PERFORMING ANDALUSIAN IN POLITICAL SPEECH: POLITICAL PARTY AND SOCIOPHONETIC PATTERNS ACROSS PRODUCTION AND PERCEPTION

Matthew Pollock

Submitted to the faculty of the Graduate School
in partial fulfillment of the requirements for the degree
Doctor of Philosophy in the Department of Spanish and Portuguese
and the Department of Linguistics, Indiana University

July 2023

Accepted by the Graduate Faculty, Indiana University, in partial fulfillment of the requirements for the degree of Doctor of Philosophy.

Doctoral Committee

_____

Committee Co-Chair, Manuel Díaz-Campos, PhD

_____

Committee Co-Chair, Samuel Obeng, PhD

_____

Erik W. Willis, PhD

_____

Stuart Davis, PhD

_____

Tracy A. Hall, PhD

June 16, 2023

To Kathy and Theo, who saw me through this journey.

ACKNOWLEDGEMENTS

I would like to express my gratitude to all of those who have been key in helping me to conclude this doctoral journey. First and foremost, I would like to express my sincerest thanks to my committee co-chair, advisor, and friend, Manuel Díaz-Campos, for many years of advice, instruction, collaboration, and encouragement. Here's to all the linguistic *milagros*. I would be remiss not to also offer my heartfelt appreciation to my other co-chair, Samuel Obeng, as well as the members of my committee, Erik Willis, Stuart Davis, and Tracy A. Hall. They have all served an important role in developing my theoretical interests and contributing to my scholarly identity over these years.

This dissertation project has been some six years in the making, extending back to my first discovery of Juan Manuel Hernández-Campoy's articles in 2017 and my original forays into the speech of Susana Díaz in Manuel Díaz-Campos' class on Spanish sociolinguistics. Only thanks to his thoughtful feedback, as well as the dedicated help and support of my co-author, Jamelyn Wheeler, was I inspired to continue building the project. My many thanks go out to Julie Auger, Jon Forrest, Kelly Berkson, Kenneth de Jong, César Félix-Brasefer, and Stuart Davis, in whose courses I built on my corpus of speakers and repertoire of regional phenomena from a sociolinguistic, phonetic, pragmatic, and phonological perspective.

My linguistic journey truly began back in August of 2013, in Robin Sabino's classroom at Auburn University, where I first learned about sociolinguistics, William Labov, and linguistic analysis. By the end of the semester, as I was playing with data in an excel sheet, running chi square tests, I was hooked. I would not be where I was today without her tutelage, and many hours spent chatting about language in her office. I also need to extend my thanks both to Robin Sabino and to my undergraduate research advisor, Craig Bertolet, for whom I wrote my first 20-page

research project, got my first IRB approval, and presented my first linguistic research at a conference. Finally, I would also like to thank James Goldstein, whose patient tutelage instructed me in the history of English and in Old English, further stoking my fascination with language.

My research interests were shaped and supported at Indiana University by an amazing group of scholars. Kari Gade and Tracy A. Hall helped to shape my knowledge of historical Germanic linguistics and Generative Theory. Erik Willis has always been willing to lend an ear, has been a wonderful co-author, and introduced me to the world of perceptual research, leading to my first publication. Susan Herring stoked my passion for computer-mediated discourse analysis and provided me with my earliest feedback on political speech. Although I did not end up using a Bayesian analysis for the dissertation, John Kruschke's instruction was foundational in my understanding of statistical modelling. And of course, my work as research assistant to César Félix-Brasdefer and, particularly, Manuel Díaz-Campos was invaluable in teaching me how to revise academic texts, develop book proposals, set up peer-review, and generally navigate academia.

This project could not have had anywhere near its current scope without the generous assistance of the Indiana University College of Arts and Science. Their Grant-In-Aid allowed me to reimburse participants of the perceptual study for their time, making it doubtless the only reason I was able to find 76 willing and generous souls in Seville. They also allowed me to pay my colleagues Santiago Arroniz and Juan Manuel Martínez a small stipend for their help in conducting inter-rater reliability for my production coding. Meanwhile, it was only thanks to the College of Arts and Sciences Dissertation Completion Fellowship that I was able to spend the academic year 2022-2023 focusing fully on my dissertation, allowing me to take a week away in Spain, code and analyze tens of thousands of tokens, and write all the words before you in this dissertation. I would also like to thank all of those who have offered feedback on my project, including the members of

FONEMI, PHLEGME, the Spanish Brown Bag audience, the Linguistics Round Robin attendees, and the Sociolinguistics Discussion group, who allowed me to present preliminary results. I also appreciate the scholars at the 2019 NWAV and 2022 Workshop on Spanish Sociolinguistics conferences who gave useful feedback on my sociolinguistic and phonetic analysis, much of which is implemented here.

Across the Atlantic is a collection of scholars who were extremely generous with their time, their advice, and in some cases, their classes. Juan Manuel Hernández-Campoy, who visited Indiana University in 2017 and 2019, and with whom I have had the privilege of collaborating with in the development of my articles for the ENCODES project, offered me formative feedback in the first stages of my interest in this topic and provided me access to an array of his research that has become an integral part of my understanding of sociolinguistic style and political speech. At the university of Seville, I have to extend my deepest gratitude to María del Mar Torreblanca López, who helped arrange my visit, including reserving the perfect space for deploying my perceptual instrument. She was also a generous interlocuter, taking me out to lunch to talk about Seville, political speech, and the dissertation project. I would also like to thank Isabel Iñigo Mora, Manuel Padilla Cruz, Julia María Fernández Cuesta, Juana Santana Marrero, Ángel Luis Jiménez Fernández, Lola Pons Rodriguez, Marta León-Castro Gómez, and Catalina Fuentes Rodríguez, who allowed me to present my research to their classes, shared information about my survey with their students, and made me feel right at home in Andalusia. I could not conclude this without thanking Erik Willis for lending me the iPads that made my study possible and talking through the data collection process with me, the members of the IU IRB who provided feedback on the layout and design, as well as all 76 of the participants who participated. You have my deepest thanks.

Last, but most certainly not least, I want to offer a final thanks to my family, including my mother, Kim, my father, Brian, my brother, Joshua, and most especially my dear wife, Kathy, for their many years of support. Between the defense of Kathy and my dissertation proposals and the writing of this dissertation, our son Theodore Pollock was born. Particularly in these last months leading up to the defense, Kathy has been an enormous help in ensuring peace and tranquility in the house while I write. She has also offered thoughtful discussion on the theoretical implications of my results, and provided vital help in the proofing of this lengthy document. All remaining errors you may find within these pages are, of course, my own.

Matthew Pollock

PERFORMING ANDALUSIAN IN POLITICAL SPEECH: POLITICAL PARTY AND SOCIOPHONETIC PATTERNS ACROSS PRODUCTION AND PERCEPTION

I examine Peninsular Spanish politicians' use of performative style, a sociolinguistic variable permitting identity formation, using a sociophonetic approach. While all speakers style-shift between prestige and non-prestige variants, politicians in particular use language both to reflect social position and to appeal to voters. Using quantitative and qualitative methodologies, I examine variation in Andalusian and Madrid Spanish, determine how style-shifting occurs at the individual level, and consider how Andalusian voters perceive political speech. Using a composite approaches, I determine that speakers' use of linguistic resources differs by political affiliation and gender.

Stage one of the analysis tracks how 32 peninsular politicians produce ten regional phenomena associated with Andalusian Spanish. While geographic and linguistic factors condition variation, additional social factors including speaker gender, political party, interlocutor gender, and age also explain variation. In Stage two, politicians were examined using *Lectal Focusing in Interaction*, tracking regional variation and style-shifting over time. Liberal politicians used moments of regional peaks as a means of emphasizing working-class solidarity, while conservatives used regionalisms more performatively to convey southern and friendly indexical meaning. Finally, the perceptual instrument in stage three showed how Seville listeners applied different criteria to community and political speech, evaluating regional variants positively, and associating them with female liberal and male conservative voices.

The composite results suggest the rising populism in Spain is leading to a change, whereby conservative voices produce more Andalusian features than liberals, and young listeners associate

regional speech with the political right. Meanwhile, female politicians navigate a web of indexical meaning, avoiding the stigma of overly vernacular speech while using regionalisms to craft unique identity. While there is an automaticity to the unmarked register of political speech that follows pre-established norms and expectations, politicians can also agentively sidestep them at times to perform identity work. This finding deepens our understanding of political discourse and Andalusian Spanish, presenting a methodology for in-depth examination of a speech community. This dissertation offers a means of generalizing beyond geographical and linguistic contexts, offering insight into stance accumulation and the connection between perception and production.

Committee Co-Chair, Manuel Díaz-Campos, PhD

Committee Co-Chair, Samuel Obeng, PhD

Erik W. Willis, PhD

Stuart Davis, PhD

Tracy A. Hall, PhD

Table of Contents

List of Figures

List of Tables

1 Introduction

There has been a recent push in sociolinguistic research to examine style and identity formation in political discourse using a sociophonetic lens to determine the degree of intention that politicians manifest when speaking publicly (e.g., Hernández-Campoy & Cutillas-Espinosa, 2012, 2013; Hall-Lew, Friskney & Scobbie, 2017; etc.). Sociolinguist theory suggests that speaker agency plays a role even at the phonetic level in the service of identity construction (Ahearn, 2001; Eckert, 2012). While all speakers style shift between a multitude of linguistic features, including regional and normative variants, and those that exist across a variety of registers and communities, politicians in particular are expected to appeal to constituents and supporters using carefully-crafted personae based on these linguistic choices. Combining community-driven (Labov, 1972), discourse analytical (Sharma & Rampton, 2015) and perceptual methodologies (Podesva, Reynolds, Callier, & Baptiste, 2015), this dissertation examines the relationship between party politics, gender, age, and speech context among speakers from four major cities in Spain: Madrid, Seville, Córdoba, and Málaga. It contributes to existing literature on political discourse, public speech, and style-shifting by analyzing variable regional norms within the peninsular Spanish political speech community, determining how individuals use stylistic variation to perform regional identity, and tracing how variants are perceived by voters in politicians' local community.

By focusing on political discourse, this dissertation examines individuals who spend a large amount of time and effort in crafting an identity for variable audiences with differing linguistic expectations, and facing potential perceptual consequences based on that self-presentation. Speech that meets the expectations of certain members of an electorate will help develop an image of a politician as local and reliable, as educated and worldly, or as any of a host of other identities to

help sell their message. Speech that over- or under-shoots this mark may be seen as voters as indicative of an individual who is, among other adjectives, uneducated, pretentious, or disingenuous, and may have job-ending results for a politician, as Hernández-Campoy and Cutillas-Espinosa (2013) and Pollock and Wheeler (2022) describe for female southern Spanish politicians. These women are both applauded as indexing regional belonging and appealing to a working-class and rural voting base, while at the same time denigrated as being unprofessional and uneducated, based on the same linguistic behavior. Importantly, production is not uni-dimensional, nor do voters represent a monolith of perception, and politicians are in a constant act of juggling these concerns. Fundamental to improving our understanding of this genre of speech are the dual issues of production and perception, both *how* politicians style-shift (in the context of their peers), and how that behavior is perceived by the community.

Several strains of theory are combined in this dissertation, including third-wave sociolinguistics, theories of performative identity through style-shifting, (perceptual) sociophonetics, and peninsular Spanish political discourse. Sociophonetic analyses of political speech in Spanish (Cruz-Ortiz, 2019; Hernández-Campoy & Cutillas-Espinosa, 2012, 2013; Pollock & Wheeler, 2022) and English (Hall-Lew, Coppock, & Starr, 2010, 2012; Hall-Lew, Friskney, & Scobbie, 2017; Holliday, 2014, 2017) have begun to address difficulties of this approach. When making stylistic claims, it is necessary to combine nuanced analysis of group production with the knowledge of community norms, such as the expected height of Scottish and Southern English vowels (Hall-Lew et al., 2017), or the consonant norms of Murcia and northern-central peninsular Spanish (henceforth referred to as NCPS; Hernández-Campoy & Cutillas-Espinosa, 2012, 2013). Geographical and social variables are also key in understanding stylistic choices; Villena-Ponsoda (2013) argues that prestige emanates from local norms, nearby cities,

and national standards, influencing speakers based on the identity they intend to construct. Only with the comparison of community norms and axes of prestige does the stylistic meaning of individual politicians come to light, contributing to our understanding of the complex issues surrounding identity construction and social perception that guide individuals' speech within a certain context, and yielding generalizable results that support broad theoretical claims.

As all of these aforementioned studies show, allophonic variation associated with regional linguistic difference is a key means of distinguishing how sociophonetic linguistic resources are used and social meaning is conveyed. In the course of this project, ten phenomena were selected to represent Andalusian Spanish, serving as some of the most salient markers of this variety. This included the following stop, sibilant, liquid, and vowel processes: (1) intervocalic /d/ reduction and elision, (2) fronting of the affricate [t͡ʃ] to [t͡s], (3) aspiration and elision of coda /s/, (4) aspiration and elision of onset /s/ in word-final position, (5) *seseo*, or the process whereby peninsular [θ] is produced as the alveolar [s], (6) *ceceo*, or the process whereby peninsular [s] is produced as the interdental [θ], (7) variable production and elision of the lateral [l], (8) variable production and elision of the trill [r], (9) variable production and elision of the tap [ɾ], (10) and a vowel harmonic process whereby [a e o] are produced as the lax [æ ɛ ɔ] preceding an elided coda /s/. Previous research has identified a distinction between the NCPS variety, which tends to be conservative and experience less elision, and Andalusian Spanish, which is marked for its innovative tendencies.

There is a long tradition in Spanish of dialectological approaches describing regional production in geographic terms, defining and adjusting dialect boundaries based on variable production tendencies (e.g., Navarro Tomás, 1939; Navarro-Tomás, Espinosa, & Rodríguez-Castellano, 1933). Perhaps in part because of this approach, Spanish linguistics often relies on

community-based analyses to describe broad macro-differences across portions of the population, as applied by Labov (1972) to English. More recently, sociological approaches related to what Bell (2016) and Eckert (2012) refer to as the third-wave have come to Spanish, examining individual-level variation and identity construction through ethnographic and perceptual approaches (e.g., Chappell & Barnes, in preparation; Delgado-Díaz, Galarza, & Díaz-Campos, 2021). Through a three-tiered approach, the present dissertation attempts to combine these methodologies into one whole, from a Labovian examination of community norms (Stage 1), to a third-wave consideration of style-shifting variation using Sharma and Rampton's (2015) Lectal Focusing in Interaction (LFI) that tracks and visualizes regional productions from moment to moment (Stage 2), to an examination of perception and social values of these same phenomena in the speech community (Stage 3). This combination of methodologies provides the benefits of each approach, both addressing and accounting for the oversights of any individual one.

Overall, this dissertation interrogates speaker style-shifting and performance through a closer look at the identity work of peninsular Spanish politicians, based on community norms and individual behavior, as well as the social meaning the community receives upon evaluating their speech. While current theory on stylistic variation is advancing, and there has been increased interest in recent decades regarding its application to public speech, there is still much that we do not know about politicians and their performance of identity.

Of particular interest is what Schilling-Estes (2013: 15) refers to as the "automaticity" of "less self-conscious" style, what Hernández-Campoy and Jiménez-Cano (2003) see as the normative "script" of social expectations for public speech, or what Sharma (2018) sees as the spaces in between the 'real me' of her speaker. In all of these moments, a type of speech is being identified in which speakers are less focused on their linguistic performance than on fulfilling the

expectations of the medium (e.g., radio talk, political speeches, etc.). However, despite these norms, research in political discourse in western cultures has suggested that individuals have the ability to deviate from this script, making agentive linguistic moves when the situation is right. Understanding *when* this deviance from the norm is permissible, and the extent to which it is socially permitted, provides results that are applicable beyond the specifically Andalusian or even Spanish context, as this behavior could transcend individual languages. These results would provide broad applicability across linguistics by documenting variant usage at the individual level to paint a picture of the linguistic repertoire available to speakers, as well as explaining how usage varies based on complex social phenomena.

In what follows, the sections of this chapter lay out important background information to contextualize the project that is to follow. In the next section, a brief presentation of terms and definitions is offered to explain several key concepts in the dissertation. Following that, the two target varieties are described in greater detail to help explain the reasons for their selection, and the complex prestige relationship between NCPS and Andalusian Spanish. Afterwards, goals and contributions of this dissertation are outlined, followed by a concluding section describing the structure of the rest of the document.

1.1 Key concepts

In the two sub-sections that follow, two important concepts, central to the analysis in this dissertation, are examined in detail. The first deals with the concept of *style* and several related ideas, including prestige, stance, and bricolage, all of which serve a key role in understanding

third-wave sociolinguistic approaches and individual analyses of variation. The second section, in turn, considers the relationship between *perception* and *production*.

1.1.1 Sociolinguistic style

In order to understand sociolinguistic style, we must move beyond the initial descriptions used by Labov (1972) of "attention paid to speech." It is not merely the extent to which speakers focus on their language, or even the general genre of discourse that is analyzed, but rather a complex contextual picture that changes across multiple contexts (e.g., Podesva, 2007). In order to best understand the way that speakers change between different linguistic styles based on contextual and social cues, analyses should focus on carefully defined contexts, and treat speakers' motivations as manifold in driving their behavior. Several social variables must be included in this portrait of usage.

First, both covert and overt prestige are important influences on style-shifting behavior. Woolard (1985: 744) describes a connection between the use of vernacular forms and central membership in community groups, in which covert prestige leads to social pressures, pushing speakers to use the vernacular in order to fit in. In order to carry out identity work, Giles, Coupland, and Coupland's (1991) Communication Accommodation Theory (CAT) also allows for speakers to address prestige by choosing to converge with, diverge from, or maintain their use of regional variants. In this way, underlying systematic power structures inform speakers' performative identity decisions.

Second, in order to show one's orientation toward a certain concept or group, the concept of stance comes into play. Arguing it to be at the heart of interaction, Bucholtz and Hall (2005: 595) describe it as the means by which a speaker describes their relationship with a certain object to their interlocutor using language and stylistic variation. While a speaker's stance may shift from moment to moment, the combined speech decisions contribute to a single unified whole that go on to help inform further choices the speaker takes in a process called stance accumulation. Through stance accumulation, one's perspective on this object is concreted, contributing to the construction of a speaker's coherent identity. While no individual stance-related move is decisive, when combined, like brushstrokes, they paint a portrait of the individual that can both shift the perception of a speaker on their own identity, as well as that of an interlocutor.

Third, the concept of bricolage serves as a means of combining linguistic phenomenon alongside other social phenomena that contribute to the development and performance of identity. Coined by Eckert (2008), bricolage serves to vest all of a speakers' social decisions with indexical meaning. Alongside linguistic decisions, speakers establish their place in society through various methods, including the clothes they wear (Eckert, 2000), the jobs they work (Zhang, 2005, Schilling-Estes, 1998), and even the music they perform and listen to (Guy & Cutler, 2011). As with stance, individual instances of bricolage may not carry inherent meaning on its own – however, when combined with overarching power structures, instances of indexical stance, and a specific social context in which variation possesses social meaning, these instances of bricolage can serve as part of the performative portrait of identity.

Finally, geography is an important feature in dictating how social prestige affects the social value of stylistic choices (e.g., Villena-Ponsoda & Ávila-Muñoz, 2014; Hernández-Campoy and Cutillas-Espinosa, 2013). As a result, in order to understand a speaker's style-shifting behavior,

linguists must also understand how social stratification and community norms influence local speakers' evaluations of sociophonetic variants.

On the whole, third-wave sociolinguistic approaches aimed at understanding how style-shifting behavior contributes to identity construction serve as an important means of clarifying the connection between social meaning and linguistic behavior. With that said, it is also important to specify the degree to which claims can be made surrounding style-shifting behavior. Individual speakers can make agentive moves with linguistic resources. However, with that said, the extent to which they have conscious control over this implementation, especially at the phonetic level is not clear. Rather than controlling which sounds will be elided or retained, in the Andalusian case, it is more likely that speakers have a type of speech in mind – perhaps a variety that they speak, perhaps simply a way of sounding, that they can choose to approximate. By tracking actual linguistic performance and comparing it to known social variables in the community, linguists can then gain insight into the specific groups or social meanings that speakers target.

1.1.2 Production and perception

The field of phonetics can be subdivided into three categories; articulatory, acoustic, and auditive/auditory research (Thomas, 2011). The first of these three, articulatory phonetics, focuses on the actual gestural movements in the lungs, throat, and mouth, using tools like x-ray, MRI, and electropalatography to track the link between movement and sounds. The second, acoustic phonetics, examines speech from the perspective of sound-waves, mapping correlates such as formant height (Hz), center of gravity (COG, Hz), intensity (dB), and duration (ms) onto related

articulatory gestures (e.g., F1 reflects the height of the tongue in the mouth). Acoustic phonetics also has a tendency to rely on characterizations of articulatory configurations that are based, in part or in whole, on researchers' impressions. Finally, auditive phonetics examines the way that listeners react to sounds, including determining the type of boundaries that exist between phonological categories (e.g., how /i/ differs from /e/ among hearers).

In this dissertation, acoustic and auditive phonetics are examined in depth, using sociolinguistic methodologies to probe results. From an acoustic perspective, linguistic research deals with speech production, a very common way for phoneticians and phonologists to examine speech (e.g., Martínez-Celdrán, 1998). An acoustic sociophonetic methodology might examine how productions differ, based on an acoustic measure like F1, by social group in a community (e.g., Labov, 1972). If vowel raising were to have social meaning, such a study could determine which members of a community are using raised variants and extrapolate what that might mean. Meanwhile, from an auditive perspective, linguistic research deals with speech perceptions, a means of identifying how listeners evaluate productions (Reetz & Jongman, 2008). An auditive sociophonetic methodology would then assess how perceptions reflect social values in the community, potentially tracking the attitudes or biases listeners use to classify speech sounds (e.g., Campbell-Kibler, 2009, 2010). In the case of vowel raising, a perception study could determine how listeners respond to tokens, allowing them to pin down social meaning of the variant – but without identifying if these tokens are frequent, or exist at all, in a given community.

This situation presents a bit of a conundrum. Ostensibly, neither research into speech perception nor production is complete without findings from the other. Production data can only tell what is happening, not what listeners think of speech, and perception data tells the opposite. On the one hand, work in perception tends to base itself in robust production research, often

including data collection from the community to develop matched-guised-style instruments where informants respond to stimuli produced by members of the target community (Lambert et al., 1960). On the other hand, production research rarely addresses perceptual questions, focusing instead on previously-observed trends associating the speech of women, young speakers, or upper-class speakers with prestige variants changing from above, or working-class speech contributing to change from below, for example (Labov, 1972). It is only in recent years that perceptual work has begun to gain greater traction, particularly in sociolinguistic research.

