-based encoding. If the most commonly encountered items are encoded and robustly represented in memory, then why would an infrequent, nonlocal variety such as BE in Sumner and Kataoka (2013) demonstrate recognition equivalence? And why might experienced listeners show recognition equivalence in immediate priming tasks but not in the long-term, as found in the case of the Covert-NYC group (Sumner & Samuel, 2009)? Finally, why would a local variant experience processing costs when compared to a more standard production, despite equal lexical acceptability (López Velarde & Simonet, 2019)? To account for such findings, Sumner et al. (2014) proposed a dual-route model in which social and linguistic information are processed simultaneously. This model introduces the idea of *social weighting*, referring to encoding not only based on frequency, but also as a result of the social meanings associated with the speech signal. Sumner et al. (2014) theorize that there are multiple ways in which a variant may be strongly encoded and robustly represented in the lexicon: by having high frequency (i.e., a large number of exemplars), by being atypical, but idealized (e.g., the [t] in 'flute'), and by being socially salient (i.e., associated with stigma or a standard).

While Sumner et al. (2014) have proposed expansions to the theoretical landscape of auditory lexical processing, remaining questions persist. For example, 'social salience' can be derived from different contexts, and it is unclear how salience from different levels of perceived standardness interact in the lexicon for experienced listeners. Moreover, Clopper (2021) called for an inclusion of the interaction between dialect classification and speech processing, as Sumner et

al. (2014) were primarily concerned with processing and did not account for the acquisition of social stereotypes, which have an effect on recognizing sociophonetic variants.

To this point, it is evident that sociophonetic variation lends an important context to examine how well certain varieties are processed, leveraging variables such as listener experience and perceived standardness of the varieties. One notable gap in the literature involves examining immediate and long-term cross-dialectal word recognition in languages other than English. The following section therefore outlines and motivates the specific research context of my dissertation.

2.4 Variable /s/ Production in Peninsular Spanish

My dissertation will contribute to a robust literature on /s/ phenomena in Spanish, as both perception (e.g., Chappell, 2016a, 2019; Regan, 2019, 2022; Walker, García, Cortés, & Campbell-Kibler, 2014) and production (e.g., Iglesias, 2003; Samper Padilla, 2011; Santana Marrero, 2016-2017) studies on myriad related variants are frequent in the field of Hispanic Linguistics. One of the most pervasive dialectal features of study in the field is word- and syllable-final weakening processes of /s/ (Colantoni, 2011). Instead of producing a voiceless sibilant, weakening occurs and can be mapped to three primary phonetic variants: [s] sibilant, [h] aspiration, and [ø] elision/deletion (Samper Padilla, 2011). The trait of coda /s/ weakening is highly salient and its presence is capable of driving perceptual shifts as the only manipulated trait (e.g., Mexican Spanish was evaluated as being more Caribbean when presented with [h] instead of [s] [Chappell, 2019]). While coda /s/ weakening processes occur throughout Spain, there are differences between north-central and southern varieties regarding production rates. For example, Iglesias (2003) discovered that in the Andalusian city of Córdoba a full, sibilant realization of coda /s/ was only employed at

a rate of 3%, with aspiration (54%) and deletion (43%) being the norm. On the contrary, Samper Padilla (2011) found a tendency to maintain sibilant coda /s/ productions in north-central regions of Spain, though aspiration and deletion still occur to a lesser extent.

While coda /s/ aspiration and deletion processes robustly occur in southern Spain, other Peninsular /s/ variants have not received ample attention in the literature. For instance, one feature that differentiates Andalusia from central and northern zones of the Peninsula pertains to the place of articulation of /s/ (Cano, 2015). In the central and northern regions of Spain, the standard production of /s/ is an apico-alveolar sibilant, typically transcribed as [s̺] (Penny, 2004). However, other allophones of /s/ exist in southern Spain. In many Andalusian cities, the place of articulation of /s/ is dental, transcribed as [s̪] (Penny, 2004; Villena-Ponsoda, 2008). Another variant, [s̪$^{\theta}$], is produced with interdental mixing and geographically associated with the small towns outside of Seville capital, other rural areas of Andalusia, and major coastal cities such as Cádiz and Málaga (Penny, 2004).

While the place of articulation of /s/ varies according to specific regions of the Peninsula, each variant also possesses different social associations. The hierarchy of these variants relates the apico-alveolar [s̺] to the national standard, granting it overt prestige. Overt prestige often corresponds with groups of high socio-economic status, with linguistic traits being represented in the media and reinforced in educational contexts (Labov, 1972, p. 249). Castilian Spanish, which is currently maintained in central and northern regions of Spain, has historically been associated with prestige (Hernández-Campoy & Villena-Ponsoda, 2009). Central and northern varieties of Peninsular Spanish have been the standard for centuries, and Andalusian speakers are likely to associate these standard varieties with “correct” speech (Santana Marrero, 2016-2017). The fact that prestige is less related to southern Spain is also partially motivated by politics, industry, and

geographic location, which can partially explain why standard traits are being more frequently incorporated into southern speech (Fernández de Molina Ortés & Hernández-Campoy, 2018; Hernández-Campoy & Villena-Ponsoda, 2009).

Despite the overt prestige associated with the national standard, some divergent pronunciation traits are considered to be locally prestigious in western Andalusia. Seville, associated with [s̪], is viewed as the epicenter of the regional standard in western Andalusia due to its historical importance (Hernandez-Campoy & Villena-Ponsoda, 2009). Seville is also well-known for having its own speech norm, often referred to as the *norma sevillana* (Hernández-Campoy & Villena-Ponsoda, 2009). Thus, Seville capital holds covert prestige, which occurs when non-standard traits possess higher value as a result of being associated with a specific community or group (Labov, 1972, p. 249). Despite the covert prestige associated with both the city and divergent linguistics traits, conservative variants are still considered to be more prestigious in comparison (Villena-Ponsoda, 2008). The national standard is apparent in Seville due to media exposure and education, fostering a constant internal battle between the prestige of the national standard and the local vernacular that resonates with the identity of the Sevillian citizens (Santana Marrero, 2016-2017). Even on southern TV networks such as *Canal Sur*, many Andalusian news anchors adjust their phonetic traits to reflect north-central varieties (León-Castro, 2016).

Though both [s̪] and [s̪$^{\theta}$] are innovative Andalusian variants, they possess distinct social implications. The use of [s̪$^{\theta}$] is highly stigmatized and does not hold any prestige at regional or national levels (Hernández-Campoy & Villena-Ponsoda, 2009; Regan, 2017a). Hernández-Campoy and Villena-Ponsoda (2009) highlight that in regions with more than one norm, the standard can often be associated with being "correct, adequate, and aesthetic," while the non-standard can be associated with "incorrect, inadequate, and unaesthetic (p. 182)." In Andalusia,

the non-standard [s̪$^{\theta}$] is often associated with the latter terms. Educated speakers will often diminish their use of [s̪$^{\theta}$] in formal situations, instead realizing the dental variant [s̪] (Penny, 2004). Cruz Ortiz (2020) confirms the masking of [s̪$^{\theta}$] in a study conducted on the speech of 35 Andalusian politicians living in Madrid from 1923-2011, only finding eight total tokens of the variant (i.e., overall rate of 0.2%).

Research on Andalusian patterns of sociophonetic variation has measured a diachronic shift towards standard variants. One difference that traditionally separated innovative varieties from their conservative counterparts is the loss of the medieval phonological distinction between /s/ and /θ/ (Hernández-Campoy & Villena-Ponsoda, 2009; Santana Marrero, 2016-2017). Conservative varieties preserve the use of /s/ and /θ/, employing the two phonemes in their respective orthographic environments. This trait is called *distinción*, 'distinction,' and can be found extensively in central and northern Spain (Penny, 2004). A speaker with *distinción* will produce [s] when associated with orthographic 's,' and [θ] for the graphemes <ce>, <ci>, and <z>. While conservative varieties maintain a distinction between [s] and [θ], a phonemic merger can be associated with innovative varieties of Peninsular Spanish (Santana Marrero, 2016-2017). Though there are numerous cases of reduction in Peninsular Spanish, two innovative phonetic traits are referred to as *seseo* and *ceceo*. *Seseo* can be defined as an extension of [s] to contexts that would traditionally require [θ], while *ceceo* is the use of [θ] in contexts that anticipate [s] (Penny, 2004). While *seseo* and *ceceo* are considered to be divergent patterns in Spain (Villena-Ponsoda, 2008), *seseo* is much more frequent outside of Peninsular Spanish, as it is the norm for Latin American varieties. *Ceceo*, however, is geographically restricted to southern Spain and has not been recorded outside of Andalusian varieties. *Seseo* and *ceceo* can lead to the loss of distinction between minimal pairs, one of the most widely referenced examples being *casar* 'to marry' and *cazar* 'to

hunt' (Penny, 2004; Villena-Ponsoda & Ávila-Muñoz, 2014). Despite these three seemingly straightforward definitions of *seseo*, *ceceo* and *distinción*, they are linguistic idealizations, meaning that some speakers do not canonically produce only one system (Regan, 2017a).

While *distinción* is predominantly employed in central and northern regions of Spain, it has steadily entered Andalusian speech. Cities in eastern Andalusia such as Málaga and Granada have been adopting conservative features such as *distinción* (Villena-Ponsoda, 2008), and *distinción* has recently been shown to compete with *seseo* in western Andalusian cities such as Huelva and Seville (Regan, 2017a; Santana Marrero, 2016-2017). While Seville, the capital of Andalusia, had historically been recognized as a zone with predominant *seseo* use, Santana Marrero (2016-2017) demonstrates the prevalence of *distinción* in *sevillano* speech. After conducting a series of sociolinguistic interviews in Seville capital, Santana Marrero (2016-2017) found *seseo* to be the leading divergent trend, with a rate of 83.5%, when compared to *ceceo*, which was employed at a rate of 15.5%. When conservative data were analyzed, results indicated that [θ] was produced in orthographic contexts of <ce>, <ci>, and <z> at a rate of approximately 56% (i.e., *distinción*), while [s̪] was produced in these same contexts at approximately 44% (i.e., *seseo*). However, it is important to note that even when *distinción* occurs in Seville, /s/ productions are still dental, separating this *distinción* from other zones in central and northern Spain that produce an apico-alveolar sibilant [s̪] (Villena-Ponsoda & Ávila-Muñoz, 2014).

2.5 Perception of Spanish Sociophonetic Variation

The perception of sociophonetic variation has been a robust area of investigation, with recent studies commonly implementing an updated matched-guise paradigm to test the impact of

specific variants on listeners' perceptions of speakers. These contemporary tasks, adapted from the traditional matched-guise paradigm (Lambert, Hodgson, Gardner, & Fillenbaum, 1960), consist of variants being spliced into the utterances of a single talker (e.g., Campbell-Kibler, 2007; Chappell, 2016a, 2019; Chappell & Kanwit, 2022; Villarreal, 2018; Walker et al., 2014). Listeners are typically asked to evaluate the speech according to a provided array of adjectives and/or complete dialect classification questions for each presented utterance. As participants are unaware that the speech signal has been modified, a comparison can be made across guises, in which the only changes are the variants themselves.

The verbal-guise technique (Markel, Eisler, & Reese, 1967) is another paradigm used for talker evaluation. While it shares many methodological similarities with the matched-guise technique, the verbal guise allows for more talker-specific variability to be present across guises. While this paradigm provides less control than a matched-guise task, an advantage is that listeners are exposed to more habitual speech from the talkers featured in the experiment (Dragojevic & Goatley-Soan, 2022). Subsequent paragraphs in this section summarize research that has implemented matched-guise techniques, as this paradigm has been used recently to explore the perception of Spanish variants; further information about the Verbal Guise in the current study is provided in chapter 3.

Within the past decade, matched-guise studies have investigated the perception of different variants in the Spanish-speaking world, ranging from Latin American contexts (e.g., Chappell, 2016a, 2019; Walker et al., 2014) to Peninsular Spanish (e.g., Barnes, 2015; Regan, 2019, 2020, 2022), even considering the behavior of L2 learners acquiring Spanish (Chappell & Kanwit, 2022). Results overwhelmingly indicate that small manipulations to the acoustic signal can influence

listeners' language attitudes; significant findings will be discussed by region in the following paragraphs.

Research on the perception of variants in Latin American locations has mainly focused on /s/ phenomena, as these contexts are often highly socially salient. In a study that included Puerto Rican and non-coastal varieties of Mexican Spanish, Walker et al. (2014) tested the influence of manipulating an aspirated vs. maintained coda /s/. Critically, a sibilant production of coda /s/ is frequent in non-coastal varieties of Mexican Spanish, while the aspirated coda /s/ variant is a salient trait of Puerto Rican Spanish. The variants also differ in standardness, as a sibilant /s/ is more associated with prestige. One significant finding from Walker et al. (2014) was that the presence of the sibilant /s/ variant led to significantly higher status ratings for both Puerto Rican and Mexican guises, but the effect was larger for Mexican talkers, regardless of whether the listeners themselves were Puerto Rican or Mexican. The authors concluded that listeners take context into consideration when evaluating the social meaning of a variant.

Chappell (2019) expanded findings from Walker et al. (2014) in a subsequent study analyzing the same variable coda /s/ production in Puerto Rican and non-coastal Mexican Spanish, but with two impactful adjustments to the experimental design. First, Chappell (2019) included a dialect classification question to examine how perception of a variant may influence listeners' perception of regional associations for the same talkers. The listener population was limited to Mexican participants, while talkers spanned Puerto Rican and Mexican locations. Furthermore, additional questions were added to the evaluation section of each guise, eliciting information about perceived profession and other traits such as 'snobbishness' and 'intelligence' (Chappell, 2019). Results supported the powerful social information that a single variant can contain, as both Puerto Rican and Mexican talkers were evaluated as significantly more Caribbean in guises featuring the

aspirated coda /s/. Moreover, corroborating significant results from Walker et al. (2014), the sibilant coda /s/ variant was more associated with status categories (e.g., higher intelligence/work ethic/confidence and snobbishness; Chappell, 2019).

Other work on Latin American /s/ processes consider a different variable context, the intervocalic voicing of /s/ in Costa Rican Spanish (e.g., [o.zo] for *oso* 'bear' [Chappell, 2016a]). Chappell's matched-guise task included questions to probe evaluation of status and solidarity categories according to the presence or absence of the salient [z] variant. She found lower perceived status ratings (i.e., social class, education level) associated with the non-standard variant [z] when produced by both men and women, but significantly higher solidarity ratings (i.e., niceness, confidence, and Costa Rican-ness) for men who used [z]. Results from Chappell (2016a) showed that the variants indexed different social meanings according to the perceived gender of the talker , exemplified by the lack of increase in solidarity ratings for [z] when produced by women.

The matched-guise paradigm has also explored Peninsular phenomena, spanning variable contexts associated with northern (i.e., Asturias; Barnes, 2015) and southern (i.e., Andalusia; Regan, 2019, 2020, 2022) regions of Spain. For example, vowel raising is a feature of Asturian Spanish that affects two particular contexts: 'o' (e.g., *el perro* → *el perru* 'the dog') and 'as' (e.g., *las casas* → *les cases* 'the houses' [Barnes, 2015, p. 219]). Barnes (2015) discovered that the use of [u] for [o] more significantly affected the perception of the talkers, though both [u] and [es] guises were associated with significantly lower status ratings (i.e., more rural). The author reasoned that the [u] variant possesses a higher degree of cognitive salience from a frequency standpoint, in that items with penultimate stress that contain word-final [u] are rare in Peninsular Spanish compared to the highly frequent -es morpheme, which exists in multiple regular contexts.

Taken together, Barnes (2015) hypothesized that the salience of the [u] morpheme was stronger and more directly related to an Asturian feature.

Shifting the focus to southwestern Spain, Regan (2019, 2020, 2022) has implemented the matched guise paradigm to examine several different Andalusian variants (e.g., *ceceo*, *distinción*, [t͡ʃ]/[ʃ] variation) and speech communities. In a matched-guise task testing the perception of *distinción* and *ceceo* in Lepe, Spain, a small town in western Andalusia in the province of Huelva, Regan (2019) created guises that only varied according to the syllable-initial /s/ variant (e.g., [s̪] or [s̪$^{\theta}$]; *distinción* or *ceceo*). As predicted, results demonstrated significant increases for *distinción* guises in status categories (i.e., perceived economic status, education, urbanity, formality). Additionally, there was a strong regional association of *ceceo* with Lepe and *distinción* with Huelva capital. In summary, listeners from the small town of Lepe evaluated their own local variant (i.e., *ceceo*) as being less prestigious than the national standard of *distinción* (Regan, 2019).

In a subsequent study involving the same variable, Regan (2022) manipulated the speech of 12 talkers from western Andalusia (i.e., Huelva and Lepe), this time including talkers and listeners from both locations. Findings mirrored trends demonstrated in Regan (2019): *ceceo* guises were evaluated significantly more lowly for perceived status, urbanity, and formality, supporting the overt prestige associated with *distinción*. Significant interactions also revealed gender effects, with women's *ceceo* guises receiving even lower status ratings than men's *ceceo* guises. Finally, participants in Regan (2022) explicitly commented on the social implications of *ceceo* and described the sounds themselves. These evaluative data provide strong evidence for both the acoustic and social salience of the word-initial /s/ variants featured in the current study.

2.6 Dialect Evaluation and Lexical Processing in Seville, Spain

While the literature presented in this section has spanned production studies on Peninsular variants as well as perception in Latin American and Peninsular sociophonetic contexts using matched-guise tasks, there is a subset of research that focuses specifically on Seville. Recently employed tasks such as dialect classification, evaluation, phonological discrimination, and semantic priming will be highlighted below.

In a study observing dialect classification and evaluation of different Peninsular varieties, Santana Marrero (2018b) surveyed Sevillian natives who were pursuing degrees at the University of Seville to examine evaluations of three varieties of Spanish: Andalusian (though talkers were specifically from Seville), Canarian, and Castilian. The author included various measures of solidarity, asking participants to evaluate talkers on a scale of 0-6. Results indicated that the students responded similarly to Andalusian and Canarian varieties, while the Castilian variety often patterned differently. For instance, in solidarity ratings, the north-central variety was evaluated as significantly more 'distant' and 'boring' than the southern varieties but received the highest ratings for 'urban' and 'clear' in status categories. The majority of the participants (52.5%) also evaluated their native variety positively when asked about pronunciation, with some explicitly identifying traits such as *seseo* and coda /s/ weakening. However, 18.37% of participants negatively evaluated Andalusian pronunciation traits, with one individual remarking that, '*no me gusta el seseo,'* 'I don't like *seseo*.' Given the ongoing expansion of *distinción* in Seville, it is unsurprising that language attitudes towards *seseo* are not uniform.

A follow-up study expanded the methodology from Santana Marrero (2018b), narrowing the participant population to western Andalusian students studying journalism at the University of

Seville. As the relationship among Andalusian variants, the attitudes associated with them, and media exposure is complex, Santana Marrero (2022) wanted to assess how future Andalusian broadcast professionals feel about their own linguistic identities. A struggle between the overt prestige of central-northern varieties and the covert prestige of Seville capital speech is apparent in the dialect evaluation patterns. Students awarded high ratings to Castilian talkers for status categories such as 'serious,' 'urban,' and 'clear.' The solidarity category of 'funny' was more significantly related to Andalusian Spanish, reflecting a stereotype that is often perpetuated by the media. Similar to the explicit mention of phonetic trends in Santana Marrero (2018b), participants in the 2022 study contributed both positive and negative evaluations of *seseo*, but Santana Marrero (2022) noted that positive comments outnumbered the negative perspectives.

Moreover, results from Santana Marrero (2022) supported the existence of a prestige hierarchy among Peninsular varieties. The Andalusian journalism students who participated in her study believed that there was a hierarchical prestige structure in Peninsular varieties, meaning that not all varieties were regarded equally (i.e., only 10.1% of participants believed that no dialects were inferior to others). More specifically, of the responses corresponding to a prestige hierarchy, 63.5% of respondents answered the question, 'What is the most prestigious variety?,' with 'Castilian' and 20.1% with 'Andalusian.' These data support the association of the Madrid variety with overt prestige by the Seville speech community, which is further compounded by media exposure of central-northern varieties.

To reference an additional investigation about the complex and dynamic linguistic landscape in western Andalusia, a recent dissertation study by Gylfadóttir (2018) explored trends in perception and production of *seseo* and *distinción* in Seville capital. Of particular relevance to the current study are two psycholinguistic tasks featured in Gylfadóttir's (2018) research: an AX

discrimination task to examine the accuracy in participants' perception of /s/ and /θ/ and a semantic priming task to compare the two forms in processing. Results from the discrimination task were high (i.e., 90%), indicating that participants have a strong perceptual *distinción* for the phonemic contrast. In a semantic priming task, Gylfadóttir (2018) included /s/ and /θ/ primes, followed by semantically related and unrelated targets (e.g., *seseo* related condition: *pozo – agua*, 'well,' 'water'; *distinción* related condition: *zumo – naranja*, 'juice,' 'orange') to test how Seville capital listeners process <z>. In other words, do both /θ/ and /s/ primes facilitate equivalent semantic activation, regardless of participants' production patterns? Results indicated that *seseo* primes facilitated faster recognition of semantically related targets and that no effect was found for primes produced with *distinción*. Gylfadóttir (2018) reasoned that this finding could be attributed to the talker, who was encouraged to produce other informal traits (e.g., r/l variation). While the talker was a Sevillian who produced both *seseo* and *distinción* in the experiment, it is possible that participants heard *seseo* and associated him primarily with the innovative trait. *Seseo* is also more frequently produced by men (Santana Marrero, 2016-2017), which could also support the findings.

2.7 Gaps in Literature and Research Questions

The research presented in this chapter has outlined significant findings in the areas of spoken word recognition, production of Peninsular /s/ variants, and perception of Spanish sociophonetic variants. The gaps in the literature that my dissertation is designed to address combine sociophonetic and psycholinguistic research methodologies, examining the relationship among dialect classification, evaluation, and auditory lexical processing using both explicit (i.e., Verbal Guise) and implicit (i.e., Long-Term Form Priming) tasks. Critical questions remain

regarding the treatment of socially meaningful variants in processing by experienced listeners, the interaction of social and linguistic information on encoding strength and representation, and the role of the local context in shaping perception. To my knowledge, the current study is the first to operationalize three Peninsular /s/ variants associated with different places of articulation, geographic locations, and levels of standardness to examine spoken word recognition. Finally, this dissertation includes diverse talker and listener populations, featuring (N = 42) participants from Seville capital and (N = 6) talkers spanning different locations, genders, and ages.

The following research questions are explored in my dissertation.

1. What are the language attitudes associated with each of the three Peninsular Spanish varieties, and to what extent do the patterns match the prestige hierarchy attested in the literature?
2. How does variation in word-initial /s/ production in the aforementioned varieties affect immediate word recognition of /s/ items for experienced listeners, and what can this tell us about the mental representation of the variants?
3. How are the /s/ variants represented in memory, and what implication do findings have for variable encoding strength as a result of social factors?

3.0 Methods

The tasks described in this chapter implement a combination of sociophonetic and psycholinguistic methodologies, yielding both explicit and implicit measures of speech perception. Before outlining the tasks themselves, the following sections provide an overview of the six talkers whose voices appear in the perception tasks (3.1) and a summary of basic demographic information pertaining to the listeners (3.2). Section 3.3 outlines the experimental sequence and testing procedure completed by participants.

Task-related information begins in subsequent sections, with information provided about the Verbal Guise in Section 3.4. The Verbal Guise contained dialect classification and evaluation questions, with the goals of measuring how well listeners associated the six talkers with their respective geographic locations, along with collecting rating data for status and solidarity categories as evidence for the prestige hierarchy of the three varieties. Subsections detail stimuli creation (3.4.1) and task creation (3.4.2).

Next, the methodology transitions from a task in which ample linguistic variation is present to a highly constrained experimental design in which the primary source of variation is the word-initial /s/ variants (Word Recognition Task; Section 3.5). The goals of the long-term form priming task were to explore immediate processing and long-term priming effects for critical /s/ items, examining how listeners familiar with all three varieties processed the /s/ words. Information regarding stimuli creation is contained in subsection 3.5.1, with the experimental creation explained in subsection 3.5.2.

The Chapter concludes with an overview of the informal interview in Section 3.6.

3.1 Talkers

The talkers featured in this project (n = 6; mean age = 47) were all born and raised in their respective dialect regions (Madrid, Seville capital, or Afueras),[3] thus assuming adequate representation of dialect traits associated with each variety. The same six talkers appear in both experimental tasks, a decision that was explicitly made (i) to ensure talker familiarization and (ii) to compile evaluative data about the talkers in the event that the social information accessed in the sociophonetic task influenced processing in the psycholinguistic task. In total, one man and one woman from each location recorded stimuli. I chose to feature both men and women talkers because (i) talker gender often yields significant effects in sociophonetic tasks, and (ii) prime-target pairs in the word recognition task must be presented across talker gender to avoid priming effects from vocal similarity (Church & Schacter, 1994). Talkers were either personal friends or contacted through mutual friends. The main selection criteria were that talkers were native to their respective regions, that their speech contained linguistic features representative of their geographic locations, and that they habitually produced the word-initial /s/ variant associated with each location. For example, Afueras talkers frequently produced [s̪$^{\theta}$] instead of the [s̪] variant present in Seville capital speech. Talker demographic information of age, gender, and occupation are included in Table 1.[4]

[3] From this point forward, I implement the following coding scheme for the three Peninsular varieties: Madrid (i.e., central Spain; one talker was from Madrid and the other from Castilla-La Mancha), Seville capital, and Afueras (i.e., Seville outskirts).

[4] The following talker-naming conventions have been adopted: S, M, or A to designate the variety (i.e., Seville, Madrid, or Afueras), and M or W to indicate talker gender (i.e., man, woman).

Table 1 Talker Demographic Information

Seville Man (SM) 57 businessperson	**Madrid Man (MM)** 50 businessperson	**Afueras Man (AM)** 23 car salesperson
Seville Woman (SW) 58 accountant	**Madrid Woman (MW)** 34 graduate student	**Afueras Woman (AW)** 60 social worker; executive

While the primary selection criteria pertained to producing the word-initial /s/ variants of each location, as well as possessing other identifiable linguistic traits associated with the regions, the talker information in Table 1 reflects the demographic diversity of the individuals. For example, a variety of ages and occupations were represented in the talkers. Though this may be a cause for exclusion in some studies, I chose to feature a diverse set of talkers in line with recent psycholinguistic work that is moving away from hyper-controlled talker conditions (e.g., Clapp et al., 2023). The support for this decision stems from the fact that linguistic variation in naturalistic settings occurs across ages, genders, occupations, and dialect regions, among other contexts, with listeners outside of laboratory settings adeptly handling this diversity in perception. Though the goal was to have a broader representation of voices in the current study, any perceptual differences that can be attributed to social factors are considered in chapters 4 (Verbal Guise Analysis) and 6 (Discussion).

After agreeing to lend their voices to the project, all talkers were sent a document containing basic information about the purpose of the experiment (i.e., that the goal was to observe speech perception of Peninsular Spanish), as well as instructions for recording the stimuli for each task (see Appendix A). Talkers were asked to complete each voice recording when they were alone and in a silent location to ensure quality sound files without any extraneous noise (e.g., buzzing from ceiling fans, street noise from open windows, etc.). As the two Madrid talkers were not

residing in Seville at the time of testing, and logistical constraints prevented the researcher from traveling outside of Seville capital to record the talkers, I requested that they self-record.[5] To avoid that talkers were overly monitoring their speech, they were instructed to speak as though they were communicating with their best friend.

Talkers were informed that while their voices would be featured in the experiment, their names would never be disclosed, thus preserving their confidentiality. Due to the large number of stimuli, approximately 30 minutes were required to complete the recordings. As a token of appreciation for their time and for permitting that their voices appear in the project, each talker was paid 30 Euro, either in the form of a gift card or in cash per their preference. Talkers used WhatsApp to record and send me their sound files.

All stimuli were recorded by each talker. The sampling rate for the recordings was 48,000 Hz, the default for the mobile devices used by the talkers. After receiving the stimuli, I converted the .mp4 files into .wav files via iTunes. All segmentation was conducted in Praat (version 6.2.14; Boersma & Weenik, 2022) and placed at the nearest zero crossing of the onset and offset of the utterance (Verbal Guise) or word (Long-Term Form Priming). The intensity was normalized to 70dB for each utterance or word, and ten ms of silence were added to the beginning of each stimulus. Additionally, the sampling frequency of each stimulus was converted to 44,100 Hz.[6]

---

5 SM requested to complete the recordings in person after I arrived to Seville capital; the other 5 participants produced the stimuli on their own.

[6] There was a minimal amount of clipping during the conversion process of the stimuli, but the frequencies required to perceive all /s/ items were unaffected. Listeners did not report difficulties in understanding the words, and the accuracy rates in the word recognition task support the quality of the stimuli.

3.2 Listener Characteristics

The listeners included in the study (N = 42) were at least eighteen years of age and were either born and raised in Seville capital or had moved there at a very young age (e.g., one month old). They were also all residing in Seville at the time of testing. I chose to examine a group of Seville capital participants in the current study to ensure familiarity with all three word-initial /s/ variants and by expansion, with the varieties themselves (i.e., Afueras is familiar due to being a local variety and Madrid is familiar due to media exposure). All participants were recruited via snowball sampling through my own personal networks or those of individuals who completed the project. The snowball sampling method also controlled for the place of residence of the participants, as only individuals from Seville capital were contacted to participate. This information was confirmed in the language questionnaire completed by participants at the end of the project; all listeners met the inclusion criteria based on place of birth and current residence.

Out of the 42 individuals who participated in the study, there were more women than men (19 men; 23 women), as is often the case in voluntary linguistic studies (Agostini & Schwenter, 2018). According to the demographic information obtained from the modified Language Experience and Proficiency Questionnaire (LEAP-Q; Marian, Blumenfeld, & Kaushanskaya, 2007), the average age of participants was 38.5 years (range: 18-64 years). Seven participants indicated that they had spent time living outside of Seville. Of these individuals, the majority had spent approximately one year living abroad in countries where Spanish is not an official language (Austria, Germany, Holland, Ireland, Morocco, and Switzerland). The average time spent living

abroad across participants was 11.4 months (range: 8-16 months). The only Spanish-speaking countries referenced by participants included a two-month stay in the Dominican Republic combined with another six months abroad in the United States for the same participant. All other listeners indicated that they did not travel for extensive periods of time outside of Seville (*N* = 35).

Regarding linguistic information, all participants (*N* = 42) indicated that their first language was Spanish. A large number of participants (*N* = 30) reported speaking an additional language. The most common responses for other languages spoken were English (*N* = 27) and French (*N* = 12). Other participants indicated having knowledge of German (*N* = 7), Italian (*N* = 4), Arabic (*N* = 1), and Japanese (*N* = 1). Socioeconomic information was measured by asking participants to indicate the highest education level they had received. Table 2 summarizes participant responses to this question, with education levels ordered from lowest to highest selections (*educación secundaria* 'high school' – *Máster* 'Master's' degree). I wanted to include a variety of listeners, mainly to avoid overrepresentation of participants with university/graduate degrees, in line with recent work that has argued for the inclusion of a broader listener population (Clapp et al., 2023). *Educación secundaria* 'secondary education,' *formación profesional* 'vocational education,' and *bachillerato*, the U.S. equivalent of junior and senior years of high school, are all pre-university categories. Finally, no participants reported having auditory processing difficulties.

Table 2 Participant Education Levels

Education Level	# of participants
Educación Secundaria	*N* = 1
Formación profesional	*N* = 6
Bachillerato	*N* = 12
Universidad (Grado/Licenciatura)	*N* = 16
Postgrado	*N* = 2
Máster	*N* = 5

3.3 Testing Procedure

All testing took place in the apartment of a friend that was centrally located in Seville capital. Upon arrival, participants were presented with a written script that explained the purpose of the project and briefly described each task (see Appendix B). The script explained that participation was voluntary and that individuals could withdraw from the study at any time. Additionally, no identifiable information about the participants was recorded during the session. Participants were informed that the experiment would require approximately one hour and fifteen minutes to complete and that they would receive thirty Euro as a token of appreciation after finishing the project. All communication with the participants was conducted exclusively in Spanish, their first language. Along with the written study script given to participants upon arrival, I provided an oral summary of the information to each individual or group before they began the experiment. Moreover, I was available to participants throughout the duration of the experiment, though the majority of the comments were clarification questions about the lexical decision task (e.g., Does slang count as a real word? [yes]; Can I change my answer after pressing the key? [no]).

Each participant was given a laptop and noise-attenuating headphones. As I had five laptops available for testing, the majority of participants came in groups of 3-5, accompanied by friends or family members. I provided each participant with a random 3-digit code to be entered at the start of each task, protecting their confidentiality but linking responses to the same individual.

The experimental sequence was purposefully ordered in this way: Verbal Guise Task, Long-Term Form Priming Task, Informal Interview, LEAP-Q. This ensured that all participants

completed the Verbal Guise first, as it provided familiarization to the same six talker voices that appeared in the psycholinguistic task. Additionally, the activation of social categories in the Verbal Guise (e.g., labels of the three Peninsular varieties in dialect classification; evaluation of voices according to status and solidarity categories) was expected to influence auditory lexical processing (e.g., Hay et al., 2006). Before participants began the Verbal Guise, I provided an oral summary of the written instructions that were presented at the start of the task in Qualtrics. Participants were told that they would hear six talkers from three Peninsular locations, and that they would need to guess where each person was from after listening to the recording as well as evaluate the voice of the person according to a series of presented adjectives. Participants could adjust the volume on their headsets and completed the task at their own pace (average completion time = 15 minutes).

After concluding the Verbal Guise, participants were offered a short break. They were informed that the second task was the longest in the experimental sequence (average completion time = 50 minutes). I also provided an oral summary of the word recognition task instructions to each participant. Individuals were told that the experiment would test their memory, and that the task was comprised of real and invented words. I explained that participants would need to decide if a word was real or not in Spanish and make a decision as quickly as possible without losing accuracy. Some participants asked for clarification about what a 'real' word meant and if they should make their lexical decisions based on whether words were valid in the *Real Academia Española* (i.e., RAE; https://www.rae.es/). I clarified that they should make their choices according to whether the word meant something to them, not necessarily if it was listed as an official term in the dictionary. This is an important distinction because many Andalusian variants are not represented in the RAE, despite being frequent in production (e.g., r/l variation is a salient feature of Andalusian Spanish and while *borso* is an acceptable variant for *bolso*, 'bag,' in Seville, a search

for *borso* in the RAE returns a warning that the word is not in the dictionary). Once participants made their first key press, the experiment began.

After finishing the word recognition tasks, participants were offered another break before completing a brief, follow-up interview. As the interview required participants to reflect on their own speech, as well as that of the talkers who had appeared in the tasks, it was completed third in the sequence. The LEAP-Q required the least amount of time to complete, which participants finished at the end.

3.4 Verbal Guise Task

The objectives of the Verbal Guise task were (i) to collect subjective information about language attitudes associated with each of the three Peninsular varieties, (ii) to measure dialect classification accuracy, and (iii) to compare ratings across the varieties, individual talkers, and talker genders to establish whether the prestige hierarchy matched previous findings from a large body of literature (see Sections 2.4, 2.5, and 2.6 for review).

3.4.1 Verbal Guise Stimuli

Talkers created stimuli for the Verbal Guise task by answering five written questions aloud and audio recording their responses. Each of the five questions contained a noun with word-initial /s/. These items were not overtly brought to the attention of the talkers (i.e., /s/ words were neither underlined nor bolded). In answering each question, talkers produced at least one instance of their

word-initial /s/ variant. There were no time limitations on recording length; talkers could offer as much or as little information as they wished. The five questions posed to the individuals have been included in Table 3.

Table 3 Questions Answered by Talkers for Verbal Guise Stimuli Selection

1. ¿Qué hiciste el sábado pasado? 'What did you do last Saturday?'
2. ¿Cuáles son algunos planes que tienes para la semana que viene? 'What are some of your plans for next week?'
3. ¿Te gusta el salmorejo? 'Do you like salmorejo?'
4. ¿Cuál es el mejor lugar para tomar el sol y por qué? 'What is the best place to sunbathe and why?'
5. ¿Qué opinas de la tradición de la siesta? 'What's your opinion of the nap tradition?'

Note: Responses to bolded questions were used for stimuli creation.

Two out of the five answer responses formed the basis for the stimuli according to the talkers' responses. For example, as MM provided the most brief answers compared to the rest of the talkers, questions that contained longer responses from him were prioritized to ensure that the utterances contained sufficient linguistic information. Additionally, responses to questions by Seville outskirts talkers that contained clear instances of *ceceo* were prioritized. Taken together, stimuli from the following two questions were included in the experiment: *¿Qué hiciste el sábado pasado?* 'What did you do last Saturday?,' and *¿Qué opinas de la tradición de la siesta?* 'What is your opinion about the Spanish nap tradition?' In total, there were twelve sound files (i.e., two per talker), each containing the talker's response to the selected questions. Transcriptions of answers provided by each talker for the two featured questions and their English translations can be found in the Appendix (see Appendix C.1).

While variable word-initial /s/ production was expected to be a feature that would guide participant evaluations of talkers, as well as assist in dialect classification, the stimuli included other traits that could be linked to geographic region (e.g., heavy coda /s/ aspiration for Seville capital and Seville outskirts talkers as compared to more maintenance of coda /s/ in Madrid talkers; see Appendix C.3). For this reason, the task measures how the talkers' voices as a whole are evaluated by Seville capital listeners.

After choosing which responses would appear in the Verbal Guise task, I extracted a short segment of each talker's answer. This was done to control for audio length, as some talkers provided more information than others while answering questions (e.g., up to two minutes of audio for one response). Selected utterances for each talker ranged from 8 - 14 seconds and included a clear answer to the posed questions. All of the utterances contained the beginning of the answer to the question until a reasonable stopping point within the designated duration (i.e., no utterances were segmented at inappropriate phrase boundaries).

The semantic content of the utterances was neutral, in that it did not contain clues that could assist listeners in dialect classification. For example, as the *siesta* 'nap' tradition is more associated with Andalusia than central Spain, both Madrid talkers indicating that they didn't like the tradition could lead listeners to guess that they were from Madrid. However, both Madrid talkers spoke positively about the tradition, with only the Afueras man responding that he didn't always take naps. Moreover, as no specific locations were included in talkers' descriptions of their Saturday plans, listeners did not have access to information that could influence their dialect classification

selections.[7] With regards to the possibility of semantic content of the utterances affecting dialect evaluation ratings, while talkers' responses to the questions were variable, listeners were instructed both in the written instructions that accompanied the dialect evaluation section and orally before beginning the task that they should evaluate the voice of each talker instead of what was said.

To compile the list of adjectives that would accompany the aural stimuli in each trial, I pretested the twelve sound files to a sample of three Seville capital listeners. I implemented this methodology because I am not from Seville capital, and I wanted the adjectives to reflect the impressions of members of the speech community. The individuals listened to each recording and wrote down the first three adjectives that came to mind upon hearing the voice. They were encouraged to focus on the voices rather than the semantic content of the talkers' answers. I selected the most frequent adjectives to be featured in the task, including a combination of status and solidarity categories. A few highly mentioned adjectives were not used in the study out of respect for the talkers (e.g., adjectives like 'repulsive' for Madrid talkers or 'hick' for Seville outskirts talkers that were strongly derogatory); I correctly anticipated that some of these adjectives would be mentioned in the comment boxes and preferred to have them reflected as optional feedback rather than a central component of the dialect evaluation task. The ten selected adjectives and their English translations can be viewed in Table 4.

[7] MW referenced visiting a beach in Murcia, a region in southeastern Spain that borders Andalusia. After consulting with a Seville capital native, the mention of Murcia was removed from the utterance, as it could indicate that the talker was from central Spain. It is more common for Seville capital residents to visit nearby beaches in Huelva, Cádiz, and Málaga, all locations within the region. As MW paused before referencing Murcia, extraction occurred without any perceivable manipulation to utterance.

Table 4 Adjectives Featured in Dialect Evaluation

alegre	'happy'	**formal*	'formal'
bonita	'pretty'	*graciosa*	'fun'
coloquial	'colloquial'	*orgullosa*	'proud'
educada	'polite'	**pija*	'snobby'
**fina*	'refined'	*tranquila*	'calm'

Note: * designates status categories

3.4.2 Verbal Guise Task Creation

After compiling the ten adjectives and twelve utterances, I built a survey in Qualtrics to administer the task. Every trial contained one stimulus, and each page in Qualtrics represented one trial. Each trial contained the recording accompanied by the dialect classification question, "In your opinion, where is the speaker from?" The question the talker was answering was provided below the recording (see Figure 1). There was no limit to how many times the recording could be played; participants were permitted to repeat the sound files if they wished.

A multiple-choice dialect classification question accompanied each recording, in which participants selected where they thought each talker were from. Three options were provided: central Spain, Seville capital, or Seville outskirts. Participants were required to make a selection before advancing to a new trial. Though some dialect classification studies feature an 'other' option, in which participants can provide an additional option if they feel that the individual is from a different location (e.g., Santana Marrero, 2018b; Villarreal, 2018), I chose not to include this methodology in the current Verbal Guise task. This decision was made to encourage participants to associate the talkers to their respective regions, thus accessing the social

information they already possessed about each location and variety. A sample dialect classification trial has been included in Figure 1.

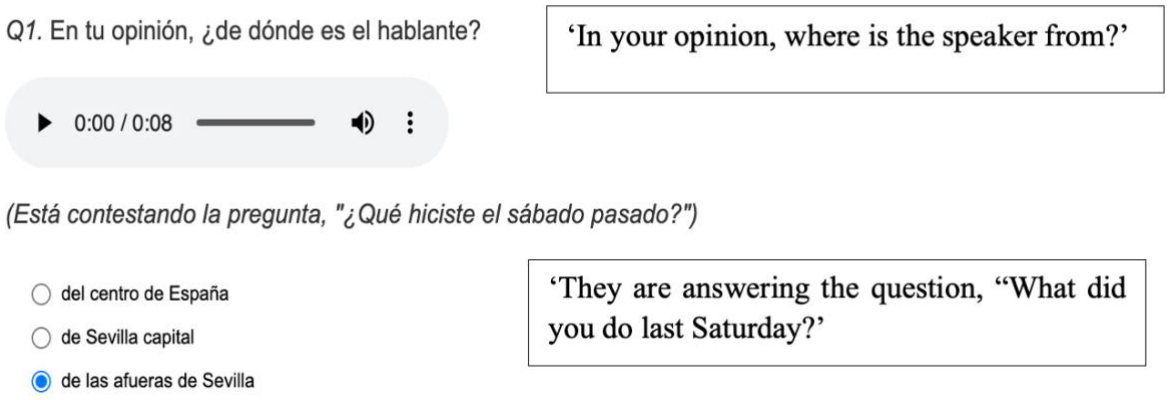


Figure 1 Sample Dialect Classification Trial

After the dialect classification question, participants completed the dialect evaluation component of the trial. A set of ten slider scales were presented, with one scale per adjective. The scales were continuous, and no values were displayed other than the endpoints. Additionally, when participants moved the scale according to their evaluations, no numerical information was provided to them (i.e., if they chose to keep the slider in the middle of the scale, they would not see '50' as the value). The possible range of the scale was 0 – 100, and only whole numbers were recorded. Forced choice logic was enabled for each slider scale, meaning that participants needed to complete all ten adjective ratings before progressing to the next trial. Figure 2 depicts the participants' view as they moved the bars on the slider scales.

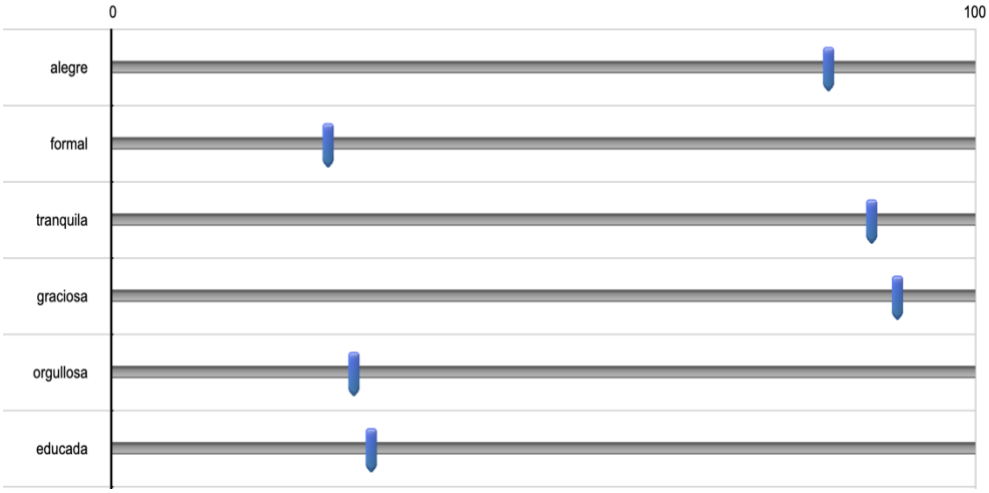


Figure 2 Sample Dialect Evlauation Scales

The last component of the Verbal Guise trial sequence was an optional, open-ended question that asked participants if they wished to comment on the voice of the talker or provide any justification for their ratings. This step captures additional social commentary on both the voices and the varieties themselves, which can enhance the interpretation of significant experimental findings. I also anticipated that some participants would comment on the linguistic features they had heard, specifically the /s/ differences. As research has shown that listeners rely on salient linguistic traits in dialect classification (Clopper & Pisoni, 2004), any commentary from the participants that named or described linguistic variants could provide valuable information about what explicit traits influenced listeners' classification and evaluation responses.

3.5 Long-Term Form Priming Task

The goal of the long-term form priming task was to examine how quickly participants recognized words in immediate processing and long-term processing. More specifically, I wanted

to determine if the recognition of /s/ words patterned differently from that of words that did not contain any salient dialectal differences, and if so, what patterns would emerge for each of the variants. In order to examine the recognition of the /s/ variants, I designed a lexical decision task consisting of 720 trials. The task was divided into two blocks, but participants were unaware of this distinction. To them, the experiment ran in the exact same manner from start to finish.

3.5.1 Long-Term Form Priming Stimuli

The long-term form priming task was comprised of real words and pseudowords. Each real word either acted as a control, a filler, or a critical /s/ word. The primary difference across conditions was that control and critical /s/ words were constrained to avoid salient phonetic traits in Peninsular Spanish (listed in Fernández de Molina Ortés and Hernández-Campoy, 2018), while fillers were permitted to have dialectal variants as a result of not being included in any reaction time analyses. Despite an explicit analysis of filler words, as some items included dialect-specific variants, listeners' regional associations for each talker could be strengthened; this was not an issue for the current study, as the purpose of the Verbal Guise utterances was to also provide listeners with regionally-specific information that they would associate with each talker. All real word controls and critical /s/ words were disyllabic nouns that ended in vowels, thus ensuring penultimate stress for all items. Fillers were also disyllabic nouns with penultimate stress, some ending in vowels and some including other coda. Salient phonetic traits were avoided in both control and critical /s/ words to ensure that any differences in reaction times were primarily

attributed to the /s/ variants (see Appendix C.3 for full list of avoided traits; adopted from Fernández de Molina Ortés and Hernández-Campoy, 2018, p. 502).[8]

Pseudo items were also divided into control and filler categories. Similar to the real word controls, pseudoword controls were disyllabic items that ended in vowels, being derived from Spanish nouns. Filler words were also disyllabic but could end in vowels or consonants. So that not all /s/ words were critical items, some real and pseudoword fillers also began with /s/. Along with the aforementioned conditions, an additional variable was manipulated in Block 2 control and filler items so that half of the words would be new and half would be repeated from Block 1. The second presentation of half of the items in Block 2 measures a repetition effect (i.e., priming; participants should be faster the second time they hear the word). All of the critical /s/ words were repeated in Block 2 but always by a different voice and talker gender so that any priming effects were not a result of vocal similarity (Church & Schacter, 1994). A breakdown of items in each block has been provided in Table 5.

Table 5 Counts for Experimental Conditions in Long-Term Form Priming Task

	Block 1: **Primes and Fillers**	**Block 2:** **Targets and Fillers**
Critical /s/ words	*48 /s/ primes (i.e., 24 produced by AM, MM, or SM; , 24 by AW, MW, or SW; all primes in one variety)	48 /s/ targets (16 per variety; 8 by gender) e.g., of 16 [s̪] primes, 8 men and 8 women
Control Words	96 control non-/s/ words 48 real (16 per var.; 8 per talker) 48 pseudo	96 control non-/s/ 48 real (16 per var.; 8 per talker; half new, half repeated)

[8] These salient traits were avoided as much as possible in critical items, there were minimal instances of tokens with coda /l/ given frequency constraints. However, none of these items were produced with the Andalusian variant /r/.

	(16 per var.; 8 per talker)	48 pseudo (16 per var.; 8 per talker; half new, half repeated
Filler Words	216 fillers 84 real (fewer for var. that presented /s/ primes) 132 pseudowords (44 per var.; 22 per talker)	216 fillers 84 real (fewer for var. that presented /s/ primes; half new, half repeated) 132 pseudowords (44 per var.; 22 per talker; half new, half repeated)
	TOTAL: N = 360 (180 real, 180 pseudo)	TOTAL: N = 360 (180 real, 180 pseudo)

*= variety was manipulated between subjects

I used the Spanish lexical database *Español Palabras* (EsPal; Duchon et al., 2013) to select the real words used in the experiment by inputting a series of criteria for each condition. Critical /s/ words were chosen from a list of mid to high frequency disyllabic nouns ending in vowels. Any /s/ words that had potential minimal pairs with word-initial /θ/ (i.e., orthographic <z>, <ce>, or <ci>) were excluded from the list so that an interdental realization from Seville outskirts talkers would be recognized as *ceceo* and not as a production with *distinción*. In total, there were a total of *N* = 48 critical /s/ words (average log frequency = 3.48; range = 1.908 – 5.176). The number of phonological neighbors was balanced across items (see Appendix C.2).

A similar methodology was implemented for selecting the real control words. I conducted a search in EsPal for disyllabic nouns that ended in vowels and limited the frequency to the range of the /s/ words. I worked in frequency bands to choose the control items so that they were as similar to the /s/ words as possible. For example, I had taken the average of the frequency of /s/ items ranging from 1.908 to 2.99. When selecting control words, I maintained a similar count of items within the first frequency band and also matched the average log count of the items (e.g., there were *N* = 14 /s/ words in the first frequency band, mean log count = 2.48; there were *N* = 14

real control items in the first frequency band, mean log count = 2.48). I repeated this process for the 3 – 3.99 frequency band, as well as for 4 – 5+. This careful selection process ensured that the critical /s/ words and real controls were as similar as possible, barring the salient word-initial /s/ and other identifiable regional variants.

Though there was more flexibility in the selection of the real filler words regarding linguistic variability, I followed a similar methodology to choose these items by working in frequency bands that mirrored those of the /s/ and control words (e.g., average log count = 3.75; range = 1.903 – 5.15). The last step in real control and filler item list creation was to compile the new words that would be featured in the second block. To do this, I followed the same methodology, working in frequency bands for each of the respective categories and matching average log counts both within frequency bands and across the whole item list. Table 6 represents a sample of real words represented in the word recognition task.

Table 6 Real Word Stimuli Examples from Word Recognition Task

Real Spanish Item	English Translation	Condition
sapo	toad	/s/
sangre	blood	/s/
finca	farm	control
vuelo	flight	control
ajo	garlic	filler
mangas	sleeves	filler

The creation of pseudowords followed a parallel methodology to that of real words. As the nonwords were derived from real Spanish words, I compiled a list of disyllabic nouns that fell within each frequency band of the real control words. All of the nouns for the pseudo controls ended in vowels, while the nouns used to create pseudo fillers were more diverse in coda; contexts

of phonetic variability were also included. After compiling a balanced list of nouns for the pseudo control and filler lists, I used Wuggy (Keuleers & Brysbaert, 2010), a multilingual pseudoword generator that creates items given a pre-established set of parameters, to generate non-words for the task. I set the language of Wuggy to Spanish so that the pseudowords would maintain phonological patterns of the target language. To generate each pseudoword, I uploaded my list of real nouns to the platform and made a selection out of the ten potential candidates provided by Wuggy for each item. I chose minimal pairs wherever possible, but other candidates varied by two phonemes from the original noun. Finally, to confirm that all pseudowords were indeed not viable Spanish words, I searched each item in the dictionary of the RAE. Indeed, none of the pseudoword items returned any results for real words in Peninsular Spanish.

To avoid production differences between real and pseudo items (e.g., slower realizations of pseudowords because they were invented and unfamiliar), talkers were instructed that tokens should be read as though they were real lexical items (i.e., without laughing or insinuating through intonation that the tokens were not legitimate). All pseudowords were derived from real Spanish words, making them phonologically plausible competitors in the lexical decision task. A sample of pseudowords can be viewed in Table 7.

Table 7 Pseudoword Stimuli Examples

Pseudoword	Real Spanish Word	English Translation	Condition
truña	*trufa*	truffle	control
plauta	*flauta*	flute	control
ucas	*uvas*	grapes	filler
cuerjo	*cuerpo*	body	filler
fúnbol	*fútbol*	soccer	filler
sirra	*mirra*	myrrh	filler

Three experimental versions were created in order to implement the between-subjects priming condition. In one version, all critical /s/ items in Block 1 were produced by Seville talkers, with each talker producing $N = 24$ words. I distributed the /s/ words evenly across talker gender, striving for a similar average log count and average number of phonological neighbors. This same methodology was implemented in the other two versions, with the difference being the prime presentation by Madrid talkers in the second version, and Afueras talkers in the third. The rest of the conditions in Block 1 (i.e., real and pseudo fillers and controls) were also distributed equally across talkers. To control for intonation across words and individuals, all items were produced by talkers in a declarative way (i.e., with falling intonation). Finally, the presentation of all items in the experiment was randomized within each Block.

In Block 2, half of the control and filler words were repeated, while the remaining items were replaced with new words. To select repeated items, I sorted the words in each category by talker according to their average log counts. Every other word was replaced with an item of a similar frequency to maintain a balance across new and repeated words. Repeated items in Block 2 were also presented across gender. A sample of real control items in Block 1 presented by SM has been included in Table 8.

Table 8 Example Distribution of Block 1 Controls for Seville Man

Talker	Item	Log Count	# Phon Neighbors
SM	*cacto* 'cactus'	1.93	19
SM	*cuenco* 'bowl'	2.80	3
SM	*peine* 'comb'	3.17	11
SM	*luna* 'moon'	4.59	30
SM	*gripe* 'flu'	3.58	9
SM	*feria* 'fair'	3.74	9

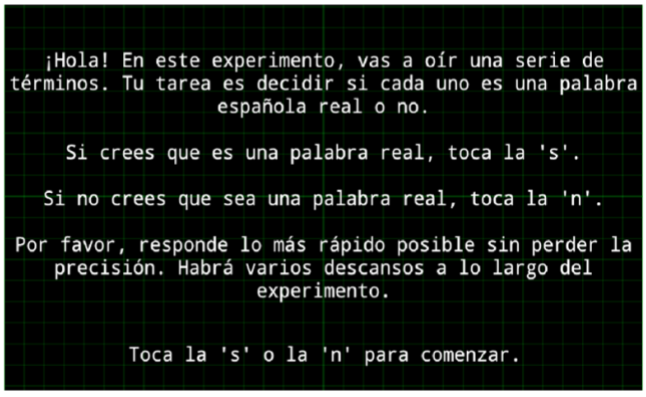
SM	*ritmo* ‘rhythm’	4.14	7
SM	*norte* ‘north’	4.59	14

For example, the highlighted real control words in Block 1 were repeated in Block 2, while the unmarked words were replaced by new items. Additionally, all of the highlighted words were produced by women in Block 2, as a result of the presentation of prime and target items across gender.

3.5.2 Long-Term Form Priming Creation

OpenSesame (Mathôt, Schreij, & Theeuwes, 2012) was used as a platform to administer the long-term form priming experiment. Upon entering the program, participants were directed to a previously assigned experiment version (Seville, Madrid, or Afueras primes) and entered their participant code.

The experimental sequence began with a set of written instructions. The information displayed to participants is depicted in Figure 3, along with an English translation of the text .



“Hello! In this experiment, you will hear a series of terms. Your task is to decide if each one is a real Spanish word or not. If you think that it is a real word, press the ‘s’. If you do not think that it is a real word, press the ‘n’. Please respond as quickly as possible without losing any accuracy. There will be various breaks throughout the experiment. Press the ‘s’ or the ‘n’ to begin.”

Figure 3 Welcome Screen in OpenSesame

Upon making their first key press, the sequence of 720 trials began in OpenSesame. The trial sequence for the prime block and target blocks was identical, with the only difference being the stimuli in each block. Items within each block were randomized, and all stimuli were presented only one time. Thus, there were N = 360 trials in each block. In each trial, participants were presented with a display that asked them if the item was a real word or not. The keyboard response required pressing the ‘s’ or ‘n’ keys to both make a lexical decision and record the reaction time. If a decision was not made within 4000 ms from the onset of the target word, a new trial was presented. There was a 1000 ms pause after the end of each trial to allow the participant to reset and prepare for the next item. The display that accompanied each trial was kept simple, depicted below in Figure 4. No feedback was given to participants after the completion of each trial or at the end of the experiment (i.e., participants never saw their accuracy scores).

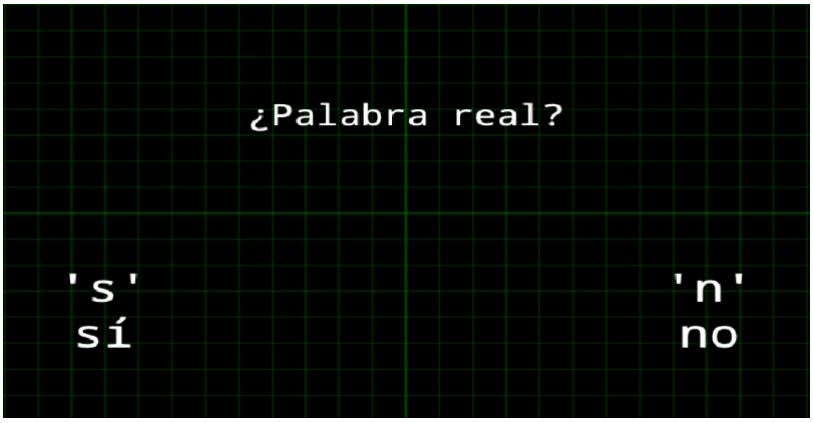


Figure 4 OpenSesame Trial Display

Finally, I built an optional break into the trial sequence so that participants could rest every sixty trials. Given the total of N = 720 trials, there were twelve total break opportunities. The breaks were implemented to reduce participant fatigue, given the tediousness of the experimental task. After the last trial was completed, participants saw a closing display thanking them for participating and informing them that their responses had been recorded.

3.6 Informal Interview

After completing the word recognition task, participants were offered a short break before continuing the experimental sequence with the brief informal interview. All participants (N = 42) agreed to verbally answer eight questions and have their responses audio recorded. The interviews were recorded using a Sony UX560 device. Before beginning the question sequence, each individual stated their participant code. During the interview, participants spoke for as long as they wished, and I interjected as minimally as possible. Given the overall length of the experiment, I designed the interview to be brief, functioning as a way to collect additional qualitative information after participants completed both the Verbal Guise and Long-Term Form Priming tasks. For this reason, the interview was less conversational in nature and was structured so that all participants were asked the identical questions. However, participants could share as much or little information as they wished.

I chose to have the interview follow the perception tasks so that participants would not try to guess the purpose of the experiment as a result of any of the posed questions, some of which mention stereotypes and dialect perception. To gain additional information about participants' attitudes towards their own speech, the dialects that were presented in the experiment, media representation of Andalusian varieties, and reactions to stereotypes, I wrote the eight interview questions listed in Table 9. Interviews took place in an adjacent room to the testing location so that participants who were still completing the experimental tasks were undisturbed.

I ordered the questions in a specific way, beginning with participants' thoughts about their own speech and experiences when traveling outside of the region. The next questions asked about the varieties represented in the experiment, additionally exploring topics such as media

representation and stereotypes. Finally, the last question was a hypothetical scenario to explore if, having to live in another Peninsular city, participants would choose another Andalusian city instead of locations outside of the region. While the interview data will not be extensively discussed, pertinent trends will be referenced in chapter 6.

Table 9 Informal Interview Questions

Question	English Translation
1. *¿Cómo te sientes sobre tu manera de hablar?*	How do you feel about the way that you speak?
2. *Cuando viajas fuera de Sevilla capital, ¿la gente normalmente sabe que eres sevillano? Si es el caso, ¿cómo crees que lo saben? Y si no, ¿de dónde piensan que eres?*	When you travel outside of Seville capital, do people usually know that you are Sevillian? If so, how do you believe that they know? And if not, where do they think you are from?
3. *¿Qué opinas de los dialectos que han aparecido en este experimento?*	What do you think about the dialects that have appeared in this experiment?
4. *¿Crees que el acento sevillano ha cambiado a lo largo de tu vida? Explica por qué sí o por qué no.*	Do you think the Sevillian accent has changed throughout your life? Explain why or why not.
5. *¿Es común que escuches un acento parecido al tuyo en los medios de comunicación o en las películas o series españolas?*	Is it common that you hear a similar accent to your own in the media or in Spanish movies or series?
6. *¿Te molestan los estereotipos sobre los andaluces/ los acentos andaluces?*	Do the stereotypes about Andalusian people/ the Andalusian accents bother you?
7. *¿Quiénes son los que más tratan de perpetuar los estereotipos?*	Who are the people who most try to perpetuate the stereotypes?
8. *Si pudieras vivir en otra ciudad española, ¿cuál seleccionarías y por qué?*	If you could live in another Spanish city, which would you select and why?

4.0 Verbal Guise Analysis and Results

This analysis chapter addresses the questions posed in the first research question of the dissertation, which are focused on (i) exploring the language attitudes associated with the three Peninsular varieties (i.e., Seville capital, Madrid, Afueras) and (ii) comparing to what extent the findings relate to the attested prestige hierarchy of the varieties (see Sections 2.4, 2.5, and 2.6 for review).

To begin, Section 4.1 focuses on the dialect classification portion of the Verbal Guise, as participants were presented with this question at the beginning of each guise. Predictions are detailed in subsection 4.1.1, followed by an analysis of classification accuracy in 4.1.2. Subsection 4.1.3 provides an interpretation of the results, making connections to previous dialect classification literature (e.g., Santana Marrero, 2018b, 2022). Next, Section 4.2 examines dialect evaluation results, presenting predictions in subsection 4.2.1. Results from the evaluation data are detailed in subsection 4.2.2., with consideration of rating data for incorrect dialect classification responses in subsection 4.2.3, providing an exploration of the relationship between regional association and social evaluation. Optional commentary provided by participant at the end of each guise are summarized in subsection 4.2.4 (see Appendix D for the full set of comments). Interpretation of the dialect evaluation results is provided in subsection 4.2.5. Finally, as the Verbal Guise was completed immediately before the psycholinguistic task, section 4.3 outlines the implications of the sociophonetic task and how they could influence processing.

4.1 Dialect Classification

4.1.1 Dialect Classification Predictions

Listeners' dialect classification responses were coded as '1' (correct) if they matched the talker's self-identification and '0' (incorrect) if they did not. As indicated in chapter 3, I implemented the following naming conventions when referring to each talker: S, M, or A to designate the variety (i.e., Seville, Madrid, or Afueras), and M or W to indicate talker gender (i.e., man, woman). So, the Seville man is written as SM, the Afueras woman as AW, etc.

In alignment with previous research findings on Peninsular dialect classification (e.g., Santana Marrero, 2018b, 2022) and keeping in mind the prestige hierarchy of the three varieties (i.e., Madrid > Seville > Afueras; e.g., Hernández-Campoy & Villena-Ponsoda, 2009; Penny, 2004), I made the following predictions. First, I expected that participants would be successful in recognizing the Seville talkers, as it is their native variety. One caveat to this prediction pertains to SM, as his utterances were more formal/carefully produced than those of SW. Any perceived formality could lead to SM being more associated with Madrid instead of Seville. For the out-of-dialect talkers, I expected participants to be more accurate in identifying Madrid talkers than Afueras talkers, given that central-northern varieties are conservative, while western Andalusian varieties are considered to be innovative (Penny, 2004; Santana Marrero, 2018b). Finally, I expected more confusion between Seville and Afueras talkers, as well as potentially between Seville and Madrid voices, minimal confusion between Afueras and Madrid talkers given the prestige hierarchy and phonetic differences between innovative and conservative varieties (e.g., Santana Marrero, 2018b, 2022).

Regarding overall dialect classification accuracy, I expected that participants would perform well on the task. Varieties that are closer together geographically have been found to yield more successful dialect classification (Clopper, 2021), which would support the prediction of high accuracy for the task as a whole. The geographic distance between all three varieties is not expansive, as Seville and Madrid are separated by approximately a one-hour plane ride, or a 2.5-hour-high-speed train. There is daily movement between Seville capital and Afueras, with commutes as short as ten minutes by train. A popular commuter train with five lines (i.e., *Cercanías*) connects the capital to many small towns outside of the city, with the company *Red Nacional de Ferrocarriles Españoles* 'Spanish National Railway Network' (Renfe) estimating the movement of over 6 million annual travelers, demonstrating an increase of 40% more commuters between 2021 and 2022.

General trends of variation in the Verbal Guise stimuli have been outlined in Table 10. This phonological variability can assist listeners in making accurate dialect classification assessments, as all talkers possess myriad traits that provide regional association with Andalusia or Madrid. Previous research on regional dialect classification has demonstrated that listeners often rely on a limited array of salient or stereotypical acoustic cues to assist in classification (e.g., r-dropping in non-rhotic varieties of New England English; Clopper & Pisoni, 2004). In the stimuli of the current study, each talker produces the word-initial /s/ variant associated with their respective variety as well as other salient dialectal traits that can be used as indicators of geographic location (e.g., heavy coda /s/ aspiration for Andalusian talkers, coda /s/ maintenance with Madrid talkers, etc.). Taken together, I expected that participants have access to sufficient acoustic information, even in short utterances, to predict with which locations talkers are associated.

Table 10 Phonetic Traits Present in Verbal Guise Stimuli

Talker	Traits
AM	• Coda /s/ aspiration • r/l variation • both [s̪$^{\theta}$] and [s̪]
AW	• Coda /s/ aspiration • loss of intervocalic /d/ • loss of coda -r • both [s̪$^{\theta}$] and [s̪]
SM	• Coda /s/ aspiration • *seseo* • [s̪] • loss of coda /d/,
SW	• Coda /s/ aspiration • *distinción* • [s̪] • loss of coda -r
MM	• Maintenance of coda /s/ • loss of intervocalic /d/ • *distinción* • [s̪]
MW	• Maintenance of coda /s/ • *distinción* • [s̪]

4.1.2 Dialect Classification Accuracy

Participants evaluated each talker during two separate trials (N = 42 participants x 2 = 84 trials per talker), leading to a total of N = 504 dialect classification trials. The mean accuracy across trials, range by participant, and standard deviation by participant are shown in Table 11.

Table 11 Descriptive Statistics for Dialect Classification Across All Participants

Mean Accuracy	$\bar{x}$ = 82.1% correct
Accuracy Range	41.7 % – 100% correct
Standard Deviation	38.3%

As noted in Table 11, the average accuracy rating across all trials was reasonably high, with 82.1% of responses to the dialect classification questions being correct. However, it was evident that there was a great deal of variability among participants, reflected by the large range and standard deviation. Participants 002 and 492 had the lowest accuracy scores, only being correct 41.7% of the time (i.e., 5/12 correct responses for each participant). Despite four participants having accuracy scores below 60% correct, their errors patterned similarly to overall trends in the data. Thus, no participants were excluded from the analysis of the Verbal Guise task. Table 12 represents the number of participants with accuracy scores falling within each percentage range.

Table 12 Distribution of Dialect Classification Accuracy

% Correct	**# of Participants**
> 45 %	*N* = 2
45 – 59.9 %	*N* = 2
60 – 69.9 %	*N* = 3
70 – 79.9 %	*N* = 9
80 – 89.9 %	*N* = 9
90 – 99.9 %	*N* = 9
100 %	*N* = 8
Total:	*N* = 42

While the majority of participants had at least 75% accuracy at correctly identifying each talker's origin (N = 35), there was still variability, indicating that many of the participants were unsuccessful in recognizing the geographic origin of at least one talker. To better understand the effects of individual talkers, I turned to accuracy scores by variety and talker. Dialect classification averages by variety are outlined in Table 13.

Table 13 Dialect Classification Accuracy by Variety

Variety	% correct trials
Afueras	79.8 %
Madrid	90.5 %
Seville	76.2 %

As depicted in Table 13, participants were actually most accurate in identifying Madrid voices (90.5 % correct trials), followed by Afueras (79.8%). According to the by-variety results, they were least accurate in recognizing Seville capital voices (76.2 %). While participants proved to be adept at regional classification, a logistic regression exploring accuracy as a function of Variety revealed significant differences in accuracy for both Afueras and Madrid when compared to Seville as the reference level. Participants were more accurate in dialect classification for Afueras talkers ($p = 0.013$) than Seville, and even more accurate in identifying Madrid talkers ($p < 0.001$) compared to their own native variety.

Average accuracy scores by talker are represented in Table 14 to further tease apart any individual differences.

Table 14 Dialect Classification Accuracy by Talker

Talker	Average Accuracy (% correct)
Afueras Man (AM)	73.8 %

Afueras Woman (AW)	85.7 %
Madrid Man (MM)	84.5 %
Madrid Woman (MW)	96.4 %
Seville Man (SM)	61.9 %
Seville Woman (SW)	90.5 %

The accuracy results in Table 14 reveal additional patterns. Comparing dialect classification on a by-Talker basis (treatment coding; ref = SW) demonstrated differences associated with the six individuals. When compared to SW, accuracy was significantly lower for SM ($p < 0.001$), AM ($p < 0.001$), AW ($p = 0.003$), and MM ($p < 0.001$). Participants were only significantly better at identifying the voice of MW ($p < 0.001$) than SW. The combination of dialect classification accuracy results by Variety and Talker indicate that while participants generally performed quite well on the task, some varieties and talkers were significantly more difficult to identify than others.

Referring to the by-Variety results in Table 15, it is clear that the difficulty in recognizing SM led to an overall lower accuracy rating for the native variety of participants. Another pattern that is evident upon comparing average dialect classification accuracy by variety and by talker (Tables 15 and 16, respectively) is that participants appear to be more accurate in placing women's voices than men's voices. Inferential statistics indicate that this difference is significant (Talker Gender = contrast code; $p < 0.001$), with higher accuracy scores attributed to women talkers. Dialect classification results by Talker Gender can be viewed in Table 15.

Table 15 Dialect Classification Accuracy by Talker Gender

Talker Gender	**% correct**
Woman	90.9 %

Man	73.4 %

Looking at the combined results of dialect classification accuracy by variety, talker, and talker gender offers a clearer overview of trends, but more exploration is necessary to explain the patterns. The aforementioned tables have considered correct responses, yet isolating incorrect answers can allow for an understanding of which varieties were confused with each other. To explore these trends, I created a confusion matrix of responses. Underneath the percentages are the response counts, out of a total of *N* = 84 trials per talker. This data can be found in Table 16.

Table 16 Dialect Classification Confusion Matrix

The percentages for each correct category are highlighted in green, while those associated with the most confused variety are highlighted in yellow.

	Afueras	**Seville**	**Madrid**	**Total:**
Afueras Man	73.8 % *N* = 62	25.0 % *N* = 21	1.2 % *N* = 1	100 % *N* = 84
Afueras Woman	85.7 % *N* = 72	13.1 % *N* = 11	1.2 % *N* = 1	100 % *N* = 84
Seville Man	14.3 % *N* = 12	61.9 % *N* = 52	23.8 % *N* = 20	100 % *N* = 84
Seville Woman	7.1 % *N* = 6	90.5 % *N* = 76	2.4 % *N* = 2	100 % *N* = 84
Madrid Man	2.4 % *N* = 2	13.1 % *N* = 11	84.5 % *N* = 71	100 % *N* = 84
Madrid Woman	1.2 % *N* = 1	2.4 % *N* = 2	96.4 % *N* = 81	100 % *N* = 84

As observed in Table 16, talker confusion largely followed a predictable pattern, with most incorrect answer choices occurring with a variety that was one step above or below on the prestige

hierarchy. For example, the majority of incorrect guesses for Afueras talkers were Seville selections instead of Madrid. Errors in identifying Madrid talkers followed the reverse pattern, with misidentifications being largely associated with Seville instead of Afueras. The two Seville talkers trended differently, with errors tending to associate SM with Madrid, and SW with Afueras.

4.1.3 Dialect Classification Discussion and Interpretation

To briefly summarize the primary results from the dialect classification portion of the Verbal Guise, participants were successful in identifying the geographic origin of talkers across the task as a whole ($\bar{x}$ = 82.1% correct). Participants with lower accuracy scores followed similar patterns with their misidentifications, either confusing Seville and Afueras talkers or Madrid and Seville talkers. There were minimal cases of confusion of the most innovative (i.e., Afueras) and most conservative (i.e., Madrid) varieties (N = 5 errors out of 504 trials; less than 1% of data). By variety, participants were most accurate in identifying Madrid talkers ($\bar{x}$ = 90.5%), and Afueras talkers ($\bar{x}$ = 79.8%), when compared to Seville ($\bar{x}$ = 76.2%). Talker gender also yielded significant effects, with participants being more accurate in identifying the geographic location of women's voices compared to men's voices (i.e., $\bar{x}$ = 90.9 % vs. $\bar{x}$ = 73.4%, respectively). Finally, regarding accuracy by talker, participants were least successful in classifying the origin of SM ($\bar{x}$ = 61.9% correct), most often confusing him with a Madrid talker.

The dialect classification predictions I had made prior to conducting the experiment were somewhat supported, with a few unexpected caveats. The broad prediction regarding dialect classification accuracy was that participants would be successful overall as a result of the acoustic information present in the speech samples, combined with the familiarity of the varieties to the Seville capital listeners, the predominant media representation of Madrid Spanish, and the fact that

the geographic distance between the three locations is not expansive. This prediction was correct, in the sense that participants were accurate in 82.1% of the trials. Accuracy results in the current study for dialect classification were higher than those observed in Santana Marrero (2018b, 2022), though this is likely attributed to a difference in response format (i.e., forced-choice selection vs. free response). Listeners in her 2018 study correctly identified Madrid talkers 55.5% and 71.7% of the time in the 2022 study compared to 90.5% in the current study. The accuracy rates were more similar for Seville talkers, with 67.8% in Santana Marrero (2018b), 74.7% in Santana Marrero (2022), and 76.2% in the current study. Some participants in Santana Marrero (2018b, 2022) did not specify a region and instead wrote 'Spain,' which were not counted as exact correct responses. However, it should be emphasized that listeners in the current study spanned multiple age ranges and education levels, whereas listeners in Santana Marrero (2018b, 2022) were all young Andalusian students pursuing their college degrees at the University of Seville. The dialect classification accuracy rates in the current study are a strong indicator that everyday Sevillian listeners are familiar with the three varieties and are adept at recognizing regional phonetic differences, especially with limited options in a forced-choice response setting.