One last matter in this definition, regarding the frequent references to phonetics. Although they share a similar level of linguistic approach, phoneticians and phonologists analyze language with differing goals (RAE, 2011: 6143; Hualde & Colina, 2014: 16). Phonetics deals with differences in sounds that do not reflect changes in meaning (e.g., allophones [a:] and [ai̯] reflecting regional differences in the production of /ai/ for Southern U.S. English), while phonology looks at the overall system of sounds, distinguishing between units that convey separate meanings (e.g., /t/ 'toe' and /d/ 'doe' are separate entities in English, distinguished by voicing). While much of the research in this dissertation focuses on phonetic differences, where meaning is not explicitly involved, I follow Ohala's argument (1990, 2008, 2010) that studying either field requires an acknowledgement of aspects from both. While the regional variation at issue here often does not affect meaning, in cases like /s/ elision and vowel harmony, meaning distinctions can be lost, requiring considerations of the effect on the resulting sound system.

Phonetics and phonology interlace to help contextualize the data at hand. This can mean that phonetics supports a phonological analysis, allowing for a nuanced claim about Optimality Theoretic systemic trends, such as what Soriano (2012) does with eastern Andalusian Spanish vowel harmony. However, it can also mean that phonological underpinnings help support a

phonetic analysis that considers a host of speaker- and context-related variables like geography, gender, age, or socioeconomic level (e.g., Henriksen, 2017). This dissertation leans more toward the latter, using a quantitative approach as a means of collecting and analyzing community norms in order to make a nuanced but complicated claim about the degree to which social and lexical factors condition variation in the overarching system, and in turn speak to the identity of the individuals that employ these variables.

1.2 Overview of target varieties

This dissertation examines two major varieties of peninsular Spanish. The first is spoken in southern region of Andalusia, represented by Córdoba, Málaga, and Seville in this study. The second is found in the central and northern regions of the country, as represented by Madrid. These sites can be seen in Figure 1. In the next two sections, social prestige and linguistic norms in these two regions are discussed.

Figure 1: Map of Andalusia and Madrid

1.2.1 Northern Central Peninsular Spanish (NCPS)

The capital city of Madrid was chosen to represent NCPS, a variety described in previous research as a way to combine Northern Castilian Spanish (including parts of Castile and Leon, La Rioja, Aragon, Navarre, Castile-La Mancha, Basque Country, and Cantabria) and Southern Castilian Spanish (including parts of Extremadura, Castile-La Mancha, Valencia, and Murcia), as well as Madrid, which is somewhat split across the two regions (Henriksen, 2014). See Figure 2 for a general idea of the extension of the northern and southern Castilian varieties.

Figure 2: Map of Northern and Southern Castilian Spanish

While near a dialectal boundary, Madrid Spanish tends to be rather conservative. Intervocalic /d/ deletion is one of the few elision processes recorded in the city. While that reality should not prevent politicians from the city from using style-shifting in order to construct identity, as they have the same concerns of representation before their voters as Andalusian politicians, they will not have the same resources available to them as Andalusian speakers. While Madrid speakers do occasionally elide syllable-final /s/ (e.g., *es que* 'it's that'), it is highly uncommon for them to delete word-final, onset /s/ (e.g., *es un* 'it's a' [eØ.un]), a phenomenon that has emerged in Andalusian Spanish as an analogy to coda /s/ elision. This group is being included in this

dissertation to serve as a control, so as to understand how Andalusians use peninsular variables in contrast to those in the capital, and demonstrating what resources are in fact uniquely a part of the southern variety.

Additionally, as the capital variety, the Spanish spoken in Madrid possesses a special degree of social power and prestige for politicians. As the seat of government, the Spanish spoken in Madrid could be seen as the standard norm for national politicians, and as Hall-Lew et al. (2017) show for a group of National Party politicians from Scotland in London, this central variety could provide a point of contrast from which Andalusian politicians wishing to emphasize their regional identity could differentiate themselves.

1.2.2 Andalusian Spanish

While the Spanish of Andalusia will, at times, in this dissertation be treated as a monolithic entity, nearly a century of dialectology has documented divisions within the region. Linguists tend to bisect the region vertically into an Eastern and Western Andalusian (e.g., in the discussion of vowel harmony, a phenomenon unique to the east; Jiménez & Lloret, 2007), as depicted in Figure 3. Less frequently, a Northern Andalusian is also identified, including Córdoba (e.g., to distinguish *seseo* tendencies in the north from *ceceo* tendencies in the south; Díaz-Campos & Pollock, submitted).

Figure 3: Map of Andalusia with dialectal divisions between east and west

Andalusian Spanish has historically been the recipient of stigma in contrast with the normative Castilian Spanish found in Madrid and the northern part of the peninsula (Cano-Aguilar, 2001, 2006). By tracing linguistic consciousness over time, Cano-Aguilar discusses perceptions of the variety as an uneducated or mal-formed version of Castilian Spanish dating back to at least the early 1400s in texts. It has only been in the last 150 years that the variety has slowly begun to receive acknowledgement from folklorists and writers as a valid, if marginal, way of speaking. More recently, Valeš (2014) finds positive attitudes toward Andalusian Spanish have continued to thrive in the modern day, positioning it as a distinct variety from Castilian based on phonetic, grammatical, and attitudinal variation.

As such, Andalusian Spanish holds a degree of covert prestige for locals, while still being seen as lesser in comparison to the northern Castilian variety, which has greater overt prestige throughout the country. Melguizo-Moreno (2010) discusses the way that these two realities combine with a third, hyperlocal variety among rural citizens, who additionally have access to local variants that possess yet another type of covert prestige, setting them in opposition not only

to the northern Castilian norm, but also to the "general" Andalusian variety most identified with urban areas in the south. As a result, Andalusian politicians are at times armed with at least three types of linguistic resources, including those associated with NCPS, urban Andalusian Spanish, and local norms. Tagliamonte and Molfenter (2007) have pointed to the difficulty of dialect acquisition and the possibility that dialect performance can fail to reach or to maintain the target norms. In their examination of young children under aged five who moved from Canada to England for seven years, they find that these speakers develop a system medial to the Canadian and British ones, showing that second dialect acquisition is unstable, and does not necessarily result in the complete acquisition of phonetic norms. Podesva et al. (2015) and Molina-Ortés (2020) show, in turn, that communities can perceive this failure as intentional deception, reacting negatively to an inaccurate portrayal of one's identity, adding consequences to unsuccessful style-shifting behavior.

Alongside these general social considerations, there is also an ongoing process of standardization that has been identified in numerous Andalusian variants, sometimes tied to the universality of Spanish education and media in recent years. While there is a tendency toward elision and reduction in Andalusia for all ten of the phenomena described above that are discussed in this dissertation, recent research has also shown in some cases that these variants are normalizing toward the NCPS standard. Regan (2017b) has shown that *ceceo* (i.e., /s/ → [θ], /θ/ → [θ]) is receding in western Andalusia in favor of the Castilian standard of distinction (i.e., /s/ → [s], /θ/ → [θ]), part of a rare demerger process away from the socially-stigmatized sibilant phenomenon. Meanwhile, in Granada, Valeš (2013) finds that the stigmatized local deaffricated variant of /t͡ʃ/, [ʃ], is being abandoned for the normative alveolar production [t͡ʃ]. In eastern Andalusia as a whole, Villena-Ponsoda (2008) argues that city- and town-based variation is decreasing in favor of

broader “Andalusian” norms that are gaining greater social prestige. Individual identity, in particular, is important to understand the use and spread of regional variation over local variation (Villena-Ponsoda, 2013).

1.3 Goals

This dissertation combines several theoretical and methodological strains to approach the concept of identity construction in public speech. These include ideas from third-wave sociolinguistics related to style and identity, perceptual sociophonetics, and political discourse in Andalusian Spanish. The work of Hernández-Campoy and Cutillas-Espinosa (2010, 2013), focused on a former Murcian female politician, provided a model for examining political communities of practice in peninsular Spanish. Hall-Lew and colleagues’ (2010, 2012, 2017) work in U.S. and U.K. politics emphasized the importance of partisanship in the examination of political speech at the phonetic level, with speakers often tending to differ by political party in moments when linguistic agency was most key. Sharma’s (2018) analysis of regional variation through discourse analysis provided a visually illustrative means for considering individual variation through an analysis of specific regional features. And Podesva and colleagues’ combination of production and perception analyses provided a clear guide for conducting an attitudinal studies of political speech.

As these aforementioned sources show, projects in this vein have only begun to emerge in the last decade, meaning that every additional one will provide foundational contributions to our understanding of how regional variation and social meaning contribute to the ways that speakers present themselves in the public sphere, allowing for greater generalizability across communities

of practice and language contexts. This in turn will lead to two key contributions, addressing current gaps in the sociolinguistic literature. First, this study addresses a political community of practice that has not previously been addressed in depth using a three-stage approach to provide a wealth of information about the inner workings of linguistic norms for this group from a variety of perspectives. This includes norms within the community, individual style-based moves that deviate from that norm, and perceptions of politicians by listeners from part of the Andalusian speech community. Through this study of variation, a vision of Andalusian variation in the present day is presented that addresses various types of linguistic meaning, including Labovian (1972) community variation, Bell's (1984) perceived audience, and Coupland (2001) and Eckert's (2008) performative individual variation.

Second, designing a project that can permit comparison to previous and future research is extremely important for the overarching goals surrounding understanding the connection between regional variation and identity work. In theory, the findings of any one study can provide generalizations about behavior and strategies that speakers use that transcend their speech context and language. One of the reasons the Andalusian context proves so interesting is due to the confluence of covert and overt prestige norms. Politicians in this context have an abundance of linguistic forms accessible to them for use in style-shifting, while at the same time needing to balance perceived professionalism with regional solidarity. Being able to compare speakers' behavior in this context with those in both similar situations (e.g., Hernández-Campoy & Cutillas-Espinosa, 2010, 2013 in Murcia) and far-removed ones (e.g., Sharma, 2018, in American TV news) can help answer questions regarding the generalizability of identity-adjacent behavior, the efficacy of studying political speech, and the ways in which speakers break the expected linguistic "script" of their medium to agentively carry out identity work through linguistic choices.

Current theory on stylistic variation is advancing, and there has been increased interest in the last decades regarding its application to public speech. However, at the same time, there are still many questions regarding political speech, identity, and the generalizability of third-wave sociolinguistic findings based on style[1]. This dissertation aims to shed light on matters of style and performance through a closer look at the identity work of peninsular Spanish politicians, as well as the evaluative norms the community employs when perceiving their speech.

1.4 Contributions

As a result of the various areas of linguistic inquiry intersecting in this project, it makes contributions in several areas, including sociolinguistics, phonetics, phonology, political discourse analysis, and even dialectology. Here, I highlight four of the most important ones, based on findings from all three stages of the dissertation, including insight into female speech, the spread of populist conservative politics, the means by which speakers navigate political speech norms, and the composite approach this dissertation employs.

First, this dissertation contributes to our understanding of how female voices operate in politics, navigating a complicated field of social meaning where vernacular variants are often more stigmatized in female than male speech (e.g., Chappell, 2016). The perceptual results suggest that women are seen as more likeable and liberal, but also less educated, when producing regional variants. Both Pollock and Wheeler (2022) and Hernández-Campoy and Cutillas-Espinosa (2010, 2013) discuss the disadvantaged situation of women in positions of political power, using regional

[1] As is discussed in greater depth in Chapter 2, recent conceptualizations of sociolinguistic style envision it in the specific social, geographical, and cultural context in which it is produced, making it difficult to extrapolate elsewhere.

variation to reinforce their solidarity with their community and to emphasize their role as politicians. This places greater social pressure on women to comply with expected norms, even as they must distinguish themselves from their peers to win and maintain voter support.

Second, the results speak to changing norms in southern Spain, at a time when political power is shifting from the left to the right[2]. *Vox*, a new offshoot from the right-leaning PP with a more populist, far-right ideology, has recently become the third largest peninsular party (Rama, Zanotti, Turnbull-Dugarte, & Santana, 2021). As the perceptual results show, younger and less politically experienced Seville listeners are coming to associate regional speech with both left- and right-leaning ideologies. When combined with production tendencies on the right, and their political success, this suggests that their political message, combined with their increasing regional usage, may have struck a chord with voters. Conservative male voices and socialist female voices are associated with regional speech and rural identity, although analysis of style-shifting patterns show that these speakers differ in their implementation.

Third, linguists have discusses the "automaticity" of less self-conscious style (Schilling-Estes, 2013: 15), or the normative "script" of social expectations for public speech (Cutillas-Espinosa and Hernández-Campoy, 2007) as a set of norms that public speakers use. However, based on Sharma and Rampton's (2015) approach to visualizing and analyzing moment-by-moment variation (i.e., Lectal Focusing in Interaction, LFI), we can see how norms of production can be occasionally set aside as a means of contributing to speakers' identity. Through the analysis of both conservative voices and female speech patterns, we can see how politicians agentively sidestep the script to carry out linguistic work, as well as how individuals differ in respect to the

[2] The socialist PSOE lost in the 2019 Andalusian elections, marking the first time in decades that the conservative PP has been the governing party in the region.

moments they prioritize to present the ‘real me’ that Sharma (2018) describes as coming through LFI analysis.

The final contribution of this dissertation comes from the combined methodology itself. Third-wave sociolinguistic research has provided greater insight into linguistic variation and helped to refine the conceptualization of sociolinguistic style; at the same time, it is not without critique. The approach can at times lose sight of the greater community in favor of zooming in on the individual leading to questions of generalizability, comparability, and applicability of results. This dissertation offers a middle path, implementing three diverse methodologies to gain access to as many different perspectives of phonetic variation as possible. This includes the community-focused analysis of the Labovian approach, the individual-centric methods of LFI analysis, and the interlocutor-guided design of perceptual research. This composite view of variation combines individual, style-shifting, and attitudinal results to offer insight into speech communities from a variety of different angles, helping each to address deficiencies in the other approaches. Altogether, these three stages of analysis combine to present a whole that is more than the sum of its parts.

1.5 Structure of Thesis

The rest of this dissertation is divided into six chapters, focusing on previous research discussing style shifting and peninsular Spanish, then three stages of quantitative, qualitative, and perceptual analysis, and closing with a chapter examining these results in greater depth.

In Chapter 2, I present a description of sociolinguistic style and public speech. This chapter begins with a description of the concept of sociolinguistic style over time, then moves to discuss the concept of style in Spanish-language linguistics, and problematizes the conceptualization of style. After that, the focus turns to public speech, examining sociophonetic studies that take a closer look at variation among radio hosts, politicians, heads of states, and others. The chapter closes with a discussion of research goals in the dissertation explicitly related to stylistic variation.

Chapter 3 examines key phonological and phonetic features of Andalusian Spanish under consideration in the study. The chapter first describes research focused on production for each of the ten phenomena under consideration, then turns to studies using a perceptual approach, before turning to compare the relative benefits of production and perception research. The chapter ends with a discussion of those research goals related to speech production and perception.

Next, Chapter 4 focuses on the analysis of general norms among the Andalusian and NCPS speech communities following a Labovian (1972) framework. This chapter begins with a presentation of research questions, posing hypotheses based on previous findings, then turns to a description of the methodology used for analysis, including discussion of the corpus, dependent and independent variables, manner of data collection and analysis, and inter-rater reliability. Next, the chapter turns to a presentation of results for each of the ten phenomena under consideration, including affricate fronting, word-final /s/ aspiration and elision, word-final onset /s/ aspiration and elision, *seseo*, *ceceo*, vowel raising, intervocalic /d/ elision, and lateral, tap, and trill variable production and elision, followed by a discussion of regional trends. The chapter closes with a discussion of the results, returning to the previous literature and the research questions.

After that, Chapter 5 presents a qualitative analysis of variation using an instrument developed by Sharma and Rampton (2015) called Lectal Focusing in Interaction (LFI), which

permits an examination of individual regional variation across time. This chapter begins by presenting research questions and hypotheses based on previous research, then describes the methodological approach for this analysis, including identifying the corpus, describing the LFI approach, identifying how data was collected and analyzed, and discussing inter-rater reliability. Following that, the chapter examines ten speakers based on LFI analysis, and presents real-time analyses for interviews and scripted speeches. The chapter wraps up by discussing these results, looking back to previous studies and the hypotheses presented at the start of the chapter to determine generalizations about individual style-shifting in Andalusia and Madrid.

Chapter 6 presents a quantitative attitudinal analysis of perception of political speech on the part of listeners from the city of Seville, following a similar perceptual methodology to that which was employed by Podesva et al. (2015). The chapter starts off with a description of research questions and hypotheses, then discusses the methodology of the perceptual survey, including the participant base, the format of the instrument itself, and the means by which the data was collected and analyzed. Next, the results are presented in three parts for stimuli (1) not identified as coming from politicians, (2) explicitly identified as coming from specific politicians, and (3) comparing "preferred" speech from specific politicians. The results section concludes with a brief description of the main takeaways of regional perception that can be gleaned from the data. The chapter concludes by connecting these results to the research questions and previous literature in a discussion section focused on how Andalusian perceptions coincide with the social meaning identified in previous literature.

Finally, Chapter 7 discusses findings in light of the dissertation's contributions to sociolinguistic theories of identity and social meaning. The chapter begins by looking at each of the three stages' specific contributions to linguistic understanding, then combines these results to

obtain an overall idea of the project's main takeaways. To conclude, limitations of the current study are discussed, recommendations are made for future research, and concluding remarks are offered to provide ideas for next steps to continue developing on and improving our understanding of identity construction in political speech.

2 Sociolinguistic style and public speech

From the early days of sociolinguistics, when Labov (1981) described style as the "amount of attention paid to speech," the concept of sociolinguistic style has often been the focus of controversy and discussion in variationist literature. Given the variability inherent in speech production, from the individual producing speech sounds to the contexts where these productions are captured, the elusive concept of style has served, under varying definitions as an anchoring point to explain the speaker-internal variation that occurs across utterances and contexts. Early variationists, described as working in first-wave sociolinguistics associated with a Labovian approach, focused on sizable communities, broad social categories, and generalizations that could hold across all speakers sharing the same gender, categorical age, socioeconomic background, etc. As such, style developed under the first-wave of sociolinguistics served as a shorthand to account for variation in the very speech tasks employed in data collection: reading tasks required greater attention than semi-guided interviews, and yielded fewer vernacular productions due to that attention. This first-wave sociolinguistic approach to style, as well as the perspectives taken from the two waves that have since followed it (in the estimation of several major members of the field, e.g., Bell, 2016; Eckert, 2012), are discussed in greater detail in the following sections. Thus, these early definitions of style served the data being analyzed at the time, but were found to be increasingly at odds with the more naturalistic and varied sources of speech used in modern variationist research.

Within the English-speaking sociolinguistic tradition, there has been a move away from analyzing speech communities as a coherent whole, instead coming to regard smaller communities of practice, and even individuals, based on shared and individual identity goals. As a result, the

definition of sociolinguistic style has been forced to adapt to help explain intraspeaker variation that occurs without a marked change in task-type. Style-shifting has come to be seen more as a means of showing one's stance toward other speakers who speak in a certain way, defining one's relationship to the communities of practice with which one interacts. This move away from describing communities as a whole, as well as the use of naturalistic data that often does not (or cannot, as in the case of politicians) include differentiation between word lists and interviews, places a greater focus instead on the performative goals of speakers.

This interpretation of style places a degree of agency in the hands of speakers with respect to their self-representation, rather than making in an unconscious inevitability related to the type of speech being studied. As a result of this updated approach to style, considerable variability has been identified within individual speech communities and communities of practice at the intersection of macro-social factors like socioeconomic class with gender (Eckert, 1989, 2000), with age (Snell, 2018), or with urbanity (Zhang, 2005), as well as across pragmatically-distinct interactions (Holmes-Elliott & Levon, 2017) and based on variable use of regional and normative forms (Guy & Cutler, 2011). These findings not only document individual variation, but also paint a picture of a community's linguistic repertoire as a whole and serve to contextualize speakers' selections of linguistic forms based on complex social phenomena in their community.

Although the English variationist tradition has perhaps been the quickest to embrace these developments in theory, similar shifts are beginning to take place in work with other languages. In Peninsular and American Spanish, some of the first linguistic works documenting variation are attributed to the early twentieth century dialectologist Tomás Navarro-Tomás (e.g., Navarro-Tomás, Espinosa, & Rodríguez-Castellano, 1933). This historical interest in dialectology and geolinguistics has contributed to the rapid development of sociolinguistics in Spanish, albeit with

a continued interest for dialectologically-relevant research. Given the focus Labovian sociolinguistics places on the documentation of community-wide variation, this approach has continued to thrive in Spanish sociolinguistics, continuing into the present day. It is only since about 2010 that more recent English sociolinguistic trends have begun to gain traction in Spanish, with greater focus being placed on small communities of practice (Molina-Ortés, 2020) or even comparing these small communities to case studies of individual speakers (Hernández-Campoy & Cutillas-Espinosa, 2010, 2013). While English-language sociolinguists developed much of the theoretical underpinning of more recent conceptualizations of stylistic variation, Juan Manuel Hernández-Campoy has been a vital contributor in the Peninsular Spanish context, developing several projects that have helped push the field to apply a performative understanding of sociolinguistic style to the analysis of political speech (e.g., Hernández-Campoy, 2016; Hernández-Campoy & Cutillas-Espinosa, 2012). As a result, Spanish provides both theoretical and experimental examples of political style-shifting that are ripe for comparison, as this dissertation intends to do.

2.1 The evolution of sociolinguistic style

The discussion of speech style has long been of interest in discussions of language, dating back thousands of years (e.g., Aristotle, 2006; Cicero, 1853). However, only since around 1970 has it begun to be treated scientifically through linguistic analysis. The field of variationist sociolinguistics, in particular, has integrated it into its consideration of linguistic change across three waves of study. While early research described it as a fixed unit (Labov, 1972), more recent

work has viewed it as multidimension (Cutillas-Espinosa & Hernández-Campoy, 2007) and related to a performative interpretation of identity. Research focusing on speaker agency and a constantly-evolving external identity inform some of the most recent interpretations of style, yielding incisive intrapersonal analyses (e.g., Zhang, 2005) while, nonetheless, suffering from a lack of the very type of generalizability that was a major contribution of earlier Labovian projects.

The concept of "style" itself has existed in some form since at least the time of Aristotle, whose early discussion of rhetoric divides it into three branches: *deliberative* for politicians to propose or condemn action, *forensic* to investigate law claims, and *epideictic* for ceremonial purposes to laud or excoriate (Aristotle, 2006: 1.3). These three branches, naturally, had differing purposes, audiences, and goals. While stylistic variation is implied by the different fields to which these styles pertain (e.g., a politician could use an epideictic style to laud), this possibility is not discussed. Cicero (1853: 1.7.9) broadens the meaning of style (i.e., *elocutio*) further, calling it "the adaptation of suitable words and sentences to the topics so conceived." As he broadens the conceptualization of style beyond a single specific genre, the Roman writer begins to advance it past the Aristotelian definition in another way, rooting "style" in both lexical and syntactic decisions and indicating the need for speech to be appropriate to a given context.

In the intervening years between early descriptions of rhetoric and the development of modern linguistics, style manifested itself in fiction. For example, Geoffrey Chaucer's *Canterbury Tales* (Burnley, 1983), written in Middle English, and Miguel de Cervantes' *Don Quijote de la Mancha* (Egido, 2018), written in Early Modern Spanish, both demonstrate an awareness of stylistic differences, with characters shifting and varying across interactions in playful ways.