To focus on specific regional dialect classification success, I had anticipated that participants would be most accurate in identifying Seville capital talkers due to it being their native variety. This prediction was partially incorrect, as participants were actually least accurate in classification of the Seville talkers. This inaccuracy was largely driven by incorrect regional association of SM (i.e., 61.9% correct trials for SM; 90.5% correct trials for SW). Difficulties in dialect classification of SM can partially be attributed to the effect of the observer's paradox (Labov, 1972). While recording, SM became more formal and shifted his pronunciation to a more careful speech style, losing some of the heavy coda /s/ aspiration that he had previously produced.

It should be noted that despite formality, participants still observed Sevillian traits, with one writing a comment that '[SM] seems to be from Seville. The last time that he pronounces siesta, he aspirates the second /s/.'

While listeners struggled with dialect classification of SM, participants were largely accurate in associating SW with Seville, yielding minimal errors overall. The quantitative success in dialect classification was qualitatively supported by comments made by participants about SW's voice. For example, participants remarked that SW was fully identifiable as being from Seville, with some specifying that it was because of her accent or pronunciation. Participants explicitly mentioned coda /s/ aspiration in SW's voice, with one individual commenting that aspiration in the word siesta (i.e., [sjéh.ta]) was 'the most determining factor to assess the person's origin.'

Another prediction that I had made regarding dialect classification was that participants would easily recognize the Madrid talkers as a result of media exposure and the presence of conservative dialect traits. This prediction was indeed true, as participants were most accurate in recognizing Madrid talkers. Accurate dialect identification of Castilian talkers is corroborated by Santana Marrero (2022), who recently found that Andalusian journalism students were adept at recognizing voices of Madrid talkers as a result of overt prestige and constant media representation. Along with quantitative success in dialect classification, qualitative comments supported participants' selections of Madrid talkers in the current study, with some referencing coda /s/ maintenance and the association of the voice with central and northern zones.

Another dialect classification prediction that I had made concerned potential confusion between Seville and Afueras talkers due to sharing many dialectal traits as a result of both being innovative Andalusian varieties. This expectation was correct, as Seville and Afueras varieties had the lowest overall accuracy. Referring to the confusion matrix in Table 7, both Afueras talkers

were most commonly associated with Seville on incorrect dialect classification trials, and SW was most frequently linked to Afueras on inaccurate responses. SM was the only exception to this pattern, being more often confused with Madrid, though there was a portion of incorrect trials that linked him to Afueras as well (14.3% of responses).

The more frequent confusion between the two Andalusian varieties is supported by Santana Marrero (2022), who found that Andalusian journalism students made significantly more errors in dialect classification of Seville talkers as a result of internal variation within the region. Santana Marrero (2022) also remarked that there is much variability at the individual level regarding which linguistic features are expressed, which was also represented in the 12 speech samples presented to participants in the current study. Along with the word-initial /s/ variant associated with each variety, talkers had access to numerous other forms of variation in the speech signal that could combine to allow for correct dialect classification or confusion (see Table 10).

Incorrect dialect classification responses in the current study followed a logical pattern, with the majority of errors occurring one step above or below on the prestige hierarchy (i.e., there were minimal cases of confusion between Afueras and Madrid talkers). In summary, the success of the dialect classification portion of the Verbal Guise allows for sufficient analysis of the language attitudes associated with talkers and varieties, as there were enough correct trials to consider associations with each talker. The high accuracy scores also validate the familiarity of the varieties to the Seville capital listeners, which was a central assumption of the current study.

4.2 Dialect Evaluation

4.2.1 Dialect Evaluation Predictions

I expected to find evidence for the prestige/standardness hierarchy attested in the literature for the three Peninsular varieties (Hernández-Campoy & Villena-Ponsoda, 2009; Regan, 2017a, 2019, 2022; Santana Marrero, 2016-2017, 2018b, 2022; Penny, 2004). This should manifest as higher evaluations of Madrid speakers for status adjectives such as 'refined' and 'formal,' followed by Seville. Given the covert prestige of Seville in Western Andalusia, and the fact that all participants are native Seville capital listeners, I also anticipated that Seville voices would be awarded the highest ratings for solidarity categories such as 'pretty' and 'proud.' Regarding Afueras voices, I predicted that these talkers would be associated with the lowest status ratings as a result of representing the least prestigious of the three Peninsular varieties (e.g., Penny, 2004; Regan, 2017a), and the fact that there is *ceceo* present in the speech samples. Despite the presence of the nonstandard variant, I expected Afueras voices to receive higher ratings than Madrid voices for solidarity adjectives, given that Afueras and Seville capital varieties are considered to be innovative and are geographically very close in western Andalusia (Penny, 2004).

I anticipated women to be evaluated differently than men, as research on perception of Peninsular dialect traits has shown varying results according to (perceived) talker gender (Barnes, 2015; Chappell, 2016a; Hadodo & Kanwit, 2020; Regan, 2017a, 2019, 2022; Santana Marrero, 2022). Following work that has examined *ceceo* perception (Regan, 2017a, 2019, 2022), I expected AW to receive lower status ratings than AM due to the presence of the stigmatized [s̪$^{\theta}$] variant. For Madrid and Seville varieties, it has been shown that women with *distinción* receive higher ratings than men with the same trait (Regan, 2022). Both Madrid talkers and the SW have

distinción, while SM has *seseo*. This could lead to higher status ratings for the Madrid and Seville women compared to the Madrid and Seville men, respectively. Other work has shown a solidarity benefit for men who produce non-standard dialectal variants (e.g., Barnes [2015] for [u] and [es] in Asturian Spanish; Chappell [2016a] for voicing of intervocalic /s/ in Costa Rican Spanish), which could affect ratings for AM. These predictions have been made according to specific traits present in the recordings, but it is likely that listeners employ a combination of acoustic cues that lead to their ratings, along with additional perceived social information.

4.2.2 Dialect Evaluation Ratings

The following section presents descriptive statistics regarding listeners' dialect evaluation. I focus on descriptive statistics because the evaluation task did not contain any direct experimental manipulations; the goal of this section is to link any qualitative trends to current literature on the perception of north-central and Andalusian varieties and to complement findings from the psycholinguistic task.

4.2.2.1 Dialect Evaluation by Variety

Dialect evaluation averages by variety are depicted in Table 17. Note that the data in Table 17, along with Figures 5, 6, and 7, are true representations of data for each variety, as only listener-talker pairings that corresponded to correct answers in the dialect classification task were included in the calculations. While dialect classification was mostly accurate across talkers, this step ensures that language attitude trends in the overall ratings are not affected by participants who associated the talker with a different variety. Adjectives in Table 17 are ordered by status and solidarity categories, with the highest scores highlighted in green for each adjective.

Table 17 Average Dialect Evaluation Ratings by Variety

		Afueras	Seville	Madrid
***fina**	'refined'	30.9	44.7	65.9
***pija**	'snobby'	24.3	35.0	63.6
***formal**	'formal'	43.8	56.3	62.0
educada	'polite'	58.5	69.8	69.1
tranquila	'calm'	68.2	62.1	67.9
alegre	'happy'	57.8	66.4	50.1
bonita	'pretty'	51.4	65.0	57.0
coloquial	'colloquial'	64.0	62.0	49.3
graciosa	'fun'	56.0	56.3	35.7
orgullosa	'proud'	42.4	55.1	49.8

Note: an * designates status categories, while no * indicates solidarity categories

At first glance, it is clear by observing the trends in Table 17 that participants evaluated the three varieties differently. Some adjectives have more of a range in ratings across the three varieties (e.g., 'snobby,' 'refined'), while ratings for other adjectives were less disparate (e.g., 'calm,' 'proud'). In a few cases, average ratings for two of the varieties were nearly identical (e.g., 'polite' for Seville and Madrid; 'fun' for Seville and Afueras). Another trend that can be observed in Table 17 is that participants evaluated their own variety quite favorably, with Seville earning the highest ratings for 50% of the adjectives, most of which are positively oriented.

To better visualize the data in Table 17, Figure 5 shows the same average adjective ratings by variety, with adjectives ordered according to the highest ratings associated with each variety.

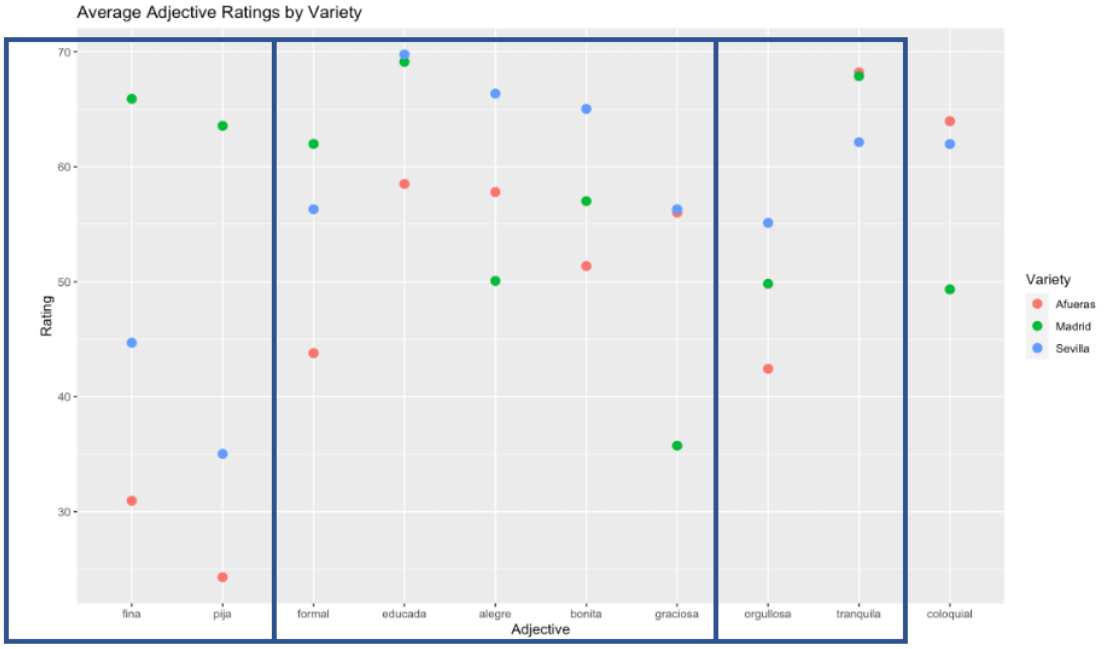


Figure 5 Ordered Adjective Ratings Across Varieties

As depicted in Figure 5, the adjectives most highly associated with Madrid were mostly linked to status categories. Madrid voices were rated as being the most 'refined' ($\bar{x}$ = 65.9), with average evaluations 21 points higher than the next closest variety, Seville ($\bar{x}$ = 44.7). Madrid voices were rated as the most 'formal' ($\bar{x}$ = 62.0), though the spread was smaller than that of 'refined,' with Seville only being approximately six points lower than Madrid. The adjective with the greatest distance in ratings between Madrid and the next closest variety (Seville; $\bar{x}$ = 35) was 'snobby,' with Madrid voices receiving an average of 29 points higher ($\bar{x}$ = 63.6).

Moving to adjectives most associated with Seville, the top-rated adjectives were solidarity categories. Participants rated voices of their own variety as being the most 'polite' ($\bar{x}$ = 69.8), closely followed by Madrid ($\bar{x}$ = 69.1), with less than one point in between. Seville listeners rated their own variety as being the 'happiest' ($\bar{x}$ = 66.4), with Afueras talkers awarded an average score

approximately nine points lower ($\bar{x}$ = 57.8), and Madrid voices being regarded as the least happy ($\bar{x}$ = 50.1). Seville voices were also evaluated as being the 'prettiest' ($\bar{x}$ = 65.0), followed by Madrid with an average score of eight points lower ($\bar{x}$ = 57.0), and Afueras being viewed as the least 'pretty' ($\bar{x}$ = 51.4). Seville and Afueras voices were rated almost identically for 'fun,' but Seville talkers earned a marginally higher score ($\bar{x}$ = 56.3) than Afueras ($\bar{x}$ = 56.0). There is a large spread between Seville and Afueras compared to Madrid voices for 'fun,' with Madrid voices receiving average ratings approximately twenty points lower ($\bar{x}$ = 35.7). Finally, Seville voices were regarded as being the most 'proud' ($\bar{x}$ = 55.1), followed by Madrid ($\bar{x}$ = 49.8), and then Afueras ($\bar{x}$ = 42.4).

The adjectives most associated with Afueras voices were both solidarity categories. Afueras voices were regarded as the most 'calm' ($\bar{x}$ = 68.2), though Madrid voices followed closely behind with an average difference of less than a half point ($\bar{x}$ = 67.9). Seville voices were evaluated as being the least 'calm' ($\bar{x}$ = 62.1). 'Calm' has a fairly small average spread, with only approximately six points separating the three varieties. Finally, Afueras voices received the highest average ratings for 'colloquial' ($\bar{x}$ = 64.0), closely followed by Seville ($\bar{x}$ = 62.0), and then Madrid ($\bar{x}$ = 49.3).

4.2.2.2 Dialect Evaluation Ratings by Talker Gender

The results depicted in Figure 6 show differences in average dialect evaluation scores from correct dialect classification trials, collapsed by talker gender.

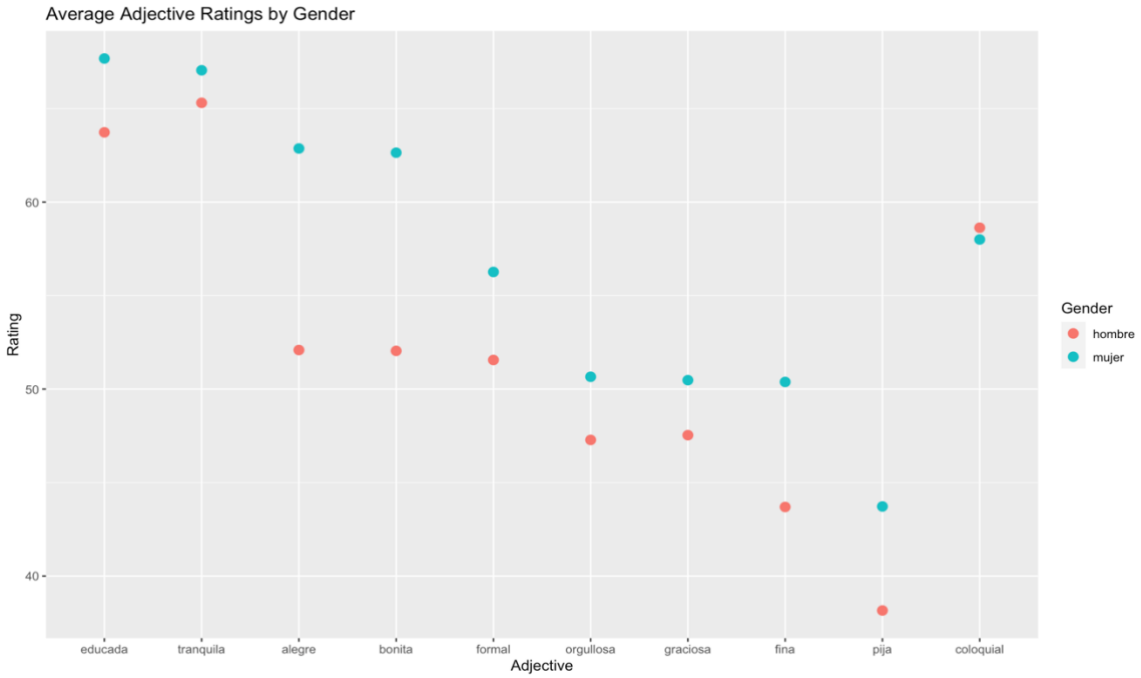


Figure 6 Ordered Adjective Ratings by Talker Gender

Perhaps the most obvious result is that women received higher average scores for all adjectives except 'colloquial,' in which men were evaluated to a marginally greater degree ($\bar{x}$ = 58.6 for men; $\bar{x}$ = 58.0 for women). The adjectives in which women talkers received higher scores included both status and solidarity categories. Some adjectives had a small spread between average ratings across men's and women's voices (e.g., 'colloquial,' 'calm,' 'fun,' 'proud,' 'polite,' 'formal), while others showed greater distance between ratings (e.g., 'happy,' 'pretty,' 'refined,' 'snobby'). Overall, the higher scores awarded to women's voices reflect positive attributes, with the exception of 'snobby.'

Similar to the dialect classification scores, I expected variation in average ratings at the talker level. Figure 7 depicts average ratings for men's and women's voices, organized by variety.

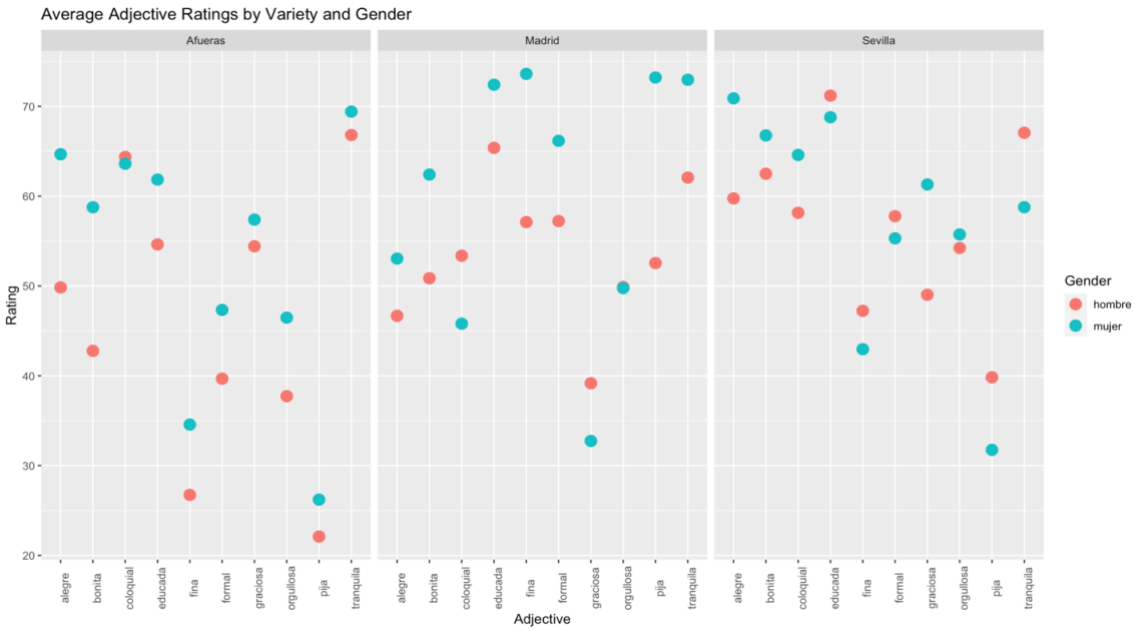


Figure 7 Average Adjective Ratings by Variety and Talker Gender

Considering both the role of variety and gender in dialect evaluation, there are a few notable differences when comparing Figure 7 (by talker; i.e., by variety and gender) to Figures 5 (variety) and 6 (gender). First, the distribution of scores reveals that AM received some of the lowest average scores (e.g., 'snobby' and 'refined'), while MW had the highest average ratings for multiple adjectives (e.g., 'refined,' 'snobby,' 'polite,' 'calm'). Another pattern is evident upon looking at the Seville talkers. Without collapsing gender, some of the trends that were observed in Figure 7 are reversed for Seville talkers. For example, SM's voice received higher average ratings for 'polite,' 'refined,' 'formal,' 'snobby,' and 'calm,' while SW's voice received higher rating for 'colloquial,' all of which were the reverse trend when averages were collapsed across gender. For Afueras and Madrid talkers, the gender patterns hold with the exception of MM's voice being rated as more 'fun' and just barely more 'proud' than that of MW.

4.2.3 Dialect Evaluation for Incorrect Dialect Classification Responses

The dialect evaluation results presented in the previous section were all calculated only considering talker-listener pairs where participants had correctly identified the talker in the dialect classification portion of the trial. However, looking at incorrect trials can offer insight on how dialect classification and language attitudes may be related.[9] This analysis provides something similar to a naturally occurring top-down processing task (Hay & Drager, 2010; Hay et al., 2006a; Niedzielski, 1999), in that the overt presentation of the three dialect regions as answer choices, and any resulting activation of regional associations, could influence listeners' perception of talkers' traits. Any differences in ratings between correct and incorrect responses of the same talker can be surmised to be a result of differential language attitudes that participants associate with each variety. Moreover, comparisons can be drawn between findings of the current study and the robust literature on matched-guise phenomena in Spanish (e.g., Barnes, 2015; Regan, 2019).

Before exploring the relationship between the language attitudes and adjective ratings in the current study, it is important to note the bidirectional nature of the perception of social and linguistic information (Barnes, 2015; Campbell-Kibler, 2009; Chappell, 2016a; Walker et al., 2014). In other words, it is possible that a participant rated MW's voice as being more snobby due to the fact that they associated her with a central talker, or that they chose Madrid in dialect classification due to the voice sounding snobby. Given that the dialect classification question always occurred first in the trial sequence before participants completed any ratings of the voices,

[9] This comparison includes ratings from participants who were correct in one trial but not in another, as well as trials where the talker was incorrectly classified both times.

it could potentially be hypothesized that the selection of variety (i.e., regional association) would have a slightly larger role in influencing the ratings than the opposite. However, no causality can be definitively inferred.

In order to highlight trends between language attitudes and adjective ratings, I considered talkers who had at least ten incorrect dialect classification trials. Following this criterion, I was able to examine results of four talkers: SM, AM, MM, and AW. Participants were highly accurate in identifying MW and SW; there were not enough trials to justify an analysis of ratings for incorrect dialect classification. SM had enough trials to consider confusion with both the Afueras and Madrid varieties, but the other three talkers were only predominantly associated with one other variety. Table 18 outlines the results of the incorrect classification analysis for SM, with adjectives ordered by most to least amount of overall change for Madrid.

Table 18 Language Attitude Analysis for the Seville Man Talker (SM)

Talker: SM	**Guessed Seville** (*N* = 52)	**Guessed Madrid** (*N* = 20)	**Guessed Afueras** (*N* = 12)	**Overall change** (Madrid-Seville)	**Overall change** (Afueras-Seville)
pija ‘snobby’	39.8	56.5	39.5	+ 16.7	- 0.3
fina ‘refined’	47.2	56.0	45.8	+ 8.8	- 1.4
formal ‘formal’	57.8	64.6	53.9	+ 6.8	- 3.9
orgullosa ‘proud’	54.2	57.7	44.8	+ 3.5	- 9.4
educada ‘polite’	71.2	68.8	63.5	- 2.4	- 7.7

coloquial ‘colloquial’	58.2	54.7	58.8	- 3.5	- 0.6
tranquila ‘calm’	67.1	61.9	66.1	- 5.2	- 1.0
alegre ‘happy’	59.8	50.4	60.2	- 9.4	+ 0.4
bonita ‘pretty’	62.5	52.3	58.8	- 10.2	- 3.7
graciosa ‘fun’	49.0	36.0	47.7	- 13.0	- 1.3

As observed in Table 18, average ratings appear to change according to participants’ evaluation of talker origin. Starting with individuals who thought SM was from Madrid, the most notable differences in average ratings are associated with the adjectives ‘snobby,’ ‘fun,’ ‘pretty,’ and ‘happy.’ On average, participants who thought SM was from Madrid rated him 16.7 points higher for ‘snobby.’ On the contrary, SM received a lower rating for ‘fun’ (-13.0 points) from being associated with Madrid instead of Seville. Ratings also decreased for ‘happy’ (-9.4 points) and ‘pretty’ (-10.2 points) in the context of a Madrid dialect classification. Other differences between Seville and Madrid ratings for SM resulted in the talker being evaluated as more ‘refined,’ ‘formal,’ and ‘proud,’ yet less ‘calm,’ ‘colloquial,’ and ‘polite’ when linked to a Madrid talker.

Confusion between Seville and Afueras for SM was also associated with a change in the average adjective ratings, but to a slightly lesser extent (i.e., the largest difference was a change of -9.4 compared to + 16.7 for Madrid confusion). When confused with Afueras, SM was rated as less proud (-9.4 points) and less polite (-7.7 points) than his average ratings associated with Seville. There were other changes, though occurring to a smaller degree. SM was also evaluated to be less ‘formal,’ ‘pretty,’ ‘refined,’ ‘fun,’ and ‘calm’ by those who selected Afueras in the dialect

classification phase. To a minimal extent (i.e., less than a one point difference), SM was rated as less colloquial and snobby, but happier when associated with Afueras.

Along with SM, MM was occasionally confused with a Seville talker during the dialect classification portion of the task. The confusion matrix for MM has been outlined in Table 19, with adjectives ordered by most to least amount of overall change.

Table 19 Language Attitude Analysis for the Madrid Man Talker (MM)

Talker: MM	**Guessed Madrid** (*N* = 71)	**Guessed Seville** (*N* = 11)	**Overall change**
tranquila ‘calm’	62.1	77.8	+ 15.7
educada ‘polite’	65.4	77.5	+ 12.4
bonita ‘pretty’	50.9	63.2	+ 12.3
coloquial ‘colloquial’	53.4	63.3	+ 9.9
alegre ‘happy’	46.7	52.1	+ 5.4
graciosa ‘fun’	39.2	44.6	+ 5.4
formal ‘formal’	57.2	61.5	+ 4.3
orgullosa ‘proud’	50.0	44.1	- 5.9
fina ‘refined’	57.1	47.0	- 10.1
pija ‘snobby’	52.5	41.3	- 11.2

As depicted in Table 19, participants also evaluated MM differently according to perceived geographic origin. There are many parallels to rating shifts in Table 18 between SM and MM guesses. For example, the majority of adjectives that exhibited a decrease in ratings when associated with MM instead of SM in Table 18 decreased in Table 19 when MM was confused with SM. For example, MM’s voice was rated as more ‘calm’ (+ 15.7), ‘polite’ (+12.4), ‘pretty’ (+12.3), ‘colloquial’ (+ 9.9), ‘happy’ (+ 5.4), and ‘fun’ (+ 5.4) when participants thought he was from Seville. The only mismatch between these adjectives between Tables 18 and 19 pertains to

'formal,' which increased when SM was confused with Madrid (+ 6.8), but also slightly increased when MM was confused with Seville (+ 4.3). Ratings for adjectives demonstrating decreases in Table 18 also followed a similar pattern, with average scores rising in Table 19. MM was rated as less 'snobby' (- 11.2), 'refined' (-10.2), and 'proud' (-5.9) by participants who thought he was from Seville.

There were also enough incorrect responses in dialect classification of AM to compare ratings. Participants who were unsuccessful in identifying AM as an Afueras talker most frequently confused him with a Seville talker. Ratings from this misidentification can be observed in Table 20, which have been ordered from most positive to negative change.

Table 20 Language Attitude Analysis for the Afueras Man Talker (AM)

Talker: AM	**Guessed Afueras** (N = 62)	**Guessed Seville** (N = 21)	**Overall change**
bonita 'pretty'	42.8	56.7	+ 13.9
orgullosa 'proud'	37.7	48.2	+ 10.5
formal 'formal'	39.7	50.1	+ 10.4
educada 'polite'	54.6	60.9	+ 6.3
fina 'refined'	26.7	30.9	+ 4.9
pija 'snobby'	22.1	27.0	+ 4.9
tranquila 'calm'	66.8	66.5	- 0.3
alegre 'happy'	49.8	48.6	- 1.2
graciosa 'fun'	54.4	49.5	- 4.9
coloquial 'colloquial'	64.4	59.3	- 5.1

When comparing the results in Table 20 (i.e., listeners who thought AM was from Seville) with those in Table 18 (i.e., listeners who thought SM was from Afueras), there are some similar trends. For instance, while an Afueras guess for SM yielded lower voice ratings for 'proud,'

'polite,' 'formal,' 'pretty,' 'refined,' and 'calm,' these adjective ratings all increased when AM was thought to be from Seville with the largest increases being associated with 'pretty' (+ 13.9), 'proud' (+ 10.5), 'formal' (+ 10.4), and 'polite' (+ 6.3). Moreover, there was a small increase for 'happy' when SM was associated with Afueras (+ 0.4), which slightly decreased when AM was associated with Seville (- 1.2). The remaining adjectives did not pattern similarly between the two tables, as AM was rated as less 'colloquial' (- 5.1), 'fun' (- 4.9), and 'calm' (- 0.3) when associated with Seville in Table 20. It should be noted that the amount of overall change seems to be greater when confusion occurs with the more prestigious variety (i.e., AM said to be from Seville) than the more stigmatized variety (i.e., SM said to be from Afueras).

The last talker for whom there were sufficient incorrect responses in the dialect classification task is AW. Participants who did not correctly guess that she was from Afueras most frequently selected Seville instead. The comparison of adjective ratings for AW can be observed in Table 21, with adjectives sorted by overall change.

Table 21 Language Attitude Analysis for the Afueras Woman Talker (AW)

Talker: AW		**Guessed Afueras** (N = 72)	**Guessed Seville** (N = 11)	**Overall change**
bonita	'pretty'	58.8	69.0	+ 10.2
alegre	'happy'	64.7	73.8	+ 9.1
graciosa	'fun'	57.4	66.2	+ 8.8
fina	'refined'	34.6	43.1	+ 8.5
pija	'snobby'	26.2	34.5	+ 8.1
educada	'polite'	61.8	69.5	+ 7.7
formal	'formal'	47.3	50.6	+ 3.3
orgullosa	'proud'	46.5	49.1	+ 2.6
coloquial	'colloquial'	63.6	57.5	- 6.1

tranquila	'calm'	69.4	59.9	- 9.5

As demonstrated in Table 21, there is also a shift in ratings for AW according to the geographic origin participants assigned to her. Participants who thought she was from Seville rated her voice as being more 'pretty' (+ 10.2), 'happy' (+ 9.1), 'fun' (+ 8.8), 'refined' (+ 8.5), 'snobby' (+ 8.1), and 'polite' (+ 7.7). Ratings for 'formal' (+ 3.3) and 'proud' (+ 2.6) also increased, though to a lesser extent. The two remaining adjectives decreased in average ratings due to dialect confusion, with AW receiving lower scores for 'colloquial' (- 6.1) and 'calm' (- 9.5) when confused with Seville. Comparing prior trends can allow for an observation of gender differences associated with dialect confusion. The adjective that showed the most positive increase was shared for both AM and AW (i.e., 'pretty'). Many of the other adjectives shared common patterns as well across genders. A different directionality of overall change was associated with two adjectives: 'happy' and 'fun.' Confusion with Seville for AM led to a small decrease in ratings for these two adjectives, while on the contrary, AW was rated as more 'happy' and 'fun' with a Seville association.

4.2.4 Participant Commentary on Talker Voices

At the end of each trial, participants had the option to leave feedback regarding their answer choices or the talkers' voices in general. Many participants wrote comments, which have been compiled for each talker. Trends for each variety will be briefly summarized, with more in-depth conclusions drawn in the next section. Unaltered comments left for all talkers are provided in Appendix D.

The most noteworthy trends from the comments about Afueras voices pertain to specific phonetic traits possessed by the talkers, regional associations, and characteristics of the voices. Many comments explicitly referenced the presence of *ceceo*, but it was mentioned more frequently for AW ($N = 6$) compared to AM ($N = 1$). Participants reinforced their dialect classification choices of Afueras by commenting on how the voices were representative of small towns outside of Seville capital, with some explicitly naming locations such as El Coronil or Utrera. Adjectives such as happiness, kindness, and ruralness were associated with the voices. Along with *ceceo,* coda /s/ aspiration was referenced by a participant who remarked for AM, '*se come muchas letras'*, literally translating to 'many letters are eaten.' A final comment of interest pertains to the diversity of voices present in Seville capital, as one participant explicitly mentioned the plurality of dialects present in the city.

Similar to feedback left for Afueras talkers, commentary surrounding the Madrid talkers highlighted specific phonetic traits, regional identification, and characteristics of the voices. Participants referenced /s/ production in Madrid talkers, with one explicitly noting coda /s/ maintenance. Anther comment linked the /s/ pronunciation to snobbiness, though it is unclear whether the person is referring to coda /s/ maintenance or the apico-alveolar place of articulation. Another participant indicated that production of /x/ was an indicator that MW was not from Andalusia, as the aspirated variant tends to be realized in innovative varieties of southern Spanish (e.g., [xen.te] vs. [hen.te] for *gente* 'people'; Fernández de Molina Ortés & Hernández-Campoy, 2018). Regarding regional identification, numerous participants commented that both MW and MM seemed to be from outside of Andalusia, connecting them to either central or northern Spain. Two individuals remarked that MW could be from Seville, but that she was adopting a more neutral

accent to seem like she was from central Spain. Finally, adjectives used to reference Madrid voices included snobby, simple, monotonous, serious, colloquial, pretty, pretentious, and repulsive.

As was the case for Afueras and Madrid comments, patterns from feedback about Seville talkers similarly referenced pronunciation traits, regional identification, and adjectives to describe the voices. Overall, there were more comments confirming SW's geographic origin than that of SM, as participants remarked that while he seemed to be southern, there was an element of forced neutrality. On the contrary, participants noted that SW's voice was natural and clearly identifiable as being from Seville. Coda /s/ aspiration was explicitly mentioned for both talkers, with one participant indicating that it was the determining factor in dialect classification for SW. Other comments did not denote specific traits, but remarked that the accent, pronunciation, and tone were clearly identifiable as being Sevillian. Finally, adjectives such as typical, warm, and natural were used to describe SW, with familiar, neutral, and forced being associated with SM.

4.2.5 Dialect Evaluation Discussion and Interpretation

To briefly summarize the key results from the dialect evaluation phase of the Verbal Guise, there were clear differences among average ratings for the three varieties. Results followed a prestige hierarchy with status predominantly associated with Madrid, followed by Seville, and then Afueras. Solidarity ratings were highest for Seville, followed by Madrid or Afueras depending on the adjective. Afueras received the lowest ratings for all status adjectives, emphasizing the stigma associated with innovative dialect traits such as the use of [s̪$^{\theta}$]. Women's voices received more positive average ratings across both status and solidarity categories than men's voices. In an analysis approximating a naturally occurring top-down processing task, participant ratings differed according to their guesses in the dialect classification phase.

One expectation was that the prestige hierarchy attested in the literature (e.g., M > S > A; Santana Marrero, 2022) would be reflected in the evaluative data of the current study. This prediction was confirmed via predictable differences in average ratings of the voices according to the region they were associated with. For instance, Madrid voices were the most highly rated for status adjectives (e.g., 'refined,' 'formal,' and 'snobby'), corroborating the fact that Andalusians view central varieties as the national standard and associate them with overt prestige. Both Seville capital and Afueras voices were evaluated as being less refined, formal, and snobby compared to Madrid, with Seville capital always receiving higher scores than Afueras. The evidence of overt prestige of Madrid talkers by southern listeners is corroborated by Santana Marrero (2022), who found that Madrid speakers had higher ratings for urbanity and clarity.

Matching the covert prestige held in Seville, I expected to find high solidarity ratings for Seville capital talkers. These predictions were upheld, as Seville listeners rated their own variety the highest for solidarity categories such as 'happy,' 'pretty,' 'proud,' and 'fun.' This result corroborates Santana Marrero (2022), who also found that Seville voices were rated the highest for solidarity adjectives such as 'pretty' and 'pleasant.'

Another anticipated result was more similarity between Seville capital and Afueras voices for solidarity ratings as a result of dialect innovation. I also expected Afueras talkers to receive the lowest ratings for status categories as a result of the stigma associated with [s̪$^{\theta}$] production, paired with Afueras talkers possessing the least amount of social prestige. These predictions were partially upheld, as Afueras voices received the lowest ratings for all status adjectives (e.g., 'refined,' 'snobby,' and 'formal'). Afueras voices were highly evaluated in some solidarity categories (e.g., 'fun,' 'calm,' 'colloquial'). They were also rated as being happier than those of Madrid talkers but less happy than Seville capital voices. Despite similar patterning for some

solidarity categories, the hierarchy was not upheld for others, possibly being attributed to the stigma associated with *ceceo* productions. For example, Afueras voices were evaluated as being the least pretty, with Seville capital listeners rating Madrid voices higher than those of their close geographic neighbors. Ratings for proudness followed a similar pattern, with Afueras voices receiving the lowest average scores despite the adjective falling into a solidarity category. Taken together, while the prestige hierarchy attested in the literature is predominantly upheld in the rating data present in the current study, it is also clear that in some cases, overt and covert prestige may interact to alter the ordering for certain adjectives.

Moving away from the raw data of the dialect evaluations, the language attitudes associated with each talker and variety are complementary to attested trends in the matched-guise literature (see section 2.5). Recall that results from matched-guise studies on numerous varieties of Spanish have indicated the perceptual shifts that can occur when small changes to the acoustic signal are made. For example, adding coda /s/ aspiration to non-coastal Mexican Spanish led to the talkers being rated as 'more Caribbean,' despite all other acoustic cues remaining the same (Chappell, 2019). Matched-guise results from the perception of Andalusian Spanish demonstrated that *ceceo* guises led participants to evaluate talkers as being more rural and less educated than those with *distinción* (Regan, 2022). While the Verbal Guise in the current study is more flexible in methodology than a matched-guise task in terms of how much linguistic variability is contained in the each guise (i.e., the manipulation of one sociophonetic variant in matched guise compared to no acoustic manipulations across verbal guise trials), ratings from incorrect dialect classifications of talkers can be considered as a naturally occurring top-down processing task.