Despite this popular awareness, early texts in modern linguistics tended to only acknowledge the presence of style without examining it in greater depth. In his foundational

Course in General Linguistics, Saussure (1966) refers to a divide between *langue*, or the underlying system, and *parole*, or the production of speech. While almost exclusively discussing the former, he does at one point reference *parole* as "an individual act" that "is wilful and intellectual," which includes "the combinations by which the speaker uses the language code for expressing [their] own thought" and "the psychophysical mechanism that allows [them] to exteriorize those combinations" (14). Sociolinguistic style is hinted at by this "individual" nature of speech, as well as the "wilful" (or, perhaps, agentive) use of language that *parole* represents. At the same time, sociolinguistic style is enmeshed with one of Saussure's complaints about *parole:* namely, that each speaker is "wilful" in a separate way, making comparison and universal claims difficult. Similarly, Chomsky and Halle (1968: 3) describe speech performance (i.e., *parole*) as "what the speaker-listener actually does," an entity "based not only on [a speaker's] knowledge of the language, but on many other factors as well – factors such as memory restrictions, inattention, distraction, nonlinguistic knowledge and beliefs, and so on." Style is thus encapsulated, among other unmentioned social phenomena, in the "nonlinguistic knowledge," such as the social value of certain speech variants, as well as the "so on."

In this way, influential modern linguists like Saussure, Chomsky, and Halle left the Aristotelian and Ciceronian descriptions of style uninterrogated, seeing them as a messy feature of speech production that could be separated out from speech competence and *langue.* It is at this point that the nascent field of sociolinguistics began to puzzle out a different way to approach style.

2.1.1 First-wave sociolinguistic style

Around the same time Chomsky and Halle were writing about performance and competence, a group of scholars were publishing the first entries in what Eckert (2012) identifies as first-wave sociolinguistic research. Among them, Dell Hymes (1974), William Labov (1966, 1972), and Peter Trudgill (1972) served as the vanguard of this nascent field, using sociological theory to examine linguistic questions. These early works set a precedent for social analyses of linguistic variation using macro-social factors including age, gender, and socioeconomic status to help describe patterns. Many of them also shared a similar interpretation of speech style.

Bell (2016:397) describes these early works using the Labovian approach as choosing to make the concept of style a "central issue on the sociolinguistic agenda," even if their definition was often "reductionist, positing that it was a direct reflex of varying amounts of attention paid to speech." While this linear view of style does not accurately represent speaker behavior, it does reflect the Ciceronian view of genre, and makes sense as a starting point for the field – speakers change the way they speak based on the context in which they are speaking. For Labov (1981: 2-3), though, this was represented through an inverse correlation between use of the vernacular and attention paid: the former rose as the latter fell, and vice versa. This implementation of style was well-suited for large groups of speakers, such as in a study examining the English of NYC, where it was determined that hierarchical social differences across socioeconomic background and gender were maintained even as attention increased (Labov, 1972). Stylistic variation was, thus, being tracked, albeit without accounting for any agency on the part of the speaker – all productions in response to certain interview questions or in word lists were expected to share the same broader style category.

Labov was not alone in his treatment of the term. Trudgill (1972: 181) defines style explicitly as represented through formality, with word list style being more formal than reading passage style, which is in turn more formal than formal speech, which, finally, is more formal than casual speech. In his examination of the gerund -ing, production of the standard progressive form [ɪŋ] was highest among all social groups for word lists, while production of the regional [ən] was highest in casual speech, while intervening steps followed a strict order across socioeconomic status. When examining the perception of Received Pronunciation (i.e., standard British English), Stewart, Ryan, and Giles (1985) focus on speech style, which they define as the difference between fast, understandable, and comfortable speech. Style was found, possibly due to negligible differences between these three levels, to have little effect as compared to socioeconomic status. In these cases and many more from that time, persuasive social and dialectological claims were reached regarding community norms, but the static interpretation of style concealed considerable variation. Speech communities, groups and individuals within those communities, and language itself, were all being painted as fixed entities – change the contextual stimulus, and style changes in lockstep. Working class men, for example, were treated as a monolith: they were expected to use high rates of stigmatized variants, which increased as speech style became more "formal." Women proved harder to classify, leading to the so-called paradox in their linguistic behavior, as they both favored prestige variants and somehow tended to lead innovative change. Later considerations of style as fluid, based on a web of contextual and individual motivations, showed that this is not always the case. Men can be variable, if their speech is analyzed dynamically (Kieslin, 2007), and women are motivated by complex power structures that serve to explain their seemingly paradoxical behavior (Eckert, 1989; Chappell, 2016).

In these early years of what Gordon (2012) terms the Labovian Paradigm, other definitions of style beyond the first-wave approach also developed. Giles and Powesland (1975) considered the relationship between style, speech message, and topic, foreshadowing the second-wave "Communication Accommodation Theory." At the same time, they also discuss the concept of style-as-formality when introducing the concept, suggesting the ubiquity of the Labovian definition. Hymes (1974: 59), remaining more theory-agnostic, termed sociolinguistic style a "selection of alternatives with reference to a common frame or purpose," which "can be applied at any level of analysis" and is personal to each individual. Although he does reference the Labovian idea of style, he situates it as a just another subcategory of the concept, placing it alongside linguistic variation across registers, dialects, and in code-switching. Thus, while many first-wave studies established style as attention paid to speech, some maintained that it was larger than just that, relating to the contextual moment.

As time passed, the shift from first- to second-wave sociolinguistic research did not fully imply the end of the Labovian paradigm. While it had reached its peak by the early 1980s, the continued work of Labov (1990, 2001) and others indicates that this approach to style continues to hold currency in modern sociolinguistics. Nonetheless, the Labovian focus on this linear continuum between standard and stigmatized language presents a false binary that researchers beginning in the 1980s would begin to problematize. While the New York and Martha's Vineyard speech communities of Labov's earliest studies provided a compelling case for style-as-attention-paid-to-speech, not all language situations are so easily described. Instead of classifying variant use as either negative or positive, some speech communities have access to a multitude of stylistic options with varying associations depending on the context in which they are used, and the

audience that hears them. It is particularly with regard to these two concepts – context and audience – that second-wave sociolinguistic research came to define itself.

2.1.2 Second-wave sociolinguistic style

There are varying perspectives on the line between first- and second-wave research in sociolinguistics, as is the case with many post hoc groupings. Schilling-Estes (2013) sees the dividing line between the waves as a matter of theory: while first-wave style correlated to attention paid to speech, the second-wave vision of style must be understood based on audience concerns, as understood through Bell's (1984) Audience Design. Meanwhile, Eckert (2012) sees speaker agency as an important addition to second-wave research: unlike in the first-wave studies that considered large social groupings alone, the second wave saw speakers as having a role in alternating between vernacular and standard forms to express identity, as exemplified through social network research (e.g., Milroy & Milroy, 1985). Rickford and McNair (1994: 240), alternately, draw attention to the work of Labov, Cohen, Robins, and Lewis (1968), in which these authors distinguish group and single styles, as a precursor to Bell's (1984) audience-focused approach. Bell (2016) himself, who defines second-wave research as using ethnographic methods to identify local-level variation, references Eckert (1989, 2000) and Milroy and Milroy (1985) as central to the second-wave. Giles' (1980) Accommodation Theory is also an important theoretical precursor to the second-wave, as it describes how speakers converge toward or diverge away from linguistic norms based on expected interlocutor responses. Aspects of this theory also foreshadow

the transition into third-wave, speaker-centric approaches that probe the goals of the speaker[3]. Given the importance of the work by Bell, Milroy and Milroy, Eckert, and Giles to second-wave sociolinguistics, as well as the key ways in which they differ with respect to style, the four are discussed in the paragraphs that follow.

As a means of expanding the Labovian definition of style in favor of the Hymesian (1974) one, Bell (1984) proposed the Audience Design model. Rather than a fixed entity correlating to attention, Bell's examination of radio speech in New Zealand led him to argue that style is a subject's awareness of their addressee. In his style axiom, he describes stylistic variation "within the speech of a single speaker" as deriving from and echoing "the variation which exists between speakers on the 'social' dimension" (151). Under this view, style reflects shared evaluation of regional and national linguistic variants across a speech community: if a newscaster were to use a stigmatized variant on a local radio program, but be overheard by viewers from elsewhere, a different set of evaluations of the stigmatized variant might apply. For that reason, Bell (1984: 159) represents the speech situation in a series of concentric circles moving outward in the following way: Speaker → Addressee → Auditor → Overhearer → Eavesdropper. Sociolinguistic analysis of style cannot only take into consideration the direct interlocutor, under Audience Design, or even assume that style is constant across levels within the speech situation. Instead, every role in the speech situation shapes stylistic choices, as a speaker must account for every level of interlocutor that may interact with their language. In the case of public speakers, for example,

[3] These few studies are meant to serve as representative examples of work performed in second-wave sociolinguistics. The edited volume by Eckert and Rickford (2001) provides a nice collection of authors writing about the first-wave school of style (e.g., Labov), the second-wave school (e.g., Bell, Finegan & Biber), as well as the third-wave school (e.g., Giles, Eckert). Meanwhile, the more dialectologically-influenced European volume from Auer, Hinskens, and Kerswill (2005) provides a contact perspective rooted in a second-wave interpretations of Accommodation Theory.

there is a high likelihood that overhearers and eavesdroppers may originate from far beyond one's own speech community.

Rather than listener focused, Milroy and Milroy (1985) focus much of their attention on social networks, tracing the interconnected web of social bonds between members of a society. Through this approach, linguists can trace the movement of linguistic variation through and between communities. This movement, it turns out, is dependent on the strength of a speaker's ties to their local network. The normative impulses set forth in a community encourage speakers to use variants their listener group (i.e., speech community) would expect to hear. Milroy and Milroy determine that certain local variants are stigmatized in the greater community, while in West Belfast they have covert prestige, which comes not from the national center but from the margins. While Milroy and Milroy refer to style in almost Labovian terms, distinguishing between the "conversational" and the "careful," the variation in Belfast vowel patterns they track across social networks resembles what Bell finds in newsrooms. In both cases, the authors identify speakers who are aware of audience expectations and alter their linguistic production to match those expectations (i.e., style-shift) as a result. In Bell's case, that reason is so a single anchor can differentiate between the expected norms of two news channels, whereas for Milroy and Milroy, it is so that Irish-English speakers can avoid social repercussions in their community.

Through her ethnographic work in a suburban Detroit High School, Eckert (1989, 2000) defines linguistic style as part of a broader lifestyle in which students choose to participate. This includes what she later defines (Eckert, 2008: 456) as bricolage, or a set of social variables that combine to construct a complex individual identity. Eckert (2000: 66) is interested not only in stigmatized regional forms, but also in bell bottom width and other visible signs of identity performance within two communities of practice that she defines. These communities, following

the ethnographic set-up, do not stem from Labovian macro-social categories (e.g., elderly women), but rather come from close observation of behavior and evaluative norms: “burnouts” have a negative attitude toward school and suburban life while “jocks” are personal-image focused and driven to leave the city after public schooling. This perspective combines Audience Design, where speakers make speech decisions based on expected members of a speech situation, with Social Network theory: Eckert’s jocks and burnouts are networks that exist side-by-side within a school yard (and intermix), yet have very different evaluations of speech norms. In contrast to the Labovian approach, social categories alone would not provide a coherent picture of this community. Examining five vowel shifts across male and female jocks and burnouts, Eckert (1989: 261) finds differences both across gender lines and between communities of practice. Burnouts, for example, have much more advanced tokens of (uh) than jocks, while girls in general had more advanced tokens of (ae): only by looking at style as a performance of identity through social networks does Eckert’s more nuanced distinction emerge.

Unlike the previous three approaches to second-wave sociolinguistics, Giles’ (1980) Communication Accommodation Theory (CAT), later updated by Giles, Coupland, and Coupland (1991), offers few concrete examples, instead relying on intuitive and adaptive theory. First-wave “attention paid to speech” can be explained under CAT as convergence toward an expected speech norm when speakers are more aware of their speech, while second-wave local variation geared toward audiences can be described as a desire to accommodate with relation to a listener. Although speakers are expected to converge toward listeners, real (e.g., in Belfast communities; Milroy & Milroy, 1985) and expected (e.g., among news reporters; Bell, 1984), social norms and personal agency can instead bring about divergence (e.g., between jocks and burnouts; Eckert, 1989). In this sense, converging, diverging, and maintaining stylistic norms is an act of speaker agency,

serving the goal of identity formation. Thus, although Bell (1984) references CAT within the second-wave approach, the versatility of this theory to be applied at the attentional, audience, and speaker level makes it applicable to an understanding of style across all three waves.

Despite the positive developments that the second-wave understanding of style represents in comparison to the more static Labovian definition, it still presents certain problems. Schilling-Estes (2013: 8) notes that Audience Design theory does not explain why speakers do not carry accommodation to its logical conclusion, aiming to sound exactly as their listeners' expect and want in every situation. Similarly, it is hard to explain how a speaker could have an imagined audience in mind at all for a wholly unfamiliar context, and why we must fully dismiss the possibility of speaker-centric agency that goes beyond a slavish fixation of satisfying audience desires. Eckert (2012: 93) also critiques these second-wave works for continuing to advance the static approach to identity that was first established in the first-wave: speakers vary across speech contexts, meaning that the linguistic behavior analyzed in one situation does not necessarily imply anything about the speaker across speech settings. As sociolinguistic research continued to probe variation at the small-group level, yet another level of stylistic variation came to light that showed exactly how dynamic and consistent linguistic behavior was at the individual level.

2.1.3 Third-wave sociolinguistic style

The focus has shifted in third-wave sociolinguistic research to focus more on individual decisions and agency, often through small-scale case studies designed to tease out stylistic choices. This theoretical push is also well-documented in anthropological and sociological literature from

the turn of the century. Johnstone (2000), for example, in linguistic anthropology argues that sociolinguistic research should be supplemented by individual case-studies. Any given individual, subject to numerous cultural influences over their lifespan in variable ways, has a unique background that must necessarily be reflected in language, making a study of parole incomplete without an intimate knowledge of the speakers that produce it. This emergent linguistic identity can be reinforced through the continued performance of certain norms, serving to create an internally coherent narrative of the self (e.g., Bucholtz & Hall, 2005), and through the use of specific lexical items, vested with social meaning, to indicate one's stance within interaction (Bucholtz, 2009). Stance, which Du Bois (2007: 173) defines as "an act of evaluation owned by a social actor," serves as an indicator of one's underlying position, accumulating over time to build a coherent picture of an individual's identity. Stance shows a position with relation to certain cultural values, making style a linguistic representation of stance based on one's knowledge of local, regional, and national social value vested in sociolinguistic variants.

Sociolinguists have shown consensus in separating second- and third-wave research. Bell (2016: 402) references the study of variation involved with the construction of social meaning, sociolinguistics proper, in which "styles are the focus," taking Eckert's (2008) more recent work as exemplary of this shift. Meanwhile, Schilling-Estes (2002) describes the third wave as one guided by the theoretical framework of Speaker Design (e.g., Coupland, 2001), where speakers' linguistic resources develop identity, societal structure, and beliefs. Finally, Eckert (2016: 94) describes this wave as examining how "speakers place themselves in the social landscape through stylistic practice," describe both her own recent work and Zhang's (2005) discussion of variation in Chinese as exemplary of this shift. Style, to Eckert (2016: 98), "is at its foundation ideological, and the stylistic form of propositions is very much a part of their meaning. The third wave locates

ideology in language itself, in the construction of meaning." In this way, style has made a full transition from a peripheral to a fundamental component in sociolinguistic analysis.

The concept of bricolage has a far-reaching theoretical application across third-wave studies. Eckert (2008) sees linguistic variables as part of a broader meaning-based system that is highly performative. As variables take on meaning through context, they can then index identity meaning among communities of practice. Rather than an Audience Theoretical approach, in which these variants are used because they are expected by a listener, they are used agentively on the part of the speaker to perform identity work and build a coherent persona through social bricolage. Reexamining jocks and burnouts, Eckert (2008: 459) demonstrates another key component of third-wave analysis: increasingly complex social variables that are not always applicable outside of the given speech context. In her analysis, that becomes the sub-category of "burned-out burnouts," who are the most dramatically alienated, wild, and urban-oriented of the burnouts. The stylistic extremity of these girls, as compared to "regular burnouts," is taken to be part of their identity work, alongside other social bricolage including clothing, attitude, and stance toward school, which allows them to intentionally separate themselves from their peers.

This development of emergent linguistic identity, reinforced through continued performance of certain norms, serves to create an internally coherent narrative of the self. As Buccholtz (2009) shows for the Mexican slang term güey 'dude,' or Schilling-Estes (1998) demonstrates in performative speech to index local solidarity, lexical and phonetic variation can become vested with social meaning and allow speakers to make indexical choices within interaction. Language, along with other bricolage, like clothing (Eckert, 2008), then serves as a performative means of indicating stance and vesting social decisions with indexical meaning. This

also can be seen through the claims of Silverstein (2003), who argues that micro- and macro-social variation combine to influence social meaning in speech.

In Zimman's (2017) ethnographic study of timbre among a San Francisco Bay Area transgender community, the amount of testosterone participants received became a variable in a linguistic study, alongside F0 and center of gravity (COG) of /s/. While overall timbre increased little for most speakers over the months of hormone treatment, increased masculine physical appearance made speakers feel less need to perform masculinity in dress and language, showing a stylistic shift toward less normative gender performance based on increased individual self-confidence. In both cases, individual or micro-group identity decisions require ethnographic knowledge to understand stylistic choices. First-wave macro-social factors would not suffice to tease out these differences in a large-scale community study, while the second-wave focus on audience ignores the agentive performances that these speakers make to demonstrate their stance toward gender identity, or any other object, beyond the expectation of converging with the interlocutor.

The third-wave of sociolinguistic analysis emerges from earlier lines of theory that saw speech productions as a means of enacting identity, such as that work performed by Le Page and Tabouret-Keller (1985:5), who see language as more than just tied to communities – instead, it serves as "the means by which individuals both identify themselves and identify with others." These authors describe speakers as existing in a multidimensional space, in which they use socially-marked linguistic resources as a means of establishing and building on their identities both as part of social groups and as individuals in their own right. Building off the ideas in Giles' (1980) CAT, Le Page and Tabouret-Keller (1985: 181) see speakers as making intentional decisions that allow them to accommodate with certain groups while distancing themselves from others as a

means of establishing their own position in society. Through their examination of Pidgin- and Creole-speaking communities, including sociolinguistic analyses from the West Indies and among British immigrant communities, these researchers show how speakers across a wide range of multilingual contexts adopt the patterns and styles of certain groups and varieties to both fit in and project a certain linguistic behavior. This theoretical background, although initially rooted in creole studies, experienced great uptake by sociolinguists, who have also shown an increasing interest for examining how linguistic resources serve the larger object of social representation.

Coupland's work on Speaker Design has been a key part of the sociolinguistic third-wave conception of stylistic variation (Schilling-Estes, 2002). It begins with a case-study Coupland (1980) published of a travel agency assistant, many years before the third-wave conceptualization of style had been normalized. The various speech styles that this assistant moves through are described as contributing to her control of the encounter, shifting to meet the needs of each speech setting. Coupland goes on to problematize the small sample size of the study (a common difficulty with third-wave approaches), while also acknowledging that the style-shifting behavior he has recorded may not be generalizable, but that it stems from the interactional context (going beyond a focus on the audience alone). In a similar approach, Coupland (2001) performs an analysis of a Cardiff radio DJ, who indexes stereotypical local identity through increased regional variant production when talking about Cardiff people and events, but also uses more standard productions outside of these contexts and employs stylized American dialectal features when cuing up rock songs. These tendencies, Coupland (2001: 209) argues, are not just part of his script: instead, "they are ways of subtly activating multiple simultaneous dimensions of meaning potential." This speaker demonstrates, like the travel agency assistant, that there is an agentive capacity to make

intentional linguistic choices, employing socially meaningful variants to provoke humor and encourage solidarity with his listeners.

Numerous recent sociolinguistic research has taken Johnstone's (2000) call for individual analysis seriously, focusing on stylistic tendencies in case-studies and small groups. Schilling-Estes (1998), for example, focuses on the speech of a resident of Ocracoke, North Carolina, who uses a single performative phrase among tourists to demonstrate the stereotypical raised /ay/ of historical island speech. This speech, combined with the speaker's adoption of a so-called "jovial" fisherman persona, serves a performative purpose, showing local belonging and fulfilling expectations about the rurality of local speech, despite it overshooting actual vowel norms on Ocracoke (63). Podesva (2007), on the other hand, performs a case-study on a gay medical student who uses a falsetto in a casual setting, but avoids this identity work around his father and patients. The falsetto helps him create a "diva" persona around his friends, allowing him to voice opinions and demonstrate stance in a performative way without backlash. In a different instance, Guy and Cutler (2011) use a variationist model of morphological constraints for coronal stop deletion (Guy, 1991) to show performative manipulation away from community norms among white members of the hip-hop community. Grammar, Guy and Cutler (2011: 159) conclude, "turns out to be a vital part of identity," as these artists at times shift toward African American Language (AAL) norms, showing alignment with speakers of AAL. Notably, this convergence is temporary, rather than the ongoing trend that Audience Design suggests. In these three cases, individuals are shown by third-wave research to employ local variants in unique ways to perform identity work that contributes to coherent identity construction within constrained social contexts.

Another approach, integrating discourse analytical considerations into an analysis of sociophonetic variation, is that of Lectal Focusing in Interaction (LFI), designed by Sharma and

Rampton (2015) and exemplified by Sharma (2018). LFI analysis quantifies lectal variation over time, representing differences on a horizontal axis to demonstrate style-shifting. Sharma (2018) uses this tool to consider the speech of an Indian American media personality who has spent 35 of his 52 years in the U.S. The subject, Fareed Zakaria, was found to employ his primary Indian American dialect as a means of indicating his deeply held stance toward an issue (Figure 4). By mapping variation over time, this approach allows for a quantitative representation of variation, which can be paired with qualitative evaluation of instances of style-shifting in action.

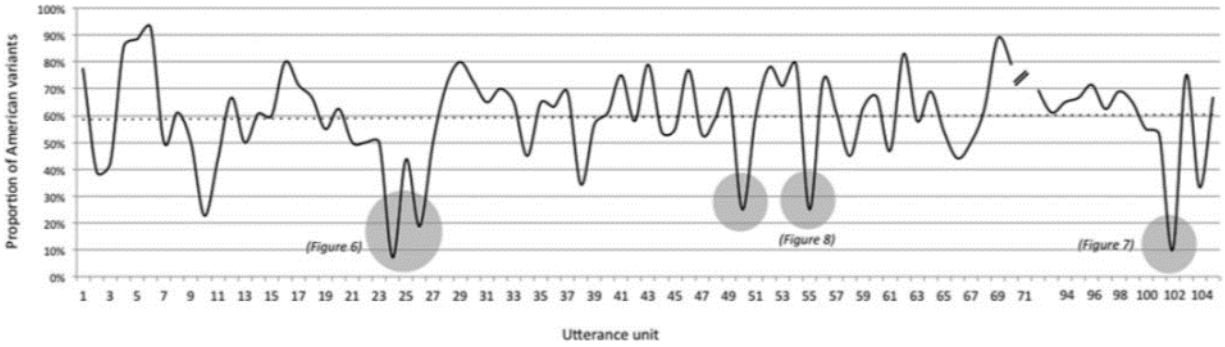

Figure 4: LFI analysis for Fareed Zakaria (Sharma, 2018:14)

Finally, yet another strain of third-wave research builds on the concept of communities of practice, identifying social boundaries that transcend the macro-social Labovian ones. In a discussion of cosmopolitan yuppies in Beijing, Zhang (2005) first sketches the social state of early twentieth century China in a fashion reminiscent to Labov's (1972) introduction to Martha's Vineyard. While this sociological lens is familiar, the depiction of style is wholly different, with variation not based on community-level factors like socioeconomic status, but rather by connecting regional variants and social meaning. Speakers in international companies, Zhang finds, adopt a "cosmopolitan" style based on their role in a global market, seeing themselves as part of multiple

business worlds and projecting an identity to reflect that, whereas workers in smaller national companies maintain regional norms. Snell (2018) determines something similar in her work with Teesside English, finding that socioeconomic status relates back to class distinctions: class informs the type of evaluative claims children made. Working class sensibilities, Snell suggests, determines how students use language and frame evaluations. Snell then goes perhaps further than many other authors of third-wave research, suggesting that her speakers are not just using linguistic resources differently, but rather, their very goals for speech production are fundamentally different, based on the lives they live. In both cases, style intersects with specific communities of practice, providing resources for demonstrating one's identity.[4]

While these third-wave conceptualizations of style allow for a better understanding of individual motivation, these very benefits also come with very real drawbacks. Third-wave sociolinguistic research tends to have severely reduced sample sizes, making generalization difficult even within a single speech community or community of practice, much less across national borders or languages. Additionally, by identifying nuanced micro-communities of practice, comparability across context-specific social variables becomes impossible (e.g., Zhang's "smooth operator," Schilling-Este's "jovial fisherman," or Podesva's "gay diva"). It is unclear how universal these factors could be, given how rooted they are in specific contexts. Schilling-Estes (2013) also notes that even those speakers that attempt to reject social hierarchy and meaning by reappropriating variants often reinforce the power of that hierarchy just by using them. She also points out that, while agency is important to an understanding of style-shifting and language use,

[4]As with second-wave research, numerous other authors have done interesting research under a third-wave framework. This includes Holmes-Elliott and Levon's (2017) consideration of stance in response to pragmatic face-threatening acts, Hernández-Campoy's (2016) volume on sociolinguistic style, and Rickford and McNair-Knox's (1994) work with Foxy, an African American girl in Oakland, during a longitudinal study of identity performance. That research which is referenced in-text is meant only as a representative example of this broad pattern in recent scholarship.

there is still the possibility of an automatic language style, or “script” that can be employed to avoid drawing attention. Thus, the results of any single case study cannot be treated as exemplary of anything but the context in which it is found. However, continued accumulation of results may begin to paint a broader picture of those norms that are universalizable.

2.2 Style in the Spanish-speaking world

While recent research in Spanish sociolinguistics has begun to employ second- and third-wave approaches, the community-focused Labovian framework remains important, in large part due to the strong impact of the dialectological tradition in the history of European linguistics. One of the earliest systematic treatments of dialectological variation in Spanish comes from Tomás Navarro Tomás, who documented phonetic inventories across the Spanish-speaking world in the first half of the twentieth century. His was also the first linguistic work to preserve vital social information about varieties like Andalusian Spanish in a robust fashion that has provided a useful point of comparison to modern linguists (as pointed out by Cano-Aguilar [2006]). In his studies, he determined that the political borders of Andalusia did not, in fact, perfectly reflect the dialectal variation that both extended into other provinces and was absent from certain parts of the region (Navarro Tomás et al., 1933: 226). His work continues to receive citation for establishing dialect boundaries and providing exhaustive phonetic descriptions of geographical variation. His goal was not to record specific speaker tendencies, however, or to analyze community variation quantitatively; instead, his approach is concerned with the description of sounds in a documentarian and impressionistic fashion. This tradition led, in part, to the continued popularity

of the Labovian sociolinguistic approach, which most closely approximates the framework of Tomás and historical dialectology.

Branches of the Spanish government have continued to support Labovian research throughout the 1990s, from Carbonero's (1992) study of speech trends in Jerez de la Frontera to Moya-Corral and García's (1995) look at the speech of urban and suburban Granada and García-Marcos' (1990) analysis of the Granada coastline near Motril. These examples serve as a small sampling of the many studies focusing on sociolinguistic variation in urban centers in Spanish, both in book-length publications following in the footsteps of Labov's (1994, 2001) work in New York, and in articles analyzing socially-predicted variation.

As Schilling-Estes (1998) and Eckert (1989) wrote about individual variation and communities of practice, the Labovian tradition remained strong in Spanish. For example, in an examination of the Spanish spoken in the suburban Madrid town of Getafe, Martín-Butragueño (1995) studies /s/ aspiration using 8500 tokens from a representative sample of speakers. Following the Labovian paradigm, Martín-Butragueño (1995: 17) describes style as static, collecting over 5000 tokens of /s/ that "proceden del estilo de conversación" (proceed from the conversational style), and over 3000 tokens "del estilo de preguntas" (from a question style). The two types of style, defined by the conversational setting rather than audience or individual concerns, are considered alongside linguistic and social variables, including speaker gender and immigration status. The study finds style to be predictive of variation, but cannot offer further insight into the social meaning of /s/ aspiration. Samper-Padilla and Hernández-Cabrera (1995) write about the primacy of linguistic variation over the social with relation to /s/ production in the Canary Islands. They only mention style to explain that their data uses a "semi-informal" one, without further explanation. While they observe variation in their 8400 tokens from 24 speakers,

the linguistic rather than social variables are found to condition usage, a reality that may stem in part from the variation buried by the lack of nuance in their conceptualization of style.

In the early 2000s, the work of Bell (1984), Eckert (1989), and Milroy and Milroy (1985) began to gain traction in peninsular Spanish sociolinguistics.[5] To a large extent, Juan Manuel Hernández-Campoy and Juan Antonio Cutillas-Espinosa served as an early locus of second-wave analysis in the peninsular context, with their work further developing over time into a third-wave approach. Cutillas-Espinosa and Hernández-Campoy (2006, 2007) provide an early example of applying Coupland's (2001) ideas into Spanish sociolinguistics. Their descriptions of Murcian Spanish establish it as a variety perceived by its speakers as having covert prestige, requiring that regional variation be based in audience perceptions and an understanding of speakers' performative gestures.

Cutillas-Espinosa and Hernández-Campoy (2006) analyze the production of a Murcian radio presenter and 20 call-in listeners' speech to determine how closely the presenter is following the community norms established by the Murcian listeners. Interestingly, the authors reject Audience Design, as well as the Speaker Design model, saying that the radio host (1) did not try to sound like his audience members, and (2) that while there are clearly performative decisions involved in the speech performance, this is not sufficient to explain the observed behavior. Instead, Cutillas-Espinosa and Hernández-Campoy argue that, in certain performances, there are "scripts" (i.e., what Schilling-Estes (2013: 15) refers to as the "automaticity" of "less self-conscious" style) that serve as an unmarked manner of projecting acceptable and expected identity. In a follow-up study, Cutillas-Espinosa and Hernández-Campoy (2007: 139) include data from an interview with the radio presenter, finding that the speaker's qualitative discussion about the Murcian dialect

[5] Villena-Ponsoda (1996) does make a social network-based inquest into convergence and divergence in Málaga prior to this timeframe, but this appears to be an outlier that still shows trappings of first-wave approaches.

aligned with his medial use of Murcian variants, maintaining the local markers of /s/ deletion (e.g., casas 'houses' [ka.sæ:]) and consonant reduction (e.g. cansado 'tired' [kan.sao]), while also utilizing standard forms of /l/ and /r/ without deletion or variable production. Their conclusion, which sets aside the aforementioned theoretical frameworks as problematic in their unidimensional and universal treatment of style, argues instead that style must be treated as a multidimensional concept, to be set in context alongside the relevant social and linguistic norms of the place where it is performed. These norms, in turn, define the permitted "script" of speech, especially public speech, and help to define what sorts of deviations are permitted (i.e., what style-shifting resources and moves might be available to speakers).

This series of articles seems to respond to Coupland's (2001) work with the Cardiff radio DJ, breaking from previous dialectological peninsular theory. While these articles cite many English sociolinguists, they are some of the first in the peninsular tradition that give few nods to Spanish linguistics, outside of regional atlases and the authors' own previous work. It would make sense, then, that the multidimensional conceptualization of style they propose serves as a break from the Spanish literature, with its focus on attention paid to speech, as well as an extension of third-wave sociolinguistic theory. Despite the forward-looking nature of these articles, there is little sign that Hernández-Campoy and Cutillas-Espinosa's nuanced view of style had an immediate impact on Spanish sociolinguistics.

Over time, other peninsular Spanish researchers have adapted second- and third-wave methodologies into the Iberian context. In the style of Milroy and Milroy (1985), Melguizo-Moreno (2010) contrasts the local prestige forms of the city of Granada and a small town a few miles past its suburbs, Pinos Puente. Through an analysis of the Andalusian sibilant pattern of *seseo* (i.e., where /s/ → [s] and /θ/ → [s]), he finds that this variant, though common in western

Andalusia, has little prestige in Granada in Eastern Andalusian Spanish (EAS). However, the phenomenon is used by speakers with higher social status from Pinos Puente as a means of showing belonging to their rural home and solidarity with their audience. These findings can be seen not only as a show of solidarity, employing community-enforced norms, but also as a means for individuals to opposition themselves to urban spaces through an intentional performance of rurality. Almeida (2019), on the other hand, researches Canary Island Spanish using an ostensibly Labovian perspective of style – either formal or informal – while also discussing style-shifting in a fashion reminiscent of third-wave theory. She describes speakers mixed use of regional and normative forms as an interdialectal "hybrid," contributing to a dual identity construction that shows speakers to have access to both the standard and regional variety. Overall, she identifies a negative social evaluation of Canary Island Spanish, while determining that it also is the recipient of positive social attitudes, suggesting that it possess a degree of covert prestige. Speakers thus have access to a hybrid repertoire of both regional and normal forms with covert and overt prestige, respectively, which in turn provides tools for identity construction. This study is not the only one to take a third-wave conceptualization of style while still maintaining many of the trappings of Labovian analysis, such as the organization of macro-social factors and focus on community-level norms. There is a simultaneous move in Spanish sociolinguistics to begin considering individual identity work, while also continuing to embrace community-based dialectological classification.

With that said, there is still considerable evidence to suggest that the move toward third-wave theorization (and beyond) is underway in the Spanish context. Alongside Hernández-Campoy and Cutillas-Espinosa, who have contributed to sociolinguistic theory both within and beyond the Spanish context, Juan Andrés Villena-Ponsoda has done significant work with social networks, community-based accounts of variation, and ethnographic studies. Villena-Ponsoda

(2013) develops relational social networks to map dialectal variation in Málaga, including elision of /x/ and *heheo* (i.e., /θ/ → [h]), within the family and community spheres. By focusing on the way that individuals diverge from regional norms or else choose to maintain them, he finds individual identity, as expressed through style, to be the key factor in describing variable production. In this discussion, he also draws on geolinguistics, a sociolinguistic adaptation of dialectology that sees geographic spaces as contributing to variability, with urban hubs controlling the perceived prestige of variants based on speakers' relative distance to those areas (e.g., a rural speaker would likely be more compelled to converge with, diverge from, or maintain the norms of nearby Granada than those of Seville or Madrid). This study shows the variety of norms surrounding a speaker, and centers identity construction. Villena-Ponsoda and Ávila-Muñoz (2014) consider similar data in Málaga, but further claim that there must be a tripartite approach to variation, focusing on national, regional, and vernacular-local norms. Only through an analysis of all three waves of sociolinguistic study, including community stratification, social network norms, and an interpretation of individual performance within these networks, can claims be brought forward regarding style, social identity, and performative moves.

Within American Spanish academia, Bucholtz (2009) uses a sociopragmatic perspective to analyze indexical stance and gender identity connected to the use of the Mexican slang term *güey* 'dude.' While the term was used in previous generations as an insult, it has become a marker of solidarity, gender, stance, and address among immigrant Mexican-American populations in the current generation. Through a qualitative analysis of conversations between high school students, Bucholtz (2009: 165) argues that slang, which is grounded in local norms, serves as a type of bricolage that combines with "prosody, gesture, posture, clothing, topics of discourse, and material objects" to lead to stance accumulation of coolness, which in turn links to masculinity. Under her

view, stylistic choices are not directly 'indexing' a social category – instead, they are part of a more complex process of identity creation that, in turn, can be compared to pre-existing social categories. This perspective also cautions against employing claims that make direct associations between variants and social meaning, such as "*güey* is a stylistic indicator of masculinity," in favor of a more nuanced approach, in which variants do not perform identity work on their own, nor are they directly associated with social meaning (e.g., the use of *güey* could be sarcastic, ironic, or come from an outsider like a parent or teacher – these are all unique situations, and the meaning of usage differs in each case). Most importantly, context is vital: it is only in combination with other social bricolage that a coherent identity can be constructed.

Both perceptual and attitudinal studies of speech are also growing in popularity in peninsular sociolinguistics, allowing a focus not only on speaker identity creation, but also the perception of this identity work on listeners. For example, in a perceptual study in Asturias, Barnes (2016) looks at social values of the code-switched variety Amestáu, which combines Castilian Spanish with the Asturian language. Respondents classify the code-switched variety as stigmatized and rural, but also perceive it as demonstrating solidarity, geographical belonging, and covert prestige. Mixing a qualitative and quantitative approach, Barnes discusses both the audience perspective of certain sociolinguistic variants and the agentive behavior available to speakers, who can craft an individual identity that is neither rural nor standardly urban. Focusing on four television presenters from Andalusia, Molina-Ortés (2020) emphasizes the importance not only of micro-focused case-studies of production and perception, but also of integrating modern social media into this approach, performing a perceptual analysis with data from Twitter. This study is grounded in both second- and third-wave approaches, looking at how presenters' identity moves are interpreted by online audiences. Based on these perceptions, Molina-Ortés (2020: 25) classifies

speakers based on standard, Andalusian, or exaggerated hyper-vernacular usage of variants, and determines that speakers with a coherent identity projection are perceived most positively. Even speakers who regularly employ regional norms are seen positively by listeners, whereas dialect-switching is critiqued as "bipolar" and even disloyal to one's region. In this way, the speaker's performative goals are combined with audience responses, showing that consistent stylistic positioning is viewed more positively than channel-appropriate variation.

2.3 The future of Spanish sociolinguistic style

As more than fifty years of sociolinguistic research shows, there is no single best way to approach stylistic variation. Speech style is an entity with close ties to a single context and setting, is associated with prestige norms identifiable to both the speaker and interlocutor, and truly requires ethnographic study to understand. Third-wave case studies of sociolinguistic variation can contribute to our understanding of style in a given context, but they cannot make strong claims about overall community norms or generalize regarding the style-shifting behavior they observe. Instead, a multitude of approaches is necessary to begin constructing an idea of how variation takes place in identity work across manifold contexts. This suggests that there remains considerable work to be done in Spanish sociolinguistics: while community stratification has been well-documented across a variety of urban and rural spaces, there remain many exciting ways to document and study variation at the small group and individual level.

One of the first changes necessary to address this gap is for the Labovian paradigm to be set aside, at least in part, among Spanish sociolinguists in order to place a greater focus on speaker

agency. There are signs that this is beginning to happen; take for example Hernández-Campoy and Cutillas-Espinosa's (2012) third-wave text on style in public speech, or Chappell and Barnes' (in preparation) edited issue of the journal Languages focusing on third-wave sociolinguistic variation in the Spanish-speaking world, as well as several other dissertations and articles that focus on variation below the community level (e.g., Cruz-Ortiz, 2019; Pollock & Wheeler, 2022).

Next, just as second- and third-wave studies began to address community variation from a new perspective in English, there is a need for these same studies to return to parts of the Spanish-speaking world that have gone underserved for decades, updating and filling in gaps in linguistic knowledge. This is not only the case for understudied varieties in Spain; the Spanish of much of Central and South America, outside of certain major urban areas, is little discussed outside of linguistic atlases and dialectological accounts dating back to the early twentieth century. Díaz-Campos and Hernández-Campoy's (in preparation) edited sociolinguistic encyclopedia on varieties of Spanish shows just how sparse recent linguistic research in these areas has been, and how infrequently they have been subjected to any form of sociolinguistic study, much less second- and third-wave research specifically. In one example, Pollock (in preparation) discusses the Spanish of Panamá, describing sociolinguistic variation from the 1940s to 1990s, over two decades out of date at the time of writing. He points to the need not only to update our understanding of community variation across the country, but also to examine how this rich bilingual context, with influences of English, creole, Spanish, and indigenous languages, provides speakers with just the resources necessary for performative identity work. At the moment, sweeping claims about variation and usage are all that exist in the linguistic record for many varieties, both peninsular and American, which suggests a wide range of opportunities for the next generation of linguists.

Part of the dearth of research stems back to issues of access. Though Hernández-Campoy and Cutillas-Espinosa (2012) and others have adopted and propagated the study of sociolinguistic style in certain parts of the Spanish-speaking world, oftentimes it is the case that scholars working under a second- and third-wave framework do not have access to the regions that have long gone underexamined. Molina-Ortés' (2020) study of stylistic variation offers an answer for future scholars: perceptual and production data can be collected through social media sources like Twitter, allowing for access to distant communities through minimally-invasive and publicly-available sources. While not perfect – rural sites are still much less accessible than urban ones in computer-mediated communication (CMC) – linguistic analysis of CMC offers at least a starting point to begin examining variation in novel ways, and among speakers who have long gone unheard. This process must, at the same time, be combined with careful ethnographic study, however, so as to fully capture and understand the cultural and social realities of the community under examination.

The geolinguistic and regional identity research of Villena-Ponsoda (2013) points to a similar problem raised by studies in CMC: there is considerably more research, across the Spanish-speaking world and likely across many other languages, of urban centers than rural speech. Regan (2017a/b, Huerta), Melguizo-Moreno (2010, Pinos Puente), and Barnes (2016, Gijón) serve as examples of studies that have shed light on spaces that had previously received little attention, but they are in the minority. Villena-Ponsoda's (2013) description of cities as centers of overt prestige and towns as centers of covert prestige demonstrates that speakers are affected by national, regional, and local norms. In English, Eckert (2000) sees rural and urban spaces in Detroit as providing students with a diverse package of bricolage to bring into speech situations. Rickford and King (2016) discuss African American Language (AAL) in the courtroom, where speakers of

AAL lose credibility based on the stigma associated with their variety, which combines with attitudinal responses from juries to cause credible evidence to go unheeded. However, given the urban-focus of much linguistic research, especially in the peninsular context, little is known about how urban sociolinguistic style and social meaning influences speech in rural Spanish spaces.

Yet another concern related to style involves generalizability and comparison. Hernández-Campoy and Cutillas-Espinosa (2006) argue that linguistic analyses of style must be taken up in their specific context, whereas Eckert (2008) and Bucholtz (2009) show that concepts like stance accretion and bricolage can apply, at the level of theory, to various contexts. However, there is, as yet, little move to examine style cross-linguistically or through meta-analysis to determine how identity moves translate across language and cultural contexts, if at all. It is also important to note the increasing dynamicity of speech style over recent decades. Each sociolinguistic wave has lasted around two decades, suggesting that a new understanding of style and sociolinguistics is already underway. The study of perception and attitudinal research, as well as public speech, have all gained support in recent years. For the sake of comparability, it is also important, where possible, to bridge the gap between waves, interweaving understandings of style and theoretical approaches to allow future research to continue to find relevance and comparability in our current approach.

Iterating on our operationalization to sociolinguistic style is an important part of addressing some of the foundational aspects of the field of sociolinguistic inquiry. Key to the field has been the documentation of the parole or performance of speakers, finding meaning through social and linguistic variation in the complex phenomenon of human speech. As Eckert (2008, 2016) has argued, style holds a central role in our understanding of social behavior, including language. If sociolinguists across all languages were to evaluate style as the multidimensional entity Hernández-Campoy and Cutillas-Espinosa (2006) and others have argued it to be, examining

shifting behavior using theoretical frameworks drawing from all three sociolinguistic waves, the results would be two-fold. First, by indicating the linguistic options available to speakers performing identity work and providing a record of community norms with which later perceptual and production research could compare, our understanding of language variation and change would become more nuanced. Second, a field-wide push to understand stylistic variation would allow theory to be tested and refined, both for universal claims and for our predictive ability, helping to usher forward more precise descriptions of social behavior.

2.4 The sociophonetic analysis of public speech

In the pursuit of understanding sociolinguistic style, several key theoretical and practical research projects related to English public speech were influential in defining and establishing second- and third-wave understandings of the concept. Bell's (1984) work with New Zealand radio hosts and Coupland's (2001) study of a Cardiff radio DJ both rely on public-facing speech to advance an understanding of variation with relation to audience and speaker goals. Sharma (2018), applying the LFI analytical methodology, examines the minute-by-minute variation in the speech of a style shifting Indian-American media personality. Linguists have also studied the speech of major television personalities, such as Oprah (Hay, Jannedy, & Mendoza-Denton, 1999, 2010). These are but a representative sample of recent examinations of public speech. This focus on public speakers, whose output is easily accessible and available across various contexts, allows researchers to use attitudinal and quantitative approaches to identify and track innovative stylistic patterns in these communities of practice.

Studies of public speech in Spanish have also yielded interesting results. In Latin American Spanish, Flores (2014, 2017, 2018) has examined the production of the Chilean dialectal variables /t͡ʃ/ and /tɾ/ by radio hosts, finding variation to have both socioeconomic and gender-based implications that explain complex differences in variable production. Meanwhile, in the peninsular context, Hernández-Campoy and Jiménez-Cano (2003) show that politicians speaking on the radio use sociolinguistic variation similar to what is found in the communities to which they belong, allowing for the analysis of both real- and apparent-time trends. This contextual "script" that public speakers follow, Hernández-Campoy and Cutillas-Espinosa (2007: 149) argue, allows them to build "an acceptable image and identity," giving weight to the expected norms that guide speakers' sociolinguistic behavior. More niche still than public speech is the sub-category of political speech, among which seem to exist speakers who have greater freedom to set aside scripts in favor of individualized identity work. Not all such speakers aim to embrace expectations: after all, more so than radio and television hosts, politicians need to stand out in a crowded field and appeal to voters.

2.4.1 Sociophonetics and political speech

As a subgenre of public speech, the linguistic analysis of political speech provides access to individuals who spend significant time performing linguistically for audiences to achieve specific identity goals. Often times, audiences have certain expectations they bring to the table when listening to political speech, which cause politicians to face consequences based on their linguistic successes (e.g., media praise, winning an election) and failures (e.g., media condemnation, losing an election). Speech that meets the expectations of supporters in an

electorate may contribute to stance accumulation, helping politicians to develop an image as local or worldly, as reliable or educated, or as a variety of other identifiers that contributes to their brand. Meanwhile, speech that over- or under-shoots may lead a different type of identity to solidify in the minds of voters, branding the speaker as perhaps uneducated, pretentious, or disingenuous. Importantly, voters do not represent a monolith of perception; instead, politicians must constantly juggle perception to ensure that they are perceived as they would most like to be, based on ever-changing social value associated with manners of speech.