Matched-guise results from Regan (2019) showed the effect of *ceceo* on dialect evaluation, finding that speech samples with *ceceo* were rated significantly lower for perceived education level

by participants from Lepe than those containing *distinción*. In a follow-up experiment including listeners from both Huelva and Lepe, the result was replicated in that *distinción* guises had significantly higher perceived status ratings than *ceceo* guises (Regan, 2022). A parallel finding has been demonstrated in the current study through observing confusion of the Afueras talkers with Seville in dialect classification. Despite the presence of *ceceo* in the recordings, participants who thought that both AM and AW were from Seville rated their voices as sounding more polite than those who correctly guessed that they were from Afueras. Though SM's speech samples had a dental [s̪] instead of interdental [s̪$^{\theta}$], participants who thought that SM was from Afueras instead of Seville rated his voice as sounding less polite (i.e., an average change of -7.7 points). It is intriguing that the pattern was upheld for SM, given that the stigmatized trait was not even present in the speech samples.

Regan (2019, 2022) also observed significant results for perceived formality, with *distinción* guises being evaluated as more formal than *ceceo* guises. The current study mirrors these findings, as AM and AW voices both received higher average ratings for formality when participants associated them with Seville instead of Afueras (an increase of + 10.4 and +3.3 points, respectively). On the contrary, SM received slightly lower ratings for formality when associated with Afueras (an average decrease of -3.4 points).

Research on other matched-guise studies on Spanish dialectal phenomena (Barnes, 2015; Chappell, 2016a, 2019; Regan, 2020; Walker et al., 2014) have found similar patterning of perceived status according to the standardness of the variants. Though the work spans multiple contexts of variation and geographic locations, in all of the aforementioned studies, the non-standard variant is evaluated as having lower perceived status than the standard variant. These results, combined with those found in Regan (2017a, 2019, 2022) for *ceceo* in Andalusian Spanish,

support the association of [s̪θ] with non-standardness and the downgrading for status categories that occurs when listeners in the current study either hear the stigmatized dialect trait or associate a talker with Afueras.

Along with matched-guise work reliably supporting the perceived status shifts according to the standardness of the variants, there have also been findings for solidarity benefits of the non-standard variants, though less consistently across contexts. For instance, results from Regan (2020) indicated that participants from Western Andalusia rated guises with the non-standard, regional variant [ʃ] as being more friendly than guises with standard [t͡ʃ]. Despite the positive association of some non-standard variants in matched-guise research, other non-standard variants have not been linked to solidarity increases. For example, Regan (2019, 2022) found no significant effects of perceived friendliness by variant when comparing *ceceo* and *distinción* guises. In the current study, results do show evidence for increased solidarity ratings when talkers were perceived to be from Seville instead of Afueras. While these results are not just due to the presence or absence of one linguistic trait, they do generally corroborate attested trends in the matched-guise literature and also support the overall patterns of status and regional prestige found in the rating data from correct dialect evaluations.

Another finding that is evident in the dialect evaluation results of the current study, along with a large body of literature on sociolinguistic variation, is the effect of gender. In the current study, women's voices were generally evaluated more positively than men's voices, though results should be interpreted with caution as a result of a small number of women talkers (N = 3). Santana Marrero (2022) also found evaluation differences by gender and variety in their data, with Seville women talkers' voices being rated as simpler, softer, and more beautiful than Seville men's, but Madrid men being rated as more pleasant, beautiful, and funny than Madrid women. The current

study found a similar result for Seville, with SW receiving higher solidarity ratings (e.g., pretty, happy, proud, funny) than SM. However, contrary to Santana Marrero (2022), solidarity ratings were also higher for MW when compared to MM. This might be explained by the fact that results from Santana Marrero (2022) included both spontaneous and read speech samples in the gender breakdown by variety, while results from the current study are all based on spontaneous speech.

A number of the matched-guise studies referenced above have also found evaluation differences according to talker gender and the perceptual category being analyzed. For example, Regan (2019) observed gender differences in the perception of socioeconomic status, which only affected the *distinción* guises. While both men and women with *ceceo* guises received lower ratings for perceived socioeconomic status, men who produced *distinción* received higher ratings than women. In a subsequent study, Regan (2022) found that women with *distinción* guises were rated as more formal than men with *distinción* guises, but less formal than men with *ceceo* guises. Chappell (2016a) found noteworthy differences in solidarity ratings of the non-standard voicing of intervocalic /s/ according to talker gender. Non-standard guises for men were associated with significantly higher ratings for niceness, confidence, and Costa-Rican-ness. However, this was not the case for women, as their guises did not receive any solidarity benefit for producing the stigmatized variant. In analyses of Andalusian *ceceo* (Regan 2019; 2022), there were no significant effects of pleasantness according to the guise variant (i.e., *ceceo* or *distinción*), so the category was excluded from further analysis. Lastly, Barnes (2015) found significant effects for masculinity ratings of [u] and [es] guises produced by men, while Asturian guises produced by women were rated as less feminine. It is clear from the aforementioned studies that gender differences can be associated with shifts in status and solidarity evaluations, but that changes are linked to specific speech communities and variants.

One last component of the Verbal Guise that can be analyzed and connected to the current body of literature on Peninsular language attitudes are the optional comments written by participants at the end of each trial. These qualitative data support the quantitative trends for both dialect classification and evaluation sections of the task. Noteworthy patterns will be analyzed both by talker and variety.

Starting with commentary for Seville talkers, as was the case in both dialect classification and evaluation trends, participants responded differently to SM when compared to SW. Listeners were sensitive to the fact that SM's voice was different than that of casual, Seville capital speech, with some explicitly remarking that SM could have been adjusting his speech patterns to adopt more conservative dialect traits. Participants stated that SM had 'southern touches but a more neutral voice'; was 'a Sevillian trying to speak in a refined way'; and 'a Sevillian with not much of an accent or a forced neutral accent.' While some participants took this neutrality as an indication that SM was from Madrid instead of Seville, others correctly classified him as a Seville talker, though noting the fact that he 'doesn't always pronounce in an Andalusian way.' It has been well documented that Andalusian news anchors and politicians frequently mask their innovative speech traits when appearing in public spheres or speaking in more formal settings (Cruz Ortiz, 2020; León-Castro, 2016; Santana Marrero, 2022). Thus, Sevillians are already aware of production shifts by talkers who wish to converge to north-central varieties, which is indeed the case in the current study.

Contrary to comments associated with SM's recordings, participants remarked that SW's voice was 'not at all forced' and 'very natural.' It is possible that the contrast between Seville capital talkers encouraged such feedback, as some participants likely heard a trial from SW before SM. Moreover, unlike comments associated with SM's more neutral southern accent, multiple

participants emphasized that SW's voice was fully identifiable as being from Seville capital. The qualitative remarks support the quantitative trends in dialect classification accuracy for the two Seville talkers. Taken together, SW was associated with high accuracy, while SM's voice was linked to more incorrect responses and a perceived element of forced neutrality by some listeners.

Moving to Madrid talkers, general trends in participant commentary also match quantitative patterns in dialect classification. For instance, remarks about the geographic origin were largely consistent with central-northern regions of Spain. The majority of participants who referenced the locations of both MM and MW associated them with places outside of Andalusia. However, similar to feedback for SM, two individuals remarked that MW could be a Sevillian forcing a more neutral accent. Two participants remarked that MM's voice was serious yet colloquial, while MW's voice was said to be pretty, yet snobby. Similar to rating data from Santana Marrero (2022), Madrid voices were associated with adjectives such as 'simple,' 'plain,' and 'monotonous,' while these descriptors were not present in commentary regarding the Andalusian talkers. Finally, the different ideologies that citizens from the same region can associate with talkers and varieties is present in these data, as MW's voice was perceived as 'pretty' by one listener but 'repulsive and pretentious' by another.

Lastly, for Afueras talkers, dialect classification comments largely supported the association of both AM and AW with the small towns outside of Andalusia. It appears that *ceceo* was a primary identifying feature that aided in classification of these talkers, as it was explicitly referenced for both individuals. Contrary to some negative commentary present in rating data from Regan (2019, 2022) for *ceceo* guises, adjectives associated with both AM and AW were positive, with no participants explicitly speaking about the voices or presence of *ceceo* with any condescending attitudes. For example, some participants described the voices of talkers that

appeared in *ceceo* guises in Regan (2022) as 'crude' or 'uncultured.' However, rural, kind, and happy were adjectives used to describe Afueras talkers in qualitative data in the current study. Participants in Regan (2019, 2022) were from Huelva and Lepe, which are both locations in Western Andalusia that are associated with *ceceo*, though Lepe has a stronger link to the local variant. It is possible that some participants had stronger language attitudes regarding *ceceo* as a result of being more frequently exposed to it when compared to the Seville capital participants in the current study.

4.3 Verbal Guise Results Implications for the Long-Term Form Priming Task

As a result of the completion of the Verbal Guise before the word recognition task, results from the former can be used to establish predictions for patterns in the Long-Term Form Priming task. First, the fact that participants were largely accurate in dialect classification ensures that the activation of language attitudes for each talker are representative of each variety/geographic location. In other words, when participants hear /s/ tokens produced by Afueras talkers, the social information about the voices that was explicitly evaluated in the Verbal Guise will generally be available to the listeners. However, as the form priming task probes implicit responses, any interaction of linguistic and social information in word recognition will occur subconsciously.

The confirmation of a prestige hierarchy with the three varieties, and the variable evaluation of the talkers according to status and solidarity categories, can help support any significant patterns in immediate and long-term processing present in the word recognition data. In immediate processing, due to the listeners' familiarity with all three varieties, it is likely that the /s/ variants will be recognized equally. While a body of literature has established patterns of

prestige and stigma of Peninsular variants, having dialect evaluation results for the exact same talkers that appear in the psycholinguistic task provides a higher level of support for the interpretation of any significant findings.

Implications from the Verbal Guise dialect evaluation results can also lend themselves to connections for long-term processing results. The literature surrounding north-central varieties of Peninsular Spanish, coupled with the results from the Verbal Guise in the current study, support the overt prestige associated with the Madrid voices. If the Madrid [s̪] variant is privileged in the long-term, I would expect to see evidence of effective priming via faster reaction times in Block 2 for the apico-alveolar variants. It would also be expected that the Seville capital /s/ would be encoded in the long-term, given that it is the native variety of the participants, the variety's covert prestige within Andalusia, and the high affective scores awarded to Seville capital voices in the Verbal Guise. Lastly, referring to both the literature and Verbal Guise results for Afueras voices, I would not expect Afueras voices to be associated with priming effects. Previous research (Sumner & Kataoka, 2014) did not find evidence of priming for the NYC variety, which would be most similar to Afueras in the current study. Any lack of priming could be interpreted as a result of the nonstandardness associated with the [s̪$^{\theta}$] production, which despite solidarity between Andalusian varieties in explicit dialect evaluation, could be more present at the implicit level of processing (Sumner et al., 2014). However, there are critical differences between the Afueras variety to the Seville capital listener and the NYC variety to GA participants in Sumner and Kataoka (2014) who were unfamiliar with non-rhotic varieties of English. If there is priming for Afueras, this could be interpreted as the familiarity of the variant to the listeners, as well as the shared solidarity between innovative Andalusian varieties, overtaking any potential processing cost that might be associated with the stigma of the [s̪$^{\theta}$] production.

Regarding other notable trends in the Verbal Guise results, one key finding that could affect word recognition patterns is the differential treatment of talker gender. As women's voices were more identifiable to listeners (i.e., higher dialect classification accuracy) and also evaluated more positively on average than men's voices, this could lead to disparate effects according to gender in processing results. Regarding specific talkers, it is possible that SM will pattern differently in processing as a result of being evaluated as more formal, and/or due to his weaker association with the Seville variety.

5.0 Word Recognition Analyses

The current chapter focuses on the processes of auditory word recognition (i.e., the matching of variable surface forms to existing lexical categories), encoding (i.e., the updating of the lexical category to include the new token), and representation (i.e., the makeup of tokens that have previously been recognized and encoded), all of which are areas of particular interest in cross-dialectal speech perception (Clopper et al., 2016). This word recognition experiment explores both immediate and long-term processing via a long-term form priming task (see section 2.1.2 for review).

The word recognition task in the current study consists of two blocks composed of half real words and half pseudowords, and on each trial listeners must indicate if the item presented is a real or made-up word. Half of the real word items are repeated in Block 2, while half are new words. This long-term form priming paradigm offers multiple benefits in exploring the research questions of the current study. First, an analysis of Block 1 presentations of items for accuracy and response time can reveal how real control words and the surface forms of the three /s/ variants are recognized by Seville capital participants. Considering Block 2 items provides information on encoding and representation of the /s/ variants, as prime-target pairs occur with a gap of 20-30 minutes. Thus, the processes of cross-dialectal auditory word recognition, encoding, and representation can all be probed within the design of the current study.

Current literature reflects an asymmetry between immediate and long-term processing, with evidence suggesting that dialectal variants can be recognized equally well by experienced listeners in the short term, but observe storage (i.e., encoding) of only a canonical variant in the

long-term (Sumner et al., 2014). Recent research has worked to update the theoretical framework of the mechanisms that can explain this asymmetry, given that purely frequency-based models of speech perception fail to capture the range of attested findings (Sumner et al., 2014; Clopper, 2021). This chapter relates the findings from the current study to processes of word recognition, encoding, and representation, interpreting them in the context of the existing literature.

5.1 Data Cleaning and Task Accuracy Trends

In total, there were N = 30,240 observations (i.e., N = 42 listeners x 720 trials). Overall accuracy for the task as a whole was high, with a rate of 86.0% across both experiment blocks (i.e., N = 26,019 correct trials). Participants were more accurate in recognizing real words (89.8% correct) when compared to pseudowords (82.3% correct). This trend also held across all conditions, with participants being most accurate with real fillers and control words when compared with pseudo fillers and pseudo control words. Participant accuracy ranges across all conditions were from 68.6% correct (i.e., Participant 514) to 93.9% correct (i.e., Participant 041). The data from one participant (Participant 514) was excluded from the analysis due to having less than 75% accuracy across real word items.

I also excluded four items, all of which were intended as pseudowords (i.e., *cuapa*, *palgas*, *vieta*, and *porbo*). *Porbo* is a legitimate variant for *polvo*, 'dust,' in Andalusian Spanish, despite not being included as a lexical item in the *Real Academia Española*. Participant selections confirmed this regional acceptability, as 86% of the responses indicated that it was a real Spanish word. *Cuapa* and *vieta* were acoustically very similar to *guapa, '*pretty woman*'* and *dieta*, 'diet,' the real words from which they were derived. Finally, *palgas* sounded similar to *valgas*, the

second-person singular subjunctive of the verb *valer*, 'to be worth.' All four excluded tokens had an accuracy rate of less than 50% when they were treated as pseudowords, thus being classified as real words by the majority of participants.

After participant and item exclusions, I calculated the accuracy for the task as a whole. Average accuracy by condition and by talker can be viewed in Tables 22 and 23. Accuracy rates across experiment blocks remained stable (e.g., 86.5% correct in Block 1 for all items, 87.7% in Block 2; 90.4 % for Block 1 real words; 90.1% for Block 2).

Table 22 Accuracy Results by Condition

Condition	% Correct
Real Fillers	91.7%
Real Controls	90.7%
Critical /s/ Words	87.4%
Pseudo Controls	84.1%
Pseudo Fillers	83.7%

Table 23 Accuracy Results by Variety

Variety	% Correct
Afueras	86.0%
Madrid	87.7%
Seville	87.6%

All statistical analyses were conducted in R (R Core Team, 2021). To determine whether accuracy for /s/ words was significantly different by Variety, I conducted a logistic regression of accuracy, including an interaction between Variety and Condition (see Appendix E.1). I set the reference levels to the Seville capital variety and real control items. The model indicated no significant effects of Variety or Condition alone, but there was a significant interaction between

Variety (Afueras) and Condition (/s/) (β = -1.03, $p < 0.001$). In other words, the /s/ word effect is significantly different for Afueras when compared to Seville talkers. When compared to accuracy rates for Seville, scores for critical /s/ words produced by Afueras talkers were lower (80.1% correct Afueras; 91.8% Seville). There was no significant interaction for Madrid /s/ items, indicating that /s/ word accuracy was equivalent for the Madrid and Seville talkers. Accuracy rates for /s/ items by Variety appear relatively stable across experiment blocks, with small increases in Block 2 for Madrid and Seville talkers. Table 24 depicts average lexical acceptability by Block, in which it is apparent that the difficulty in recognizing Afueras /s/ words does not seem to lessen as the experiment progresses.

Table 24 Average Accuracy Rates for /s/ Words by Variety and Experiment Block

	BLOCK 1	**BLOCK 2**
AFUERAS	80.3 %	80.7 %
MADRID	88.1 %	91.5 %
SEVILLE	90.7 %	92.3 %

5.1.1 Durational Analysis of Stimuli

One final analysis that was conducted to prepare for analysis of immediate and long-term processing involved investigating word duration differences. Given that the stimuli were produced in a naturalistic manner and no manipulation of word durations was conducted, I anticipated that there would be variation in the speed at which talkers produced the items, which could then become confounded with RT results. To test this, I ran a mixed-effects model in R with lme4 (v 1.1-31, Bates, Mächler, Bolker, & Walker, 2015) and lmerTest packages (v 3.1-3, Kuznetsova, Brockhoff,

& Christensen, 2017) for word durations of real controls and /s/ words. I log transformed duration, the dependent variable, and included Variety, Talker Gender, and Condition as independent variables and random intercepts by Target. A three-way interaction among the independent variables was also tested.

The results of the model revealed significant effects and interactions among all predictors, indicating that duration was highly variable according not only to the talker that produced the item, but also to the condition and gender. Durations for /s/ words were longer than for real controls, which could be a result of longer duration of the initial segments (i.e., fricatives) when compared to a combination of other word-initial segments (i.e., stops, fricatives, nasals, and vowels). Despite the variability in the initial phoneme of the control words, they were all short, disyllabic nouns ending in vowels. Regarding Talker Gender, women produced the stimuli significantly more slowly than men for both real controls and /s/ words, though there was a greater difference for /s/ words (i.e., an average difference of 32 ms for controls compared to 49 ms for /s/ words). As a result of the differences in word durations across contexts, I calculated reaction times for the processing data according to the offset of the word, ensuring that any results are not confounded by duration effects.

5.2 Immediate Processing Analysis

5.2.1 Predictions

Based on results by Sumner and Samuel (2009) for experienced listeners of non-rhotic NYC English, I expected to observe evidence of recognition equivalence for the three varieties, manifesting as equal rates of recognition via accuracy and response times for /s/ items across all varieties in Block 1. This prediction would match Sumner and Samuel (2009)'s explanation of functional equivalence for variants in immediate processing as a result of listener experience, which has already been established in the current study through high accuracy in Verbal Guise dialect classification.

Additionally, similar to immediate processing results from an earlier study that I conducted in Seville in 2019, I expected that word-initial /s/ items would require more time to recognize than control words that did not contain /s/. This prediction was based on the salience of /s/ as a variant in Peninsular Spanish, and the many variants which can be associated with the phoneme in numerous contexts (e.g., variable place of articulation, coda /s/ deletion, *seseo*/*ceceo*/*distinción* processes, gradient voicing, etc.). In the prior study, I had hypothesized that /s/ words were accompanied by a slight processing delay as a result of the variability associated with the variant and anticipated to replicate the finding in the current study, despite slight differences in task type (i.e., semantic relatedness vs. form priming).

5.2.2 Immediate Processing Models

The purpose of the first analysis of the task was to observe immediate processing, meaning how quickly participants made a lexical decision in Block 1 trials. All reaction times were calculated from the stimulus offset and log transformed, and only accurate responses to real-word trials were included; 9.7% of real-word trials were excluded due to inaccurate responses. Any reaction time trials that were above or below three standard deviations of each participant's mean were removed (N = 147; 1.1% of correct real-word data).

The analysis of Block 1 reaction times implemented contrast coding for Talker Gender and Helmert coding for Variety, such that the model intercept corresponds to the average response time for all three Varieties combined. The contrasts for Variety correspond to 1) the average RT for Afueras talkers compared to the average of Madrid talkers, and 2) the average RT for Afueras and Madrid talkers combined compared to the average of Seville talkers. This coding allows for a comparison of RTs between the two non-native varieties, as well as the native to the non-native varieties. A linear mixed-effects model was fit to the data, predicting log RTs for real control words as a function of Variety, Talker Gender, and their interaction. Random intercepts by Participant and Target Word were also included; this was the most complex random effects structure that converged. Statistical significance in this and all other models was assessed using the p values for individual predictors, with an alpha-level of .05. All main models are available in Appendix E.

Descriptive trends in RTs indicated that participants were fastest in recognizing real control words by Seville talkers ($\bar{x}$ =617.3 ms), followed by Afueras ($\bar{x}$ = 625.8 ms), and lastly Madrid ($\bar{x}$ = 634.7 ms). The model indicated a statistically significant effect by Variety (β = -0.02, p = 0.036), in which RTs for Seville controls were significantly faster than the average of Madrid and Afueras.

RTs were not significantly different between the non-native varieties ($p > 0.05$). Additionally, there was a significant interaction between Variety and Talker Gender, indicating that the gender effect is different for SW[10] when compared to the average of the other varieties ($\beta = -0.04$, $p = 0.015$). The plot in Figure 8 depicts this significant interaction, showing that RTs for SW tended to be faster than those observed for SM.

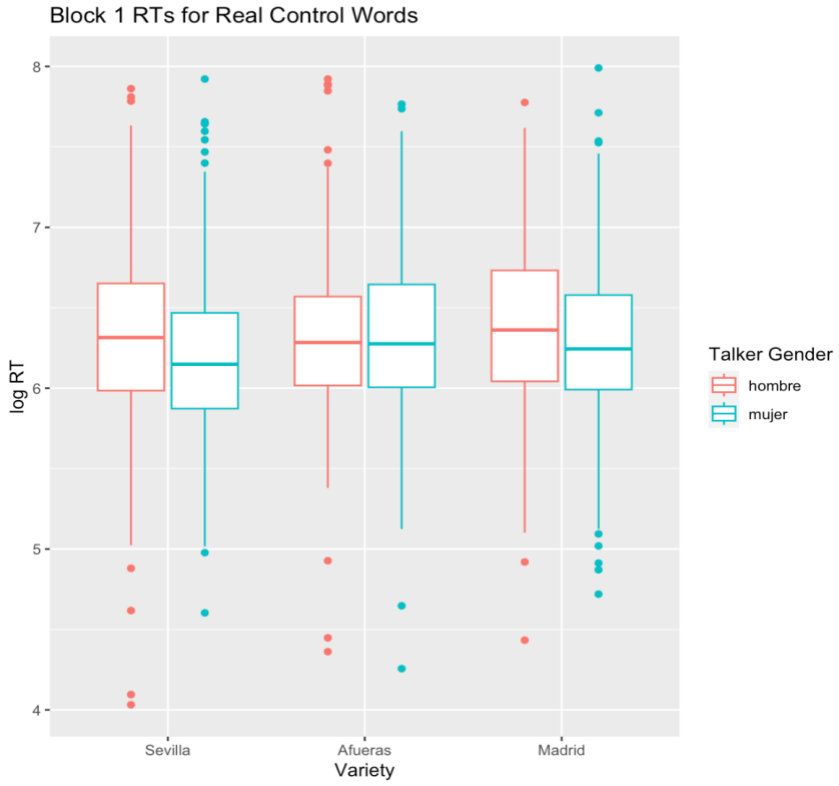


Figure 8 Word Recognition Times for Block 1 Real Control Words

Average response time data indicated that participants were fastest to recognize real control words produced by SW ($\bar{x} = 570.81$ ms) when compared to all other varieties. I conducted a follow-

[10] The coding scheme implemented for talkers in the Verbal Guise will be upheld in this chapter, with the key as follows: letter 1 = A (Afueras), M (Madrid), S (Seville); letter 2 M (Man), W (Woman); e.g., SW = Seville Woman.

up regression in which SW was set as the reference level, thus comparing average response times for each talker to SW. The model demonstrated a significant effect for four talkers: SM, AW, MW, and MM (ps ≤ 0.019), with participants being significantly faster in recognizing controls produced by SW. Only AM was not recognized significantly slower than SW. These results suggest that SW is likely driving the significant effect of Variety present in the control words; participants were slower to recognize controls produced by SM, and descriptively, average response times for his items were more similar to those of MM.

In the analysis of critical /s/ words, one response (0.065% of the data) was removed because it occurred before the target word offset. The /s/ model yielded a different pattern of results than the real control model; there were no significant differences by Variety or Talker Gender for /s/ items. Moreover, there were no significant interactions between Variety and Talker Gender. The similar processing of /s/ words by Variety and Talker Gender is represented in the boxplot in Figure 9.

.

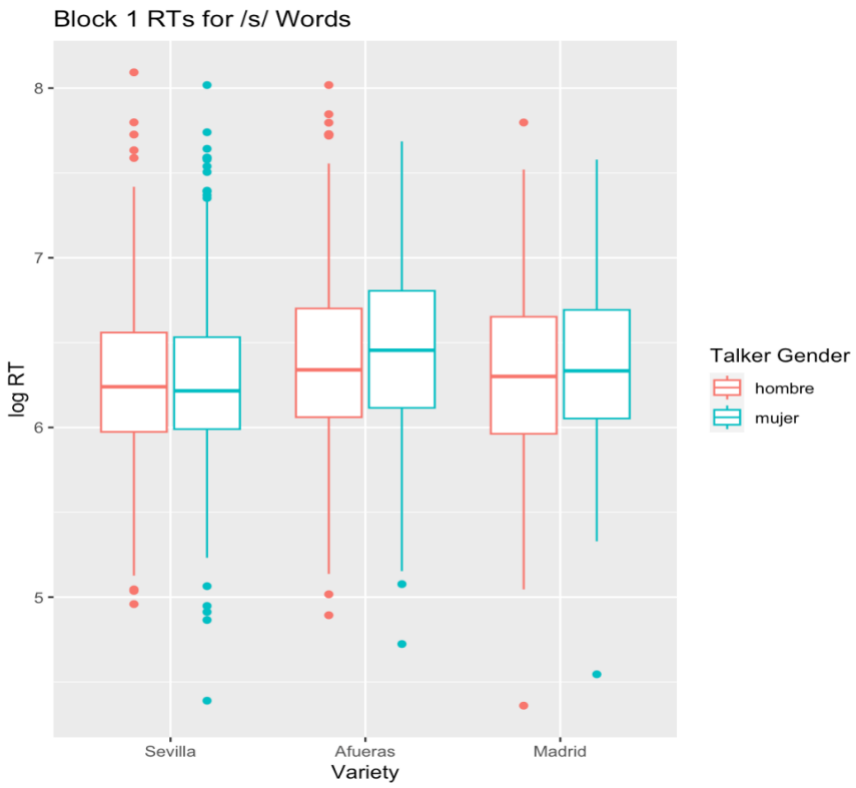


Figure 9 Word Recognition for Block 1 /s/ Words

To directly compare response times for /s/ tokens and real word control items, I then fit a model that included both sets of items, with control words as reference level (log(RT) ~ Variety * Condition + (1|Participant) + (1|Target)). The results of the model do not show a significant main effect of Condition, meaning that participant RTs for /s/ words are not different than real controls. Similar to the real control model, there was a significant difference by Variety (β = -0.02, p = 0.034), in which Seville items were recognized faster than the average of Madrid and Afueras words. There were also two significant interactions, one occurring between Variety and Talker Gender (i.e., SW; β = -0.04, p = 0.014). The effect of SW persists in the combined model. Similar to the Block 1 real control analysis, I ran another model (ref = SW) to see if participants responded to SW and SM, given the significant effect of Variety and significant interaction. With both /s/

words and real controls in the model, SM is not processed differently than SW ($p > 0.05$). Thus, the Variety effect in this model is not just a result of fast RTs for SW.

The combined model also showed another significant interaction between Variety and Condition ($\beta = -0.07$, $p < 0.001$). More specifically, the effect of /s/ words was significantly different for the two non-native varieties (i.e., Madrid and Afueras). The /s/ words produced by Afueras talkers yielded a larger processing cost (i.e., slower RTs) as compared to Madrid talkers. However, when averaged together, the size of the /s/ effect for the two non-native varieties was not significantly different from that of Seville. The significant effect of Variety and the two interactions are depicted in Figure 10.

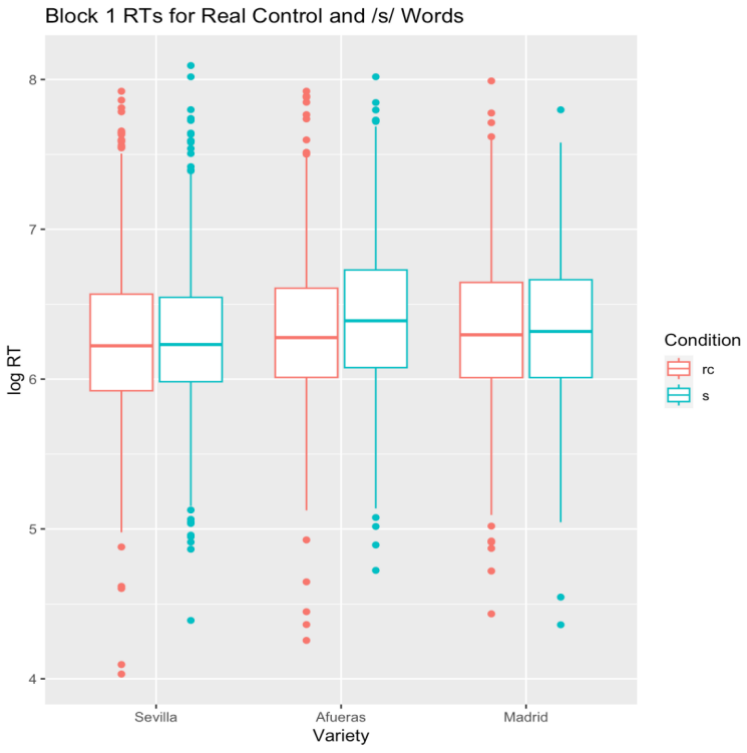


Figure 10 Block 1 Recognition of Real Control and /s/ Words

5.2.3 Immediate Processing Summary

To summarize the key results for immediate processing, /s/ words were generally recognized on par with real control words, evidenced by a lack of a significant effect of Condition in the combined model. When comparing /s/ items across varieties, all three variants were associated with equally efficient word recognition speed in the /s/ model. There were no statistically significant differences in reaction times by Variety or Talker Gender, with no significant interactions. Despite equivalent processing speeds, participants experienced more difficulty in lexical activation for the Afueras variant in the combined model, shown by lower accuracy scores when compared to Madrid and Seville /s/ words. There was also a significant interaction for Afueras and Madrid /s/ words, with the Afueras variant being more costly to recognize when compared to controls.

Regarding the independent variables of Variety and Talker Gender, a processing benefit for the native variety of the listeners was observed in the combined model, with Seville words being recognized significantly faster. Talker-specific effects were observed for SW in both the real control and combined models. Differential patterning was observed for SW when compared to the average of Madrid and Afueras response times, with words produced by SW recognized the fastest. In the real control model, participants are significantly faster to recognize items produced by SW when compared to SM, AW, MW, and MM.

5.3 Long-Term Processing

The following section will focus on Block 2 reaction times, allowing for an observation of long-term processing and priming effects. Due to the experiment length (i.e., N = 720 tokens; average completion time of 50 minutes), prime-target pairs occurred approximately 20 minutes apart.

5.3.1 Long-Term Processing Predictions

Regarding predictions for long-term processing, similar to Sumner and Samuel (2009) with General American and NYC English and Sumner and Kataoka (2013) with General American and British English, I expected to find similarities between Seville and Madrid, given that Seville is the native variety of the participants and Madrid is a familiar variety associated with a high level of overt prestige. I anticipated that both Seville and Madrid voices would exhibit priming effects, meaning that participants would be faster to recognize the repeated Block 2 items when they had previously been presented in either a Seville or Madrid voice in Block 1. I did not expect Afueras voices to facilitate priming in the long-term, as a memory advantage for a nonstandard, cross-dialectal variant has not yet been attested in the literature to my knowledge. However, if Afueras voices would also lead to priming effects, this could be attributed to the familiarity of the variety and dialectal similarity to that of the listeners (i.e., both are innovative Andalusian varieties).

Regarding the condition of the words, I anticipated that both real control words and /s/ words would be associated with repetition effects. The long-term form priming task was designed to be difficult (i.e., pseudowords closely resembled Spanish lexical items), which has previously been linked to stronger priming effects (McLennan et al., 2003). Moreover, Sumner and Samuel

(2005; 2009) observed significant repetition effects for repeated items in long-term processing. Taken together, I expected both conditions to be susceptible to repetition effects.

There are multiple questions to consider in the analysis of Block 2 items. First, did participants get faster in responding to real words the second time they heard them, and are there any differences according to either the prime or target variety? Additionally, are there differences in Block 2 RTs between repeated control words and new control items? Moving to /s/ items, did participants experience priming for the Block 2 /s/ words, and did these effects vary according to either the prime or target variety? Lastly, are there differences in Block 2 between real control and /s/ words?

5.3.2 Long-Term Processing Models

To begin, I provide average RTs for real control words by Condition and Variety in Table 25.

Table 25 Average Reaction Times for Real Control Words by Block, Variety, and Condition

	Block 1	Block 2 (repeated)	Block 2 (new)
Afueras	625.79 ms	654.06 ms	719.29 ms
Madrid	634.69 ms	684.69 ms	723.81 ms
Seville	617.25 ms	618.05 ms	667.31 ms

Surprisingly, the trends in Table 25 do not appear to reflect a repetition effect for Block 2 real control words, as average RTs are actually slower for each variety the second time participants heard the items. However, upon comparing average response times to new items in Block 2 to real controls in Block 1, it seems that participants slowed down as the experiment progressed. If experiment fatigue were not a factor, it would be expected that response times to control items in

Block 1 and new controls in Block 2 would be similar. To test whether experiment fatigue yielded significant differences in response times, I conducted a mixed-effects model comparing real controls in Block 1 with new real controls in Block 2, with contrast coding for Block. An interaction between Variety and Block was included, along with participant and target word as random effects. The model indicated a marginally significant effect for Block ($p = 0.099$). There was an overall effect for Variety, with participants faster to recognize items produced by Seville capital talkers when compared to the average of Afueras and Madrid ($p < 0.001$). This result indicates that while participants appeared to slow down over the course of the experiment, which is understandable given the length, the effect of fatigue was not a significant factor in influencing response times.

It should be noted that Table 25 does not take into consideration the nine possible prime-target pairings by Variety and instead presents a collapsed summary of the data. To observe whether priming effects occurred for Block 2 repeated real control items and account for the prime-target pairs, I ran a mixed-effects regression that examined Block 2 reaction times, measured from the word offset. The model included an interaction between prime and target varieties, along with random intercepts by participant and target item. Results from the model indicated no statistically significant effects of prime or target variety on Block 2 RTs, and there were also no significant interactions.

As findings from the immediate processing of real control words revealed a significant interaction between Variety and Talker Gender, I fit another model that included this interaction for Block 2. The same random effects were maintained. This model yielded a significant interaction between Variety and Talker Gender ($\beta = -0.05$, $p = 0.0299$), more specifically the same

pattern that was observed in Block 1 with faster processing of SW. Figure 11 depicts the significant interaction for SW on the processing of Block 2 real control words.

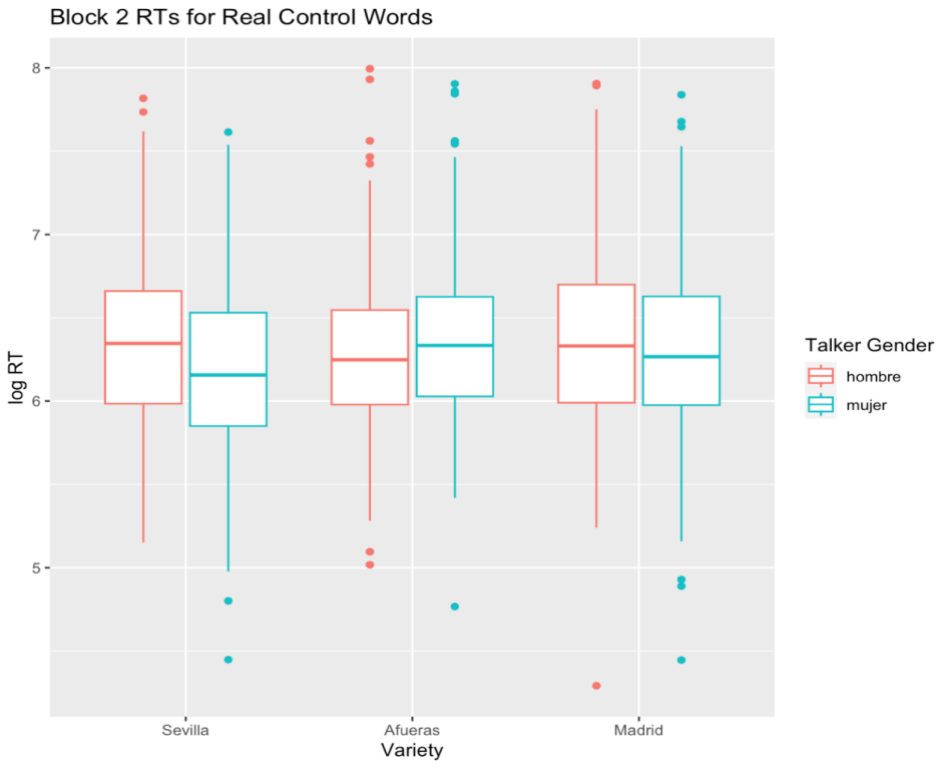


Figure 11 Processing of Real Control Words in Block 2

In addition to testing the effect of repeated real control words in Block 2, I also wanted to determine whether there was a difference between Block 2 repeated and new controls. This measure essentially serves as a gauge for the repetition effect of the control words, as it would be expected that participants respond more quickly to items they have already heard before. As demonstrated in Table 25, there appears to be a repetition benefit in long-term processing, as participants responded more quickly on average to the repeated items when compared to new words for all varieties. To test whether this pattern yielded significant results, I ran a mixed-effects regression for log RTs as a function of Repetition and Variety, with contrast coding for Repetition

and an interaction between the two variables. Target and participant were included as random intercepts. The model did not demonstrate a robust repetition effect, as there was no significant effect on reaction time according to whether a word was repeated or new. However, there was a significant effect for Variety, with Seville words being recognized more quickly than those produced by Afueras and Madrid talkers ($p < 0.001$).