Considerable theorizing has gone into the sociolinguistic and sociological validity of examining politician behavior. Hernández-Campoy and Jiménez-Cano (2003) establish the validity of analyzing politicians in radio speech, showing that speakers use sociolinguistic variation found in their community, although they are subject to different social norms than the average speaker. With respect to partisanship, Feldman (2003) argues that political attitudes can serve as a shorthand for individuals' underlying values, making political party affiliation a socially-salient membership signifier for communities of practice. Passionate members of national partisan groups often make decisions based on emotional grounds, Huddy (2013) argues, further giving politicians the power to steer political identity of a party through language. As a result, politicians help define and shape a community's values, granting them a degree of power that not all public speakers possess.

While certain commentators and media hosts may hold power comparable to politicians due to their influence on popular culture, many public speakers' capacity for broad societal change is limited. Oftentimes, speakers in the public realm are restricted to a script constructed of societal norms and expectations, what Schilling-Estes (2013: 15) refers to as "automaticity." Considerable research using third-wave conceptualizations of style have examined political figures such as

members of U.S. Congress (Hall-Lew, Starr, & Coppock, 2012), 2000s-era U.S. politicians (Podesva et al., 2015), Queen Elizabeth (Harrington, 2007), former UK Labour Party leader Ed Miliband (Kirkham & Moore, 2016), and peninsular Spanish politicians (Hernández-Campoy & Cutillas-Espinosa, 2013; Pollock & Wheeler, 2022). These studies show that change can both occur over time as speakers' identity and linguistic standards shift (e.g., Harrington, 2007; Holliday, 2017), and that partisan and individual differences allow politicians to differentiate themselves from peers (e.g., Hall-Lew et al., 2012; Kirkham & Moore, 2016). As a result, while politicians face similar cultural and linguistic norms that establish expectations for their speech, research suggests that they have greater latitude to set aside automaticity and make sometimes risky agentive decisions in the name of adhering to their identity construction goals.

In the following two sections, I examine fourteen articles and one dissertation, published since 2007, that focus on the linguistic agency of politicians from a third-wave sociophonetic perspective. Six examine a U.S. American English context, while three others discuss UK English, one looks at Bulgarian, and six delve into peninsular Spanish. The following two sections examine first studies focusing on English and Bulgarian, then those in the Spanish context.

2.4.2 Sociophonetics and political speech in English and Bulgarian

To begin, the most fruitful context for analysis of political speech over the last fifteen years seems to have been American English, with various examples of third-wave style being analyzed among a host of political speakers. This includes examinations of partisan variation in U.S. Congress (Hall-Lew, Coppock, & Starr, 2010, 2012), a pair of case studies on coronal stop deletion

(CSD) in former President and First Lady Barack and Michelle Obama's speech (Holliday, 2014, 2017), and a dual study of production and perception of released /t/ by six major 2000s-era politicians (Podesva et al., 2015). These six studies exemplify a variety of third-wave sociolinguistic approaches, from the examination of specific communities of practice among politicians in a single setting, to speaker case-studies with an ethnographic twist, to studies of attitudinal responses to political speech.

First off, through an examination of speeches given in the House of Representatives in 2007 to voice support for President Bush's proposal to send more troops to Iraq, Hall-Lew and colleagues (2010, 2012) find partisan divisions relating to identity performance. The variant at hand here is the <a> vowel from the words Iraq, Iraqi, and Iran, which the authors describe as having two allophonic productions: [a], the positive, "European" production, and [æ], the derogatory production. Hall-Lew et al. (2010) had previously found that no social division other than political party could accurately predict speakers' use of the two variants: regardless of their stance on the war and liberal beliefs, Republicans were most likely to use [æ] (e.g., [ai.ræk] or [i.ræk]), while Democrats were favored [a] (e.g., [ai.rak] or [i.rak]). Hall-Lew et al. (2012) analyze variable production over time during this session of Congress, finding that Democrat rates of [a] use increase over time, whereas what [a] use existed among Republican falls off early as differences are established. This finding permits a qualitative analysis to determine how topic influences variant production, leading the authors to suggest that the [a] allophone is associated with voters' evaluation of speakers as elitist, "Washington insiders," reflecting a "foreign" sound more common in Europe. As a result, Republicans avoid it as a means of emphasizing their own local American identity that indexes both patriotism and party solidarity. Here, style can be seen to operate at the party level to establish a phonetic "script" of sorts that influences usage.

In a case study comparing the speech of former U.S. president Barack Obama with that of his wife, former first lady Michelle Obama, Holliday (2017) finds marked differences in usage. Barack's coronal stop deletion patterns more closely mirror Modern Standard US English (MUSE), while Michelle trends more closely with African American Language (AAL). Both speakers, however, demonstrate stylistic variation that prevents them from perfectly mapping onto either MUSE or AAL. Holliday (2017: 479) describes Barack and Michelle's CSD rates as "informal," following guidelines set out in Guy's (1991) exponential model of morphological constraints, a pattern that helps to contribute to their apparent approachability on public television. To explain Michelle's stronger AAL patterns, the author suggests that her role as First Lady causes her to act as a sort of link between the president and the populace, requiring her to perform likeability and down-to-earth-ness. These two political figures develop a coherent identity through consistent linguistic performance throughout their time in the White House, the data shows, resulting not only from agentive behavior, but also their respective sociolinguistic backgrounds, while at the same time employing CSD variation in a meaningful way to construct identity.

The final study from an American context, by Podesva et al. (2015), examines coronal stop release tendencies by George Bush, Hillary Clinton, John Edwards, Barack Obama, Nancy Pelosi, and Condoleezza Rice. The authors consider previous correlations between released /t/ and "articulateness," analyze production in speeches, and gather attitudinal perception data using adjectival continua (e.g., friendly-unfriendly, inarticulate-articulate). The authors find perceptions to vary based on the position of released /t/ in a word, suggesting that simple classification of a variant as indexing a complex attribute like "articulateness" is reductionist, and that speech evaluation is subject to nuance that requires study of the context, such as the ethnographic approach taken by other sociolinguists (e.g., Eckert, 2000). Nonetheless, one key finding by Podesva and

colleagues relates to the meaning that variants possess. As Eckert (2008) argues for bricolage, it is the entire package that matters, rather than any single unit. Released /t/ serves as a socially salient marker, but it requires more context for listeners to explicitly connect it to a single social meaning.

The second most-frequent context for the study of English political speech comes from the UK. The first, a case study of Queen Elizabeth, examines longitudinal change in a single speaker. The next, examining a former UK party leader, exemplifies style-shifting based on speech setting. The final article examines members of the Scottish National Party, determining how partisanship manifests in vocalic variation in parliament.

Both of the first two articles serve as case-studies, focusing on changes in speech norms within a single individual. First, spanning a fifty-year period at the time of his writing, Harrington (2007) used Queen Elizabeth's Christmas speeches to examine how a well-documented, gradual shift in Received Pronunciation (RP) manifested itself in her speech. Data showed that Elizabeth's vowel height and backness shifted to mirror RP norms over time, documenting how speaker-specific changes occur long after L1 acquisition has ceased. Rather than change over time, for Kirkham and Moore (2016) the focus is on change based on one's audience. The authors examine /t/-glottalling and verbal variation in the speech of the former UK Labor Party leader, Ed Miliband. They find that Miliband's phonetic variation differed based on the ideological orientation of his audience, showing an awareness of ideological and linguistic differences between listeners.

Next, in a study of the interaction between Scottish English and RP, Hall-Lew, Friskney, and Scobbie (2017) investigate the performative goals of members of the northern Scottish National Party (SNP) as evidenced through their production of the CAT vowel. Members of the SNP were found to eschew the Southern English low vowel system and RP, while members of the Scottish Labor Party (SLP) converged toward this norm. Given the ideological goals of the SNP –

to secede from the UK – and the lack of a southern branch of the nationalist party composed of RP speakers (as existed for the SLP), the authors argue that these politicians are less likely to converge. By using the more northern variants, they were able to symbolically align themselves in opposition to southern values and unity. Voters likely also expect these politicians to represent Scottish identity, rather than embracing southern norms. The study shows that political parties can serve as communities of practice, explaining variation more reliably than regional factors.

Finally, examining a Bulgarian political context, Kementchedjhieva (2016) combines a sociolinguistic and pragmatic analysis that considers style-shifting as code-switching. The author analyzes production of the Slavic vowel (Ѣ) in both a recorded speech and casual conversation between the Bulgarian prime minister and his electorate. Kementchedjhieva finds regional variation in face-to-face communication, while standard variants are favored in scripted speeches. This is a rare study that draws on speech data from a private context not intended for a national audience, examining conversations with small groups of voters, and shows that politicians can differ considerably when in and out of the limelight. This study is also noteworthy, in that it represents the sole third-wave analysis of style outside of English and Spanish, further suggesting the merit of expanding the scope of sociolinguistic analysis to include novel and understudied contexts to continue developing our understanding of identity construction.

2.4.3 Sociophonetics and political speech in Peninsular Spanish

Within the peninsular Spanish sociolinguistic context, Juan Manuel Hernández-Campoy and Juan Antonio Cutillas-Espinosa's examination of María Martínez, a former president of Murcia, serves

as one of the first third-wave analyses of political speech (Cutillas-Espinosa, Hernández-Campoy, & Schilling-Estes, 2010; Hernández-Campoy & Cutillas-Espinosa, 2010, 2013). Hernández-Campoy and Cutillas-Espinosa (2010) analyze five speech contexts with varying levels of formality, attempting to measure style across a variety of genres, including interviews, press conferences, parliamentary debates, investiture speeches, and statements hearings. Martínez's speech is compared to that of male and female Murcian politicians with a similar socioeconomic status, as well as male politicians from outside of Murcia and male community members from the middle and working class. In their production study, they describe Martínez's use of regional variants as "hyper-vernacular," since she uses regional variables, including the weakened coda /s/ and intervocalic /d/ as well as elided final /r/ and /l/, and consonantal assimilation, more frequently than any other politician, most closely resembling male Murcian community members. She maintains this level of production across four years, using the most regional variants when speaking in Murcia (i.e., 49.4% standard variants in Murcia versus 64.6% in Madrid). This high frequency of stigmatized Murcian variants serves to reinforce her regional identity, as well as her association with the working class and socialist ideals.

By performing an interview with Martínez and including perceptual evaluations of her speech from the time of her presidency, Hernández-Campoy and Cutillas-Espinosa (2013) follow up on their previous study with a more nuanced look at her identity work and its perceived success. In the interview, Martínez identified a connection with the Murcian working-class, having ostensibly grown up as a part of it, and was attempting to show solidarity with that community. Local newspapers and politicians describe Martínez as uneducated and unskilled at public speech, while at the same time branding her hard-working and trustworthy, which the authors identify as a typical part of the "schizophrenic" of Murcian identity that both rejects and embraces the covert

prestige indexed by the regional variety. Hernández-Campoy and Cutillas-Espinosa (2013: 96) argue that this over-use of Murcian variants is a performative overshoot by a speaker not completely familiar with the norms of the variety, while at the same time becoming part of a successful performance of authentic regional identity, allowing her "to project an image of a working-class person who achieved positions of great responsibility and, in spite of that, remained working-class." In this series of articles, the authors provide a roadmap for study of political discourse, as well as results that can generalize to female and minority speakers accessing progressive political identity beyond the Murcian context.

Interestingly, Hernández-Campoy and Cutillas-Espinosa (2013) also make the case for comparability of their results across linguistic contexts. They reference a speaker in Labov's (1972) research on style in New York City, Steve K., who insisted on his membership in a stigmatized variety, despite producing language more comparable to the standard. Martínez's speech appears similar, with her move toward greater normative variant production in Madrid, and over-use of working-class variants in Murcia in a type of reversed cross-over pattern. Instead of the typical Labovian case of cross-over, in which middle-class speakers unfamiliar with upper-class norms overuse prestige variants and hypercorrect, Martínez accommodates in the opposite direction, toward the group with lower socioeconomic status with whom she associates (for more on cross-over patterns, see Labov [1972: 245]).

Three other studies, published in the wake of Hernández-Campoy and Cutillas-Espinosa's work in Murcia, have examined sociophonetic regional variation in other Spanish autonomous communities. Molina-Ortés (2021) examines three former presidents of Extremadura, finding that regional variant use across three speech contexts differed considerably by speaker. Although politicians were not found to be hypervernacular in their application of regional variants, like

Martínez, the extent to which inter-speaker variation distinguished the three individuals suggests that politicians employ language in highly individualized ways that reflect strategic political goals.

Cruz-Ortiz (2019) comes to a similar conclusion in a historical look at variation in political speech in Andalusia. Intervocalic /d/ was found to be elided in only 5.8% of cases across 32 former members of the Andalusian government, dating from 1923-2011, with elision occurring more often when /d/ was part of a suffix or a participle, especially in <-ado>. Overall, Andalusian politicians were found to converge with Madrid linguistic norms while speaking in the capital, providing evidence that both linguistic and stylistic factors influence elision, and that Andalusians modify their phonetic features to conform with national norms. However, as in Extremadura, certain Andalusian politicians use more regional variants intentionally, emphasizing the importance of individual agency and identity goals in regional variant use among politicians.

Finally, applying Hernández-Campoy and Cutillas-Espinosa's methodology to Andalusian Spanish, Pollock and Wheeler (2022) examine how a former president of the region used regional variation as a means of navigating southern peninsular indexicality in scripted and unscripted speech. The authors determine that intervocalic /d/ and coda /s/ elision increased in scripted speeches late in Susana Díaz's presidency, deviating from her peers. They draw on third-wave sociolinguistic research to discuss the dynamic nature of gender performance, extending it beyond a fixed Labovian category (e.g., Eckert, 1989; Podesva, 2007; Zimman, 2017). As a woman in a position of power, Díaz is shown to receive scrutiny and negative evaluations in social and news media, influencing her linguistic decisions. In the particularly salient case of intervocalic /d/ elision in the participial ending <-ado>, Díaz is more extreme than politicians and community members of both genders, eliding almost categorically. The authors argue that the politician uses this context as a means of reflecting her militant status in the Socialist party, and as the first female president

of Andalusia. Díaz's self-representation both proved successful enough to merit broad support from constituents, while at the same time being the subject of mockery perceived hypervernacularity.

All told, the growing number of studies on political speech since 2019 suggests that the topic may gaining broader interest among Spanish linguists. The interesting findings that have built on Hernández-Campoy and Cutillas-Espinosa's original studies reveal a wealth of variation and style-shifting behavior that both follows from theoretical research, while at the same time offering insight into real-world examples of variation that emphasize the numerous implementations in the process of identity work.

2.5 Research Goals

In the face of the current conceptualization of sociolinguistic style, this dissertation aims to address several issues raised in this chapter. While theoretical definitions of the term continue to gain nuance, uncovering and exploring further social variation in language production, there are also gaps that merit further attention. First off, linguistic analyses of public speech have been vital to the development of sociolinguistic theory, and promising findings suggest that political speech offers just as much rich breadth of variation for study as public speech. Additionally, there is a need to continue advancing third-wave theoretical approaches, such as those used by Zhang (2005) or Snell (2018), to improve the generalizability of results and theoretical findings beyond a single speech context.

Following that, the sociolinguistic variable of political party has been shown to have great predictive power with relation to variation in certain contexts, but has seen very little application as of yet in the field. Additionally, despite some of the problems inherent with a Labovian conceptualization of speech communities, which do not reflect nuanced social differences, there is still a need in modern sociolinguistic research to understand broad community norms in order to best contextualize and understand individual and small-group variation. Finally, advancing the study of style-shifting offers the chance for further enlightenment related to the broader phenomenon of language variation– the accumulation of information about individual variation leads, in turn, to an understanding of how social meaning gains prestige and influence linguistic change.

Of the aforementioned studies that examine style-shifting and identity performance in political discourse, only three were published since 2018. Furthermore, articles dealing with large groups of speakers tend to be co-written (e.g., Cutillas-Espinosa et al., 2010; Hall-Lew et al., 2010, 2012; Hall-Lew et al., 2017; Hernández-Campoy & Cutillas-Espinosa, 2010, 2013; Podesva et al., 2015; Pollock & Wheeler, 2022) or presented as a dissertation (Cruz-Ortiz, 2019), a likely side-effect of the amount of data analyzed to present a reliable picture of multi-speaker variation across various speech contexts or time periods. This dissertation aims to provide a model and point of comparison for future research, offering up a three-stage process that can be more easily replicated, in part or in whole. The political speech publicly available across a variety of digital media sources serves as a vast trove of data, waiting to be examined. Hopefully, this project can provide a road map to others interested in exploring it.

The multidimensionality of style as a concept, which Hernández-Campoy and Cutillas-Espinosa (2006) refer to, leads to numerous questions about the generalizability of style-shifting

research in the broader field of linguistics. While theoretically-focused research has been able to draw out results that ring true across language contexts (e.g., Bell, 1984; Coupland, 2001), there is a need for more comparative research to show how findings from any single community of practice or individual can provide similarly meaningful results. I argue, as a result, that there is a need for replication and comparative studies that offer comparisons to pre-existing work, allowing for an increased chance of connecting individualized quantitative findings to the point that they can contribute to theoretical understandings of concepts like language, agency, and identity.

In part as a means of offering a resolution to the question of generalizability, it may be possible to reuse pre-existing social classifications that have not been introduced to sociolinguistic research in a rigorous way. While it may not be possible to reconcile Eckert's (1989, 2000) "burnout" or "jock" with Zhang's (2005) "smooth operator," Hall-Lew et al. (2010, 2012) and Hall-Lew et al. (2017) have determined similar effects of political ideology across two English-language contexts an ocean apart. Despite ostensibly stronger regional bonds, individuals in both the U.S. and U.K. government have based stylistic choices in specific contexts on partisan matters. In order to identify these social groupings, the variants are key – as Sharma (2018) finds using LFI analysis, regional variant use is changeable from moment to moment without speech, with a concentration of regional features emerging in specific moments. Some Andalusian politicians leverage variation to emphasize identity differences, as Pollock (2022) and Pollock and Wheeler (2022) find, while others mainly stick to the baseline "script" established for the public speech venue (Cutillas-Espinosa & Hernández-Campoy, 2006). Political affiliation, especially as it increasingly becomes central to social identification and differentiation, is a compelling variable emphasized in this dissertation.

While third-wave linguistic approaches emphasize important variables in speech, including speaker agency and identity, an inevitable downside to many implementations of this approach involves the common problem of case studies: low numbers of tokens and speakers. In addition to problematizing generalizability across linguistic contexts, this also raises questions as to how a given speaker's behavior is cast against the backdrop of their broader speech community. While static interpretations of macro-social and stylistic factors need not be revived from the Labovian approach, the robust and balanced samples of speakers from across a community can serve an important role in bolstering third-wave individual-focused research (i.e., the other half of Johnstone's [2000] call – for both case-studies and traditional sociolinguistic research). Acquiring this data would help provide further context to case studies, rather than relying on previous studies that may not reflect the realities of the specific context under consideration. For example, in Holliday's (2017) examination of Barack and Michelle Obama, a community examination might include both speakers from the politicians' native Chicago, as well as other politicians and public figures who use MUSE and AAL as a means of determining how novel (or normal) the Obama's variation is alongside their peers. In the current dissertation, in order to account for broad and narrow variation, I combine a community-wide examination of sociophonetic variation with LFI and attitudinal analyses to not only determine individual identity work, but also understand and contextualize it with relation to the behavior of peers.

Finally, the main goal of sociolinguistic research is often to better understand language variation and change. Public speech is not just suggestive of linguistic norms in a community—given that it reflects the normative "script" expected by community members—it also serves as an important indicator of linguistic shift. As speakers who are regularly recorded, host large audiences, and are aware of linguistic expectations of constituents, politicians are unlikely to be

linguistic innovators of the sort described by Milroy and Milroy (1985). Those speakers who bring novel variants into a community are often network-peripheral, rather than central, giving them access to other varieties and linguistic resources. However, politicians are both vocal and frequently recorded, with access to recorded speech very often available publicly, giving interested linguists insight into shifts among rural communities and in language where regular linguistic analysis is rare. In this way, political speech serves as a sort of linguistic thermometer, providing quick access to a variety of contexts and communities, offering a means of tracking—both in the present and historically—change patterns.

These goals offer some insight into the means by which this dissertation aims to make use of and build off of third-wave sociolinguistic research and conceptualizations of style. The next chapter examines the ways in which production and perception offer differing types of insight into variable production, as well as focusing on specific linguistic phenomena in Andalusian Spanish that merit further examination.

3 Production and perceptual studies of Iberian Spanish

In recent years, sociolinguistic analyses of political speech in both Spanish and English have often approached style-based identity performance from a sociophonetic perspective (e.g., Hall-Lew, Coppock, & Starr, 2010, 2012; Hall-Lew, Friskney, & Scobbie, 2017; Hernández-Campoy & Cutillas-Espinosa, 2012, 2013; Holliday, 2017; Pollock & Wheeler, 2022). This perspective requires an in-depth understanding of phonetic and phonological variation in a given community, such as understanding the differences in vowel height norms between Scottish English and RP (Hall-Lew et al., 2017), or distinguishing between Murcian and NCPS consonant production norms (Hernández-Campoy & Cutillas-Espinosa, 2012, 2013). By calculating how frequently politicians produce regional variants, and placing these results into the context of community speakers and other politicians, sociolinguists make claims about complex issues of identity construction and social perception. However, as Villena-Ponsoda (2013) argues for Andalusian, there are various loci of prestige that influence social meaning, emanating from local towns, nearby cities, and national standards, which analyses of speaker identity work must take into consideration.

Politicians have access to a variety of linguistic variants associated with social meaning, and as previous perceptual work shows (e.g., Hernández-Campoy & Cutillas-Espinosa, 2013; Podesva et al., 2015), their utterances are subjected to differing interpretations by listeners based on a network of attitudinal, emotional, and linguistic reasons. In order to better understand this complex identity-based web of meaning for Andalusian Spanish politicians, this chapter focuses on grounding the dissertation's analyses of production and perception in research on geographical variation, as well as differentiating between production-focused and perceptual approaches to variation. This includes a description of production-based approaches in Section 3.1, including

phonetic variation related to the 10 peninsular Spanish phenomena examined in depth in this dissertation, followed by a discussion of perceptual research in Section 3.2, with a focus on regional identification, previous perceptual research in Andalusian Spanish, and perceptual studies of political speech. The section wraps up with a comparison of production and perception-based approaches, and a statement of the research goals that follow from this research.

3.1 Production research in Andalusian Spanish

The analysis of speech production has been a key part of linguistic analysis in peninsular Spanish since the dialectological work of the early twentieth century (e.g., Navarro-Tomás, 1939), and has been integral in furthering the understanding of both articulatory and acoustic Spanish phonetics. The way speakers produce sounds, and the sounds that they produce, provide insight into the configuration of the speech organs, how individual speakers differ, and what boundaries exist to separate phonemic and allophonic categories. For the purposes of this section, articulatory phonetics focusing specifically on the movement of speech organs will be set aside, favoring Thomas' (2011:17) description of acoustic phonetics as production[6]. Acoustic research favors quantitative measures drawn from audio analysis to often make similar articulatory claims, analyzing cues such as formant height or the range in intensity as a representation of vowel backness or occlusion degree, for example. These findings can then be paired with an analysis of social factors (e.g., trill /r/ closures, reflecting tongue moment, as suggesting reduction processes:

[6] The articulatory perspective, which looks at the physical production of sounds (e.g., using x-ray during vowel production: Delattre, 1969; MRI of Mongolian vowel harmony (VH): Saito & Yurong-Maekawa, 2019; electropalatography of Eastern Andalusian Spanish (EAS) vowels: Alonso, Vicente, & de Zamora, 1950), tends not to be used in sociophonetic research.

Henriksen, 2014; sibilant frequency (Hz) to demonstrate tongue position and track variable production: Gordon et al., 2002).

Production studies, as exemplified by Martínez-Celdrán's (1998) analysis of Spanish vowels in Figure 5, focus on presenting a description of linguistic phenomena through qualitative and quantitative description of measurable acoustic difference. Although production research serves as the foundation of sociolinguistic analysis, employed in the field since Labov's (1972) earliest studies in Martha's Vineyard or New York Department Stores, this approach also requires linguists to theorize about the effect of these variants on community members. Ethnographic methodologies that allow linguists to enmesh themselves in communities and collect qualitative responses about variation may provide some insight into the social value of certain production types. Furthermore, third-wave research that examines speech contexts and individual case-studies in great detail, leveraging a linguist's increased familiarity and exposure to the context to better represent differences in style, goes a long way toward answering questions about identity and speaker agency that community-driven research alone can ignore.

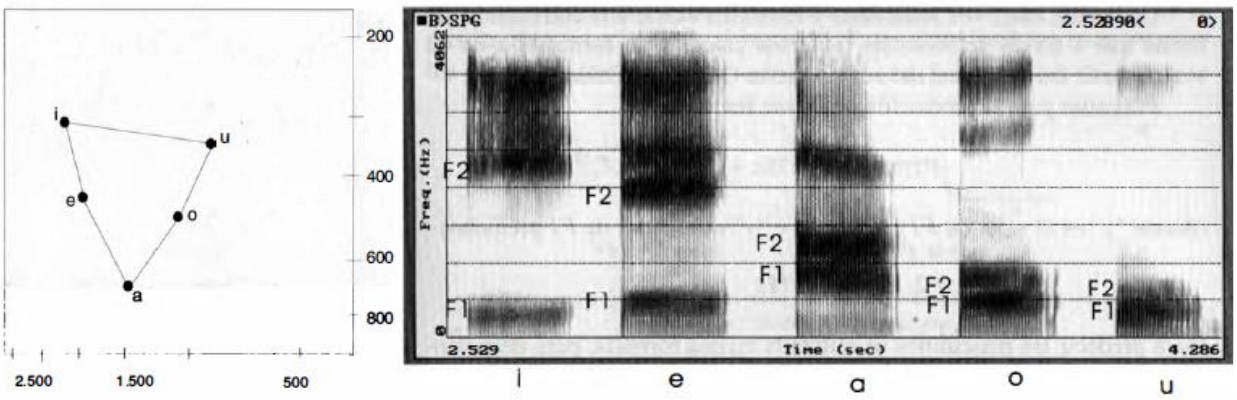

Figure 5: Representation of the Spanish vowel space (Martínez-Celdrán,1998: 42)

One of the major tools used in production analyses, as identified by Thomas (2011: 17), is the analysis of fundamental frequency (F0) for intonational patterns, formant 1 (F1) for height, formant 2 for backness (F2), and other formant values to indicate rounding, rhotacization, and additional variability. Another measure, Center of Gravity (COG), averages out noise vertically across formants, providing a suggestion of tongue position for the analysis of occlusives like sibilants, while a measure like Intensity Range measures the average rise and fall in speech intensity over a determined span, serving as a means of identifying occlusions in sounds like stops and trills. These measures, identified through spectrographic analysis and quantified using tools like Praat, allow linguists to make articulatory claims without requiring more rigorous and time-consuming articulatory methods like MRI or ultrasound (Boersma & Weenink, 2023).

As this dissertation focuses on identifying a profile of regional speech phenomena that can be used to distinguish Andalusian and NCPS Spanish varieties, ten processes are discussed below in greater detail, sub-divided into four categories: production studies of sibilants (i.e., affricate fronting, syllable-final /s/ elision, *ceceo*, and *seseo*), of voiced stops (i.e., intervocalic /d/ elision), of liquids (i.e., variable production and reduction of /r/, /ɾ/, and /l/), and of vowels (i.e., vowel harmony processes). The Andalusian Spanish (AS) phenomena, at times subdivided into both western (WAS) and eastern (EAS) regions, are positioned in opposition to normative NCPS productions.

3.1.1 Sibilant production

The moniker sibilant refers to fricatives like [s, z, ʃ, ʒ], as well as affricates like [t͡s, t͡ʃ, d͡ʒ] that employ said fricatives. Sibilants have greater acoustic energy "at a higher pitch than other fricatives" (Ladefoged & Johnson, 2011:175). Several acoustic measures are effective in examining and distinguishing these sounds. Gordon, Barthmaier, and Sands (2002), in a consideration of voiceless fricatives across seven unrelated languages, find that duration and formant transitions are often unreliable measures to contrast fricatives. Instead, they find that COG and spectral contours are both consistent cross-linguistically in the identification of phonemic distinctions. In a comparison of the sibilant fricatives /s/ and /ʃ/, their examination of languages from the Americas (e.g., Aleut, Apache), Europe (i.e., Gaelic) and Asia (i.e., Toda) shows a tendency for the COG of /s/ to be 400Hz higher than that of /ʃ/ on average. Similar acoustic work analyzing differences in sibilant production has been done in Chinese (Li & Li, 2019), Japanese (Amano & Yamakawa, 2011), Polish (Nowak, 2003), Korean (Shertz, Kang, Kochetov, Kong, & Han, 2015), and Spanish (Pollock, 2022).

While modern Spanish has only two sibilant phonemes (i.e., /s/ and /t͡ʃ/), the medieval Spanish was three times larger, as depicted in Figure 6. These six sibilants underwent separate processes in NCPS (Figure 7) as compared to the variety spoken in Andalusia and carried by colonizers to Latin America (Figure 8). While NCPS experienced reduction, resulting in /x/, /θ/, and the apico-alveolar /s̪/, Andalusian and Latin American Spanish saw variable production of sibilant contrast, leaving only the predorso-alveolar/dental /s̪/. Meanwhile, modern Spanish /t͡ʃ/ derived mostly unaltered from early reductions of the Latin orthographic <ct> cluster, as in *factu>hecho* 'done' (Penny, 2002: 104). Notably, later Latin borrowings were not subjected to the same process of affrication (e.g., *effectus>efecto* 'effect', *luctu>luto* 'mourning').

Phoneme	Classification	Graphemes	Example	IPA
/s/	Voiceless apico-alveolar fricative	*s-, -ss-*	*saco* 'I take'	/sáko/
/z/	Voiced apico-alveolar fricative	*-s-*	*cosa* 'thing'	/kóza/
/ts/	Voiceless dental affricate	*ç, c*$^{(e,i)}$	*caça* 'hunt'	/kátsa/
/dz/	Voiced dental affricate	*z*	*dizia* 'said' (imp)	/didzía/
/ʃ/	Voiceless prepalatal fricative	*x*	*dixo* 'said' (pret.)	/díʃo/
/ʒ/	Voiced prepalatal fricative	*j, i, g*$^{(e,i)}$	*ojo* eye	/óʒo/

Figure 6: The sibilants of medieval Spanish (Hualde & Colina, 2014:151)

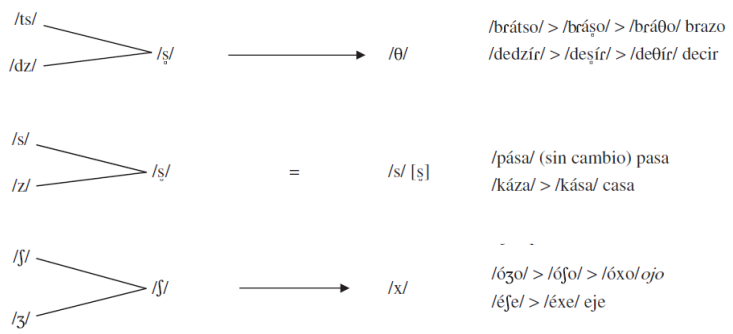

Figure 7: Sibilant changes in NCPS (Hualde & Colina, 2014:153)

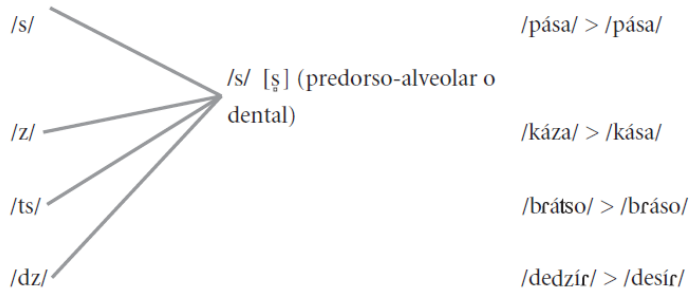

Figure 8: Sibilant changes in Andalusian and Latin American Spanish (Hualde & Colina, 2014:154)

Out of this sibilant system, three linguistic phenomena are selected for discussion in the following sub-sections. First, the Spanish affricate /t͡ʃ/ receives little attention in recent peninsular sociolinguistic literature, in part due to the steady decline in use of the deaffricated fricative allophone /ʃ/, which tends to be found mainly in the speech of elderly rural speakers of EAS (Melguizo-Moreno, 2006, 2007; Samper-Padilla, 2011:116). Following that, syllable-final /s/ deletion is one of the most-discussed topics in Spanish phonetics, with early descriptions hailing back to Navarro-Tomás' work in the early twentieth century. Frequent reduction phenomena include so-called aspiration, or production as [h], elision, and the compensatory process of vowel doubling (Herrero de Haro, 2017b). Finally, the last sub-section discusses the three possible outcomes for /s/ and /θ/ production in peninsular Spanish, including distinction (i.e., where the phonemic distinction is maintained), and two mergers: *seseo* (i.e., where /θ/ → [s]), and *ceceo* (i.e., where /s/ → [θ]) (Alonso, 1951; Samper-Padilla, 2011:115).

3.1.1.1 Affricates

Most discussion of the peninsular affricate in linguistic research tends to focus on the deaffricated variant /ʃ/. Since the 1960s, phonetic and sociolinguistic studies of the affricate tend to refer to a shrinking population of rural, elderly speakers who produce a deaffricated variant /ʃ/, finding this variation across Andalusia in Seville, Granada, and Málaga (Herrero de Haro, 2017b:343; Samper-Padilla, 2011). Melguizo-Moreno (2006, 2007) finds that that [ʃ] is mainly produced word-initially and almost exclusively among older, working class men with little education in Granada. Meanwhile, Villena-Ponsoda (2008:148) argues that the fricative variant

experiences a regional divide across Andalusia: while it still experiences use in southwestern cities like Jerez de la Frontera (83% of cases), it is stigmatized due to its association with elderly rural speech, causing its use to decline precipitously in the north and east of the region. In Granada and Málaga, the fricative production occurs in less than a quarter of tokens. (i.e., Granada and Málaga). Thus, while the fricative allophone is essentially unused outside of Andalusia and it is stigmatized in the west, it still enjoys regional prestige in parts of the west.

There has also been a move in recent years to view deaffrication as a continuous rather than categorical process, with some researchers analyzing periods of frication and closure in affricates to examine change from full affrication to fricative production in a continuous fashion. Díaz-Campos, Cole, and Pollock (2023) use this process in Caracas, Venezuela, finding an ongoing process of retiming that falls short of full deaffrication, but which nonetheless suggests an ongoing change in progress moving toward fricative production.

Another affricate variant, observed mainly outside of Andalusia in the Murcian town of Ricote, involves a voiced affricate allophone. Torrano-Moreno (2017) finds the voiced [dʒ] in this community, ten miles outside the capital of Murcia, mainly in use by older male speakers, reflecting a similar sociolinguistic profile to the fricative production in EAS. Normative production of [t͡ʃ] (i.e., uso estándar) and regional production of [dʒ] (i.e., uso no-estándar) are depicted in Figure 9, showing a steep drop-off in regional variant use among younger speaker groups. Almeida (2019) describes a similar voiced variant in the Canary Islands, which he defines as having post-palatal tongue contact, rather than the pre-palatal contact Torrano-Moreno finds in Murcia. Other than these two studies, discussion of a voiced affricate variant in peninsular Spanish is rare.

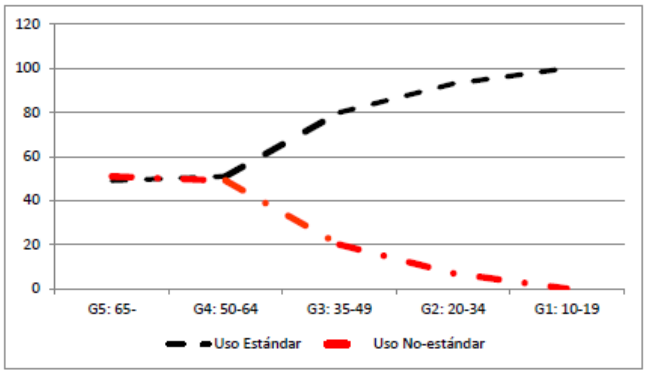

Figure 9: Production of /t͡ʃ/ as [t͡ʃ] (estándar) and [dʒ] (no-estándar) in Ricote (Torrano-Moreno, 2017:16)

Finally, identified in Chile, a third affricate variant involves a process of fronting, whereby the pre-palatal affricate becomes the alveolar [t͡s]. In her study of Chilean radio speech, Flores (2014) compares a fronted alveolar affricate that has a high starting frequency (Figure 10a, around 7000 Hz) with the pre-palatal variant, in which aperiodic noise starts at a lower frequency range (Figure 10b, around 5000 Hz). In a follow-up study, Flores (2018) finds that speaker gender and age influence production, with middle-aged women being most likely to produce the fronted variant. Lexical frequency and speech context also significantly predicted variation, with front vowels and stressed syllables favoring fronting. Use of the fronted affricate is common in Chile, potentially resulting from contact with aboriginal languages like Huilliche, which have a contrastive distinction between /t͡ʃ/ and /ʈʂ/ (Henríquez-Barahona & Fuentes-Grandón, 2018).

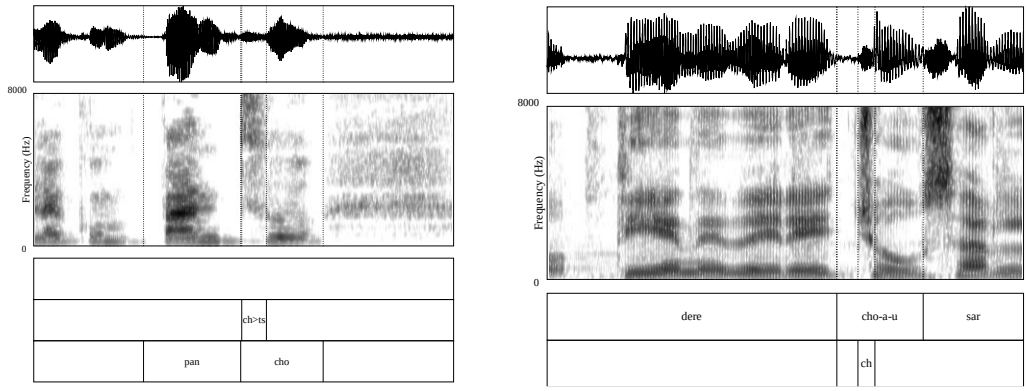

Figure 10a/b: Fronted affricate in pancho 'hot dog' [pantso] and pre-palatal affricate in derecho 'right' [deret͡ʃo] (Flores, 2014:76-77)

/s/ Duration (ms)	/s/ COG	/ʃ/ Duration (ms)	/ʃ/ COG
61.8	4783.4	54.5	3978.4

Table 1: Acoustic correlates of frication in AS affricates [t͡s] and [t͡ʃ]

In an acoustic study of the Andalusian affricate, Pollock (2022) argues that peninsular Spanish also has a fronted alveolar affricate. Similar to the differences in COG between the Chilean examples, he performs a quantitative analysis of the COG and duration of over 3,000 affricates to demonstrate notable differences across affricate productions in AS. These quantitative differences are presented in Table 1, with the higher frequency being assigned to the more fronted variant. After normalizing COG by individual to reduce variation based on vocal track size, Pollock determines that a heightened COG (i.e., the fronted variant) is favored in Madrid (NCPS) and Málaga (EAS), by left-leaning politicians, in formal speech, and preceding front vowels. While language contact may help explain this phenomenon in Chile, it cannot be thought to apply to peninsular Spanish. As such, Pollock argues that these results could follow from a phenomenon described by Torreira (2006), Ruch (2012), and Del Saz (2019) related to the alveolar fricative /s/, whereby it undergoes elision in /st/ clusters, then further affricates to [t͡s]. This phenomenon,

which has also been identified in Andalusian Spanish, suggests a process of analogy between /st/ and /t͡ʃ/, opening up questions about possible neutralization in minimal pairs like hecho ‘fact’ [etso] and esto ‘this’ [etso]. Given that the acoustic differences in Andalusian Spanish are of a similar magnitude to those in Chilean Spanish, where [t͡s] and [t͡ʃ] contrast in social meaning, there are exciting possibilities for continued research to determine differences in social meaning between the forms in Andalusia.

3.1.1.2 Syllable-final /s/

Reduction of syllable-final /s/ is one of the most frequently documented phonological phenomena in Spanish. In an overview of Peninsular Spanish, Samper-Padilla (2011:100) describes three common allophones of /s/: the innovative elided[7] variant [Ø], the intermediate aspirated variant [h], and the conservative retained variant [s]. He maps these three variants onto geographical areas: eastern and western Andalusia are described as using the innovative variant, Northern Andalusia and the Canary Islands use the intermediate one, and NCPS uses the conservative one. Samper-Padilla is not alone: aspiration and elision are often described as characteristic of Andalusian cities like Granada (e.g., García-Marcos, 1990), Linares (e.g., Gómez-Serrano, 1994), and Málaga (e.g., Vida, 2004). Madrid, the political center of the country and a representative of NCPS, has been identified with equal frequency as using the conservative variant, with various studies finding negligible rates of /s/ weakening (e.g., Blanco, 2004; Gil-Peña, 2004; Ruiz-Martínez, 2003).

[7] Herrero de Haro (2017) also suggests that vowel doubling can be a compensatory result of elision, taking the place of the now-absent /s/.

Pollock (2023) examines political speech in southern Spain, providing examples of NCPS (Figure 11) and WAS (Figure 12) productions of the elided /s/ and intervocalic /d/. In the example from NCPS, the speaker produces the plural /s/ morpheme, as seen in the band of high energy near 4500Hz, as well as aperiodicity in the wave form – typical of sibilant production. Meanwhile, in the WAS production, there is a general lack of energy in the spectrogram or aperiodicity in the waveform, other than background noise. It is also worth noting, for the later discussion of vowel harmony, how the two productions of /o/ in *podemos* differ; the first has relatively flat formant values across the production, while the second shows a lowering of F2 and raising of F1 (i.e., backing and falling) and an overall increased duration, suggesting a change based on elided /s/.

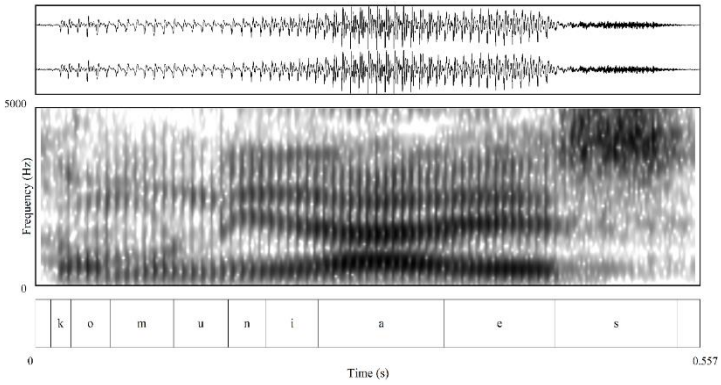

Figure 11: NCPS production of Comunidades 'communities' with word-final coda /s/ production (Pollock, 2023)

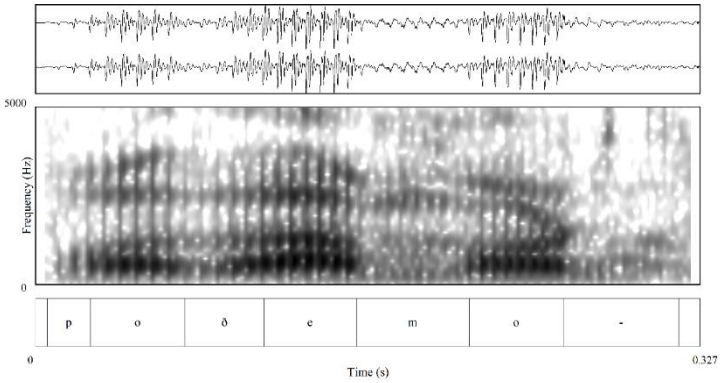

Figure 12: AS production of Podemos 'we can' with word-final coda /s/ deletion (Pollock, 2023)

Examinations of aspiration, sibilant duration, and voicing often focus on geographical (or dialectological) differences rather than social ones. Bishop (2007) finds that manipulated [ahta] with reduced [h] is more likely to be identified as *hasta* 'until' in EAS, while longer [h] leads to identification as *apta* 'apt.' In word-medial contexts in Andalusian Spanish, Gerfen (2002) identifies a compensatory phenomenon where shorter productions of /s/ favor longer following stops. Meanwhile, using a perception test, O'Neill (2005) finds that the breathy phonation that remains after /s/ deletion leads to roughly chance (i.e., 40-55%) accuracy in identifying plurality among participants from both WAS and EAS. He argues that the breathy voice component inherent in vowels prevents listeners from distinguishing between typical vowel voicing and [h], neutralizing the distinction between singular and plural words. In a follow-up study on word-medial contexts, O'Neill (2010) finds that *pata* 'paw' and *pasta* 'pasta' are distinguished based on voicing rather than presence of [h]: Andalusian *pata* is often produced [paða], whereas underlying /s/ disfavors voicing in *pasta* [pata]. In NCPS, Samper-Padilla (2011:106-107) and Ruiz-Dominguez (1997) find that female speakers produce more [s] and less [h] than males, suggesting that [s] holds a more prestigious social interpretation. The prestige of [s] is further supported by Samper-Padilla's (2011) finding that working class speakers produce the least [s], overall, while

younger speakers favor [s]. Older speakers, meanwhile, tend to use more [h], possibly a result of having left the rigors and expectations of the job market.