After establishing the patterns for Block 2 processing of real control words, I then turned to /s/ items. The analysis follows a similar pattern, measuring words from the offset and excluding any tokens in which the participant made a lexical decision before the end of the word ($N = 2/1531$ trials; 0.13% of data). I first wanted to observe whether Block 2 reaction times would be faster as compared to Block 1, as all critical /s/ items were repeated. Table 26 depicts average RTs from Blocks 1 and 2 for critical /s/ words only.

Table 26 Block 2 Averages for /s/ Words by Variety

	Block 1	**Block 2**
Afueras	713.16 ms	711.01 ms
Madrid	632.87 ms	607.98 ms
Sevilla	608.92 ms	593.43 ms

Descriptive patterns indicate faster Block 2 average reaction times for both Madrid (-25 ms) and Seville (-15 ms) talkers, but nearly equivalent times for Afueras talkers (-2 ms). To observe any effects of prime-target dialect pairs on long-term processing, I ran a mixed-effects regression that considered prime and target varieties, as well as an interaction between them, on the Block 2 reaction times. Participant and target were included as random intercepts. The results of the model indicated no significant effects of prime variety, meaning that the variety that participants heard in Block 1 did not substantially influence processing. However, target variety

was found to be statistically significant. As Variety was Helmert coded, Madrid and Afueras were found to statistically differ from each other (β = -0.08, $p < 0.001$), with participants being faster to recognize words produced by Madrid talkers. Additionally, the average of Madrid and Afueras response times was found to be statistically different from Seville (β = -0.04, $p < 0.001$), with faster reaction times associated with Seville talkers. There were no significant interactions between prime and target varieties. The significant effect of target variety is depicted in Figure 12.

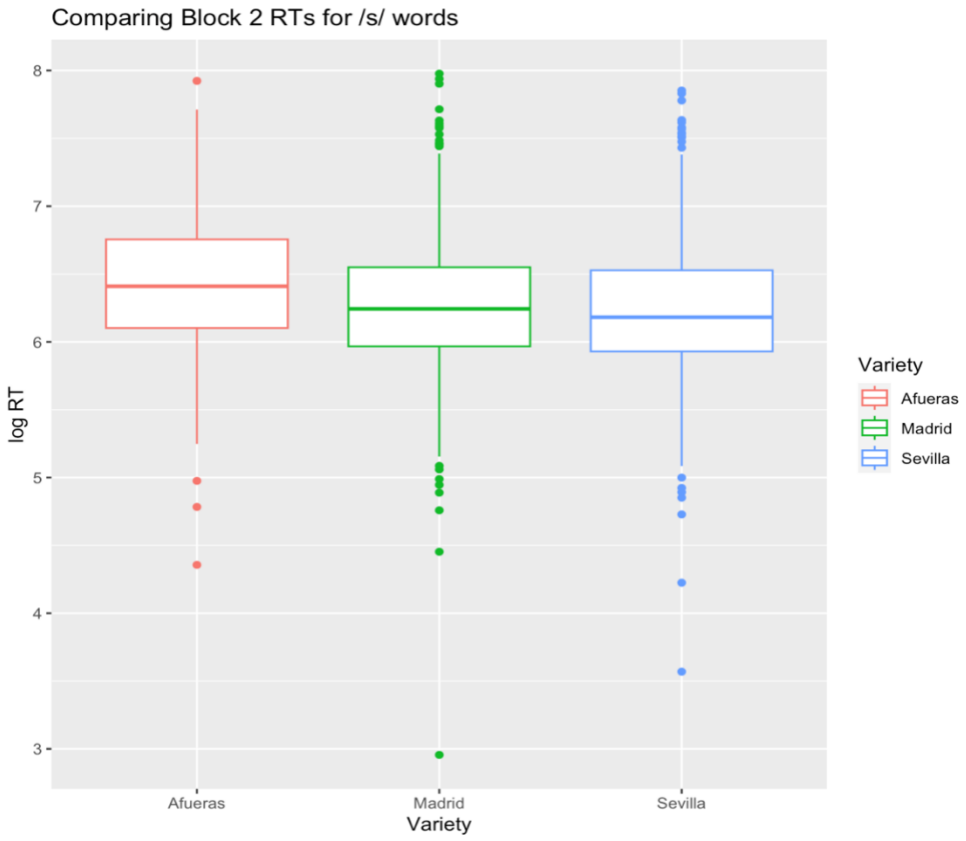


Figure 12 Block 2 Recognition of /s/ Words by Target Variety

As observed in Figure 12, Afueras response times are slower than both Madrid and Seville. However, while Helmert coding helps to differentiate the two non-native varieties from each other, as well as compare the average of both of them to Seville, it does not reveal whether RTs for Seville and Madrid voices are significantly different from each other. To explore this question, I

ran another mixed-effects model using treatment coding for Variety that included Madrid as the reference level. This model did not result in any significant differences in response times for /s/ items produced by Seville versus Madrid talkers ($p > 0.05$). The comparison between Madrid and Afueras voices was again significant in the second model, with Afueras voices being recognized more slowly than Madrid voices ($p = 0.0042$). Thus, for /s/ items, equivalent processing was observed for both Seville and Madrid, but not for Afueras. Additionally, only the target variety was found to be a significant predictor of response times, and there were no interactions between prime and target varieties.

Lastly, to compare the relationships between new and repeated Block 2 real controls as well as /s/ items, I conducted a combined model that included a three-way interaction between Repetition, Condition, and Variety (log(Block2_rt) ~ Repetition * Condition * Variety + (1|Participant) + (1|Target Word)). Repetition (old vs. new items) was contrast coded and Variety was Helmert coded. Real control words were set as the reference level for Condition. Talker Gender was initially included as a predictor, but as it was not found to be significant, I removed it from the model. Participant and target word were set as random intercepts. Similar to the analysis of real and new controls in Block 2, repetition was not found to be a significant predictor of response time, and there were no interactions. Similar to the combined model in Block 1 for real controls and /s/ items, there was no significant effect of Condition in Block 2. There was a significant effect of Variety, namely that participants recognized Seville targets faster than the average of the Madrid and Afueras targets ($\beta = -0.03$, $p < 0.001$). Finally, similar to the combined Block 1 analysis, there was a significant interaction between Condition (/s/) and Variety (non-native varieties) ($\beta = -0.09$, $p < 0.001$), with increased processing difficulty for the Afueras variant

compared to the Madrid variant. This relationship is shown in Figure 13. There were no other significant interactions.

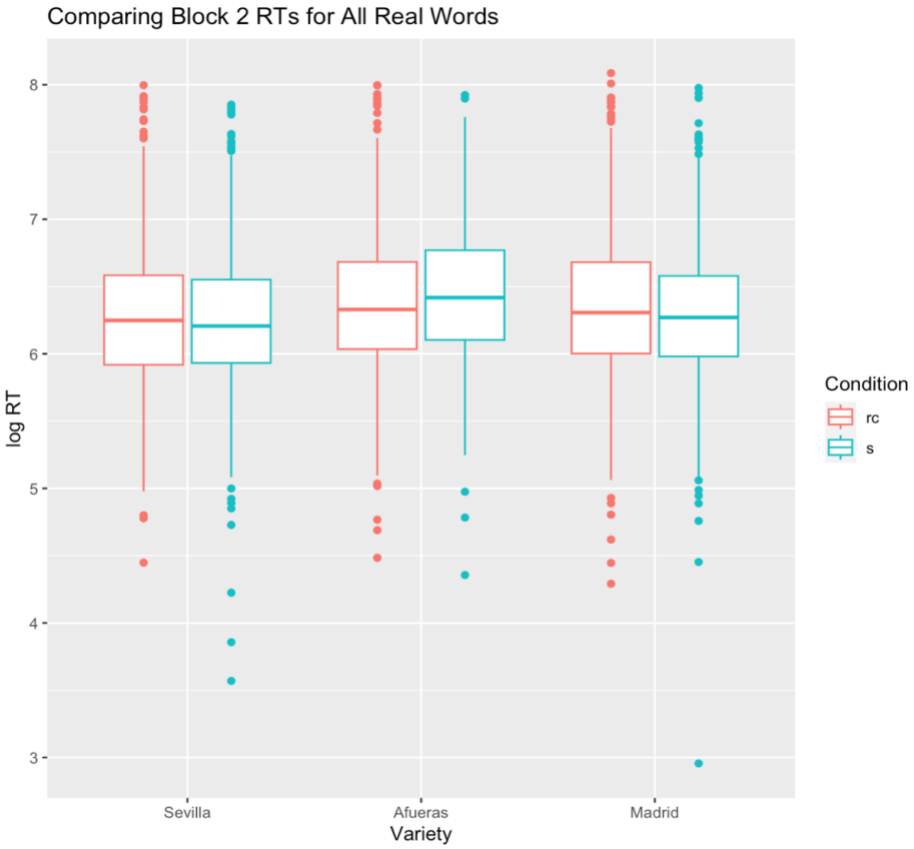


Figure 13 Word Recognition of Real Words in Block 2

To conclude, Table 27 presents a summary of the significant findings from both Blocks 1 and 2.

Table 27 A Summary of Immediate and Long-Term Processing Results

	Block 1	**20 – 30 minute lag**	**Block 2**
Real controls model	- Significant effect of Variety; Seville recognized faster than combination of Madrid and Afueras (*)		-Interaction between Variety & Talker Gender: effect of SW, recognized fastest (*)

	-Interaction between Variety & Talker Gender: effect of SW, recognized fastest (*)		
/s/ word model	- No significant effects or interactions		- Significant effects for target Variety: Madrid targets significantly faster than Afueras (***) - Seville targets significantly faster than combination of Madrid and Afueras (***) - Seville and Madrid targets do not significantly differ
Combined model	- No significant differences between /s/ words and RCs -Significant effect of Variety; Seville recognized faster than combination of Madrid and Afueras (*) - /s/ word effect is significantly more pronounced for Afueras than Madrid (***) - Interaction between Variety & Talker Gender; effect of SW, recognized fastest (*)		- No significant differences between /s/ words and RCs - /s/ word effect is significantly more pronounced for Afueras than Madrid (***) - Significant effect of Variety; Seville targets recognized significantly faster than combination of Madrid and Afueras (***)

5.4 Word Recognition Task Synthesis

5.4.1 Key Framework for Interpretation

The results of the word recognition experiment can be interpreted with respect to an ongoing body of work that has enhanced the theoretical framework of speech perception of multiple variants (Sumner et al., 2014; Clopper, 2021). Some critical challenges have emerged as researchers work to explain how listeners are seemingly able to sufficiently handle vast amounts of surface variation in immediate processing (McLennan et al., 2003; Sumner & Samuel, 2005, 2009; Sumner & Kataoka, 2013), yet prioritize a canonical or idealized variant in the long-term (Clopper et al., 2016; Sumner & Samuel, 2005; 2009; Sumner & Kataoka, 2013).

For instance, regarding within-dialect variation, McLennan et al. (2003) explored flaps in American English (i.e., the use of [ɾ] instead of intervocalic [t] and [d], as in *but<u>t</u>er*). The authors found that both variants primed each other equally well, regardless of the fact that /t/ and /d/ variants are less frequent and more widely associated with careful speech. In another study, Sumner and Samuel (2005) explored word-final /t/ variation, finding that listeners demonstrated equal semantic priming for all three variants examined. There was no processing benefit for any of the three variants (i.e., the RTs for targets primed by [t] did not significantly differ from each other), despite highly variable frequency rates between the most frequent and canonical forms. This pattern extends to cross-dialectal research, with findings demonstrating equivalent immediate processing of multiple variants by experienced listeners (Sumner & Samuel, 2009), or when a variant is associated with overt prestige (Sumner & Kataoka, 2013).

While numerous studies have observed similar processing of multiple surface variants in immediate tasks, results on long-term processing reflect an asymmetry, whereby listeners have been shown to prioritize a canonical or idealized form in memory (Clopper et al., 2016; Sumner & Samuel, 2005; 2009; Sumner & Kataoka, 2013). In one study, while all three final /t/ variants in Sumner and Samuel (2005) resulted in immediate semantic priming effects, only the less frequent, canonical variant was stored in memory. In other words, the only variant associated with a strong priming effect via faster Block 2 reaction times was the Basic [t] – Basic [t] condition. Other work has replicated this result, with Sumner and Samuel (2009) showing memory for the standard form (i.e., r-ful productions by General American talkers, as compared to r-less variants by NYC talkers) for all listener groups, as well as memory for an infrequent variant that is associated with prestige (i.e., an r-less production by a British English speaker, Sumner & Kataoka, 2013). These results present challenges for purely frequency-based models, as it would be expected that listeners would have vastly less long-term representation for an infrequently encountered variety (i.e., British English for the General American listeners in Sumner & Kataoka).

Taken together, the discrepancy between the processing of variants in the short and long-term has resulted in the proposal of socially weighted encoding (i.e., *social weighting*), the process by which socially salient tokens are encoded more strongly than both typical and atypical, non-salient tokens (Sumner et al., 2014). A key construct from this proposal is recognition equivalence, which occurs when all participants recognize variants equally well in immediate processing tasks, regardless of frequency differences in variant production or perception to the listener. Another construct is memory inequality, in which words produced with infrequent, idealized forms are remembered better in the long-term. The authors propose that while multiple surface features can be recognized equally well in immediate processing, not all variants are encoded with the same

strength, which leads to memory inequality in the long term (Sumner et al., 2014). The constructs of recognition equivalence, memory inequality, and social weighting will be referenced in the interpretation of results of the word recognition task in the current study.

5.4.2 Immediate Processing Interpretation and Discussion

Research on the perception of within-dialect variation (e.g., flapping in American English, McLennan et al., 2003) and cross-dialectal variation (e.g., rhoticity in NYC English and General American English, Sumner and Samuel, 2009) can help interpret the immediate processing results. The Seville capital participants have a three-way familiarity with the varieties, making them experienced listeners. Moreover, while Seville capital talkers are predominantly associated with [s̪], not all productions are categorical, similar to findings on production of *seseo*/*ceceo*/*distinción* (Regan, 2017a; Santana Marrero, 2016-2017).

5.4.2.1 Evidence for Recognition Equivalence in /s/ Words

A critical construct for interpreting the immediate processing results for /s/ words is recognition equivalence, the similar treatment of multiple surface variants by listeners despite disparate levels of frequency or familiarity. Sumner et al. (2014) emphasized that while evidence of recognition equivalence had been previously attested, they did not claim that all contexts and variants would demonstrate recognition equivalence in immediate processing. Partial evidence of recognition equivalence of the three /s/ variants was observed in the current study, manifested via equivalent response times in immediate processing. In other words, no statistically significant differences were present by Variety or Talker Gender in word recognition data for the three variants in the /s/-word model.

While the three surface variants are all familiar to the Seville capital listeners, they are presumably not equivalent in frequency of exposure. A frequency-based account would privilege the Seville capital [s̪] variant in processing, and it would be expected that listeners respond significantly faster to words produced by Seville capital talkers if this were the case. Instead, Seville capital listeners show perceptual flexibility in processing speed when recognizing the three /s/ variants. Along with familiarity, both of the non-native variants are socially salient, but for different reasons: the Madrid [s̪] is an idealized variant represented in the media, while the Afueras [s̪$^{\theta}$] is salient due to stigmatization.

Another consideration in interpreting the processing of the /s/ variants pertains to lexical frequency. While all disyllabic critical /s/ items were frequent (i.e., greater than the average log frequency value in EsPal), there was a range of lexical frequencies represented in the word recognition task. Differential processing according to variable lexical frequency would be supported by exemplar models, and it would be expected that the /s/ words in the lowest frequency (e.g., *sarro* 'dental plaque') band be associated with lower accuracy scores than items in the highest frequency band (e.g., *sitio* 'place). This anticipation was generally correct; the average combined accuracy for the five lowest frequency words across varieties was lower than that of the five highest frequency words (77.8% correct to 94.0 % correct, respectively).

In many cases, accuracy trends were similar across varieties. For example, the word *sorna* 'sarcasm' in the lowest frequency band was associated with lower accuracy scores regardless of the variety in which it was produced (i.e., 50% correct Afueras, 53% correct Madrid, 50% correct Seville). Patterns were parallel for some items in higher frequency bands (e.g., *sangre* 'blood'; 92.9% correct Afueras, 100% correct Madrid, 97.6% correct Seville). This similar cross-dialectal processing of /s/ words according to variable lexical frequencies can provide partial evidence in

support of recognition equivalence, but this pattern was not upheld in all instances of high and low frequency words (see 5.4.2.2. for exceptions to these trends).

The findings for recognition equivalence of multiple surface variants in spoken word recognition are corroborated by a body of work on within-dialect variation (i.e., McLennan et al., 2003; Sumner & Samuel, 2005), and cross-dialectal speech perception (i.e., López Velarde & Simonet, 2019; Sumner & Samuel, 2009; Sumner & Kataoka, 2013). In semantic priming research, numerous studies have observed that multiple variants can prime each other equally well in immediate and short-term processing experiments. For instance, McLennan et al. (2003) found that there was no difference in word recognition of carefully and casually produced variants (i.e., use of [ɾ] instead of intervocalic [t] and [d]), and each variant primed the other equally well. In Sumner and Samuel (2005), there was no processing benefit for any of the three word-final [t] variants in semantic priming, regardless of the canonical or non-canonical status of the variant. These findings support the equal recognition of /s/ words in immediate processing in the current study, regardless of the variety in which they were produced.

Considering results on cross-dialectal word recognition, Sumner and Samuel (2009) found equal semantic priming of r-ful and r-less variants for experienced NYC listeners, despite whether they actually produced the r-less variant. However, priming effects were not observed for General American listeners who were unfamiliar with the r-less variant. The Seville capital participants in the current study can be compared to Sumner and Samuel (2009)'s Covert-NYC group, in that their Covert-NYC listeners were frequently exposed to both r-ful and r-less variants through living in the city, yet predominantly produced the rhotic variant. Having conversed with each participant during the interview task, the majority of their productions seemed to be dental. Results from the perception of /s/ words in immediate processing in the current study corroborate the cross-dialectal

recognition equivalence observed in Sumner and Samuel (2009), highlighting the perceptual flexibility of listeners who have exposure to multiple variants.

While listener experience contributed to immediate processing results in Sumner and Samuel (2009), findings from Sumner and Kataoka (2013) indicated that frequency of exposure was not the only factor in allowing variants to be recognized equally well across dialects in immediate processing. Results from a semantic priming task found that listeners who were unfamiliar with non-rhotic varieties demonstrated equal priming for General American (rhotic) and British English (non-rhotic) varieties, but not for NYC talkers (non-rhotic). To explain the disconnect between the lesser exposure to British English by participants and the equivalence in priming, Sumner and Kataoka (2013) hypothesized that the atypical, idealized nature of the British English variety was leading to recognition equivalence with the native variety of the listeners. While all listeners in the current study were familiar with the three presented varieties, findings from Sumner and Kataoka (2013) support the notion that a variety with a lower frequency of exposure can still be recognized quickly if it is associated with prestige or standardness, as would be the case for the Madrid variety to the Seville capital listeners.

5.4.2.2 Processing Benefits and Costs in Cross-Dialectal Word Recognition

While the aforementioned studies have found evidence of equivalent treatment of surface variants in immediate processing, other studies have observed significant processing benefits for certain varieties, especially ones that are associated with a standard or idealized form (Clopper et al., 2016; López Velarde & Simonet, 2019; Sumner & Samuel, 2005). The exploration of Midland and Northern varieties of U.S. English has yielded surprising results, as multiple experiments have demonstrated processing benefits for a non-native, yet standard, Midland variety (Clopper &

Bradlow, 2008; Clopper et al., 2016; Dossey et al., 2023). For example, Clopper et al. (2016) found significantly faster response times in a lexical decision task for the Midland variety, even for participants who grew up in the Northern region. One caveat is that as testing occurred in the Midland region, it is possible that participant anticipation for what they would be exposed to in the experiment contributed to the benefit for the Midland variety, as listener expectations can affect language processing (Hay & Drager, 2010). Dossey et al. (2023) replicated the processing benefit for the Midland variety via better accuracy and faster response times, despite a relative lack of participant experience with this variety. Taken together, it is not always the case that participants familiar with different surface variants experience recognition equivalence in immediate processing, as standardness and local relevance was hypothesized to lead to a processing benefit for the Midland variety (Clopper et al., 2016).

Also considering standard vs. non-standard varieties, recent results on immediate processing of *shesheo* in *norteño* Spanish from López Velarde and Simonet (2019) contribute novel non-English findings to the cross-dialectal word recognition literature. In a form-priming task, the authors observed equal priming for both the standard [t͡ʃ] and non-standard [ʃ] variants, meaning that the two variants equally facilitated recognition of one another. However, processing speed was not equivalent, as participants were faster overall to recognize targets that were produced with the standard variant. This result is especially curious, given that all of the listeners were from a location (i.e., Hermosillo, Mexico) in which the regional variant was frequent. The variants of interest in López Velarde and Simonet (2019) are similar to the /s/ variants in the current study in that they occur in word-initial position, and the non-standard variant in *norteño* Spanish is also a fricative.

While a benefit was found for the standard variant in López Velarde and Simonet (2019) via response times, equal accuracy rates were observed for both [t͡ʃ] and [ʃ] variants, meaning that these variants mapped equally well onto the stored lexical items for listeners in this experiment. Despite an uncertainty in the frequency distribution of the variants to the Hermosillo listeners, the fact that both variants were equally accepted speaks to the familiarity of *shesheo* to the participants. While Seville capital listeners in the current study were also familiar with all three /s/ variants, accuracy scores for /s/ items produced by Afueras talkers were significantly lower than those produced by Seville and Madrid talkers, indicating greater difficulty in mapping the [s̪$^{\theta}$] variant onto the target lexical items. It is intriguing that despite the lesser accuracy in recognizing some /s/ words, participants in the current study still demonstrated recognition equivalence in terms of reaction times for the three variants, while this pattern was reversed in López Velarde and Simonet (2019). Sumner and Samuel (2009) observed a similar patterning of error rates for the non-standard r-less variant. Not only did all participants have fewer lexical decision errors for items that were presented in the General American (i.e., r-ful) variety, but even for both NYC participant groups, hearing a GA prime facilitated easier recognition of a proceeding target variant.

Sumner and Samuel (2009) reasoned that there was a processing benefit for a less-frequent variant if it corresponded to the orthographic representation of the word. The connection to orthography could also explain why participants in López Velarde and Simonet (2019) were faster to recognize words produced with the standard [t͡ʃ] variant, as this is represented in writing while [ʃ] is not. The accuracy results in the current study also can be related to orthographic differences, as the [s̪$^{\theta}$] variant may also be associated with orthographic <z>, <ce>, or <ci> as a result of acoustic overlap with /θ/. The other two /s/ variants have a more direct correspondence with orthographic <s>. It is possible that the multifunctionality of [s̪$^{\theta}$], according to the linguistic

context, led to significantly lower accuracy scores for /s/ items produced by Afueras talkers. This can be related to Barnes' (2015) notion of cognitive salience, in which nonstandardness of the [es] in Asturian Spanish was less salient due to the association of -es in other contexts. What is intriguing is that the difficulty in lexical access for Afueras /s/ words was constant across both experiment blocks, with participants failing to recognize words containing the [s̪$^{\theta}$] variant in 20% of the trials. As Seville capital participants are familiar with the [s̪$^{\theta}$] variant, and multiple individuals explicitly referenced *ceceo* in the Verbal Guise comment section, the difficulty in word recognition warrants additional exploration.

The fact that Seville capital listeners experienced more processing difficulty for the Afueras variant can be hypothesized to result from competing activation of other phonological forms. For a Seville capital listener, [s̪] can be associated with either /s/ or /θ/, the latter in instances of *seseo*. Similarly, the [s̪$^{\theta}$] for Afueras talkers can be linked to either /s/ or /θ/, as the variant is acoustically similar to the /θ/ employed in *distinción*. Results from Gylfadóttir (2018) confirmed accurate perception of /s/ and /θ/ in a discrimination task by Seville capital participants, supporting participants' ability to differentiate between the two phonemes when presented in isolation. However, a lexical decision task presents additional challenges, especially when the variant occurs in word-initial position.

A connection can be made to literature on phonological lexical competition, as research shows slower response times for lexical items with minimal pairs (e.g., Andruski et al., 1994; Clopper & Walker, 2017). In this literature, lexical items that share similarity with the incoming speech signal are activated, and compete with each other until one candidate is chosen in word recognition (McQueen & Cutler, 2010). While the ambiguity of the /s/ words does not result in minimal pairs in the context of the current study (i.e., none of the word-initial /s/ items have lexical

counterparts with /θ/), the activation of the initial forms may be resulting in a prelexical competition, causing slower RTs for the Afueras /s/.

The hypothesis of competing activation of /s/ and /θ/ can also explain why there is a greater cost in recognizing the Afueras variant via accuracy rates. It is possible that some participants originally mapped the [s̪$^{\theta}$] variant to the wrong phonological form, instead associating it with /θ/. If this were the case, the lexical item would be deemed a nonword, as none of the word-initial /s/ items were minimal pairs with /θ/ (e.g., an erroneous phonological mapping for *sangre* 'blood' would lead to an interpretation of *zangre*, a nonexistent word). Exploring accuracy rates according to the lexical frequency of the /s/ words supports this prediction, as some items demonstrated accuracy differences across varieties, even for high frequency target words (e.g., *sala* 'living room'). For instance, when *sala* was produced by Seville and Madrid talkers, accuracy rates were at 100% for both varieties. However, when produced by Afueras talkers, accuracy rates dropped to 50%. As participants could not change their answers after pressing the 's' or 'n' keys during the task, any initial association of [s̪$^{\theta}$] to the wrong form could lead to an inaccurate lexical identification, though some individuals may have realized the error after making the decision.

5.4.2.3 Treatment of Real Control Words in Immediate Processing

A peculiarity of the results of the current study is the behavior of the real control items in immediate processing. The selected target words were maximally dialect-neutral, avoiding instances of salient phonetic variation in Peninsular Spanish. The reasoning behind this decision was that participants would not exhibit much variation in processing speed for a frequent, unmarked word such as *pila* 'battery' regardless of the variety or gender of the talker who produced the item. In this way, a baseline could be established to compare salient /s/ items to neutral control

words, both in immediate and long-term processing. While I had predicted equivalent processing speed for real control items across variety and gender, the reality is quite different. Despite no obviously variety-specific pronunciations, a consistent significant interaction between Variety and Talker Gender in all models that include real control words showed differential patterning for SW. Words produced by her were responded to quickly, sometimes resulting in significant differences between other talkers (i.e., SM, MM, MW, AW in the Block 1 real control model).

Additionally, there was a significant effect for Variety (i.e., Seville) in both the real control model and combined model. In some contexts, participants show a benefit to recognizing words produced in their native variety. Significantly faster RTs for SW compared to SM in the real control model appear to be largely contributing to the effect of Variety, but SW and SM were processed similarly in the combined model, when both control and /s/ words were featured.

The results for the immediate processing of real control words were unexpected and had certain implications for the priming analyses in Block 2. For example, in Sumner and Samuel (2009), response times to real control words were used in priming calculations, in which responses to target items containing the variant were subtracted from repeated real control items. The authors did not report any significant differences in response times to real control words according to talker characteristics, despite the inclusion of two varieties, as well as men's and women's voices appearing in the study. As real control words had both variety- and talker-specific effects, they failed to serve as a reliable baseline for the priming analysis given these differences.

A critical difference between the current study and that of Sumner and Samuel (2009) is that in the current study, participants completed the Verbal Guise immediately before beginning the word recognition task. In addition to collecting critical dialect classification and evaluation data on the six talkers, the Verbal Guise also served as a familiarization phase for the voices. Early

research exploring variable speaker characteristics in lexical processing showed recognition benefits for familiar voices when compared to unfamiliar talkers (e.g., Bradlow et al., 1999; Goldinger, 1996, Nygaard & Pisoni, 1998). The fact that listeners in the current study not only became familiar with the voices of each talker in the Verbal Guise, but were asked to match the talkers to their respective locations and rate the voices according to status and solidarity categories possibly affected the treatment of real control words. In other words, the activation of language attitudes for each talker and variety likely carried over into the word recognition task, as all of the talkers are repeated from the Verbal Guise. In Sumner and Samuel (2009), their long-term form priming experiment was not preceded by an explicit dialect classification or evaluation task.

To summarize the immediate processing results, the current study observed partial evidence of recognition equivalence for the three /s/ variants via no significant differences when comparing RTs for /s/ words only by Variety or Talker Gender. However, while RTs were equivalent in the /s/ model, evidence of a processing cost for the Afueras variant was manifested by significantly lower accuracy scores for Afueras talkers and a greater slowing effect for the /s/ words for Afueras talkers as compared to Madrid talkers. Another finding is the lack of significance of Condition in the combined model, with no differences in RTs for /s/ words compared to controls. This result supports the perceptual flexibility of the Seville capital listeners in processing familiar surface variants. Finally, an effects of Talker (i.e., SW) and Variety (i.e., Seville) were observed in the processing of real control words, leading to the hypothesis that language attitudes are very much activated and present in the task, even for lexical items free of salient variation.

5.5 Long-Term Processing Interpretation and Discussion

The long-term processing task implemented in the current study allows for an examination of how the /s/ variants were encoded and represented in memory, given the long gap between presentation of prime and target pairs. As participants demonstrated equivalent processing speed for all three variants in immediate processing, any differences in response times for Block 2 items can potentially be related to long-term perceptual processes.

5.5.1 Memory Inequality in Long-Term Processing

Results on within-dialect variation (e.g., Sumner & Samuel, 2005) and cross-dialectal variation (e.g., Sumner & Samuel, 2009) have both observed differences in treatment of surface variation in long-term processing paradigms. While listeners accommodate variation equally well in immediate processing, patterns have indicated preferential treatment for a canonical variant when a delay is introduced between prime and target presentation (Sumner & Samuel, 2005; 2009; Sumner & Kataoka, 2013). This paradox has led Sumner et al. (2014) to propose that infrequent, but idealized (i.e., standard) variants are better remembered than non-idealized forms as a result of salience. The authors argue that additional attention allocated to the idealized variants allow them to be more strongly encoded, thus resulting in more robust representation in memory. Sumner et al. (2014) posit that an instance of an infrequent, idealized variant that was more strongly encoded would lead to a representation that was on par with an entry from a token that was represented by more frequent, less salient tokens.

Examples of memory inequality have been observed across multiple variants and contexts (e.g., Sumner & Kataoka, 2013; Sumner & Samuel, 2005, 2009). Results from these studies helped

shape the predictions established for long-term processing in the current study and will be discussed in tandem with significant findings from the three Spanish variants. Given the attested prioritization of standard variants in memory, I expected to find evidence of memory inequality for Seville and Madrid varieties in long-term processing of /s/ words. The Seville variant is most frequent to the listeners, and according to Sumner et al. (2014)'s hypothesis for representation, it would be associated with a high number of default tokens in memory. However, if there is social weighting for the Madrid /s/, it would be expected that participants also have a robust representation of the variant due to its association with overt prestige. Memory inequality for the Madrid variant would be manifested by equivalent long-term processing of Madrid /s/ and the Seville capital /s/. Similar to findings from Sumner and Kataoka (2013), I did not expect Afueras voices to facilitate long-term priming for /s/ items, as robust memory encoding for a non-native, non-standard variety has not been attested.

The findings from the Block 2 /s/ model can be related to the concept of memory inequality. While /s/ variants in Block 1 were associated with equivalent processing times, there were significant differences by Variety in Block 2: participants were significantly faster in recognizing Block 2 /s/ words presented by Seville capital and Madrid talkers, as compared to Afueras talkers. The fact that this result was not observed in Block 1 indicates that three varieties were responded to differently as the experiment progressed, which is perhaps a result of differential encoding of the variants. One possibility is that Seville and Madrid primes were encoded after presentation in Block 1, but the lack of evidence of identity priming in the Block 2 model complicates the interpretation. These options will be considered more extensively in the Discussion chapter.

5.5.2 Differential Processing Strategies for Real Controls and /s/ Words

Given previously attested trends for treatment of control words in priming paradigms, I had originally predicted that both real control words and /s/ words would be associated with significant repetition effects. However, repeated controls were not processed significantly more quickly than new real controls in the current study. These results differ from findings in Sumner and Samuel (2005) for control items, where a significant repetition effect was established between new and repeated words. Sumner and Samuel (2009) replicated this effect in their long-term form priming task, with repeated items being recognized faster than new items in Block 2. Reaction times for Block 2 repeated real controls in the current study did not reveal any significant effects for target variety, and the same interaction for Variety and Talker Gender observed in Block 1 was upheld, with an effect for SW compared to the combination of Madrid and Afueras talkers.

An explanation of the differential treatment of real controls and /s/ words in the current study relates to processing strategies, specifically the difference between top-down and bottom-up language processing (e.g., McQueen & Cutler, 2010). In top-down processing, listeners start with information from the linguistic context, knowledge of the speaker, or knowledge of the world before working their way down to surface-level features such as phonetic information. Bottom-up processing is the reverse phenomenon, where information such as sounds is the primary driver of processing effects, with morphosyntactic and semantic information exerting their influence later. I hypothesize that processing of the /s/ words was primarily affected by bottom-up processing factors, as the salient variation occurs in word-initial position. As the real control words do not contain salient phonetic information, it is possible that top-down processing had a greater effect,

as listeners accessed information about the talkers, namely regional associations and their attitudes towards them.

Recent findings from Dossey et al. (2023) indicate that both bottom-up processes (i.e., acoustic-phonetic variability) and top-down constraints (i.e., semantic predictability) can heavily influence processing, even more than lexical characteristics such as word frequency. The treatment of real controls and /s/ words in the present experiment is also congruent with Sumner et al.'s (2014) dual route model, in which social and linguistic information are processed in tandem. Given that listeners were familiarized with all six voices before beginning the word recognition task, the strong effect of talker for real controls (i.e., faster word recognition for SW) suggests a benefit for the voice most positively linked to the social identity of the participants. The fact that SM patterned differently from SW, was significantly more difficult to classify in the Verbal Guise, and was perceived to have a more neutral/forced formality, can support this hypothesis. Despite differential patterning for real control words, participants responded similarly to the two Seville talkers in the /s/ word model and combined models. As both talkers consistently realized the dental variant, it is possible that this similarity in surface variation overrides talker differences.

While the level of control present within the creation of the long-term form priming task should not be understated (e.g., the careful matching of frequency across conditions in the stimulus lists), emphasizing the contrast between talkers and listeners in the current task and those featured in many of the studies cited in this chapter can provide additional insight into differences observed in the current study. More specifically, the experiment was administered in a naturalistic setting, with testing taking place in a friend's apartment in the middle of Seville capital instead of a sound attenuated booth in a lab. Similarly, the talkers featured in the current study were not trained linguists producing stimuli; they were individuals spanning different ages, genders, and dialect

regions. Participants in the current study were not undergraduate students completing tasks for course credit, but rather individuals from Seville capital with a variety of age ranges and education levels who electively chose to complete the experimental tasks. This unique combination of listener, talker, and testing characteristics offers perhaps a more realistic context, from the perspective of a person naturally encountering variation in their daily life, for an exploration of cross-dialectal language processing. It is also possibly because of this broader combination of talker voices that no repetition priming effects were observed for repeated items, similar to Clapp et al. (2023).

5.6 Synthesis of Results and Directions for General Discussion

In summary, the results observed in the word recognition task are complementary to previous findings in the literature on cross-dialect processing and also provide avenues of discussion regarding the long-term processing of Spanish variants. Partial evidence for a three-way recognition equivalence was found for /s/ words in immediate processing via equivalent RTs across varieties, but a processing cost was also observed for Afueras via lower lexical acceptability scores and a slowing effect for the variant over the Madrid /s/ in the combined model. Thus, listeners sometimes struggled to activate the correct lexical representations for the Afueras /s/ words, and when the correct lexical items were identified, they were recognized with slower processing times relative to the Afueras control words. Results also indicate that recognizing /s/ words compared to control words in immediate processing is not more costly overall, with no significant effects for /s/ words when all three varieties are averaged together. Talker-specific effects also appear in examining recognition of real control words, with fast RTs linked to SW.

Variety effects were also observed in the combined model. Finally, no priming effects were observed for real control words, as 'repetition' was not a significant predictor of Block 2 response times.

While this chapter has related findings from the long-term form priming results to relevant research on the processing of multiple variants, additional interpretation is necessary before gaining a comprehensive view of the findings of the current study. As previously alluded to, I hypothesize that the task sequence of the Verbal Guise, followed by the Long-Term Form Priming experiment, have social implications that are embedded in the results of the processing data, namely occurring with real control words. The Discussion chapter provides a detailed account of the relationship between the two tasks and explores the mechanisms responsible for recognition equivalence of /s/ in immediate processing, memory for Seville and Madrid in long-term processing, a lack of priming effects attributed to real control words, and talker-specific effects in the control items.

Moreover, while it is not the purpose of this dissertation to make a strong argument in favor of one model of word recognition over another, I situate the findings of the current study within the literature on these models, attempting to better understand the complex workings of cross-dialect word recognition, the interactions between top-down vs. bottom-up processing strategies, the possibility of differential encoding for different dialect variants, and the intersection of social and linguistic information in language processing. Finally, I add to the conversation on the relationship between dialect classification, evaluation, and word recognition, as these areas often do not coexist within the realm of one experiment.

6.0 Discussion

This chapter outlines the research questions posed in the dissertation, the corresponding findings observed in each experimental task, and a synthesis of key results of both tasks taken together. Relevant literature and participant commentary from the sociolinguistic interviews is referenced in support of the experimental overview.

6.1 Research Question 1: Language Attitudes towards Peninsular Varieties

1. What are the language attitudes associated with each of the three Peninsular Spanish varieties, and to what extent do the patterns match the prestige hierarchy attested in the literature?

The word-initial /s/ variants were operationalized in this dissertation as a way to explore the interactions between linguistic and social information, providing a more reliable mapping between varieties than the dynamic *seseo*/*ceceo/distinción* processes that were outlined in chapter 2.0. Critically, while in the past *seseo* and *ceceo* were more reliably associated with Seville capital and Afueras, respectively, the conservative norm of distinction has spread to both locations and caused competition with the innovative regional norms. Recent production research reveals that this process is very much underway in Andalusia (see Section 2.4), and while the expansion of conservative traits is more advanced in eastern Andalusia when compared to western Andalusia, ,*distinción* has been shown to compete with both *seseo* and *ceceo* in locations such as Sevilla and Huelva (e.g., Regan, 2017a; Santana Marrero, 2018b, 2022).

As this dissertation aims to understand the intricate relationship between linguistic and social information in speech perception, primarily investigating the processes that underlie spoken word recognition, I selected a different set of variants that would provide a more stable mapping to geographic location. Thus, the three word-initial /s/ variants were featured in both the Verbal Guise and Long-Term Form Priming tasks, as they vary in place of articulation, are highly indexical, and are more consistently associated with their respective varieties.

While the prestige hierarchy of the three Peninsular varieties has been widely attested in a large body of literature, with perceived standardness driven by media representation and education, I implemented the Verbal Guise with the goal of exploring dialect classification and evaluation for the six talkers featured in this dissertation. Critically, I also needed to understand what specific language attitudes would be activated for listeners in the word recognition tasks when hearing the voice of each talker, as social information is inherently present in the speech signal, and talkers from the same dialect can be responded to differently in word recognition tasks (Floccia et al., 2006; Clapp et al., 2023).

6.1.1 Dialect Classification

The Verbal Guise featured a restricted set of three options for dialect classification of the six talkers: Seville, Madrid, and Afueras. I made this decision so that listeners would make associations between the voices and the three specific geographic locations of interest in the study, also connecting their previous experience and pre-established language attitudes for the varieties. The limitation of response options certainly boosted accuracy for the task as a whole, with an overall dialect classification rate of 82.1%. Given the limited nature of response type, in addition to the familiarity of the listeners with all three varieties and relatively small geographic distance

between all three locations, high dialect classification accuracy in the task is to be expected. While participants proved to be adept at regional classification, inferential statistics revealed significant differences across varieties, talkers, and talker gender.

There were significant differences in accuracy for both Afueras and Madrid when compared to Seville. Participants were more accurate in dialect classification for Afueras talkers than Seville, and even more accurate in identifying Madrid talkers compared to their own native variety. The higher accuracy in classification of Madrid talkers can be supported by dialect similarity, as Seville and Afueras are both innovative Andalusian varieties, while Madrid is a conservative, north-central variety (Penny, 2004).

There were also significant difference in dialect classification on a by-Talker basis. When compared to SW, accuracy was significantly lower for four talkers: SM, AM, AW, and MM. Participants were only significantly better at identifying the voice of MW. Additionally, listeners were significantly better at classifying women's voices. The combination of dialect classification accuracy results indicates that while participants generally performed well on the task, there was variety among talkers.

Creating a dialect confusion matrix depicted a systematic patterning of differences in accuracy. For instance, classification errors for Afueras voices were most frequently associated with Seville, and the pattern was stable across talkers. Errors for Seville dialect classification demonstrated different trends by talker, with SW more confused with Afueras and SM more linked to Madrid on wrong identification trials. Errors with Madrid voices were most often linked to Seville. Finally, there were minimal instances of confusion between Madrid and Afueras talkers (N = 5/504 trials; < 1% of data). The combination of results further establishes a differentiation in

accuracy between local (i.e., Seville and Afueras) and non-local (i.e., Madrid) varieties, though not all voices were as easily identifiable to listeners (i.e., SM).

6.1.2 Dialect Evaluation

The adjectives included in the dialect evaluation portion of the Verbal Guise task represented both status (i.e., 'refined,' 'snobby,' 'formal') and solidarity (i.e., 'calm,' 'polite,' 'happy,' 'pretty,' 'colloquial,' 'fun,' 'proud') categories. Associations for these categories according to patterns detailed in Chapter 2 anticipate higher ratings for status categories linked to Madrid, followed by Seville. Out of the three varieties, Afueras is expected to receive the lowest status ratings. While Madrid is expected to elicit high status ratings, solidarity ratings were expected to be the greatest for Seville, followed by Afueras, and lowest for Madrid. Finally, as Seville capital is associated with covert prestige, it is expected that some status ratings for Seville compete with those awarded to Madrid, either being equal or higher.

These predictions were largely upheld. The Madrid variety was rated highest for the status categories of refined, snobby, and formal, supporting the association the Seville capital listeners have with the standardness of the Madrid variety. Evidence of the covert prestige of Seville was manifested by the Seville capital voices always receiving the second-highest scores for status ratings, with Afueras associated with the lowest scores. As expected, Seville and Afueras were strongly evaluated in solidarity categories (Seville highest for happy, pretty, fun, proud; Afueras highest for colloquial, calm). For some solidarity adjectives (i.e., pretty), Madrid was evaluated more positively than Afueras, providing evidence for the nonstandard associations that accompany the local variety.

There were also different trends in dialect evaluation by talker and talker gender. Women's voices tended to receive higher ratings, though this finding should be interpreted with caution as all of the adjectives were represented in the feminine form. This methodological choice reflects the grammatical gender of the Spanish word for voice, *la voz*, as participants were always instructed to evaluate the voice of the talker according to the series of presented adjectives. However, it is possible that the presentation of feminine adjective forms resulted in a priming or biasing effect (i.e., higher evaluations for the three women talkers).

One noteworthy pattern evident in the consideration of the dialect evaluation ratings by talker is the differential treatment of SM and SW. While evaluation results collapsed by gender demonstrated higher ratings for women, participants reversed this trend for the Seville talkers, rating SM more highly for status categories and SW higher for 'colloquial.' Moreover, qualitative data provided by participants about the Seville capital talkers indicated a "forced neutrality" produced by SM, but a more typical Seville capital accent for SW.

The primary takeaways for the Verbal Guise are as follows: participants were successful in dialect classification of the talkers, there are significant differences in accuracy across both varieties and talkers, participant dialect evaluations generally followed expected patterns (status: Madrid > Seville > Afueras; solidarity: Seville > Afueras > Madrid), and talkers of the same variety are sometimes evaluated differently (i.e., SW and SM). These results will be referenced again after a summary of significant findings from the word recognition task.

6.2 Research Question 2: Immediate Recognition of /s/ Words (Block 1 Performance)

2. How does variation in word-initial /s/ production in the aforementioned varieties affect immediate word recognition of /s/ items for experienced listeners, and what can this tell us about the mental representation of the variants?

The second research question of this dissertation pertains to the processing of multiple sociophonetic surface variants. The three /s/ variants not only vary according to place of articulation, but they can all be linked to separate geographic locations and the language attitudes that correspond to them. The perception of cross-dialectal surface variants has been a recent topic of investigation, as results have not been fully congruent with frequency-based models of speech perception. Specifically, some studies have observed equivalent processing of multiple surface variants despite widely different rates of exposure for the listeners (e.g., Sumner & Samuel, 2009; Sumner & Kataoka, 2013). Other studies have observed costs in immediate processing even for familiar listeners, in which one of the surface variants is favored (e.g., López Velarde & Simonet, 2019). The current study joins López Velarde and Simonet (2019) and Gylfadóttir (2018) in contributing findings from the immediate processing of Spanish variants to the literature.

6.2.1 Accuracy and RTs from Block 1 /s/ Words

Accuracy and average RTs from Block 1 /s/ words were analyzed to assess the difficulty of spoken word recognition across the three Peninsular varieties. Error rates for /s/ items across the experiment indicated better accuracy for /s/ items produced by Seville and Madrid talkers when compared to Afueras talkers. Despite familiarity with all three varieties (and by proxy, the /s/ variants), the Seville capital listeners were significantly less accurate in accepting /s/ words

produced by Afueras talkers. There were no significant differences in accuracy for /s/ words for Seville and Madrid talkers. This result contrasts with López Velarde and Simonet (2019), who observed equal lexical acceptability rates for two local variants. The similarity of [s̪$^{\theta}$] to [θ] can support the difference in accuracy rates, as it is likely that some Afueras /s/ trials were erroneously interpreted as items produced with *distinción*. However, the nonstandard variant in López Velarde and Simonet (2019) does not overlap with any other phonetic or phonemic contrasts in *norteño* Spanish. For this reason, it is possible that López Velarde and Simonet's Hermosillo listeners experienced less dialect interference (i.e., competing activation) when mapping the local variant to lexical items.

An analysis of RTs for Block 1 /s/ words revealed no statistically significant differences by Variety or Talker Gender. There were also no significant interactions. Thus, while participants did not always accept lexical items beginning with the Afueras /s/ variant, when they did, the Afueras variant appeared to be mapped to its corresponding lexical items as efficiently as the other two variants, as only accurate trials are represented in the analysis of RTs. The lack of significant differences in RTs across variables for Block 1 /s/ words provides some evidence for recognition equivalence, in which variants associated with different frequencies are processed equivalently. López Velarde and Simonet (2019) did not observe recognition equivalence via RTs in immediate processing, as Hermosillo listeners responded significantly faster to the standard variant.

A hypothesis for the mental representation of the three /s/ variants in immediate processing can be made through referencing Sumner and Samuel's (2009) explanation of the 'fluent listener.' In their study, the 'Covert-NYC' English group of participants typically only produced r-ful forms (e.g., *baker* as [beɪkɻ]), but had exposure to r-less variants (e.g., *baker* as [beɪkə] through living in NYC. The Covert-NYC listeners showed perceptual flexibility in immediate processing,

manifested by equally rapid recognition of both r-ful and r-less forms. However, this flexibility did not transfer to representation, as the listeners only appeared to store the r-ful forms in memory, an effect seen in the results of a long-term priming task. Sumner and Samuel (2009) hypothesized that the Covert-NYC group mapped both variants onto a single form in immediate processing, with experience allowing the listeners to accept multiple surface forms.

The Seville capital listeners are similar to Sumner and Samuel's Covert-NYC group in that they primarily produce one of the variants (i.e., [s̪]) but have frequent exposure to the other two. However, they are also different in critical ways. For instance, is not clear whether the Covert-NYC group is affected by the nonstandardness indexed by r-dropping, as they typically do not actually produce the variant themselves. The native variety of the Seville capital participants is less standard when compared to that of Madrid, and the participants in the current study have personal experience with their native variety not being (accurately) portrayed in the media and being associated with negative stereotypes. During the sociolinguistic interviews, many of the Seville capital participants remarked that the negative stereotypes surrounding their region and variety bothered them. There was also prominent evidence of regional solidarity in their responses to the hypothetical question about living somewhere other than Seville capital: the vast majority of responses were locations within western Andalusia (e.g., Cádiz). This reality likely translates to the Seville capital listeners having a stronger activation of social categories associated with the varieties and the talkers that represent them in the word recognition task.

6.2.2 The Role of Real Control Words in Block 1

One topic of consideration that is absent from the research questions I had formulated when initially designing this dissertation is the processing of the real control words in the word

recognition task. While research has demonstrated that increasing the number of diverse talkers in an experiment can lead to greater processing difficulty (e.g., Clapp et al., 2023), the listeners in the current study were exposed to all six talkers in the Verbal Guise, allowing them to catalyze the familiarization process for the different voices. Moreover, the varieties presented in the word recognition task are familiar to the listeners; while each individual voice is unique, the speech signal from the talkers includes segmental and suprasegmental information that would be represented in memory by Seville capital listeners due to being native speakers of Peninsular Spanish. Lastly, the control words did not include salient sociophonetic variants that could result in slower processing. In summary, while I did not anticipate significant differences in the processing of control words, I expected that listeners would begin the word recognition tasks with regional associations and language attitudes activated from the Verbal Guise, which could affect processing of the items.

The processing of real control words resulted in multiple statistically significant effects when considering RTs. The model that only included real control items found a statistically significant effect for Variety (i.e., Seville) in Block 1 as well as an interaction between Variety and Talker Gender (i.e., SW). Participants were consistently faster in recognizing controls produced by SW. Despite both SM and SW being Seville capital talkers, patterns in word recognition varied. The patterning between SM and SW was more similar in models that include both /s/ words and real controls, while effects for SW were more related to real controls.

The observed Variety and Talker effects can be related to the Verbal Guise results, as participants were significantly more accurate in dialect classification for SW and stated that her voice was more identifiable as being Sevillian. Participants who did correctly associate SM with Seville tended to rate him higher for status categories and lower for ‘colloquial.’ Some participants

also referenced elements of forced neutrality in SM's voice. In summary, it is possible that prior completion of the Verbal Guise triggered particularly fast RTs for SW, as listeners had explicitly classified her as being from Seville and evaluated her positively. However, SM and SW are not processed significantly differently in the combined model; participants are exposed to the dental [s̪] variant that is shared by both SM and SW.

One interpretation that can explain these results is the possibility that different processing strategies are affecting the recognition of the real controls and /s/ words. If a more top-down approach is applied to real controls, information about the talkers will be more readily accessible to participants before lower-level surface features. In the case of Block 1 real controls, this hypothesis is supported by the significant interaction for SW. Provided that the frequency of the lexical items themselves is highly controlled, the processing benefit for one of the Seville talkers could reflect the top-down approach and the influence of the regional solidarity linked to SW.

The patterns in the processing of /s/ words can also be a result of differential strategies experienced by the listeners in word recognition. As the significant effect of Variety and the faster processing of SW disappears in the Block 1 /s/ model, it is probable that a bottom-up processing strategy is occurring, in which lower-level features such as the /s/ variants are attended to first before talker-specific regional information. This strategy is perhaps a necessity for Seville capital listeners given the task at hand, as they need to quickly map the variable speech signal to its appropriate form.

Another indication that a bottom-up processing strategy is applied to /s/ words relates to the greater slowing effect of the Afueras [s̪$^{\theta}$] when compared to the Madrid [s̪]. The first segment that listeners encounter is phonologically ambiguous, given the processes of *distinción*/*seseo*/*ceceo* that exist in Seville capital. While none of the word-initial /s/ items include

any phonologically contrastive minimal pairs, the participants in the study are naïve to this information. In the combined model, the significant interaction indicates that the Afueras [s̪$^{\theta}$] is even more costly to process for the listeners. Following the hypothesis of competing phonological activation, it is likely that some listeners initially mapped [s̪$^{\theta}$] to /θ/. Those who did not adjust this mapping upon hearing the rest of the speech signal recorded the item as a pseudoword, which occurred approximately 20% of the time for the Afueras [s̪$^{\theta}$] in Block 1.

However, even listeners who were ultimately able to correctly map [s̪$^{\theta}$] to /s/ before making a lexical decision would have encountered a cost if they first associated the variant with /θ/. While they would realize that it was an Afueras talker producing the nonstandard [s̪$^{\theta}$] variant instead of a talker producing a standard variant (i.e., [θ] for *distinción*), the initially incorrect mapping would lead to longer word recognition times. While recent speech perception models argue that linguistic and social information is processed in tandem (Sumner et al., 2014), it is not clear at what level this interaction occurs, as word recognition includes interaction between prelexical segments and suprasegemental features (McQueen & Cutler, 2010). When planning the word recognition task, my prediction was that the social information encapsulated by the three /s/ variants would be beneficial to listeners in processing, compounded by their familiarity with all three varieties and exposure to all six talkers in the Verbal Guise. However, the slowing effect for Afueras /s/ words compared to controls suggests not only that there is competing phonological activation for the first segment, but also that the association of [s̪$^{\theta}$] to an Afueras talker comes into effect later in the word recognition process. Another possibility to consider pertains to the activation of regional concepts. Similar to work by Hay et al. (2006), participants were exposed to 'labels' of the three regions during the dialect classification portion of the Verbal Guise. It could be that the activation of these

three locations, especially one that is associated with standard north-central varieties, implicitly guides participant associations of [s̪$^{\theta}$] to [θ].

6.3 Research Question 3: Priming Effects for /s/ Variants vs. Control Words

3. How are the /s/ variants represented in memory, and what implication do findings have for variable encoding strength as a result of social factors?

The third research question that I had established related to how the three /s/ variants are represented in memory. Prior work (e.g., Sumner et al., 2014) has suggested that the social value of variants can lead them to be more strongly encoded, even if they are less frequently encountered. The three socially salient /s/ variants provided an excellent context for the examination of social weighting, as one less frequently encountered variant is idealized and one is associated with nonstandardness. However, it is unclear how the social weighting proposal would affect the encoding process for listeners who have regional solidarity with a nonstandard variety and also speak a variety that is regarded as less standard than those associated with north-central zones.

The long-term form priming task provided an avenue to explore the encoding and representation of the three /s/ variants by the Seville capital listeners. As encoding is assumed to follow successful word recognition, this process would have occurred after participants recognized Block 1 /s/ words. However, encoding is not a given; not all tokens are encoded, and even if they are, encoding strength can be variable (Clopper, 2021; Sumner et al., 2014). Exploring the Block 2 presentation of /s/ words can provide answers to these questions: Are all three variants encoded by the listeners? If so, are they all encoded equally well? Equal encoding and representation of all three /s/ variants would manifest as faster Block 2 RTs for the three varieties. Provided that

encoding happens after word recognition, any significance of the prime variety, or interactions between prime and target varieties, can provide information regarding the encoding process.

For instance, if the Seville capital listeners only store their native variant, then only the Seville variety would be expected to speed word recognition times in Block 2 of the experiment (i.e., to prime them). If the Madrid [s̲] was also stored, faster facilitation for primes produced by Madrid talkers would be anticipated. Moreover, interactions between Block 1 and 2 varieties provide information about the specificity of any priming effects. Do the variants only prime themselves (i.e., identity priming), or do they also prime the other two variants? As an example, Sumner and Samuel (2005) only found a long-term priming effect for the Basic /t/ - Basic /t/ condition, leading them to hypothesize that only the canonical variant was encoded and represented in memory, despite equal treatment of the three surface variants in immediate processing.

6.3.1 Block 2 Results for /s/ Words & Differential Processing Strategies

The results from the Block 2 /s/ model revealed significant effects for both target conditions. Participants recognized /s/ words significantly faster when presented in Seville and Madrid varieties, but not Afueras. Further testing did not reveal any significant difference in RTs for Seville and Madrid Block 2 /s/ targets. There were no significant effects of prime variety and no interactions between prime and target varieties. The fact that there were no statistically significant effects in Block 1 for /s/ words and that significant effects appear in Block 2 for Madrid and Seville targets indicates a change, but more information is required to interpret this result.

Findings from Blocks 1 and 2 did not demonstrate significant effects of Condition for /s/ words compared to real controls. Despite this overall consistency, the significant interaction between Condition and Variety observed in the Block 1 combined model persists in Block 2, with a significantly slower effect in processing Afueras [s̪$^{\theta}$] words than Madrid [s̪] targets when compared to real controls. Moreover, an examination of Block 2 accuracy for /s/ words indicated that participants still encountered difficulties in lexical acceptability for Afueras [s̪$^{\theta}$] words. As in Block 1, participants in Block 2 failed to correctly identify [s̪$^{\theta}$] tokens in approximately 20% of the trials. In summary, participants are still slower and less accurate for [s̪$^{\theta}$] in Block 2, while accuracy remains high for the Seville and Madrid /s/ variants, and processing speed becomes significantly faster for the two varieties. An explanation of this result must consider the processing difficulties in successfully mapping the Afueras [s̪$^{\theta}$] to the correct representation, and the differential processing strategy hypothesis that I posed for Block 1 items can lend itself to interpreting these Block 2 results.

While participants have knowledge of the talkers and varieties within the experiment and native associations with the varieties outside of the experimental context, the lack of adaptability in efficiently recognizing [s̪$^{\theta}$] can be explained by an active bottom-up processing strategy for the /s/ words. It appears that some listeners continue to erroneously map the Afueras [s̪$^{\theta}$] variant to /θ/ in Block 2, resulting in error rates of 20%. If the top-down knowledge they already possess about the talkers and varieties were coming into play (i.e., *ceceo* is associated with Afueras, there are Afueras talkers in the experiment), the processing cost for [s̪$^{\theta}$] should be reduced in Block 2 as listeners become more effective at mapping the variable acoustic signal to lexical representations. The fact that this is not the case indicates that perhaps listeners continue to first map [s̪$^{\theta}$] to /θ/.

Those who do not adjust their mapping before the regional information is accessed reject the item, while others incur a processing delay after needing to adjust their initial association.

In addition to a bottom-up strategy winning out for the /s/ variants, it is also possible that the Seville capital listeners have a default mapping between any production similar to [θ] as /θ/, a result of the ongoing expansion of *distinción* in Seville. After listening to the interviews, the majority of the participants themselves employ *distinción*, which could mean that their default is to map the [θ] sound /θ/ in perception. Moreover, as *distinción* is also associated with standardness, it is possible that in an experimental context, their strong expectation was to hear *distinción* instead of *ceceo*. The lack of media representation of Andalusian varieties could further solidify this anticipation.

While /s/ words in Block 2 appear to be approached from a bottom-up strategy, there is evidence that a top-down approach continues to apply to the real control words. The significant interaction between Variety and Talker Gender was replicated in the Block 2 real control model, highlighting a processing benefit (via faster RTs) for SW. Despite this Block 2 significance, averages for response times for SW across Block 1 real controls, as well as Block 2 repeated and new control words, demonstrate the (nonsignificant) slowing effect participants encounter as the experiment progresses (i.e., approximately 60 ms between Block 1 real controls and Block 2 new real controls for SW), even for the most quickly recognized talker. These trends are depicted in Table 28.

Table 28 RTs of Real Controls by Block and Repetition

	Block 1 Control Words (first presentation)	**Block 2 Control Words (second presentation)**	**Block 2 New Control Words (first presentation)**
SW	570.81 ms	574.12 ms	629.87 ms

A final piece of evidence in support of the different processing strategies experienced by the Seville capital listeners pertains to patterns in significance across Condition. There are significant effects associated with SW in models that only include real controls, but SM does not appear to be linked to the same fast processing speed, despite also being from Seville capital. The addition of /s/ to the Block 1 control model leads to significant Variety effects, and no significant differences between SW and SM for RTs. Moreover, in the Block 2 /s/ model, there are significant effects by Variety, no effects by Gender, and no significant interactions. Thus, where SW and SM pattern differently for real controls, the presentation of /s/ items with a surface variant that they both share (i.e., [s̪]) potentially encourages a bottom-up strategy as the acoustic information is presented first.

6.3.2 Implications for Encoding and Representation

The purpose of the long-term form priming task in this dissertation was to explore the relationships among word recognition, encoding, and representation of the /s/ variants. Given the long lag between prime and target presentation (i.e., approximately 20 minutes), priming in this task would correspond to differences in how the variants were encoded and represented in memory. However, there are some challenges in interpreting the Block 1 and 2 results for /s/ words. First, while it is true that /s/ targets were responded to more quickly in Block 2 when produced by Seville and Madrid talkers as compared to Afueras talkers, it is not clear if this is indicative of differential encoding, as no significant effects or interactions were observed for prime or prime/target varieties. In other words, a "true" long-term priming effect as traditionally conceived should manifest as an effect of which variety produced the prime in Block 1, given that this is when encoding occurs.

Further complicating the interpretation is that the control words did not yield substantial repetition effects. In other words, repeated control words were not recognized significantly faster than new control words, contrary to repetition effects yielded for control words in other studies that implement priming paradigms (e.g., Sumner & Samuel, 2005, 2009).

Keeping in mind that evidence for differential encoding effects for the /s/ words is somewhat weak, the results do allow for multiple interpretations of the Block 2 /s/ word effects. I will overview these possibilities with caution, as no strong conclusions can be made given the lack of significance of the prime variety, the fact that participant responses in Block 2 are descriptively slower, and the absence of a strong repetition effect for new vs. repeated items that is often observed in repetition-based priming paradigms. As a reminder, Table 29 indicates the average Block 1 and Block 2 RTs for /s/ words by Variety, with significant target effects identified for Madrid (i.e., recognized significantly faster in Block 2 than Afueras), and Seville (i.e., recognized significantly faster in Block 2 than the average of Madrid and Afueras).

Table 29 A Comparison of Block 1 and Block 2 RTs for /s/ Words by Variety

	Block 1 RT, /s/ words (first presentation)	**Block 2 RT, /s/ words (second presentation)**
Seville	608.92 ms	593.43 ms (***)
Madrid	632.87 ms	607.98 ms (***)
Afueras	713.16 ms	711.01 ms

Despite participants generally slowing down in Block 2, word recognition times are notably faster for Seville and Madrid /s/ targets. This result can either be a result of a more effective adaptation to Seville and Madrid voices in Block 2 or a repetition effect for Seville and Madrid /s/

words. Floccia et al. (2006, p. 1277) expanded an account of the talker normalization mechanism (Johnson, 1997) to listener adaptation of regional accents by denoting two distinct phases: an initial period of disruption marked by decreased accuracy and response time to the stimulus, followed by an adaptation phase, in which participants either partially or fully return to baseline (i.e., unperturbed) comprehension. As there are no significant differences in response times for Block 1 /s/ words compared to the control words, and no significant effects by Variety in the Block 1 /s/ model, it is possible that the Seville capital listeners are not experiencing an initial disruption phase for /s/ words. Thus, while talker adaptation naturally occurs over the course of an experiment as listeners gain more exposure to the variable speech, the lack of disruption for Seville capital listeners could suggest that the faster Block 2 processing of /s/ words produced by Seville and Madrid talkers should not be characterized as talker adaptation *per se*, but rather as priming at some level of representation. This account seems plausible given the listeners' familiarity with all three varieties.

From a representational perspective, form repetition paradigms predict that a variant will elicit stronger priming effects for itself as compared to other possible variants.[11] Therefore, a significant priming effect would at least be expected for the Seville-Seville or Madrid-Madrid prime-target pairs to provide information regarding variant-specific encoding. While this finding is not reflected in the Block 2 /s/ model, it is possible that more statistical power is necessary to demonstrate a clear priming effect by Variety. Descriptive trends do reveal the fastest response times for Seville-Seville prime-target pairs when compared to the other conditions, with Madrid-

[11] However, so-called identity priming is not guaranteed; priming effects for the NYC-NYC condition were not observed for inexperienced listeners in Sumner and Samuel (2009).

Madrid as the next fastest pairing. These non-significant qualitative patterns can be observed in Table 30.

Table 30 Descriptive Patterns in Prime-Target Pairs

Block 1 Variety	Block 1 RT, /s/ words (first presentation)	Prime-Target Pairing	Block 2 RT, /s/ words (second presentation)
Seville	608.92 ms	**Seville-Seville**	549.62 ms
		Seville-Madrid	607.82 ms
		Seville-Afueras	684.18 ms
Madrid	632.87 ms	**Madrid-Madrid**	588.75 ms
		Madrid-Seville	616.46 ms
		Madrid-Afueras	725.11 ms
Afueras	713.16 ms	**Afueras-Afueras**	721.65 ms
		Afueras-Seville	621.91 ms
		Afueras-Madrid	629.43 ms

What is also relevant in the descriptive trends outlined in Table 30 is that the Afueras-Afueras pairing appears to be associated with particularly long response times. A speculative encoding prediction based on the difference in recognition times for Seville and Madrid target varieties in Block 2, as well as prime-target trends in the /s/ word data, is that both Seville and Madrid variants were encoded in Block 1. It would be reasonable to assume that the Afueras variants would not be strongly encoded as a result of the processing difficulty encountered in the task, as even experienced listeners still experienced costs for the Afueras /s/ variant in Block 2.

If Seville and Madrid variants were both encoded, but not Afueras tokens, it is possible that the Seville capital listeners treat the Afueras [s̪$^{\theta}$] as a sub-variant of the Seville capital [s̪]. Similar to the mechanisms proposed for the fluent listener in Sumner and Samuel (2009), this cognitive landscape would allow for perceptual flexibility in immediate processing but would not result in long-term representation of the [s̪$^{\theta}$] variant. If this were indeed the case, it could be hypothesized that social weighting would explain the storage of Madrid variants, similar to results for the British English variety in Sumner and Kataoka (2013). While not as frequent to the listeners, the variant is associated with the national standard. This possibility is depicted in Figure 14; as in exemplar models, acoustic details and social information would both be represented in the lexicon. Future research is necessary to confirm if these trends reflect how the variants are treated in memory by Seville capital listeners.

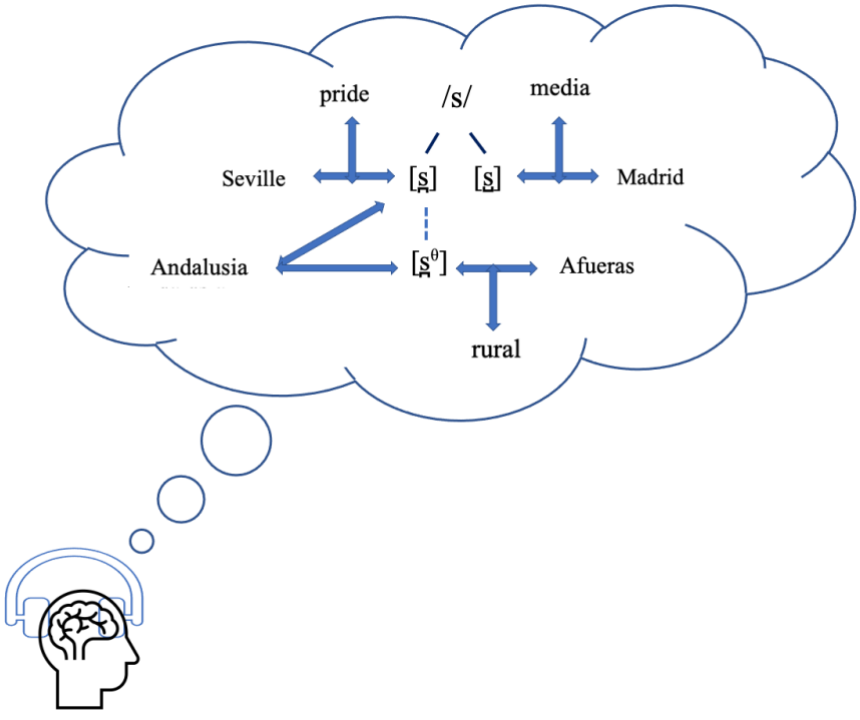


Figure 14 Possible Treatment of /s/ Variants by Seville Capital Listeners

6.4 Mixed Methodologies and the Local Context

As outlined in the Introduction chapter 1, the current study employs a unique combination of sociophonetic and psycholinguistic tasks that provide for a more intricate analysis of the results. For example, although some experimental studies in spoken word recognition gathered evaluative data about the talkers featured in their projects, these ratings were not completed by the same listeners who participated in the implicit tasks (e.g., Sumner & Kataoka, 2013). While collecting dialect evaluation information about talkers is a beneficial methodological decision in experimental work, I argue that it is even more informative when the evaluations are completed by the same listeners who participate in the implicit tasks. Listeners' own experience with the variants and varieties can highly influence perception (e.g., Clapp et al., 2023); in the current study, not all participants rated the talkers' voices in the same way. For example, one participant wrote that MW's voice was 'pretty,' while another found it to be 'elite and repulsive.' It is critical to capture these variable language attitudes, as data from sociophonetic tasks can support the interpretation of unexpected findings.

Just as many psycholinguistic studies do not integrate sociophonetic methodologies, the same pattern is upheld with sociophonetic investigation, as it is rare that dialect classification or evaluation research include experimental tasks that probe auditory lexical processing. For example, while matched-guise work has yielded robust results about the indexical field of salient variants, the task does not reveal any information about how the variants are recognized or encoded. The point of this discussion is certainly not to demand that every experimental task include sociophonetic elements in the task sequence, or insinuate that sociophonetic tasks are incomplete without psycholinguistic methodologies; prolific findings have emerged from the

individual implementation of these tasks (see chapter 2). My point is, rather, that the combination of these traditionally separate subfields (and by proxy, the methodologies and tasks associated with them) provides a more comprehensive view on the relationship between social and linguistic information in speech perception. In order to understand the immense variation that listeners encounter on a daily basis, the combination of these methodologies is paramount, as linguistic information can affect social perception (e.g., Chappell, 2019; Walker et al., 2014), and social information can lead to differential processing of linguistic variants (e.g., Hay & Drager, 2010). The use of the Verbal Guise and Word Recognition tasks in the current study served as an effective pairing to explore this relationship.

A factor that cannot be ignored in the investigation of how linguistic and social information interact in processing pertains to the local context. The language attitudes and stereotypes associated with any given variant not only depend on talker and listener characteristics, but also the broader sociopolitical landscape. One of the primary reasons that I selected the three Peninsular varieties is because of the complex dynamic between standardness and regional association.

The cultural pride that the Seville capital listeners have for their variety, place of residence, and region was established through multiple task measures. In the Verbal Guise, the Seville capital talkers were awarded high solidarity ratings, with SW receiving some of the highest scores across all talkers for solidarity categories. While differences emerged between the evaluation of SM and SW, when compared to the other two men, SM received the highest ratings in multiple solidarity categories (i.e., 'pretty,' 'proud,' and 'happy'). When considering results for incorrect dialect classification, solidarity ratings for both MM and AM were higher when listeners associated them with a Seville capital talker. These are explicit measures, as participants consciously evaluated the talkers' voices and the Verbal Guise was not timed.

Other explicit evidence of the positivity associated with Seville capital is found in participant commentary in the brief interviews following the psycholinguistic task. For example, when answering the question about hypothetically living in a place other than Seville capital, three participants remarked that they would not even entertain the idea of residing in a different location. For example, Participant 002 stated that, *en principio no me voy de Sevilla* 'in principle, I'm not leaving Seville.' In another instance, after rejecting the selection of another location, Participant 045 remarked that, *como Sevilla no hay 'na [nada]* 'there's nothing like Seville.' Participants also indicated that they felt positively about their own way of speaking, with the majority summarizing that they felt good, proud, and comfortable.

Just as participants in the current study positively evaluated Seville capital talkers and the location itself, this sentiment expanded to the region of Andalusia. Afueras talkers also received high solidarity ratings in the Verbal Guise, accompanied by more dialect confusion with Seville capital talkers. Additional evidence of strong regional solidarity was present in the interview data. For example, when answering the hypothetical moving location question, the vast majority of participants referenced cities within Andalusia. Cádiz, a coastal city in western Andalusia with frequent *ceceo*, was the most popular response. Other Andalusian cities such as Málaga, Granada, and Córdoba were provided as answers. Furthermore, numerous participants explicitly stated that they would not want to live outside of Andalusia. Taken together, the combination of Verbal Guise and interview trends support the clear prevalence of both locational (i.e., Seville capital) and regional (i.e., Andalusia) pride maintained by participants in the current study.

While ample evidence outlined the positive associations participants had with their home, the national standard actively competes with the Seville capital variety, exemplified in the literature by the expansion of *distinction* in western Andalusia and masking of Andalusian traits

in the media (see consistent evidence of the prestige hierarchy in Sections 2.4, 2.5, and 2.6). Results from the current study replicate the presence of this hierarchy, with Madrid talkers receiving the highest status ratings in the Verbal Guise along with less confusion between Andalusian and central talkers in dialect confusion. The exception to this rule was SM; listeners often classified him as a Madrid talker. Written comments provided by participants, as well as interview responses when asked about the varieties that appeared in the experiment, explicitly referenced the neutralizing of an Andalusian accent to converge towards a central variety. The lack of representation of Andalusian varieties in the media was supported in the literature (e.g., León-Castro, 2016), and when asked if they felt that their variety was accurately represented in the media or movies, the majority of the participants in the current study answered no. Thus, the themes of minimal dialect representation on national levels and convergence towards central norms in situations of perceived formality (i.e., news anchors speaking on TV) are overtly and prominently displayed by participants.

The presentation of the Verbal Guise first in the trial sequence was expected to activate the social categories associated with each of the talkers and varieties. Results from the word recognition task confirmed this anticipation, yet also served as a metaphor for the complex linguistic and social landscape present in Andalusia. Evidence of regional solidarity is associated with the (proposed) top-down processing of control words. It was not the more formal, prestigious sounding Seville talker (i.e., SM) whose voice yielded a processing benefit via faster RTs, but rather the one whose voice was more identifiable and representative of a proud Seville capital citizen (i.e., SW). However, upon presentation of a highly indexical variant, the differences between the Seville capital talkers disappear. Moreover, in Block 2 /s/ words, both Seville and Madrid targets are associated with a processing benefit via fast RTs. The implications from the

processing results suggest that the prestige hierarchy is also present on an implicit level, manifested by the differential treatment of Seville capital talkers in control and /s/ word conditions.

The results from the current study strongly support the influence of social categories on auditory lexical processing, and that this influence can be affected by differential processing strategies. Critically, even disyllabic, phonetically neutral, frequent nouns can exhibit social priming effects in experimental tasks; the processing benefit for SW could be a result of strong activation of social categories such as ‘Seville capital,’ along with positive dialect evaluation.