However, social values of /s/ are not shared throughout all of Spain. In the Canary Islands, young and upper-class speakers produce the aspirated [h] variant more than older and working-class speakers. Molina-Ortés (2015), for instance, studies Merida and Badajoz in the southwest of the country near Western Andalusia. The author finds retention, aspiration, and elision of /s/ present in interviews conducted among speakers, with aspiration being most common, and elision more frequent among elderly speakers. However, in a later analysis of real and apparent time, Molina-Ortés (2018) finds that speakers of the youngest generation now favor the innovative elided variant. These studies provide evidence that /s/ aspiration and elision are socially stratified across much of the Peninsula, while also emphasizing the regional nature of prestige norms and the influence of local perceptions on speech.

Linguists have also examined the preceding context that remains following reduction, identifying certain compensatory processes. In some varieties of EAS, vowel harmony occurs before elided /s/ (Herrero de Haro, 2017a, 2019; Lloret & Jiménez, 2009), while there is evidence that it does not extend to all parts of Andalusian Spanish (NAS: Ransom, 1991:136). Vowel duration has also been shown to change in the wake of elision (e.g., *casa* ‘house’ /kasa/ vs. *casas* ‘houses’ /kasa:/), with an average increase of 25% preceding elided /s/ (Carlson, 2012; see also Navarro-Tomás [1939]).

It is important to note that, in these descriptions of /s/ reduction, little space is dedicated to social and geographic variation. While the phenomenon is mainly confined to AS, few sources provide a more nuanced social breakdown. In his overview of the field, Samper-Padilla (2011:106-107) observes that, overall, (1) women in NCPS tend to follow the prestige standard, producing

more /s/ and less /h/ than men, (2) older speakers across Spain tend to aspirate more, while younger speakers use /s/ more, (3) upper- and middle-class speakers produce more /s/, while working-class speakers aspirate and elide, and (4) the Canary Islands upsets these generalizations, as younger and upper-class speakers aspirate, older speakers retain /s/, and working class speakers elide.

Before wrapping up this discussion of coda /s/, three under-studied phenomena are considered in slightly greater depth, beginning with post-aspiration. Also known as extended VOT, post-aspiration is the process Torreira (2006, 2007) terms "retiming," whereby [h] occurs after voiceless stops that follow an elided /s/. Ruch and Harrington (2014) determine that post-aspiration is longer among younger speakers and speakers of WAS, as opposed to EAS. Listeners from Argentinean Spanish, which does not employ post-aspiration, were additionally able to parse it alongside pre-aspiration in a perceptual task where they identified tokens of [pat:a] as *pasta* rather than *pata*. This suggests that even if post-aspiration is unique to AS as Torreira (2012) describes, it still enjoys intelligibility across varieties. Although WAS, Argentinean, and Puerto Rican Spanish all share syllable-final /s/ reduction, Torreira (2006, 2007) argues that the peninsular variety differs due to an issue of timing: in WAS, the dental closure begins and ends earlier than in other varieties, meaning that, especially word-medially, pre-aspiration is shorter and post-aspiration continues after the following obstruent (e.g., *pasta* as [pa^{h}t:^{h}a]). Torreira (2012) classifies post-aspiration as part of a unique, region-specific articulatory overlap that only occurs in all WAS [hC] clusters. Although the small size of his dataset does not allow him to make divisions based on social categories, he acknowledges the likelihood of social variation and urges future work to determine how it affects post-aspiration.

Another curious process related to /s/, which also occurs in WAS and in the same context as post-aspiration, is a process of affrication, in which word-medial /sC/ clusters become [t͡s]. In a

sociophonetic analysis of variation across age, education, and gender, Ruch (2012: 63) finds considerable allophonic variation in this context: this includes post-aspiration and affrication for the following consonant as well as pre-aspiration and breathy voicing in the preceding vowel. The affricate was favored in the "less-formal" speech style case (22%), as compared to read aloud (13.7%) and word list (6.6%) tasks, and was used most by university-educated young speakers (35.6%), in opposition to older (8.4%) and middle-aged (21.4%) speakers. In a phonetic comparison of WAS and EAS, Del Saz (2019) identifies a novel process whereby both post-aspiration [t^h] and affrication [t͡s] occur in clusters of stop + /t/, unrelated to the presence or elision of /s/ (e.g., *exacto* /eksakto/, Figure 13: [eksat^{h}o], Figure 14: [eksat^{s}o]). This phenomenon is also found to be most common among young speakers. Del Saz argues that, while pre-aspiration duration contrasts point to underlying /s/ in EAS, a neutralization phenomenon is emerging across coda consonants by phonetic analogy to /st/ that shows a divergent treatment of the same context in WAS.

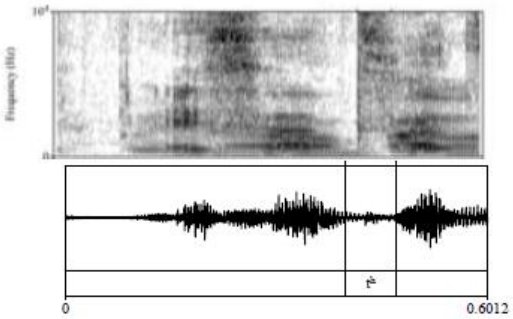

Figure 13: Aspiration in WAS exacto 'exactly' (Del Saz, 2019: 761)

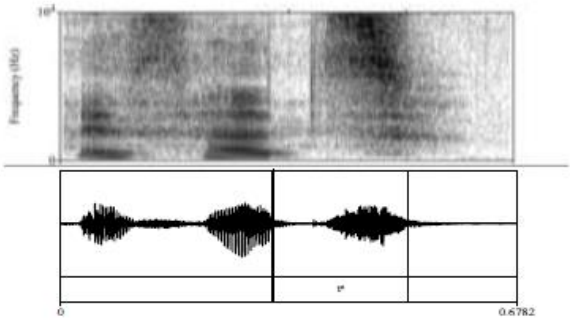

Figure 14: Affrication in WAS exacto 'exactly' (Del Saz, 2019:761)

Finally, the phenomenon of /s/ voicing is often referred to in Spanish linguistic textbooks as an example of assimilation (e.g., Hualde & Colina [2014: 155], in reference to *mismo* 'same' as [mizmo]). However, while assimilation does occur, /s/ voicing is particularly regularized in Madrilenian Spanish, especially in intervocalic position (where it is less likely to voice in other varieties). In an examination of over a thousand cases of intervocalic /s/, Torreira and Ernestus (2012:136) find that roughly a third of their tokens contain uninterrupted voicing, which is also most common in high frequency tokens (e.g., the /s/ in *nosotros* 'we' and *pues* 'well,' before a vowel-initial word, are voiced in over 50% of cases). The authors describe this as part of a larger Madrilenian process of lenition for all intervocalic voiceless obstruents, although a lack of consideration of social categories makes it difficult to determine the sociolinguistic profile of this phenomenon.

3.1.1.3 *Seseo* and *ceceo*

The final two, related sibilant fricative phenomena under discussion in this chapter include two outcomes from a merger of the peninsular /θ/ and /s̊/: the merger to [θ] is known as *ceceo*

[θeθeo], while the merger to [s] is termed *seseo* [*seseo*]. Historically, much of Latin America, the Canary Islands, and Andalusia has used *seseo*. Use of *ceceo* is mostly relegated to southern Andalusia, although it has also been documented in Honduras, El Salvador, and Venezuela (Díaz-Campos & Pollock, submitted). The NCPS normative production, on the other hand, uses the bipartite system known as distinction, which maintains both the apico-alveolar fricative /s̊/ for the grapheme <s> and the interdental fricative /θ/ for graphemes with <z> and <c>. These three sibilant patterns are exemplified in Table 2. One other optional outcome of this merger is alternation, also derogatorily referred to as *confusión* 'confusion' (Dalbor, 1980), which has been noted in much of the literature of the last century. This refers to an Andalusian tendency for speakers often raised in the south and educated in the dominant NCPS variety to alternate between [θ] and [s] without reflecting the underlying phone. In addition to this macro-geographical variation, there is also regional and individual alternation between these production norms (Villena-Ponsoda & Ávila-Muñoz, 2014).

	Pattern	<*sapo*> 'frog'	<*cielo*> 'sky'	<*zapato*> 'shoe'
a.	*Distinction*	[s̊apo]	[θjelo]	[θapato]
b.	*Seseo*	[sapo]	[sjelo]	[sapato]
c.	*Ceceo*	[θapo]	[θjelo]	[θapato]

Table 2: Spanish sibilant patterns (Díaz-Campos & Pollock, submitted)

Several additional details complicate this picture of sibilant patterns. In geographical terms, there have been some descriptions of use of both *seseo* and *ceceo* outside of Andalusia, although this is rare (Molina-Ortés, 2012: Badajoz; Villena-Ponsoda, 2007: Murcia; Sempere, 2014: Murcia). Additionally, Villena-Ponsoda (2013: 188) references an additional, stigmatized sibilant merger outcome, often associated with working-class speakers from rural areas, known as *heheo*,

in which /θ/ is produced as [h] (e.g., *hacer* 'to do' /aθeɾ/ → [aheɾ]). Regan's (2017a) discussion of distinction in WAS also raises the interesting possibility that a de-merger is in progress in Huelva, based on falling rates of formerly prestigious *ceceo* and the tendency of highly-educated and female speakers to favor distinction.

The possibility of a de-merger, based on falling frequencies of *seseo* and *ceceo* use, seems not just true for WAS, but rather to be a phenomenon occurring across Andalusia (Samper-Padilla, 2011). *Seseo*, which was more frequent historically in WAS, has seen a considerable drop off in urban spaces in favor of distinction (Villena-Ponsoda, 2008), and *ceceo*, formerly frequent throughout the southern portion of Andalusia, is now mainly only found in rural speech and can serve to situate a speaker in contrast to urban identity (Melguizo-Moreno, 2009). Despite the shift toward distinction in recent decades, Villena-Ponsoda (2008:156) argues that *seseo* continues to be a "part of the western regional standard, due to [its] social prestige," even as its usage has declined in the east. Díaz-Campos and Pollock (submitted) perform a comparative analysis of *seseo*, *ceceo*, and distinction across Andalusia based on previous studies. They show that *seseo* is frequent in Seville, Córdoba, and Jerez de la Frontera, while distinction is most frequent in other urban centers, and *ceceo* is only maintained in the rural eastern Pinos Puentes (Table 3).

City	*Seseo*	*Ceceo*	Distinction
Seville (WAS)	87%	6%	7%
Jerez (WAS)	44%	47%	9%
Huelva (WAS)	6%	23%	71%
Córdoba (NAS)	51%	0%	49%
Málaga (EAS)	31%	25%	70%
Díaz-Campos & Pollock:			
Málaga Liberals	12.5%	3.1%	84.4%
Málaga Conservatives	1.9%	0%	98.1%
Granada (EAS)	40%	5%	55%
Pinos Puente (EAS)	3%	70%	17%
Immigrants to Granada (EAS)	12%	36%	66%

Table 3: Seseo, ceceo, and distinction in AS (Díaz-Campos & Pollock, submitted)

With respect to social variation, there is evidence that factors including education, age, gender, and prestige play a role in variable production. The increasing education standards in Spain since the end of Franco's regime in 1975 coincide with the stark reduction in *ceceo*. *Ceceo* is most common among speakers with little to no formal education, *seseo* is more common among speakers with a primary or secondary education, and distinction occurs among speakers with university degrees (Ruiz-Sánchez, 2017: Seville; Ávila-Muñoz, 1994: Málaga; Melguizo-Moreno, 2007a: Granada; see also Carbonero, 1985; García-Amaya, 2008; Regan, 2017a). Related to education, age is also an important factor in sibilant pattern use, with older speakers that enjoyed less access to education being more likely to employ *ceceo*, while younger speakers tend to use greater amounts of distinction (García-Amaya, 2008; Marrero, 2016). The move to distinction also seems to be a change conditioned by speaker gender, with women being more likely in Granada (Moya-Corral & Sosinski, 2015; Melguizo-Moreno, 2007a) and Seville (Santana-Marrero, 2017) to employ distinction. Despite these changes toward distinction, *seseo* still enjoys overt regional prestige in Andalusia (particularly in Seville), while *ceceo* is seen as less normative and more rural (Dalbor, 1980), a fact that causes some rural communities to favor it (e.g., Melguizo-Moreno, 2007b: Pinos Puente).

Interestingly, as a closing point, while there has been some acoustic analysis of peninsular /θ/ (e.g., Iribar & Túrrez, 2015: Basque; Regan, 2017b: Huelva; Regan, 2019: Lepe), many studies referenced above, especially those focused on geographical and social variation, spare little thought for a description of acoustic distinctions between /θ/ and /s/. A descriptive comparative study (e.g., in the vein of Gordon et al., 2002) is sorely needed for peninsular varieties, especially

given the tendency for some impressionistic coding to attempt to identify gradient production (e.g., [θ^{s}], [s^{θ}]; Dalbor, 1980), which suggests that acoustic differences may exist to determine differences in production norms, allowing for a more repeatable methodological approach.

3.1.2 Voiced stop production in Andalusian Spanish

Elision of intervocalic /d/ is perhaps the most wide-spread phenomenon under consideration in this dissertation, having been documented across Spain (as well as throughout the Americas). In Madrid, Gil-Peña (2004) and Ruiz-Martínez (2003) describe that normative production of intervocalic /d/ as represented by a dental approximant allophone [ð̞]. Meanwhile, elision occurs throughout the country, although it is conditioned by linguistic, stylistic and social factors.

In a study of sociolinguistic interviews from Madrid, Molina-Martos and Paredes-García (2014) find retained [ð̞] to be the most frequent variant, while elision occurs frequently in the suffix <-ado> '(past participle) -ed.' The rate of elision is influenced by linguistic factors including lexical tonicity, syllable count, grammatical category, functionality, lexical diffusion of the word, and phonetic context. Elision was more frequent when speakers talked about personal topics, during exposition and narratives, in adjacent pairs, colloquial speech, near the end of the interview, and among men and members of the youngest and oldest generations. Other sociolinguistic research has shown that women tend to be more conservative in Madrid for this phenomenon, including in the Salamanca district (Gil-Peña, 2004) and in Alcalá de Henares (Blanco, 2004). Men, on the other hand, tend to favor elision, as documented in Linares (Gómez-Serrano, 1994) and Toledo (Molina-Martos, 1998).

Andalusian regional norms tend to favor elision as well, as seen in various parts of the region, including Jaén (Moya-Corral, 1979) and Linares (Gómez-Serrano, 1994). Villena-Ponsoda and Moya-Corral (2016) argue that elision is also advancing in Granada, with Málaga following. As in Madrid, rates of intervocalic /d/ elision are affected by speakers' age, gender, and socioeconomic class. Elision is favored by young speakers in Málaga (Villena-Ponsoda, Montesinos, Muñoz, & Cervantes, 2011) and Granada (Samper-Padilla, 2011:114). This may result from speakers' level of education, as younger speakers grew up in a post-Franco Spain with broader access to higher education (García-Amaya, 2008).

Socioeconomic status is an influential factor for /d/ elision in both Andalusia and NCPS. In both regions, speakers with less formal education and from lower socioeconomic backgrounds tend to elide the most (e.g., Ruiz-Domínguez, 1997; Ruiz-Martínez, 2003; Moya-Corral, 1979; Gómez-Serrano, 1994). Studies in Jaén (Moya-Corral, 1979), Linares (Gómez-Serrano, 1994), El Hierro (Pérez-Martín, 2003), and La Jara (Paredes, 2001) have found high rates of /d/ elision among speakers from lower socioeconomic classes.

With respect to political speech, in a study of historical Andalusian politicians' speech in Madrid, Cruz-Ortiz (2019) found that /d/ was elided in only 5.8% of cases, and that elision occurred more when /d/ was part of a suffix or a participle, especially in <-ado>. It also occurs more in cases with post-tonic stress and without lexical diffusion. Elision of /d/ occurred more in unscripted than scripted speeches, women were more likely to elide than men, and more recent time periods favored elision over those during Franco's dictatorship. Andalusian politicians converge with Madrid linguistic norms while speaking in the capital, Cruz-Ortiz concludes. This provides evidence that Andalusians modify their phonetic features to conform with national norms,

and more broadly this section establishes intervocalic /d/ elision as an innovative, vernacular form that exists throughout Spain but is most common in Andalusia (Samper-Padilla, 2011).

3.1.3 Liquid production in Andalusian Spanish

Grouped together in this section are phenomena related to the variable production and elision experienced by the Spanish liquids /ɾ/, /l/, and /r/. Variable liquid production and contrast neutralization have been observed between the rhotic tap and trill (Bradley, 2001), between /h/ and /r/ in Puerto Rican Spanish (Delgado-Díaz & Galarza, 2015), in Venezuelan realizations that include approximants with and without occlusions (Díaz-Campos, 2008), between /l/ and /ɾ/ in Seville (Ruiz-Peña, 2013), as well as in high-frequency words in Andalusia, where trills are more likely to reduce to taps (Zahler & Daidone, 2014). Comparative research by Pollock et al. (in press) identified similar processes of trill /r/ reduction underway in the Caribbean Spanish varieties of Caracas, Venezuela, and Caguas, Puerto Rico. Overall, these phenomena are common throughout the Spanish-speaking world, although often associated with southern peninsular varieties of Spanish.

Sociolinguistic research in southern Spain suggests that these liquid processes correlate with speaker socioeconomic status, age, gender, and regional identity. Ruiz-Peña (2013) describes variable /l/ and /ɾ/ production, in which /l/ is produced as an [ɾ]-like sound and vice versa, as being influenced by social differences; in particular, working-class Seville speakers affiliated with socialist political parties and proud of their Andalusian identity were most likely to demonstrate variable production. Henriksen and Willis (2010) and Henriksen (2014) both find that older

speakers and men tended to produce more innovative reduced variants in Jerez, Spain. In Málaga, Zahler and Daidone (2014) find that speakers under 35 years of age favor the innovative reduced variant, the sociolinguistic profile for a change in progress. With respect to regional identity, Hernández-Campoy and Cutillas-Espinosa (2010, 2013) situate post-vocalic /r/ deletion and variable liquid production as markers of regional identity in their study of the politician María Antonia Martínez from Murcia. While intervocalic /r/ and /ɾ/ are retained categorically by politicians and community members, the former female president trended with male Murcians in eliding word-final /r/.

Several linguistic factors have also been found to play a role in describing variable production: these include word position (Henriksen & Willis, 2010; Willis, 2006), lexical stress (Henriksen & Willis, 2010; Zahler & Daidone, 2014), and phonetic context (Bradley, 2006; Henriksen, 2014). Henriksen (2014) found phonetic context, in particular, to be predictive of reduction in NCPS: high back vowels (i.e., /u/) favored innovative reduced variants in León and Ciudad Real. In Andalusia, Zahler and Daidone (2014) determined that innovative reduced variants were disfavored in stressed syllables. These authors found that high-frequency tokens in Málaga favored production of the innovative reduced variant, while measures of frequency based on words' phonological neighbors show that words with more neighbors favor the non-innovative trill.

In the north of Spain, several studies have encountered variation for liquids in consonant clusters. In Barcelona, Blecua (2001) determines that tap production in consonant clusters tends to use an approximant allophone without occlusions. In La Rioja, Weissglass (2011) finds both the approximant and an assibilated production one that takes on sibilant properties similar to [ʒ], both produced without an occlusion. Both authors suggest that reduced task "formality" plays a role in

favoring productions without an occlusion, and Weissglass (2011) specifically finds that sonority of the preceding consonant in the cluster conditions allophonic variation in the tap. Hualde & Colina (2014) also identify instances of rhotic segments in consonant clusters being produced, instead, as a trill for those varieties in contact with Basque. Weissglass (2015) finds, in a comparison of monolingual Spanish speakers and Spanish-Basque bilinguals, that monolinguals tend to favor approximant production, whereas bilinguals are more likely to produce taps and, in one speaker's case, a trill in consonant clusters.

Reduction processes have also been documented in Spain. Henriksen (2014), for example, identifies cases of reduction of the trill /r/ in Spain. Furthermore, in contexts of word-final /r/ deletion, Herrero de Haro (2017a) finds that patterns in the formants of preceding vowels are sufficient to allow listeners in a perceptual task to distinguish between elided /r/, elided /s/, and word-final vowels (e.g., contrast is maintained between /amos/ [amo], /amor/ [amo], and /amo/ [amo]). This movement toward reduction maps onto similar processes underway in a variety of Romance languages, including Portuguese, where the alveolar trill has become a voiceless velar fricative (Parkinson, 1988).

3.1.4 Vowel production in Andalusian Spanish

Acoustic studies of Spanish vowels, performed by Martínez-Celdrán (1995) and Marín-Gálvez (1995), offer an idea of average formant values across male and female speakers and durations in various speech contexts, respectively. In a comparison of Peninsular and Peruvian Spanish, Morrison and Escudero (2007) find that vowels from the former variety have a greater

duration and a lower fundamental frequency, but do not differ significantly in formant height. These studies use speech data from small sets of speakers, often collected in controlled settings, to establish general production tendencies within the vowel space.

		singular		plural	
		F1	F2	F1	F2
/e/	pretonic	479.75	1830.53	588.66	1664.94
	tonic	476.07	1937.76	579.26	1791.52
	final	489.15	1867.41	568.06	1774.16
/o/	pretonic	482.44	1077.49	578.97	1113.49
	tonic	504.96	1119.66	602.31	1179.86
	final	497.03	1099.99	564.49	1147.82

Figure 15: EAS vowel formant values (Hualde & Sanders, 1995:433)

Since the early days of dialectology, a number of researchers have specifically examined AS vowels in the wake of coda /s/ deletion. Navarro-Tomás (1939) identified a curious case in which these vowels underwent a process of doubling, as well as experiencing a change in backness and height. Alonso, Vicente, and de Zamora (1950) use electropalatography and x-ray to examine /s/ deletion, finding adjustments in jaw width of preceding vowels. Using data from EAS and NCPS Spanish, Hualde and Sanders (1995) argue that Navarro-Tomás was actually referring to a tense-lax vowel contrast that results not from plural deletion, but rather is the result of a historical AS tendency where final vowels in non-plural contexts raised (Figure 15). Historically, word-final plural /s/ deletion left now-final vowels unchanged, such that they contrasted with the underlyingly final raised vowels. However, vowel raising became stigmatized and fell out of favor, causing non-plural final vowel height to fall, and leading vowels preceding elided word-final /s/ to also fall, maintaining contrast. Herrero de Haro (2017a, 2019) has demonstrated that vowel laxing in EAS is extremely nuanced, with F1 and F2 shifts before deleted /s/, /r/, and /θ/ being different enough to allow perceptual distinction for the vowels /e/ (Figure 16a), /o/ (Figure 16b), and /i/ (Figure

16c). This study, which relies on both perception and production, demonstrates the degree of laxing present in non-underlyingly final vowels.