7.0 Main Conclusions, Limitations, and Future Directions

7.1 Main Conclusions

This dissertation contributes a novel experimental context to the literature on cross-dialectal word recognition, employing variable place of articulation of /s/ to observe dialect classification, evaluation, and word recognition processes of three Peninsular varieties of Spanish. Results indicated that Seville capital listeners were successful in identifying regional associations of talkers, but some individuals were harder to identify. Patterns in dialect evaluation demonstrated that Seville capital listeners associated Madrid voices with status categories, evidence of the overt prestige associated with the national standard. On the other hand, Andalusian varieties were primarily rated highly for solidarity categories. Results also indicated that not all talker voices of the same varieties were evaluated equally. For example, Seville capital participants indicated that SW's voice was highly identifiable as being from Seville, while SM's voice was accompanied by comments of forced neutrality. Additional differences emerged according to talker gender; participants tended to rate women's voices higher than men's voices across status and solidarity categories. The exception to this pattern was the Seville capital talkers; SM's voice received higher scores for status categories than that of SW, and SW's voice yielded higher solidarity ratings than SM's voice, contrary to gender results across the other pairs of talkers.

Results from the word recognition experiment revealed talker-specific effects, differential treatment of /s/ words between Blocks 1 and 2, and increased processing difficulty in Afueras /s/ words when compared to control words. Evidence of recognition equivalence was observed in the treatment of Block 1 /s/ items, manifested via the lack of significant effects of Variety. The long-

term implications for encoding and representations are not fully clear; future research will be necessary to observe whether patterns in the data for faster recognition of Seville and Madrid /s/ targets after a long lag are truly the result of differential encoding. The interaction of processing strategies can support the differences in recognition of control and /s/ words, in which segmental information of the word-initial /s/ appears to drive a bottom-up approach. Additionally, the inclusion of diverse talker voices appeared to increase task difficulty; no significant repetition effect was observed for repeated items.

A novel methodological feature of the current study was a unique combination of sociophonetic and psycholinguistic tasks, permitting a comprehensive analysis of the results. As research continues to explore the intricate interactions of social and linguistic information in auditory lexical processing, I argue that the mixed-methods approach employed in the current study is fundamental for work that strives to apprehend the extraordinary sociophonetic variability present in naturalistic speech. Congruent patterns of results appeared across tasks with varying levels of experimental control, emphasizing the utility of exploring perception via both explicit and implicit tasks.

The results observed in the current study strongly support the role of social information in processing, highlighting the impact of the activation of social categories on spoken word recognition. This activation and its implications are clearly manifested by the differential treatment of Seville talkers across tasks. Listeners consistently showed a processing benefit via faster RTs for SW, the talker who was most identifiably Sevillian to participants and who received the highest solidarity ratings in the Verbal Guise. On the other hand, control words produced by SM were not accompanied by fast RTs despite the talker's regional background. SM's voice was met with less accurate classification, higher status ratings than SW, confusion with a Madrid talker, and

comments of forced neutrality. Thus, it can be hypothesized that SW's voice yielded stronger activation of social categories, and that this activation was consistent with the local context of the varieties (i.e., cultural pride for Seville capital and Andalusian speech traits). In the case of SM, I hypothesize weaker activation of social categories like 'Seville capital,' as listeners were less accurate in classifying his voice. However, even for some listeners who did correctly identify SM as a Seville capital talker, the linking of SM's voice with a talker trying to sound less Andalusian may very well have contributed to slower processing of controls compared to SW, associations that are reflective of discussions of a lack of media representation of Andalusian varieties that is a central component of the local context.

7.2 Limitations and Future Directions

The majority of the limitations of the current study are related to the long-term form priming task; one issue pertains to statistical power. I had estimated needing a minimum of 36 participants to complete the word recognition task, and while 42 individuals completed the task, it is possible that some of the trends were masked by a potential lack of data. Additionally, given that participants were least accurate with the Afueras /s/ trials, there are less data represented for these conditions.

Other limitations relate to the lack of a repetition priming effect for repeated words in the word recognition task. While recent research has observed a similar pattern in studies that include diverse talker voices (Clapp et al., 2023), the lack of repetition priming effects makes the interpretation of the results more challenging. Adding to the complexity of the interpretation is also the element of participant fatigue. The trade-off for establishing a long lag of 20 minutes

between prime and target presentation to observe the representational status of the variants manifested as participant RTs slowing down as the experiment progressed. Finally, while the effect of SM's more formal productions led to intriguing comparisons of listener ratings and elicited commentary from participants, it was the original intent of the study to include talkers with clearly identifiable regional associations. It is possible that the decreased association of SM with Seville capital masked some of the priming effects that could have been produced from the Seville-Seville condition in the word recognition tasks, as prime-target pairs were always produced across talker gender.

Future work would do well to test perception of the three Peninsular /s/ variants through the implementation of processing tasks that are less susceptible to inaccuracies from a variant being less widely accepted as a lexical item (e.g., cross-modal lexical decision). In these paradigms, participants are presented with an auditory prime followed by a written target word, making a lexical decision on the target word (e.g., Clopper et al., 2016). It would also be interesting to explore if processing of *ceceo* would be any different in word-medial contexts, providing a continued observation of top-down vs. bottom-up processing. Lastly, the manipulation of the experimental environment has been demonstrated to have a robust effect on speech perception (e.g., Hay & Drager, 2010). Probing testing environments that foster informality or introducing elements to condition participants' perception toward the nonstandard Afueras /s/ variant will be fruitful avenues for future research.

Finally, the perception of dialectal variants in L2 populations can reveal effects of experience, with tasks such as matched-guise (e.g., Chappell & Kanwit, 2022) and categorization (e.g., Schmidt, 2018) employed across learners of diverse proficiency levels. Future directions for this research in an L2 context could include testing learners' ability to recognize a variant that

includes multiple mappings (i.e., *ceceo*), and language attitudes they attribute to the variant. Testing learners who are studying in coastal Andalusian cities such as Cádiz or Málaga would provide the ideal context for this research, as *ceceo* is frequent and likely unfamiliar to learners before arrival.

In conclusion, Seville capital participants demonstrate impressive perceptual flexibility as listeners who are surrounded by immense social and linguistic variation.

Appendix A Talker Instructions

¡Hola! Muchas gracias por prestarme tu voz. Por favor, lee las instrucciones en cada parte. Antes de empezar a grabar, te pido que estés en un lugar silencioso, ya que las voces de otras personas o ruidos adicionales pueden bajar la calidad del audio.

Tu voz aparecerá en una serie de experimentos que estoy dirigiendo. El proyecto trata de la percepción del habla. Aunque los participantes escucharán tu voz, tu nombre no aparecerá en ningún sitio. Es posible que las grabaciones de voz o transcritos se compartan con otros investigadores en el futuro, pero aún en este caso todo será anónimo. Aunque estarás contestando preguntas escritas/ pronunciando palabras escritas, por favor, habla como si estuvieras charlando con tu mejor amigo. No son tareas formales, se busca un estilo de habla casual.

Estimo que se requiere aproximadamente 30 minutos para contestar las preguntas y decir las palabras. Cuando termines cada parte, me puedes enviar el archivo de sonido por WhatsApp. Para agradecerte, al final te paso una tarjeta de regalo de 30 Euro a una tienda de tu elección.

Si tienes cualquier pregunta, no dudes en escribirme por WhatsApp.

Parte I: Preguntas Abiertas

Por favor, contesta las 5 preguntas siguientes en voz alta. Puedes incluirlo todo en la misma grabación. Recuerda que no es una tarea formal.

Preguntas

6. ¿Qué hiciste el sábado pasado?
7. ¿Cuáles son algunos planes que tienes para la semana que viene?
8. ¿Te gusta el salmorejo?
9. ¿Cuál es el mejor lugar para tomar el sol y por qué?
10. ¿Qué opinas de la tradición de la siesta?

Parte II: Palabras reales e inventadas

Uno de los experimentos que van a hacer los participantes trata de decidir si una palabra es real o no. Por favor, haz una grabación en la que dices cada palabra en voz alta. Debes pronunciar la palabra de manera declarativa y pausar un segundo entre cada palabra para que no haya interferencia. Puedes hacerlo todo en la misma grabación o dividirlo en grupos más pequeños, lo que prefieras tú.

Como verás, todas estas palabras son reales:

Palabra							
yuca	soto	sauna	olla	cejas	barcos	lengua	pelo
sarro	mella	tapas	ciervo	nube	cuadro	costa	traje
tromba	dorso	senda	cancha	feria	salto	prensa	sexo
cacto	bucle	sorbo	menta	nido	joyas	silla	grupo
diabla	sarta	chisme	ajo	grasa	plaza	leche	fuerza
sorna	gruta	caldo	zumo	queja	túnel	ruido	sueño
tilde	cóndor	soja	nieta	susto	salsa	brazos	trato
sardo	fierro	simio	sonda	burro	móvil	hielo	cielo
naipe	silba	peine	suero	horno	templo	trampa	sitio
ubres	guiño	mangas	bota	trenes	fiebre	beso	libro
surco	tecla	hongo	monja	vidrio	moda	banda	falta
churros	fósil	plancha	frenos	copas	seres	puente	fuego
mimbre	ostra	soplo	soda	huella	venta	flores	hijos
miga	cuenco	ancla	frutas	pesca	sombra	precio	coche
equis	ocio	mugre	pausa	mancha	ritmo	ganas	sangre
grapa	silo	rabo	ruina	suma	sopa	fondo	tierra
caqui	podio	mármol	gripe	hombros	saco	vuelo	cara
facha	bache	aula	mito	postre	fecha	isla	ojos
orca	franja	charco	sexto	selva	gira	zona	agua
leño	saldo	piña	falda	brindis	bruja	radio	horas
seto	ámbar	secta	soga	tarta	plata	reina	muerte
plagio	rasgo	sable	bicho	muebles	notas	santo	miedo
trufa	helio	sabio	signo	letras	granja	luna	clase
salvia	lienzo	fresa	diente	medias	siglo	norte	suerte
gamba	viudo	loro	pila	sueldo	barrio	cárcel	
yema	calcio	sobras	beca	barba	queso	sala	
sigma	sodio	himno	zorro	lujo	lago	suelo	
mote	jerga	finca	siesta	hoja	lluvia	alma	
rizo	pera	sapo	talla	risa	techo	tienda	
pana	cono	dato	sierra	jefa	serie	centro	
cuña	prendas	suegra	raya	frase	polvo	foto	
salmo	lomo	jarra	sello	palo	ángel	mesa	

Estas palabras son las inventadas. **Solamente tienes que decir las palabras en negrita.** Las palabras no reales fueron inventadas de palabras existentes, que son las que salen a la derecha de cada ítem. Aunque no tienes que decir las palabras reales en esta sección, pueden servir como guía de pronunciación por si tienes alguna duda. Debes pronunciar las palabras inventadas como si fueran reales, siguiendo las reglas de pronunciación española. Por favor, échale un vistazo a la lista antes de empezar a grabar- debe sonar lo más natural posible.

Palabra inventada	Palabra real								
conlo	calco	**pulma**	pulga	**mafo**	mazo	**marno**	mango	**viendre**	vientre
bestro	bistro	**corjus**	corpus	**darco**	dardo	**llotos**	motos	**gaca**	hada
galmo	galgo	**lugro**	lucro	**lonte**	lente	**asla**	asma	**munos**	muros
plimas	plisas	**vermo**	verbo	**jaunas**	jaulas	**cisle**	cisne	**lorta**	torta
lostre	lustre	**febia**	fobia	**balna**	balsa	**resis**	tesis	**flaga**	plaga
sirra	mirra	**diedra**	hiedra	**crote**	brote	**dibra**	fibra	**blecha**	flecha
lirra	birra	**linte**	tinte	**bilto**	bulto	**ganro**	ganso	**nerma**	norma
fialo	fauno	**meño**	moño	**palgas**	palmas	**fuepos**	huevos	**juedos**	jueces
carfe	canje	**duerta**	huerta	**bunda**	funda	**custa**	cesta	**misca**	mosca
dedro	cedro	**muema**	muela	**liegre**	liebre	**dasto**	gasto	**fueños**	dueños
guna	duna	**dallos**	gallos	**trurra**	trucha	**ucas**	uvas	**bafos**	baños
fucto	ducto	**dactros**	filtros	**sodia**	socia	**flardo**	flanco	**milta**	multa
mada	cata	**trempa**	trompa	**musno**	muslo	**taga**	tasa	**rinta**	renta
truña	trufa	**akras**	atlas	**groche**	broche	**truelo**	trueno	**lluño**	puño
barlo	bardo	**bloro**	cloro	**balne**	balde	**nuzo**	nudo	**fruce**	cruce
calpa	caspa	**samla**	samba	**trispas**	chispas	**yona**	yoga	**huedo**	hueco
fueto	dueto	**letas**	setas	**rego**	remo	**pemos**	pesos	**ogos**	osos
ajia	apio	**baichas**	bayas	**diso**	filo	**ragas**	ramas	**cagles**	cables
belmo	bombo	**banfo**	banjo	**chijo**	chivo	**firo**	faro	**chostes**	chistes
euge	auge	**arja**	arpa	**yiso**	yeso	**rares**	mares	**mubo**	cubo
fiosas	diosas	**péster**	póster	**bomas**	bocas	**trimo**	trigo	**blobo**	globo
fanzas	danzas	**bimo**	lino	**colso**	colmo	**tarno**	tango	**mele**	coro
gesco	vasco	**mebra**	cebra	**vimio**	vicio	**folo**	velo	**pamos**	pavos
resca	rosca	**verpa**	verja	**flera**	flora	**nuemes**	nueces	**pento**	parto
afre	acre	**penza**	pinza	**licas**	ligas	**rigles**	rifles	**vieta**	dieta
ririo	lirio	**alclo**	astro	**flantas**	llantas	**yete**	yate	**duelga**	huelga
paitas	pautas	**gada**	gama	**rospa**	rampa	**plemo**	plomo	**papria**	patria
rerno	termo	**frimas**	primas	**rogre**	roble	**darras**	garras	**crito**	grito
dorro	forro	**euna**	aura	**plauta**	flauta	**chicre**	chicle	**mialo**	cuero
noya	boya	**lollo**	bollo	**suemo**	sueco	**floque**	bloque	**mundes**	mundos
hichas	hachas	**crapta**	cripta	**fuende**	duende	**chiunfo**	triunfo	**lópiz**	lápiz
cengo	censo	**custe**	coste	**legros**	logros	**álsum**	álbum	**dacho**	fallo

refto	texto	**panio**	patio	**pade**	pago	**merro**	perro
gisto	gesto	**houda**	deuda	**carres**	calles	**erden**	orden
marde	marte	**pima**	mina	**fiario**	diario	**argas**	armas
cagra	cabra	**tama**	lana	**hoches**	hechos	**cunsto**	cuarto
binta	quinta	**ausco**	asco	**piegra**	piedra	**maba**	cama
melos	celos	**raca**	pata	**cergo**	cerdo	**puebro**	pueblo
loda	gota	**tíes**	tíos	**pluya**	playa	**mosio**	medio
rarcho	rancho	**meblo**	metro	**pamas**	casas	**cuapa**	culpa
mecra	mezcla	**charga**	charla	**carsa**	carga	**diasta**	fiesta
niedra	niebla	**nunta**	junta	**lurra**	lucha	**viepta**	vuelta
glama	drama	**guka**	guía	**órbol**	árbol	**cuerjo**	cuerpo
huinza	fianza	**pados**	pasos	**numbros**	nombres	**medes**	meses
dena	vela	**rura**	ruta	**golre**	golpe	**vuerra**	guerra
vama	fama	**dumo**	humo	**manlla**	marcha	**nochos**	noches
trolo	trono	**piaza**	pieza	**tuelvos**	tiempos		
labla	lata	**clanta**	planta	**pusta**	pista		
mava	masa	**bistia**	bestia	**obla**	obra		
umas	uñas	**clape**	clave	**blogas**	drogas		
cafas	papas	**compnas**	compras	**flave**	llave		
cuepa	cueva	**tabios**	labios	**bame**	base		
vomes	voces	**retla**	regla	**cricen**	crimen		
rigre	tigre	**rorre**	torre	**olo**	oro		
mirte	monte	**nieje**	nieve	**llajas**	llamas		
ruema	rueda	**fúnbol**	fútbol	**crisa**	prisa		
nersias	nervios	**ciercia**	ciencia	**fample**	hambre		
maos	caos	**firga**	firma	**nepia**	novia		
gaxas	gafas	**juemes**	juegos	**iare**	aire		
flima	clima	**porbo**	pollo	**pono**	modo		
goco	voto	**malpa**	mapa	**reba**	ropa		
ulas	alas	**masia**	magia	**grueba**	prueba		
nalle	valle	**duente**	fuente	**giaje**	viaje		
ranque	tanque	**líser**	líder	**campue**	cambio		

Parte III. Preguntas Adicionales

Por favor, contesta estas 4 preguntas en otra grabación. ¡Esta es la última parte!

1. ¿Cómo te sientes sobre tu manera de hablar?
2. Cuando viajas fuera de tu ciudad natal, ¿la gente normalmente sabe de dónde eres? Si es el caso, ¿cómo crees que lo saben? Y si no, ¿de dónde piensan que eres?
3. ¿Es común que escuches un acento parecido al tuyo en los medios de comunicación o en las películas o series españolas?
4. Si pudieras vivir en otra ciudad española, ¿cuál seleccionarías y por qué?

¡Mil gracias de nuevo por ayudarme! Lo aprecio un montón. ☺

Appendix B Study Script Provided to Participants

Appendix B.1 Spanish Study Script

El objetivo de este proyecto es investigar los patrones del español peninsular. A tal fin busco participantes de Sevilla, España. Les pido a los participantes que completen una breve tarea de evaluación lingüística (15 minutos), una tarea de percepción (35 minutos), una breve entrevista (6 minutos), y un cuestionario de experiencia lingüística (10 minutos). No hay ningún riesgo previsible que se asocie con este proyecto, **ni hay ningún beneficio directo para usted**. Dicho esto, su participación nos ayudará a tener un mejor entendimiento de cómo se usa la lengua en España. Todas las tareas son anónimas, lo que significa que sus respuestas no se asociarán con usted de ninguna manera. Cada participante recibirá 30 Euro inmediatamente después de la terminación del proyecto. Se estima una duración de una hora y 15 minutos para realizar las 4 tareas. Su participación es voluntaria, y usted puede abandonar el proyecto en cualquier momento. Si decide dejar de completar el estudio antes de subir algunas respuestas, no usaré sus datos. Sin embargo, si decide abandonar el estudio después de subir todas las respuestas, sí usaré los datos debido a que los investigadores no sabrán cuales respuestas son suyas. Es posible que las grabaciones de voz o transcritos de la entrevista se compartan con otros investigadores en el futuro, pero su nombre no aparecerá en ningún sitio. Para mitigar el riesgo de una violación de confidencialidad como resultado de participar en este proyecto, no le pido a usted que diga su nombre en la grabación. Cualquier comentario que contenga información personal (ej. nombres de amigos, direcciones, fecha de nacimiento) será eliminado del audio y del transcrito para preservar la anonimia. Entiendo que puedo ponerme en contacto con el Defensor de la Protección de Participantes Humanos en la Oficina de Protección de la Investigación Humana, en la Universidad de Pittsburgh (*1-866-212-2668) para hablar sobre cualquier problema, inquietud, o pregunta; para obtener información; ofrecer mi punto de vista; o mencionar situaciones que ocurrieron durante la participación mía*. Este estudio es realizado por Angela Krak (amk308@pitt.edu) por si usted tiene cualquier pregunta. ¡Muchas gracias!

Appendix B.1.1 English Translation of Study Script

The purpose of this research study is to investigate patterns of Peninsular Spanish. For this reason, I will be surveying participants from Seville, Spain. I will ask participants to complete a brief linguistic evaluation task (15 minutes), a perception task (38 minutes), a short interview (6 minutes), and a language background questionnaire (10 minutes). There are no foreseeable risks associated with this project, **nor are there any direct benefits to you**. However, your participation will allow researchers to better understand how language is used in Spain. All tasks are anonymous, meaning that your responses will not be linked back to you in any way. Each participant will receive 30 Euro immediately after completing the project. The four tasks are estimated to require 1 hour and 15 minutes to complete. Your participation is voluntary, and you may withdraw from the study at any time. If you choose to withdraw before submitting any responses, your data will not be used. However, because no identifiers will be collected, if you decide to withdraw after you have submitted responses, your data will continue to be used since the investigators will have no way of knowing which responses are yours. It is possible that audio recordings or transcripts of the interview will be shared with other researchers in the future, but your name will not appear anywhere. To mitigate the potential risk of a breach of confidentiality as a result of participating in the study, you will not be asked to state your name in the audio recordings. Any mention of identifiers (e.g., names of individuals, street addresses, dates of birth) will be redacted from the transcript and deleted from the audio recording to preserve anonymity. *I understand that I may contact the Human Subjects Protection Advocate of the Human Research Protection office, University of Pittsburgh (1-866-212-2668) to discuss problems, concerns, and questions; obtain information; offer input; or discuss situations that occurred during my participation.* This study is being conducted by Angela Krak, who can be reached at amk308@pitt.edu, if you have any questions. Thank you!

Appendix C Task Stimuli

Appendix C.1 Verbal Guise Stimuli

Appendix C.1.1 Transcriptions of Verbal Guise Responses to *Sábado* Question

Talker	Response	Translation
AM	*El sábado pasado me levanté a las 12 de la mañana estuve todo el día viendo la serie de Stranger Things con mi novia.*	Last Saturday I got up at noon, I was watching Stranger Things all day with my girlfriend.
AW	*El sábado pasado lo que hice es lo que suelo hacer casi todos los sábados por la mañana que es limpiar la casa y hacer las tareas del hogar porque trabajo entonces pues la mañana la dedico a eso.*	Last Saturday, what I did is what I usually do almost every Saturday morning, which is to clean the house and do housework because I work, so I dedicate the morning to that.
MM	*¿Qué hiciste el sábado pasado? El sábado pasado fue un sábado muy bonito porque estuve viendo carreras de motos todo el sábado.*	What did you do last Saturday? Last Saturday was a very nice Saturday because I was watching motorcycle races all Saturday.
MW	*El sábado pasado estuve en la playa con unos amigos y estuvimos paseando, tomando el sol, de chiringuitos, y de relax.*	Last Saturday I was at the beach with some friends and we were walking, sunbathing, at beach bars, and relaxing.
SM	*El sábado pasado por la mañana fui a recoger a mi hija. Por la tarde nos fuimos a la piscina y después vimos una peliculita tranquilamente en la television.*	Last Saturday morning I went to pick up my daughter. In the afternoon we went to the pool and later we calmly watched a little movie on television.
SW	*Pues mira el sábado pasado no hice nada especial; me quedé en mi casa. Me quedé mi casa porque como me operaron hace un mes del pie y no me puedo mover, estoy inmovilizada. Tengo la bota está walker que no puedo caminar bien.*	Well, look last Saturday I didn't do anything special, I stayed at home. I stayed at home because they operated on my foot a month ago and I can't move, I'm immobilized. I have this boot and walker and I can’t walk well.

Appendix C.1.2 Transcriptions of Verbal Guise Responses to *Siesta* Question

Talker	Response	Translation
AM	*¿Qué opino de la tradición de la siesta? No suelo dormir siesta. Sí es verdad que en verano duermo algo más pero tampoco soy muy fanático de la tradición.*	What do I think of the siesta tradition? I don't usually take naps. Yes it's true that in the summer I sleep a little more, but I'm not really a big fan of the tradition.
AW	*La siesta es uno de los mejores inventos. Recargas las pilas, te da otro aire. Afrontas la mitad del día un poquito más descansada. En fin, es salud física y mental.*	The nap is one of the best inventions. You recharge your batteries, it gives you a second wind. You face half the day a little more rested. In short, it's physical and mental health.
MM	*¿Qué opinas de la tradición de la siesta? Si se tiene tiempo, es una tradición buenísima.*	What do you think of the nap tradition? If you have time, it's a great tradition.
MW	*La tradición de la siesta me parece necesaria porque hace mucho calor después de comer, especialmente en verano. Y creo que sienta bastante bien.*	The nap tradition seems necessary to me because it's very hot after eating, especially in the summer. And I think it feels pretty good.
SM	*¿Qué opinas de la tradición de la siesta? El mejor invento de la Península Ibérica para la humanidad. Sin siesta no se puede vivir. La siesta es el secreto del equilibrio humano.*	What do you think of the siesta tradition? The best invention of the Iberian Peninsula for humanity. You can't live without naps. The nap is the secret of human balance.
SW	*Pues mira, la tradición de la siesta eso no es un lujo, eso es una necesidad, sobre todo en verano. Así que la tradición de la siesta es una cosa que no se debería quitar.*	Well look, the siesta tradition, that's not a luxury, that's a necessity, especially in the summer. So the nap tradition is one thing that shouldn't be taken away.

Appendix C.2 Long-Term Form Priming Words

Appendix C.2.1 Critical /s/ Word Stimuli

Word	Translation	Log Count	#Phon Neighbors
sarro	plaque (dental)	1.908	33
sorna	sarcasm	1.944	12
sardo	Sardinian	2.076	19
surco	groove	2.155	10
seto	hedge	2.446	40
salvia	sage	2.482	8
sigma	sigma (number)	2.521	5
salmo	psalm	2.562	23
soto	thicket	2.572	40
sarta	string	2.648	16
silba	whistle	2.725	19
silo	silo	2.826	39
saldo	balance	2.866	22
sodio	sodium	2.982	8
sauna	sauna	3.075	8
senda	track	3.112	13
sorbo	sip	3.112	16
simio	ape	3.158	13
soplo	breath	3.171	9
secta	sect	3.237	7
sable	saber	3.242	14
sabio	sage; intellect	3.251	16
sapo	toad	3.319	34
suegra	mother-in-law	3.331	2
sonda	probe	3.473	16
suero	serum	3.48	13
soda	soda water	3.534	32
sexto	sixth	3.635	27
soga	rope	3.643	25

signo	sign	3.669	12
suma	sum	3.799	35
selva	jungle	3.817	13
sueldo	salary	3.883	5
salto	jump; skip	3.996	28
salsa	sauce	4.016	17
sombra	shadow	4.132	10
sopa	soup	4.146	33
saco	sack; bag	4.149	43
siglo	century	4.201	12
serie	series	4.286	13
santo	saint	4.574	21
sala	room	4.631	65
suelo	floor	4.642	14
sexo	sex	4.761	7
sueño	dream	4.796	9
sitio	place	4.869	13
sangre	blood	5.064	6
suerte	luck	5.176	6
Average freq:	3.481104167		

Appendix C.2.2 Real Control Words

Word	Translation	Log Count	# Phon Neighbors
cacto	cactus	1.929	19
diabla	devil	1.944	2
tilde	accent mark	2.072	8
miga	crumb	2.271	38
leño	log	2.408	18
trufa	truffle	2.477	5
gamba	prawn	2.484	14
pana	corduroy	2.539	50

cuña	crib	2.545	17
bucle	loop	2.629	6
fierro	iron	2.715	6
cuenco	bowl	2.8	3
helio	helium	2.899	22
viudo	widow	2.944	2
lomo	loin	3.072	33
caldo	broth	3.12	22
peine	comb	3.165	11
hongo	mushroom	3.171	8
ancla	anchor	3.181	10
loro	parrot	3.277	34
himno	hymn	3.309	4
finca	farm	3.317	12
dato	fact	3.33	32
menta	mint	3.415	22
nieta	granddaughter	3.471	9
gripe	flu	3.579	9
mito	myth	3.615	58
diente	tooth	3.67	11
pila	battery	3.67	47
nube	cloud	3.738	12
feria	fair	3.742	9
burro	donkey	3.77	24
tarta	cake	3.837	15
palo	stick	3.965	65
templo	temple	4.071	8
moda	fashion	4.101	34
ritmo	rythym	4.141	7
plata	silver	4.19	13
barrio	neighborhood	4.22	12
polvo	dust	4.292	6
puente	bridge	4.444	8
luna	moon	4.592	30
norte	north	4.593	14
tienda	store	4.669	7
grupo	group	4.764	8
libro	book	4.87	16
cara	face	5.096	79

agua	water	5.123	8
tromba	whirlwind	1.914	4
naipe	card	2.097	2
grapa	staple	2.294	20
orca	killer whale	2.396	25
mote	nickname	2.534	47
tecla	keyboard key	2.751	8
podio	podium	2.834	13
cono	cone	3.054	46
aula	classroom	3.209	35
piña	pineapple	3.229	27
bota	boot	3.491	50
ruina	ruin	3.578	12
falda	skirt	3.643	9
beca	scholarship	3.673	37
nido	nest	3.744	26
horno	oven	3.772	27
barba	beard	3.884	28
fiebre	fever	4.077	4
venta	sale	4.127	22
lengua	language	4.308	4
banda	band	4.432	21
foto	picture	4.689	41
fuego	fire	4.923	8
tierra	land	5.074	6

Appendix C.2.3 Pseudo Control Words

Pseudo	Real Word	Translation	Log Count
conlo	*calco*	tracing	1.875
galmo	*galgo*	greyhound	2.179
lirra	*birra*	beer	2.33
fialo	*fauno*	faun	2.34
guna	*duna*	dune	2.441
mada	*cata*	tasting	2.476

truña	*trufa*	truffle	2.477
fueto	*dueto*	duet	2.534
ririo	*lirio*	lily	2.706
vermo	*verbo*	verb	2.812
linte	*tinte*	dye	2.827
meño	*moño*	bun	2.851
bloro	*cloro*	chlorine	2.888
bimo	*lino*	linen	2.964
darco	*dardo*	dart	3.1
rego	*remo*	oar	3.215
plauta	*flauta*	flute	3.268
fuende	*duende*	elf	3.279
marno	*mango*	mango	3.294
dibra	*fibra*	fiber	3.34
truelo	*trueno*	thunder	3.385
nuzo	*nudo*	knot	3.415
plemo	*plomo*	lead	3.547
gaca	*hada*	fairy	3.592
milta	*multa*	fine	3.66
huedo	*hueco*	hole	3.692
blobo	*globo*	globe	3.722
cagra	*cabra*	goat	3.78
niedra	*niebla*	fog	3.804
glama	*drama*	drama	3.811
dena	*vela*	candle	3.835
mirte	*monte*	mountain	3.907
ruema	*rueda*	wheel	3.937
flima	*clima*	climate	3.947
panio	*patio*	patio	4.038
pima	*mina*	mine	4.053
raca	*pata*	paw	4.061
meblo	*metro*	metro	4.067
dumo	*humo*	smoke	4.127
clape	*clave*	key	4.173
rorre	*torre*	tower	4.207
fiario	*diario*	diary	4.326
piegra	*piedra*	stone	4.347
obla	*obra*	work	4.509

iare	*aire*	air	4.765
reba	*ropa*	clothing	4.822
puebro	*pueblo*	town	4.919
vuerra	*guerra*	war	5.112
bestro	*bistro*	bistro	2.107
dedro	*cedro*	cedar tree	2.396
fucto	*ducto*	pipeline	2.442
barlo	*bardo*	bard	2.528
belmo	*bombo*	drum	2.594
afre	*acre*	acre	2.689
dorro	*forro*	cover	2.761
euna	*aura*	aura	3.05
crote	*brote*	outbreak	3.111
flera	*flora*	vegetation	3.249
firo	*faro*	lighthouse	3.439
trimo	*trigo*	wheat	3.451
papria	*patria*	homeland	3.746
crito	*grito*	shout	3.748
binta	*quinta*	country house	3.781
cuepa	*cueva*	cave	3.886
rigre	*tigre*	tiger	3.897
houda	*deuda*	debt	4.043
rura	*ruta*	route	4.126
retla	*regla*	rule	4.185
malpa	*mapa*	map	4.28
fample	*hambre*	hunger	4.712
nepia	*novia*	girlfriend	4.756
merro	*perro*	dog	4.86

Appendix C.2.4 Real Filler Words

Word	Translation	Log Count	#Phon Neighbors
yuca	yuca	1.903	14
churros	churros	2.25	9
mimbre	wicker	2.27	6
equis	x	2.276	8
facha	look; figure	2.389	24
plagio	plagiarism	2.467	6
yema	yoke	2.509	23
rizo	curl	2.538	32
dorso	back	2.622	6
gruta	grotto	2.688	12
guiño	wink	2.729	22
fósil	fossil	2.788	3
ocio	leisure	2.813	19
bache	pothole	2.841	40
ámbar	amber	2.874	8
lienzo	canvas	2.922	2
jerga	slang	3.015	7
prendas	garments	3.069	8
chisme	gossip	3.115	5
mangas	sleeves	3.167	11
plancha	iron	3.171	12
rabo	tail	3.206	37
mármol	marble	3.208	0
fresa	strawberry	3.257	13
sobras	leftovers	3.306	12
jarra	jar	3.335	32
olla	pot	3.372	40
cancha	court	3.398	18
ajo	garlic	3.441	51
frutas	fruits	3.546	8
pausa	pause	3.568	15
bicho	beatle	3.649	25
zorro	fox	3.675	25
siesta	nap	3.681	5
sierra	mountain range	3.708	9

raya	stripe	3.713	38
grasa	fat	3.758	17
queja	complaint	3.759	27
trenes	trains	3.772	11
vidrio	glass	3.775	5
copas	cups	3.786	29
huella	fingerprint	3.796	8
mancha	stain	3.799	25
postre	dessert	3.811	10
muebles	furniture	3.844	5
letras	letters	3.856	7
hoja	leaf	3.886	36
risa	laugh	3.919	34
frase	phrase	3.965	13
barcos	boats	3.973	14
cuadro	painting	3.981	6
joyas	jewels	4	7
plaza	plaza	4	13
túnel	tunnel	4.009	6
seres	beings	4.115	23
fecha	date	4.15	17
gira	tour	4.152	39
notas	grades	4.191	26
granja	farm	4.193	4
queso	cheese	4.229	32
lluvia	rain	4.261	6
ángel	angel	4.298	7
costa	coast	4.315	28
silla	chair	4.386	26
ruido	noise	4.412	10
brazos	arms	4.413	8
hielo	ice	4.422	33
trampa	trick	4.427	6
flores	flowers	4.46	14
fondo	bottom	4.52	19
vuelo	flight	4.524	15
reina	queen	4.571	22
cárcel	jail	4.631	8
alma	soul	4.651	34

mesa	table	4.709	43
traje	suit	4.72	20
fuerza	force	4.788	7
trato	treaty	4.829	21
cielo	sky	4.866	6
falta	lack	4.886	18
hijos	kids	4.986	14
horas	hours	5.125	54
miedo	fear	5.141	3
clase	class	5.15	17
ubres	utters	2.152	5
caqui	khacki	2.34	26
mella	dent	2.59	31
cóndor	condor	2.713	4
ostra	oyster	2.789	16
franja	fringe	2.846	4
rasgo	trait	2.874	7
calcio	calcium	2.978	2
pera	pear	3.043	54
tapas	tapas	3.089	35
soja	soy	3.13	27
mugre	dirt	3.19	5
charco	puddle	3.227	19
cejas	eyebrows	3.354	24
ciervo	deer	3.38	6
zumo	juice	3.468	16
monja	monk	3.5	6
frenos	brakes	3.531	6
talla	size	3.692	43
sello	stamp	3.715	34
susto	scare	3.763	14
pesca	fishing	3.798	17
hombros	shoulders	3.809	2
brindis	cheers	3.827	4
medias	tights	3.881	10
lujo	luxury	3.884	18
jefa	boss	3.96	14
móvil	cell phone	4.043	6
bruja	witch	4.153	9

lago	lake	4.239	44
techo	roof	4.277	26
prensa	press	4.34	11
leche	milk	4.391	31
beso	kiss	4.428	45
precio	price	4.508	14
ganas	desire	4.517	35
isla	island	4.525	8
zona	zone	4.56	21
radio	radio	4.568	14
centro	center	4.682	8
pelo	hair	4.718	52
coche	car	5.002	40
ojos	eyes	5.117	15
muerte	death	5.137	8

Appendix C.2.5 Sample of Pseudoword Fillers

Pseudo	Word	Translation	Log Count
plimas	*plisas*	breezes	2.297
carfe	*canje*	exchange	2.358
ajia	*apio*	celery	2.573
euge	*auge*	boom	2.622
fanzas	*danzas*	dances	2.646
gesco	*vasco*	Basque	2.665
hichas	*hachas*	axes	2.773
cengo	*censo*	census	2.78
corjus	*corpus*	corpus	2.798
lugro	*lucro*	profit	2.809
duerta	*huerta*	orchard	2.866
dallos	*gallos*	rooster	2.87
dactros	*filtros*	filters	2.873
trempa	*trompa*	horn	2.875
samla	*samba*	samba	2.893
banfo	*banjo*	banjo	2.947
péster	*póster*	poster	2.955

Appendix C.3 Salient Variants Avoided in Control and Critical /s/ Stimuli

Table adapted from Fernández de Molina Ortés and Hernández-Campoy (2018), p. 502

Context	Standard Castilian		Conservative dialects (north-central)		Innovative dialects (south)		Gloss
Coda /s/	*las casas*	[-s]	[lasˈkasas]	[-s]	[læʰˈkæsæ]/[læˈkæsæ]		the houses
Velar production	*gente*	[x-]	[ˈxéɳte]	[x-]	[ˈʰeɳte]	[h-]	people
Coda /r/	*coger*	[-r]	[koxer]	[-r]	[koˈʰɛ]	Ø	to take
Coda /l/	*piel*	[-l]	[ˈpjel]	[-l]	[ˈpjɛ]	Ø	skin
Coda /θ/	*luz*	[-θ]	[ˈluθ]	[-θ]	[ˈlu]	Ø	light
Intervocalic /-l-/	*soldado*	[-l-]	[solˈðaðo]	[-l-]	[solˈðaðo]/[sorˈðaðo]	[-l-]~[-r-]	soldier
Intervocalic /-s-/	*casa*	[-s-]	[ˈkasa]	[-s-]	[ˈkasa]–[ˈkaθa]	[-s-]~[-θ-]	house
Intervocalic /-θ-/	*caza*	[-θ-]	[ˈkaθa]	[-θ-]	[ˈkaθa]–[ˈkasa]	[-θ-]~[-s-]	hunting
Affricate production	*cacho*	[ʧ]	[ˈkaʧo]	[-ʧ-]	[ˈkaʃo]	[-ʃ-]	bit
Voiced palatal fricative	*rayo*	[ʝ]	[ˈraʝo]	[-ʝ-]	[ˈraʒo]	[-ʒ-]	lightning
Voiced palatal lateral	*callo*	[ʎ]	[ˈkaʎo]~[ˈkaʝo]	[ʎ]~[ʝ]	[ˈkaʒo]	[-ʒ-]	corn

Appendix D Unaltered Participant Commentary from Verbal Guise

Appendix D.1 Feedback Provided for SM

Participant comments: SM	
small dejes del sur pero voz muy neutra	Small southern touches but a very neutral voice
Parece un sevillano tratando de hablar "fino"	It seems like a Sevillian trying to speak in a refined way
Creo que es Sevillano, con muy poco acento, o acento neutro forzado.	I think that he's Sevillian, with not much of an accent or a forced neutral accent
El tono de voz me resulta muy familiar	I find the tone of voice to be very familiar
Parece de Sevilla. La última vez que pronuncia "siesta" aspira la segunda s	[SM] seems to be from Seville. The last time that he pronounces siesta he aspirates the second 's'
No siempre pronuncia como andaluz	He doesn't always pronounce in an Andalusian way

Appendix D.2 Feedback Provided for SW

Participant comments: SW	
La identifico con una ciudadana totalmente natal de Sevilla	I identify her as a completely native Seville citizen
Se le nota el acento pero no cecea ni se come muchas letras al hablar	The accent is noticeable but [SW] doesn't have *ceceo* and doesn't aspirate many letters when speaking

Voz nada forzada	A voice that's not at all forced
Cambia el uso de las "s" a final de las palabra dependiendo de la palabra	[SW] changes the use of the word-final s's depending on the word
típico acento sevillano	typical Sevillian accent
cálida	warm
muy natural	very natural
Se trata de una ciudadana plenamente sevillana	She's completely a Sevillian citizen
en este caso, el elemento mas determinante para valorar el origen de esta persona es la pronunciacion de la s de siesta como h inhalada	in this case, the most determining factor to assess the person's origin is the pronunciation of the 's' in siesta with aspiration
Acento sevillano por el deje en la pronunciación	A Sevillian accent because of the pronunciation
Al final del audio se nota más el acento andaluz.	The Andalusian accent is more noticeable at the end of the recording

Appendix D.3 Feedback Provided for MM

Participant comments: MM	
Del norte de España?	From northern Spain?
hombre basic	Basic/simple man
Parece ser del centro, incluso norte de España	[MM] seems to be from central or even northern Spain
Esta voz me recuerda a una zona procedente del norte	This voice reminds me of a northern zone
Serio, aunque coloquial. Parece que no es sevillano	Serious, but also colloquial. It seems like [MM] is not a Sevillian
simple	simple/plain

Su voz es parecida a la de Chicho Ibañez Serrador, director español. Es educado, serio y coloquial	His voice is similar to that of Chicho Ibañez Serrador, a Spanish director. He is polite, serious, and colloquial
Pronuncia bien los finales de frase en plrual, pero no las terminaciones de los verbos en participio.	[MM] pronounces all of the ends of plural phrases, but not the ends of verbal participles.

Appendix D.4 Feedback Provided for MW

Participant comments: MW	
Parece ser que este hablante no es de Sevilla capital, ni de la provincia. Parece ser de fuera de Andalucía	It seems like this speaker is neither from Seville capital, nor the province. They seem to be from outside of Andalusia
mi principal razon para pensar que esta persona no es de Sevilla es la forma en la que pronuncia la G	my primary reason for thinking that this person is not from Seville is because of how the ‘g’ is pronounced
Voz bonita, muy pija. Los sevillanos pijos también hablan así, pero sin ese acento	A pretty voice, very elite. The snobby Sevillians also speak like this, but without that accent
Debido a la pronunciacion de las “s” y por las palabras que usa da la impresión de ser mas "pija" que el resto	Due to the pronunciation of the s’s and because of the words that [MW] uses give the impression of being more snobby than the rest
Puede que sea alguien de Sevilla capital intentando parecer del centro	It could be someone from Seville capital trying to seem like they’re from central Spain
Chica del norte del país	A girl from northern Spain
repipi y repelente, muy fisna ella	pretentious and repulsive, she’s really ‘refined’ (sarcasm)

Voz monótona, parece de fuera de Sevilla, pero ya he dicho que hay muchos sevillanos que esconden su acento y en una respuesta corta podría ser de cualquier sitio	A monotonous voice, it seems to be from outside of Seville, but I've already said that there are many Sevillians that hide their accent and in a short response they could be from anywhere

Appendix D.5 Feedback Provided for AM

Participant comments: AM	
Este chico cecea también , quizás es de un pueblo o de la zona de Málaga	This guy also has *ceceo*, maybe he's from a small town or from Málaga
La identifico como my de la tierra	I identify it/him as being very rural
Tiene una voz caracteristica de las afueras de Sevilla, como si de un pueblo se tratase.	[AM] has a voice that's characteristic to the outskirts of Seville, as if he were from a small town.
Parece de un pueblo. Se come muchas letras	He seems to be from a small town. A lot of letters are not pronounced
esta voz me resulta muy de los pueblos de sevilla	This voice is very typical of the outskirts of Seville
acento andaluz	Andalusian accent
pueblo profundo	From the deep countryside
Se nota bastante que es de Sevilla capital, por su manera de hablar y expresarse.	You can really tell that [AM] is from Seville capital because of his way of talking and expressing himself.
Parece que tiene un acento más cerrado que el anterior, por eso me inclino a pensar que no es de la capital, a pesar de que en la capital hay gran confluencia de acentos por la diversidad de personas que viven en ella y la mezcolanza de sus hablas.	It seems like [AM] has a stronger accent than the previous speaker, which makes me think that [AM] is not from Seville capital, even though in Seville capital there's a large confluence of accents because of the diversity of the people that live there and the mixture of dialects.

Appendix D.6 Feedback Provided for AW

Participant comments: AW	
Esta señora parece de un pueblo de Sevilla, tal vez Utrera, cecea	This woman seems to be from a small town outside of Seville, maybe Utrera. [AW] has *ceceo*
simpatica	kind
Posee un acento que se identifica plenamente con las afueras de Sevilla.	[AW] possess an accent that's completely identified with the outskirts of Seville.
Como la anterior, el ceceo la hace parecer de fuera de Sevilla capital	Like the last talker, the ceceo makes her seem from outside of Seville capital
muy voz de personas mayores de Sevilla, sobre todo senoras	a voice similar to elderly people from Seville, specifically women
La chica de Utrera de nuevo	The lady from Utrera again
Tiene una voz bastante alegre, parece ser de Sevila capital aunque tiene un poco de ceceo	[AW] has a rather happy voice, she seems to be from Seville capital even though she has a little bit of *ceceo*
Parece de fuera de la capital por el ceceo	[AW] seems like she's from outside of the capital because of the *ceceo*
Su acento es igual que el de una amiga, que es de un pueblo de Sevilla, de El Coronil. *Aunque la "j" es aspirada, se produce el "ceceo" propio de la provincia de Sevilla.*	Her accent is the same as that of a friend of mine, who is from a small town outside of Seville, from El Coronil. Even though the "j" is aspirated, the typical "*ceceo*" from the Seville province is produced.
El ceceo da normalmente un caracter mas "cotidiano" a las cosas que se dicen	*Ceceo* normally gives an "everyday" feel to the things that are said

Appendix E Accuracy and Reaction Time Models from Long-Term Form Priming Task

Appendix E.1 Accuracy for /s/ Words and Controls Across the Experiment

```
Call:
glm(formula = correct ~ Variety * ConditionCode, family = "binomial",
    data = Tarea2_Error2)

Deviance Residuals:
    Min       1Q   Median       3Q      Max
-2.2293   0.4172   0.4437   0.4830   0.6669

Coefficients:
                             Estimate Std. Error z value Pr(>|z|)
(Intercept)                   2.26868    0.09358  24.243  < 2e-16 ***
VarietyAfueras                0.01764    0.13282   0.133    0.894
VarietyMadrid                -0.17916    0.12787  -1.401    0.161
ConditionCodes                0.12921    0.13601   0.950    0.342
VarietyAfueras:ConditionCodes -1.02552    0.17900  -5.729 1.01e-08 ***
VarietyMadrid:ConditionCodes  -0.05898    0.18470  -0.319    0.749
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

(Dispersion parameter for binomial family taken to be 1)

    Null deviance: 5712.8  on 8063  degrees of freedom
Residual deviance: 5603.4  on 8058  degrees of freedom
AIC: 5615.4

Number of Fisher Scoring iterations: 5
```

Appendix E.2 Block 1 Real Control Model

```
Linear mixed model fit by REML. t-tests use Satterthwaite's method [
lmerModLmerTest]
Formula: log(Block1_rt_adj) ~ (1 | Target) + (1 | subject_nr) + Variety *
    Gender
   Data: Tarea2_Block1_testmerge_rc

REML criterion at convergence: 2323.3

Scaled residuals:
    Min      1Q  Median      3Q     Max
-4.2018 -0.6101 -0.0924  0.5368  3.8077

Random effects:
 Groups     Name        Variance Std.Dev.
 Target     (Intercept) 0.02231  0.1494
 subject_nr (Intercept) 0.05188  0.2278
 Residual               0.19002  0.4359
Number of obs: 1792, groups:  Target, 48; subject_nr, 41

Fixed effects:
                   Estimate Std. Error         df t value Pr(>|t|)
(Intercept)       6.313e+00  4.288e-02  6.210e+01 147.233   <2e-16 ***
Variety1          6.008e-03  1.269e-02  1.705e+03   0.473   0.6361
Variety2         -1.528e-02  7.289e-03  1.702e+03  -2.097   0.0362 *
Gender1          -6.246e-02  4.786e-02  4.415e+01  -1.305   0.1986
Variety1:Gender1 -4.194e-02  2.539e-02  1.705e+03  -1.652   0.0988 .
Variety2:Gender1 -3.569e-02  1.458e-02  1.702e+03  -2.448   0.0145 *
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Correlation of Fixed Effects:
            (Intr) Varty1 Varty2 Gendr1 Vr1:G1
Variety1     0.001
Variety2    -0.002 -0.001
Gender1      0.000 -0.008 -0.012
Vrty1:Gndr1 -0.005  0.010  0.013  0.001
Vrty2:Gndr1 -0.007  0.013 -0.023 -0.003 -0.001
```

Appendix E.3 Block 1 /s/ Model

```
Linear mixed model fit by REML. t-tests use Satterthwaite's method [
lmerModLmerTest]
Formula: log(Block1_rt_adj) ~ (1 | Target) + (1 | subject_nr) + Variety *
    Gender
   Data: Tarea2_Block1_testmerge_onlys

REML criterion at convergence: 2090.1

Scaled residuals:
    Min      1Q  Median      3Q     Max
-4.1011 -0.6174 -0.1034  0.5592  4.2425

Random effects:
 Groups     Name        Variance Std.Dev.
 Target     (Intercept) 0.01951  0.1397
 subject_nr (Intercept) 0.05588  0.2364
 Residual               0.18048  0.4248
Number of obs: 1675, groups:  Target, 48; subject_nr, 41

Fixed effects:
                   Estimate Std. Error         df t value Pr(>|t|)
(Intercept)       6.373e+00  4.342e-02  5.582e+01 146.778   <2e-16 ***
Variety1         -4.904e-02  4.740e-02  3.819e+01  -1.034    0.307
Variety2         -3.566e-02  2.696e-02  3.780e+01  -1.323    0.194
Gender1           5.364e-02  4.561e-02  4.303e+01   1.176    0.246
Variety1:Gender1 -6.879e-03  2.633e-02  1.589e+03  -0.261    0.794
Variety2:Gender1 -1.815e-02  1.453e-02  1.588e+03  -1.249    0.212
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Correlation of Fixed Effects:
            (Intr) Varty1 Varty2 Gendr1 Vr1:G1
Variety1    -0.031
Variety2    -0.019  0.025
Gender1      0.011 -0.007 -0.006
Vrty1:Gndr1 -0.014  0.027  0.011 -0.042
Vrty2:Gndr1 -0.011  0.011  0.011 -0.032  0.066
```

Appendix E.4 Block 1 Real Control and /s/ Word Combined Model

```
log(Block1_rt_adj) ~ (1 | Target) + (1 | subject_nr) + ConditionCode *
    Variety * Gender
   Data: Tarea2_Block1_testmerge_rcs

REML criterion at convergence: 4355.2

Scaled residuals:
    Min      1Q  Median      3Q     Max
-4.2986 -0.6285 -0.0928  0.5672  4.2520

Random effects:
 Groups     Name        Variance Std.Dev.
 Target     (Intercept) 0.02073  0.1440
 subject_nr (Intercept) 0.05154  0.2270
 Residual               0.18630  0.4316
Number of obs: 3467, groups:  Target, 96; subject_nr, 41

Fixed effects:
                                 Estimate Std. Error         df t value
(Intercept)                     6.312e+00  4.236e-02  6.947e+01 149.003
ConditionCodes                  6.072e-02  3.297e-02  8.752e+01   1.842
Variety1                        6.268e-03  1.256e-02  3.331e+03   0.499
Variety2                       -1.527e-02  7.214e-03  3.326e+03  -2.117
Gender1                        -6.249e-02  4.637e-02  8.580e+01  -1.348
ConditionCodes:Variety1        -7.294e-02  2.202e-02  3.328e+03  -3.313
ConditionCodes:Variety2        -8.975e-03  1.242e-02  3.320e+03  -0.723
ConditionCodes:Gender1          1.167e-01  6.592e-02  8.747e+01   1.770
Variety1:Gender1               -4.053e-02  2.513e-02  3.331e+03  -1.613
Variety2:Gender1               -3.567e-02  1.443e-02  3.326e+03  -2.472
ConditionCodes:Variety1:Gender1  2.889e-02  3.668e-02  3.331e+03   0.788
ConditionCodes:Variety2:Gender1  1.720e-02  2.063e-02  3.328e+03   0.833
                               Pr(>|t|)
(Intercept)                     < 2e-16 ***
ConditionCodes                 0.068862 .
Variety1                       0.617930
Variety2                       0.034348 *
Gender1                        0.181329
ConditionCodes:Variety1        0.000934 ***
ConditionCodes:Variety2        0.469832
ConditionCodes:Gender1         0.080223 .
Variety1:Gender1               0.106863
Variety2:Gender1               0.013503 *
ConditionCodes:Variety1:Gender1 0.431014
ConditionCodes:Variety2:Gender1 0.404705
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1
```

Appendix E.5 Block 2 Repeated Control Model

```
Linear mixed model fit by REML. t-tests use Satterthwaite's method [
lmerModLmerTest]
Formula:
log(Block2_rt_adj) ~ (1 | Target) + (1 | subject_nr) + Variety.y *
    Gender.y
   Data: Tarea2_Block2_rc

REML criterion at convergence: 1170.1

Scaled residuals:
    Min      1Q  Median      3Q     Max
-3.7830 -0.6110 -0.0752  0.5556  3.3428

Random effects:
 Groups     Name        Variance Std.Dev.
 subject_nr (Intercept) 0.08984  0.2997
 Target     (Intercept) 0.01365  0.1168
 Residual               0.21233  0.4608
Number of obs: 798, groups:  subject_nr, 41; Target, 24

Fixed effects:
                     Estimate Std. Error        df t value Pr(>|t|)
(Intercept)           6.32671    0.05514  49.44514 114.733   <2e-16 ***
Variety.y1            0.01108    0.02025 739.00484   0.547   0.5845
Variety.y2           -0.02186    0.01164 736.71461  -1.879   0.0606 .
Gender.y1            -0.05006    0.05818  19.49023  -0.860   0.4000
Variety.y1:Gender.y1 -0.06862    0.04041 738.27306  -1.698   0.0899 .
Variety.y2:Gender.y1 -0.05072    0.02332 737.61634  -2.175   0.0299 *
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Correlation of Fixed Effects:
            (Intr) Vrty.1 Vrty.2 Gndr.1 V.1:G.
Variety.y1   0.005
Variety.y2   0.000 -0.008
Gender.y1    0.006 -0.013 -0.031
Vrty.y1:G.1 -0.008  0.058  0.016  0.009
Vrty.y2:G.1 -0.016  0.023 -0.028  0.000 -0.003
```

Appendix E.6 Block 2 /s/ Word Model

```
Linear mixed model fit by REML. t-tests use Satterthwaite's method [
lmerModLmerTest]
Formula:
log(Block2_rt_adj) ~ (1 | Target) + (1 | subject_nr) + Variety.x *
    Variety.y
   Data: Block2_s_redo

REML criterion at convergence: 2137.2

Scaled residuals:
    Min      1Q  Median      3Q     Max
-5.7193 -0.6182 -0.0784  0.5313  3.6758

Random effects:
 Groups     Name        Variance Std.Dev.
 Target     (Intercept) 0.01218  0.1104
 subject_nr (Intercept) 0.05198  0.2280
 Residual               0.20901  0.4572
Number of obs: 1529, groups:  Target, 48; subject_nr, 41

Fixed effects:
                       Estimate Std. Error         df t value Pr(>|t|)
(Intercept)           6.334e+00  4.084e-02  4.974e+01 155.102  < 2e-16 ***
Variety.x1            2.485e-03  4.631e-02  3.809e+01   0.054    0.957
Variety.x2           -1.540e-02  2.637e-02  3.789e+01  -0.584    0.563
Variety.y1           -8.161e-02  1.626e-02  1.020e+03  -5.020 6.11e-07 ***
Variety.y2           -4.084e-02  9.160e-03  1.075e+03  -4.459 9.12e-06 ***
Variety.x1:Variety.y1 -7.242e-03  2.009e-02  9.422e+02  -0.360    0.719
Variety.x2:Variety.y1 -6.203e-03  1.234e-02  4.777e+02  -0.503    0.615
Variety.x1:Variety.y2  3.348e-03  1.194e-02  7.127e+02   0.280    0.779
Variety.x2:Variety.y2 -7.496e-03  6.887e-03  4.880e+02  -1.088    0.277
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Correlation of Fixed Effects:
            (Intr) Vrty.x1 Vrty.x2 Vrty.y1 Vrty.y2 V.1:V.1 V.2:V.1 V.1:V.2
Variety.x1  -0.031
Variety.x2  -0.018  0.024
Variety.y1  -0.010  0.000  -0.011
Variety.y2  -0.006 -0.008  -0.011   0.115
Vrty.x1:V.1 -0.002 -0.006  -0.003  -0.036   0.008
Vrty.x2:V.1 -0.007  0.003  -0.019  -0.043   0.077  -0.022
Vrty.x1:V.2 -0.007  0.000   0.008   0.004  -0.022   0.043  -0.162
Vrty.x2:V.2 -0.005  0.003  -0.013   0.039  -0.061   0.160  -0.105   0.143
```

Appendix E.7 Block 2 Combined Real Control and /s/ Model

```
Formula: log(Block2_rt_adj) ~ Repetition * ConditionCode * Variety + (1 |
    Target) + (1 | subject_nr)
   Data: Block2_all_analysis

REML criterion at convergence: 4817.2

Scaled residuals:
    Min      1Q  Median      3Q     Max
-5.3638 -0.6388 -0.0840  0.5501  3.7648

Random effects:
 Groups     Name        Variance Std.Dev.
 Target     (Intercept) 0.01497  0.1224
 subject_nr (Intercept) 0.06507  0.2551
 Residual               0.21754  0.4664
Number of obs: 3444, groups:  Target, 96; subject_nr, 41

Fixed effects:
                           Estimate Std. Error         df t value Pr(>|t|)
(Intercept)              6.355e+00  4.503e-02  5.849e+01 141.131  < 2e-16
Repetition1             -4.640e-02  4.195e-02  8.904e+01  -1.106 0.271692
ConditionCodes           1.381e-02  3.636e-02  8.920e+01   0.380 0.705023
Variety1                -2.389e-03  1.389e-02  3.325e+03  -0.172 0.863437
Variety2                -2.698e-02  7.940e-03  3.320e+03  -3.398 0.000686
Repetition1:Variety1     2.964e-02  2.778e-02  3.325e+03   1.067 0.286154
Repetition1:Variety2     1.285e-02  1.588e-02  3.320e+03   0.809 0.418444
ConditionCodes:Variety1 -9.426e-02  2.518e-02  3.314e+03  -3.743 0.000185
ConditionCodes:Variety2 -1.700e-02  1.429e-02  3.360e+03  -1.190 0.234294

(Intercept)             ***
Repetition1
ConditionCodes
Variety1
Variety2                ***
Repetition1:Variety1
Repetition1:Variety2
ConditionCodes:Variety1 ***
ConditionCodes:Variety2
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1
```

Bibliography

Agostini, T. D., & Schwenter, S. A. (2018). Variable negative concord in Brazilian Portuguese. In J. MacDonald (Ed.), *Contemporary Trends in Hispanic and Lusophone Linguistics: Selected papers from the Hispanic Linguistics Symposium 2015* (pp. 71–94). John Benjamins. https://doi.org/10.1075/IHLL.15.05AGO

Andruski, J. E., Blumstein, S. E., & Burton, M. (1994). The effect of subphonetic differences on lexical access. *Cognition, 52*, 163–187. https://doi.org/10.1016/0010-0277(94)90042-6

Barnes, S. (2015). Perceptual salience and social categorization of contact features in Asturian Spanish. *Studies in Hispanic and Lusophone Linguistics 8*(2), 213–241. https://doi.org/10.1515/shll-2015-0009

Bates, D., Mächler, M., Bolker, B., & Walker, S. (2015). Fitting linear mixed-effects models using lme4. *Journal of Statistical Software, 67*(1). https://doi.org/10.18637/JSS.V067.I01

Boersma, P. & Weenink, D. (2022). Praat: doing phonetics by computer [Computer program]. Version 6.2.11, retrieved 10 April 2022 from http://www.praat.org/

Bradlow, A. R., Nygaard, L. C., & Pisoni, D. B. (1999). Effects of talker, rate, and amplitude variation on recognition memory for spoken words. *Perception & Psychophysics, 61*(2), 206–219. https://doi.org/10.3758/bf03206883

Campbell-Kibler, K. (2007). Accent, (ing), and the social logic of listener perceptions. *American Speech 82*(1), 32–64. https://doi.org/10.1215/00031283-2007-002

Cano, R. (2015, April). Tesis: Habla Andaluza. *El habla andaluza.* [Video]. YouTube. https://www.youtube.com/watch?v=3ywNlhf0gtE

Chappell, W. (2016a). On the social perception of intervocalic /s/ voicing in Costa Rican Spanish. *Language Variation and Change, 28*, 557–378. https://doi.org/10.1017/S0954394516000107

Chappell, W. (2019). Caribeño or mexicano, profesionista or albañil? Mexican listeners' evaluations of /s/ aspiration and maintenance in Mexican and Puerto Rican voices. *Sociolinguistic Studies, 12*(3-4), 367–393. https://doi.org/10.1558/sols.34154

Chappell, W., & Kanwit, M. (2022). Do learners connect sociophonetic variation with regional and social characteristics? The case of L2 perception of Spanish aspiration. *Studies in Second Language Acquisition, 44*(1), 185–209. https://doi.org/10.1017/S0272263121000115

Church, B. A., & Schacter, D. L. (1994). Perceptual specificity of auditory priming: Implicit memory for voice intonation and fundamental frequency. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *20*, 521–533. https://doi.org/10.1037//0278-7393.20.3.521

Clapp, W., Vaughn, C., & Sumner, M. (2023). The episodic encoding of talker voice attributes across diverse talkers. *Journal of Memory and Language*, *128*, 1–36. https://doi.org/10.1016/j.jml.2022.104376

Clopper, C. G. (2017). Dialect interference in lexical processing: Effects of familiarity and social stereotypes. *Phonetica*, *74*, 25–59. https://doi.org/10.1159/000446809

Clopper, C. G. (2021). Perception of dialect variation. In J. S. Pardo, L. C. Nygaard, R. E. Remez, & D. B. Pisoni (Eds.), *The Handbook of Speech Perception* (pp. 335–364). Wiley-Blackwell. https://doi.org/10.1002/9781119184096.ch13

Clopper, C. G., & Pisoni, D. B. (2004). Some acoustic cues for the perceptual categorization of American English regional dialects. *Journal of Phonetics*, *32*(1), 111–140. https://doi.org/10.1016/S0095-4470(03)00009-3

Clopper, C. G., Tamati, T. N., & Pierrehumbert, J. B. (2016). Variation in the strength of lexical encoding across dialects. *Journal of Phonetics*, *58*, 87–103. https://doi.org/10.1016/j.wocn.2016.06.002

Clopper, C. G., & Walker, A. (2017). Effects of lexical competition and dialect exposure on phonological priming. *Language and Speech* *60*(1), 85–109. https://doi.org/10.1177/00238309166437

Colantoni, L. (2011) Laboratory approaches to sound variation and change. In M. Díaz-Campos (Ed.). *The Handbook of Hispanic Sociolinguistics* (pp. 9–28). Wiley-Blackwell**.** https://doi.org/10.1002/9781444393446.ch1

Cruz Ortiz, R. (2020). Seseo, ceceo y distinción de /s/ y /θ/: El caso de los políticos andaluces en Madrid. *Nueva Revista de Filología Hispánica*, *68*(1), 137–174. https://doi.org/10.24201/nrfh.v68i1.3585

Dossey, E., Jones, Z., & Clopper, C. G. (2023). Relative contributions of social, contextual, and lexical factors in speech processing. *Language and Speech, 66*(2), 322–353. https://doi.org/10.1177/00238309221107870

Dragojevic, M., & Goatley-Soan, S. (2022). The verbal-guise technique. In R. Kircher & L. Zipp (Eds.), *Research Methods in Language Attitudes* (pp. 203–218). Cambridge: Cambridge University Press. https://doi.org/10.1017/9781108867788.017

Duchon, A., Perea, M., Sabastián-Gallés, N., Martí, A., & Carreiras, M. (2013). EsPal: One-stop shopping for Spanish word properties. *Behavior Research Methods, 45*, 1246–1258. https://doi.org/10.3758/s13428-013-0326-1

Fernández de Molina Ortés, E., & Hernández-Campoy, J. (2018). Geographic varieties of Spanish. In K. Geeslin (Ed.), *The Cambridge Handbook of Spanish Linguistics* (pp. 496–528). Cambridge University Press. https://doi.org/10.1017/9781316779194.024

Floccia, C., Goslin, J., Girard, F., & Konopczynski, G. (2006). Does a regional accent perturb speech processing? *Journal of Experimental Psychology: Human Perception and Performance, 32*, 1276–1293. https://doi.org/10.1037/0096-1523.32.5.1276

Goldinger, S. D. (1996). Words and voices: Episodic traces in spoken word identification and recognition. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 22*, 1166–1183. https://doi.org/10.1037/0278-7393.22.5.1166

Goldinger, S. D. (1998). Echoes of echoes? An episodic theory of lexical access. *Psychological Review, 105*(2), 251–279. https://doi.org/10.1037/0033-295X.105.2.251

Gow, D. W. Jr. (2001). Assimilation and anticipation in continuous spoken word recognition. *Journal of Memory and Language, 45*, 133–159. https://doi.org/10.1006/jmla.2000.2764

Gylfadóttir, G. D. (2018). The effective borrowing of a phonemic contrast. University of Pennsylvania dissertation.

Hadodo, M. J., & Kanwit, M. (2020). Menos masculino, demasiado infantil: How age and gender ideologies index sexuality in diminutive perception in Madrid Spanish. *Journal of Language and Sexuality, 9*(2), 127–151. https://doi.org/10.1075/jls.19013.had

Hay, J. & Drager, K. (2010). Stuffed toys and speech perception. *Linguistics*, *28*, 692–865. https://doi.org/10.1515/ling.2010.027

Hay. J., Nolan, A., & Drager, K. (2006a). From fush to feesh: Exemplar priming in speech perception. *Ling Rev., 23*, 351–395. https://doi.org/10.1515/TLR.2006.014

Hernández-Campoy, J. M., & Villena-Ponsoda, J. A. (2009). Standardness and nonstandardness in Spain: Dialect attrition and revitalization of regional dialects of Spanish. *International Journal of the Sociology of Language, 2009*(196-197), 181–214. https://doi.org/10.1515/IJSL.2009.021

Hintzman, D. L. (1986). "Schema abstraction" in a multiple-trace memory model. *Psychological Review*, *93*(4), 411–428. https://doi.org/10.1037/0033-295X.93.4.411

Iglesias, R. (2003). Estudio sociolingüístico de la ciudad de Córdoba. Universidad de Las Palmas de Gran Canaria dissertation.

Jakobson, R., Fant, G., & Halle, M. (1951). Preliminaries to speech analysis: the distinctive features and their correlates. Cambridge, MA: MIT Press.

Johnson, K. (1997). Speech perception without speaker normalization: An exemplar model. In Johnson., K. & J. W. Mullennix (Eds.), *Talker Variability in Speech Processing* (pp. 145–165). Academic Press.

Johnstone, B., Andrus, J., & Danielson, A. E. (2006). Mobility, indexicality, and the enregisterment of "Pittsburghese." *Journal of English Linguistics*, *34*(2), 77–104. https://doi.org/10.1177/0075424206290692

Jones, Z., & Clopper, C. G. (2019). Subphonemic variation and lexical processing: Social and stylistic factors. *Phonetica*, *76*, 163–178. https://doi.org/10.1159/000493982

Keuleers, E., & Brysbaert, M. (2010). Wuggy: A multilingual pseudoword generator. *Behavior Research Methods 42*(3), 627–633. https://doi.org/10.3758/BRM.42.3.627

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

Labov, W. (1972). *Sociolinguistic patterns*. University of Pennsylvania Press.

Lambert, W. E., Hodgson, R. C., Gardner, R. C., & Fillenbaum, S. (1960). Evaluational reactions to spoken language. *Journal of Abnormal and Social Psychology, 60*(1), 44–51. https://doi.org/10.1037/h0044430.

León-Castro, M. (2016). La presencia del andaluz en los medios de comunicación. *In* Rosalba Mancinas-Chávez (Ed.), *Actas del congreso internacional comunicación y pensamiento* (pp. 1583–1600). Sevilla: Egregius.

López Velarde, M., & Simonet, M. (2019). Spoken word recognition and *shesheo* in Northwestern Mexico: A preliminary investigation into the effects of sociophonetic variability on auditory lexical access. In W. Chappell (Ed.), *Recent Advances in the Study of Spanish Sociophonetics* (pp. 266–285). Amsterdam: John Benjamins. https://doi.org/10.1075/ihll.21.10vel

Marian, V., Blumenfeld, H. K., & Kaushanskaya, M. (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. https://doi.org/10.1044/1092-4388(2007/067)

Markel, N. N., Eisler, R. M., & Reese, H. W. (1967). Judging personality from dialect. *Journal of Verbal Learning and Verbal Behavior, 6*(1), 33–35.

Mathôt, S., Schreij, D., & Theeuwes, J. (2012). OpenSesame: An open-source, graphical experiment builder for the social sciences. *Behavior Research Methods*, *44*(2), 314-324. https://doi.org/10.3758/s13428-011-0168-7

McClelland, J. L., & Elman, J. L. (1986). The TRACE model of speech perception. *Cognitive Psychology*, *18*(1), 1–86. https://doi.org/10.1016/0010-0285(86)90015-0

McLennan, C. T., Luce, P. A., & Charles-Luce, J. (2003). Representation of lexical form. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *29*, 539–553. https://doi.org/10.1037/0096-1523.31.6.1308

McQueen, J. M., & Cutler, A. (2010). Cognitive processes in speech perception. In W. J. Hardcastle, Laver, J., & Gibbon, F. E. (Eds.), *The Handbook of Phonetic Sciences* (pp. 489–520). Wiley Blackwell. https://doi.org/10.1002/9781444317251.ch14

Niedzielski, N. (1999). The effect of social information on the perception of sociolinguistic variables. *Journal of Language and Social Psychology, 18*, 62–85. https://doi.org/10.1177/0261927X990180010

Norris, D., McQueen, J. M., & Cutler, A. (2003). Perceptual learning in speech. *Cognitive Psychology, 18*, 1–86. https://doi.org/10.1016/S0010-0285(03)00006-9

Nosofsky, R. M. (1988). Exemplar-based accounts of relations between classification, recognition, and typicality. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *14*(4), 700–708. https://doi.org/10.1037/0278-7393.14.4.700

Palmeri, T. J., Goldinger, S. D., & Pisoni, D. B. (1993). Episodic encoding of voice attributes and recognition memory for spoken words. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 19*, 309–328. https://doi.org/10.1037//0278-7393.19.2.309

Penny, R. & Sánchez-Méndez, J. (2004). Variation and change in Spanish. Madrid: Gredos.

Pierrehumbert, J. (2001). Exemplar dynamics: Word frequency, lenition, and contrast. In J. Bybee & Hopper, P. (Eds.), *Frequency Effects and the Emergence of Lexical Structure* (pp. 137–157). John Benjamins.

R Core Team (2021). R: A language and environment for statistical computing. *R Foundation for Statistical Computing.* Version 4.1.1

Radeau, M., Morais, J., & Segui, J. (1995). Phonological priming between monosyllabic spoken words. *Journal of Experimental Psychology: Human Perception and Performance*, *21*, 1297–1311. https://doi.org/10.1037/0096-1523.21.6.1297

Regan, B. (2017a). A study of *ceceo* variation in Western Andalusia (Huelva). *Studies in Hispanic and Lusophone Linguistics*, *10*(1), 119-160. https://doi.org/10.1515/shll-2017-0004

Regan, B. (2019). Dialectology meets sociophonetics: The social evaluation of ceceo and distinción in Lepe, Spain. In W. Chappell (Ed.), *Recent Advances in the Study of Spanish Sociophonetic Perception* (pp. 85–121). John Benjamins. https://doi.org/10.1075/ihll.21.04reg

Regan, B. (2020). Intra-regional differences in the social perception of allophonic variation: The evaluation of [t͡ʃ] and [ʃ] in Huelva and Lepe (Western Andalucía). *Journal of Linguistic Geography 8*(2), 82–101. https://doi.org/10.1017/jlg.2020.7

Regan, B. (2022). The social meaning of a merger: The evaluation of an Andalusian consonant merger (*ceceo*). *Language in Society, 53*(3), 481–510. https://doi.org/10.1017/S0047404521000543

Samper Padilla, J. A. (2011). Sociophonological variation and change in Spain. In M. Díaz-Campos (Ed.), *The Handbook of Hispanic Sociolinguistics* (pp. 98–117). Wiley-Blackwell. https://doi.org/10.1002/9781444393446.ch5

Santana Marrero, J. (2016-2017). Factores externos e internos influyentes en la variación de /s/ en la ciudad de Sevilla. *Analecta Malacitana*, *39*, 143–177. https://doi.org/10.24310/analecta.v39i1.5613

Santana Marrero, J. (2018b). Creencias y actitudes de jóvenes universitarios sevillanos hacia las variedades normativas de España: andaluza, canaria y castellana. *Pragmática Sociocultural*, *6*(1), 71–97. https://doi.org/10.1515/soprag-2018-0003

Santana Marrero, J. (2022). How do Andalusian journalism students perceive Andalusian and Castilian linguistic varieties of Spanish? *Spanish Journal of Applied Linguistics, 35*(2), 565–595. https://doi.org/ 10.1075/resla.20019.san

Schmidt, L. B. (2018). L2 development of perceptual categorization of dialectal sounds: A study in Spanish. *Studies in Second Language Acquisition, 40*, 857–882. https://doi.org/10.1017/S0272263118000116

Sumner, M., & Kataoka, R. (2013). Effects of phonetically-cued talker variation on semantic encoding. *Journal of Acoustical Society of America Express Letters*, *134*(6), 485–491. https://doi.org/10.1121/1.4826151

Sumner, M., Kim, S. K., King, E., & McGowan, K. (2014). The socially weighted encoding of spoken words: A dual-route approach to speech perception. *Frontiers in Psychology, 4*, 1–13. https://doi.org/10.3389/fpsyg.2013.01015

Sumner, M., & Samuel, A. G. (2005). Perception and representation of regular variation: The case of final /t/. *Journal of Memory and Language*, *52*, 273–298. https://doi.org/10.1016/J.JML.2004.11.004

Sumner, M., & Samuel, A. G. (2009). The role of experience in the processing of cross-dialectal variation. *Journal of Memory and Language, 60*, 487–501. https://doi.org/10.1016/j.jml.2009.01.001

Villarreal, D. (2018). The construction of social meaning: a matched-guise investigation of the California vowel shift. *Journal of English Linguistics 46*(1), 52–78. https://doi.org/10.1177/0075424217753520

Villena-Ponsoda, J. A. (2008). Sociolinguistic patterns of Andalusian Spanish. *International Journal of the Sociology of Language, 2008*(193-194), 139–160. https://doi.org/10.1515/IJSL.2008.052

Villena-Ponsoda, J. A., & Ávila-Muñoz, A. M. (2014). Dialect stability and divergence in southern Spain: Social and personal motivations. In Braunmüller, K., Höder, S., & Kühl, K. (Eds). *Stability and Divergence in Language Contact: Factors and Mechanisms* (pp 207–237). John Benjamins. https://doi.org/10.1075/SILV.16.10VIL

Walker, A., García, C., Cortés, Y., & Campbell-Kibler, K. (2014). Comparing social meanings across listener groups: The indexical field of Spanish /s/. *Language Variation and Change, 26*, 169–189. https://doi.org/10.1017/S0954394514000088