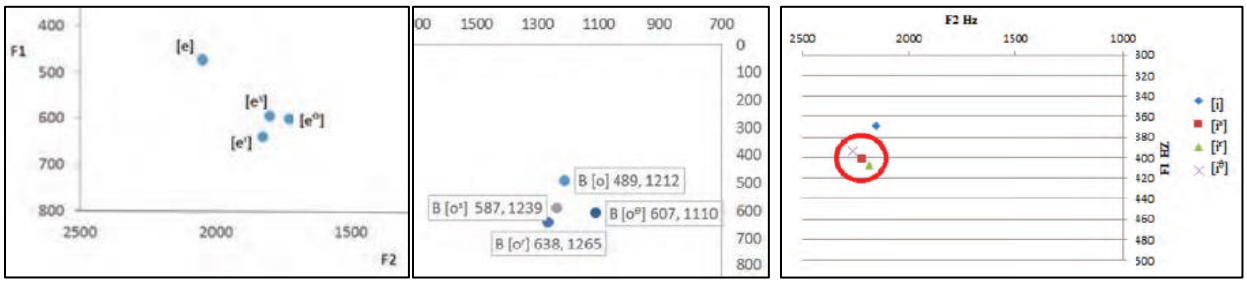

Figure 16a/b/c: Vowel space for /e/ and /o/ and /i/ (Herrero de Haro, 2017a: 99, 2019:12)

Several authors under the generative tradition have also described a process of vowel harmony (VH) resulting from this vowel height change, in which laxness spreads from the right-most vowel leftward within a word, in EAS. There is some disagreement about the extent of VH in the literature, however: some authors describe a type of harmony that occurs with minimal constraint inside of a single prosodic word, while others observe opacity in certain vowels and syllabic contexts. For example, Soriano (2012:297) describes laxing as part of a harmonic effect that takes place across all vowels in plural words with final /s/ deletion. This description proposes a case of total harmony, in which all vowels within a word where the plural morpheme is deleted take on [RTR], regardless of vowel type or phonological context. Her optimality theoretic analysis of the situation relies on the constraint MAX(GESTURE), which argues that coda /s/ deletion leaves an [RTR] feature on the preceding vowel, explaining the origin of spreading. This analysis also provides insight into means of dealing with the tendency for harmonic spread to halt at morpheme boundaries. Meanwhile, Jiménez and Lloret (2007) offer a slightly more nuanced description of acquisition of the feature [RTR], describing an EAS system in which high vowels

(i.e., [i] and [u]) do not lax, and low vowels (i.e., [a]) only lax in final position. Under their view, /s/ deletion leaves behind a [spread glottis] feature attached to [h] aspiration, which in turn elides, leaving the [RTR] feature to spread through vowel harmony. They rely on licensing and anchoring constraints to explain deletion and spreading, as well as tendencies related to syllabic stress.

There is evidence to suggest that two separate types of vowel harmonic processes occur. In a comparative analysis of VH in EAS and a variety of Cantabrian Spanish (i.e., Montañes, which experiences centralization and height harmony), Neumann and Kanwit (2018) separate EAS harmony into that which occurs prior to consonant deletion and that which occurs preceding morphological plural /s/. They argue that, compared with the three other VH processes they consider, EAS harmony can be classified as a highly morphophonological (as opposed to automatic) process. Under this definition, it is morphologically conditioned, restricted to a derived environment, and faces word-level restrictions. Like Neumann and Kanwit, Herrero de Haro (2017a) references a division between "phonological" and "phonetic" laxing, and that laxing indicates plurality.

As a result of VH processes in Andalusian Spanish, there has also been a discussion of the phonemic categories that the variety possesses. In a diachronic description of EAS, Herrero de Haro (2017b) describes the variety as having a vowel system with 8-10 vowels, resulting from doubling and laxing (i.e., /æ ɛ ɔ/), rather than the five vowel Castilian Spanish system (i.e., /a e i o u/). In a comparison of EAS and NCPS, Henriksen (2017) also finds that speakers of the former system have categories for eight vowels (i.e., /a æ e ɛ i o ɔ u/), whereas speakers of the latter have only five. Speakers of EAS were able to distinguish words with laxed or [-ATR] vowels as being pluralized forms of those with tense or [+ATR] vowels (i.e., *jefe* 'chief' [jefe] versus *jefes* 'chiefs' [jɛfɛ]), whereas speakers of NCPS were slow to identify the difference between plural and singular

tokens and responded at closer to chance. On the whole, closely related to sibilant reduction phenomena, VH is an interesting hyper-regional phenomenon that merits study under an identity-based framework to determine if it possesses social meaning in AS.

3.1.5 Overview of Andalusian production studies

Despite the covert prestige that the local varieties of WAS and EAS possesses, numerous studies in recent years have suggested that there is a process of standardization underway in southern Spain moving toward either a unified regional AS (e.g., Villena-Ponsoda, 2008, 2013), or even in the direction of the normative NCPS variety (Regan, 2017b; Valeš, 2013; Torrano-Moreno, 2017). In urban areas of southern Spain, *ceceo* is on the decline (Regan 2017a/b), voiced affricates are devoicing (Torrano-Moreno 2017), and deaffricated fricatives are affricating among younger generations (Samper-Padilla, 2011). This presents a complex indexical field of linguistic meaning to speakers, who have exposure to local urban and rural norms, regional WAS or EAS tendencies, and more overtly prestigious NCPS forms.

Based on the discussion presented in the preceding sections, sibilant phenomena, including affricates, syllable-final /s/, *seseo*, and *ceceo*, voiced stop patterns, such as intervocalic /d/ deletion, liquid reduction and variable production affecting /ɾ/, /l/, and /r/, and vowel laxing are all expected to serve as representations of AS norms, although varying in frequency between west and east. Unsurprisingly, given typical sociolinguistic profiles of vernacular variants (e.g., Labov, 1972), social meaning associated with many of these variants can be tied to rural, working class, and male identity, with *seseo*, vowel laxing, and affricate fronting serving as three possible exceptions. As

politicians select variants most fitting for their identity goals, this social meaning serves to influence their choice, as does the automaticity of the speech script available to public speakers. While *ceceo* may be available for some speakers, its increasing association with uneducated and working-class speech may make it too socially-charged for many speakers to employ. After all, while they may aim to show solidarity with their constituency by using regionalisms, politicians also have good reason to avoid the mockery aimed at those who over- and mis-use regional variants, which can affect public perception (e.g., Hernández-Campoy & Cutillas-Espinosa, 2013; Pollock & Wheeler, 2022). In these cases, speakers can receive negative classifications in the press or from peers, as Pollock and Wheeler show for Susana Díaz, which can in turn recast politicians not as indexing regional solidarity, but rather as being undereducated and worthy of mockery.

While previous research offers an idea of the sociolinguistic profile for several of these phenomena, it is less than complete. In many cases, as with the discussion of coda /s/ elision, scholars have focused more on geographical differences and linguistic contexts for reduction than on the differences between speakers, or particularly the social value associated with aspiration and elision. Perceptual studies, which could begin to address attitudinal evaluations, would also help to determine the exact acoustic criteria for differentiating between phones like /s/ and /θ/ or /ɾ/ and /l/. Much of the work that looks at these differences, as in consideration of *seseo*, *ceceo*, and variable liquid production, relies on impressionistic coding. Linguists using these approaches often admit to the difficulty of separating out edge cases, as variation often presents itself on a continuum of productions that are gradual rather than categorical. In the case of sibilant patterns in particular, while Regan (2017a, 2017b, 2019) has done a thorough job of analyzing the progression of distinction in Huelva, little descriptive acoustic and perceptual data exists to determine exactly what intermediate forms might exist in AS or how they might be represented in a spectrogram.

Flores (2014), for example, found a considerable range of COG for [t͡s] and [t͡ʃ] productions in Chile, which in turn differed from Pollock's (2022) findings for the same sounds in AS. Perceptual data can provide an important step toward moving past impressionistic coding and applying linguists' intuitions to production results, instead accessing community listener impressions and relying on the acoustic cues and boundaries that they use to evaluate speech.

3.2 Perceptual research

Auditive phonetics, targeted at the perception of sounds by listeners, provides a type of information that has only recently come to be targeted in sociophonetic study (e.g., the edited volume by Chappell [2019] on sociophonetic perception). Perceptual research is less concerned with sound production than with the way listeners identify and react to said sound. Reetz and Jongman (2008) describe one type of perceptual phonetics as focusing on equivalence categorization, determining which cues are used to group some sounds and contrast others, helping establish category boundaries with measures like formant height (in vowels), duration (in fricatives), or voice-onset time (VOT; in stops). This approach is exemplified, for example, in a perceptual instrument used by Solon (2015) to determine what causes L2 Spanish learners' lateral /l/ production to be seen as non-native. By comparing tokens of the English velar [ɫ] and the alveolar Spanish [l], the author discovers how F2 and duration influence non-native perception, allowing her to identify a boundary between Spanish-like and English-like productions of /l/.

In his description of auditive phonetics, Thomas (2011:55) describes perception as a comparatively underused tool for sociophonetic analysis, resulting from the production-centric

focus of much phonetics research. Perception need not only focus on phonemic and phonological boundaries. It can be studied with identification tasks, open-ended written tasks, as well as forced-choice questions, ratings, attitudinal surveys, and category goodness tasks. The goal of these tasks is to determine something about the classification system of the listener, whether it pertains to phonological categories (e.g., category goodness tasks), social ones (e.g., attitudinal tasks), or phonetic awareness (e.g., rating the native-ness of L2 Spanish sounds).

Studies of perception help to fill in the gaps left behind by production research. Investigations in the Labovian tradition, for example, can often present findings regarding social and linguistic variation without being able to justify these realities. Determining which social groups produce a variant often does not explain if or why that variant is associated with a certain group or identity (e.g., young, educated males, students at a certain school, etc.). In studies of production, which often use impressionistic methodologies, replication can become difficult, and linguists' intuitions are often necessary to categorize audible stimuli, bringing preconceived ideas and biases into research.

Sociolinguistic research, especially in the third-wave, has pushed in recent years to better understand social categorization, often combining perception and production. Using an ethnographic approach, Zhang (2005) tracks the local attitude of Beijing "yuppies" in interviews toward specific phonetic variants. One local persona, which the author describes as the "smooth talker," is seen as being used by speakers with a "heavy r." Zhang finds that this evaluation of a "heavy r," in an examination of speech production, results from rhotacization of syllable-final sounds and lenition of retroflex obstruents in onset position. This study shows that production and perception research can be combined, offering a better picture of how speakers produce language in response to the evaluative norms of their communities

Perceptual sociolinguistics (e.g., Campbell-Kibler, 2009, 2010) presents a means of accessing listener attitudes and evaluations that can serve an important role in more recent waves of sociolinguistic thought. Knowing how listeners react to speech decisions is important in an analysis of a speaker's behavior, as is understanding the degree of salience and perceptibility a certain variant possesses in a community and how contextual information shapes meaning. Studies using perceptual designs have shown, time and again, that speakers possess a great aptitude for identifying meaningful linguistic differences and acknowledging, if implicitly, the role of overt and covert prestige in speech evaluation. Thus, while production research is important to determine trends in a community's speech, the growing urge to examine individual style in third-wave sociolinguistic research also demands a quantitative means of examining linguistic variation at the production level with the attitudinal responses of members of the community to distinguish what indexical goals a speaker may pursue through use of certain variants.

3.2.1 The perception of regional variation

Understanding how social meaning is evaluated helps to track its progression within a community. In recent decades, research into speech perception has shown how important a role social variation plays in comprehension and cognition. Pisoni (1997) finds that listeners use voice information in perception, allowing them to identify social realities about speakers with the same facility as the identify aspects of one's appearance. Features can index geographical origin, age, socioeconomic background, and other social aspects. From a production perspective, we see how this plays out in Eckert's (2008) assessment of speakers at a high school, where she argues that linguistic features

serve a similar function to bell-bottom jean width in distinguishing students' belonging to one of two major cliques. However, not all social factors can be easily drawn from speech. While Lass et al. (1976) find that the gender of an unknown speaker could be identified with 96% accuracy, regional identification is much less successful, with Clopper and Pisoni (2004) finding that participants were only able to identify about a third of American English varieties when hearing sentences with eight to ten regional variables. Regional identification was more successful among participants who had greater contact with other varieties and geographic areas, suggesting that social mobility plays a role in perceptual identification.

Understanding how individuals make use of linguistic resources can also reveal ideological decisions mediated through indexical meaning, as speakers who converge toward certain forms may favor the social meaning and group membership they reference. Several volumes have considered how language attitudes map onto the beliefs, stereotypes, and associations found in speech communities based on dialectological data, linguistic atlases, and folk linguistics (e.g., Preston, 1999; Niedzielski & Preston, 2000; Cramer, 2016). Much of the research into perceptual dialectology asks speakers what they consider to be "correct" representations of their language, and how they feel about other production types with less linguistic prestige. As Preston (1989) finds, informants evaluated standard Modern U.S. English stimuli as being associated with greater kindness and correctness than regional variants, although informants' own variety was seen as more intelligible and correct than other non-standard varieties. Beyond that, Preston's participants also preferred certain regions over others, with stimuli from western and northern states being evaluated as more correct than those in the south.

These perceptual results show a tendency for speakers to describe language varieties better known to them with greater accuracy, and often with stronger opinions, than those at a greater

distance. For example, Preston's (1993) study of communities in the U.S. Midwest finds that participants had the basic ability to categorize English along geographical lines to a certain extent, successfully identifying northern versus southern speech. However, the precise geographical border between the two fluctuated based on whether informants were from the northern or southern reaches of the American Midwest, suggesting that in-group membership is defined based on individual linguistic experience.

Perceptual work in the Spanish-speaking world has shown similar results across a broad geographical divide. In Madrid, Moreno Fernández and Moreno Fernández (1999) found that informants from the city distinguished their variety from others with significant linguistic differences, such as the Canary Islands, Andalusia, and areas of bilingual contact. However, these participants had greater difficulty distinguishing between Madrid Spanish and that of nearby regions. This difference was explained by age. While older speakers from a lower socioeconomic background were more likely to classify non-Madrid varieties as very different from their own variety, younger, educated speakers were more likely to identify northern and central Peninsular Spanish as similar to that of Madrid. The authors suggest that this social difference may come from a generational change in the perception of the Spanish of the capital. These results could also be influenced by an increase in mobility among young Spaniards, who have greater access to travel than members of the previous generation.

Comparisons have also been made between Peninsular and American varieties of Spanish. For example, Díaz-Campos and Navarro-Galisteo (2009) consider the perception of six distinct Spanish varieties by participants from Spain and Venezuela. Informants accurately identified only one in three stimuli. Interestingly, Venezuelans, who had greater geographical proximity to many of the varieties, were considerably more successful than Spaniards in identifying Colombian and

Chilean Spanish, while both groups had success identifying Peninsular Spanish. Sociopolitical events, such as the Chilean dictatorship and instability in Colombia, led to increased immigration into Venezuela that may have increased informants' experience with these varieties. In a follow-up study, Díaz-Campos and Killam (2017) compared the aforementioned results with those of Peruvian speakers, who were best able to distinguish Spanish varieties. The authors conclude that linguistic experience is an important factor conditioning accurate identification – informants with greater contact with more varieties better distinguished among stimuli. This finding, like those of Preston and Moreno Fernández and Moreno Fernández, suggests that regional dialects within a single country, such as the Dominican Republic, may be governed by language experience.

Concerning perceptions of vernacular variants, several studies have shown the connection between social factors, geographical identification, and class-based biases. In Ecuador, García (2019) considers intervocalic /s/ voicing as a marker of the highland Spanish variety, finding that informants associate it with Quito and Cuenca. Evaluations of use of this variant differ by speaker gender; women using voiced /s/ are seen as being less pleasant, younger, and coming from a lower socioeconomic background, while no social meaning was found for male speech, with which it is typically associated. Chappell (2016) finds something similar for intervocalic /s/ voicing in Costa Rican Spanish. Referencing Eckert and McConnell-Ginet (2003), who argue that women tend to use more extreme degrees of innovative or normative variants than men as a means of differentiating themselves, she determines that women are more sensitive to indexical meaning and avoid voicing, which indexes lower social status. Rather than choosing to be more concerned with social capital, Chappell (2016: 372) argues that this is a matter of social necessity – women tend to avoid /s/ voicing when possible because covert prestige, which is afforded to men who use it, does not apply to female voices.

Regional belonging has also been explored in depth to determine how listeners associate variable production with indexicality and regional variation. In a comparison between informants from Cuban and Peninsular Spanish varieties, Loyola and Rafat (2022) find that assumptions were made based on regional belonging, with Peninsular Spanish correlating with higher education and income levels and Cuban Spanish being seen as funnier and easier to understand. Examining Mexican and Puerto Rican Spanish, Walker, García, Cortés, and Campbell-Kibler (2014) consider how word-internal, coda /s/ weakening is employed to identify social meaning. Mexican voices using the [s] allophone were rated as much more prestigious by both Mexican and Puerto Rican listeners, while that difference fell in relation to [h] for Puerto Rican voices. The authors determine that listeners consider the speakers' regional variety when selecting the appropriate social meaning to associate with speech.

The indexical meaning inherent in linguistic variation holds social meaning that can be employed in creative ways, which, in turn, reinforces the indexicality of these forms in an ongoing play of meaning creation (Eckert, 2008). Abercrombie (1967) emphasizes that indexical properties of speech differ from lexical ones, involving individual factors, group membership, and changes in emotional state, rather than purely linguistic information. This includes traditional sociolinguistic factors like gender and age, as well as regional variation and stylistic shifts related to emotions. Rather than being fixed entities used by certain pre-defined speaker groups (e.g., young speakers, working class speakers, etc.), linguistic variables are part of what Eckert (2008) calls an indexical field of ideologically linked meaning, relying on context and changing over time.

In the process of better understanding indexical meaning and style-shifting, perceptual research documents the evaluative norms of members of a speech community toward specific linguistic resources. Labov (1972: 120) describes membership as stemming back to common

norms of linguistic evaluation and production, meaning that evaluation of forms by community members should help uncover their social meaning. As Zimman (2016: 273) argues, speakers are not necessarily aware of linguistic details, such as the relationship between acoustic correlates and gender, when they make use of these resources. When speakers aim to reach an indexical goal through stylistic decisions—such as aligning with traditional gender norms—they agentively construct a target identity without necessarily being able to identify the linguistic moves they take in the process of doing so. This suggests that even if listeners could not explicitly express social meaning, their immersion in their speech community allows them to identify – through attitudes and other perceptual questions – salient differences in phonetic meaning.

It is also important to note that Andalusians have a positive evaluation of their own variety. There has been effort in recent years to paint AS as a distinct variety from NCPS based on phonetic, grammatical, and attitudinal variation, rather than just being a "bad way of speaking Spanish," as it was identified in previous centuries (Valeš, 2014; Melguizo-Moreno, 2010). This suggests that, in addition to being acutely aware of regional differences in speech norms, Andalusians will be more likely to readily identify those differences with a less positive evaluation offered for NCPS than their own variety.

3.2.2 Perceptual research in Andalusian Spanish

There have been relatively few perceptual studies in peninsular Spanish not yet referenced in previous sections of this chapter. In particular, attitudinal studies examining community evaluations of key AS phenomena are, for the moment, still few and far between. This is not to

say that there is no perceptual research available in Spanish, simply that this approach is still underused for AS specifically.[8] One of the few topics that has received considerable attention, especially due to the difficulty of determining phonological differences without perceptual data, has been vowel height and backness in the context of word-final /s/ elision and VH.

For example, Herrero de Haro (2017a, 2019) uses audio data from EAS speakers, comparing acoustic trends for /o/, /e/, and /i/ in word-final position, as well as preceding underlying elided /r/, /θ/, and /s/. After determining differences in height and backness of production types through formant analysis (Figure 17), the author presents examples for each realization to EAS listeners to see if the quantitative differences are also perceptually distinctive (Figure 17). For /i/, Herrero de Haro (2019) found that participants were able to distinguish [i] and [ir] from the other productions at a rate above chance, whereas for /e/ and /o/, the author (2017a) finds that [o], [oθ], [e], [es], and [er] can be distinguished from other allophones, suggesting that there may be more vowel categories audible to listeners than would be explained by a tense-lax division.

[8] In the language classroom, for example, perceptual studies of VOT boundaries between voiced and voiceless stops have been popular (e.g., Zampini, 1998; Schoonmaker-Gates, 2015; Pollock, 2020). These studies tend to examine phonological boundaries for language learners, focusing on differences between English and Spanish VOT norms, which differ enough to permit category overlap. Additionally, Delgado-Díaz, Galarza, and Díaz-Campos (2020) and Pollock, Willis, and Díaz-Campos (submitted) have done interesting attitudinal work examining perceptions of liquids in Puerto Rico and the Dominican Republic, respectively, to determine what social values listeners assign to reduction and variable production. Finally, in a Spanish-Galician bilingual context, Amengual and Chamorro (2015) encounter perceptual differences in the vowel space for Galician-dominant bilinguals (i.e., they have an /e/-/ɛ/ division), and similarities between Galician- and Spanish-dominant bilinguals (i.e., both have an /o/-/ɔ/ contrast).

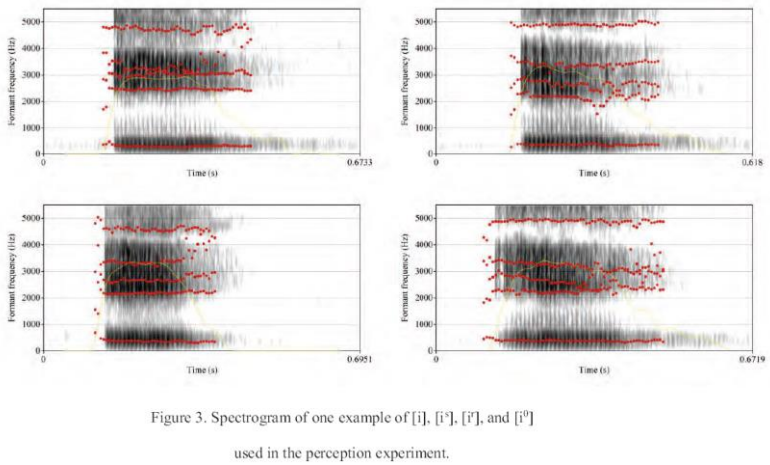

Figure 3. Spectrogram of one example of [i], [iˢ], [iʳ], and [iᶿ]

used in the perception experiment.

Figure 17: Spectrograms of [i], [is], [ir], and [iθ] (Herrero de Haro 2019:18)

In the distinction between singular nouns and plural ones with final /s/ deletion, Herrero de Haro (2018) finds that 153 Andalusian speakers identify nouns' number well above random chance (over 75% across two experiments). Underlying /s/ was identifiable both in cases of morphological and phonological deletion, although whether listeners relied solely on vowel laxing or also used another word-internal suprasegmental feature, the author was unable to determine for sure. VH, meanwhile, was a much less important secondary cue in the identification of noun number. In a similar study examining vowel harmonic spread, Henriksen (2017:114) combines production and perception to see how VH using [ɛ] influences the plural identifications of EAS and NCPS speakers in words such as /nenes/ 'kids' [nɛnɛ]. He aims to determine how necessary the leftward spread of laxness is in the morphological marking of plurality. Speakers of NCPS, it turns out, are less likely to recognize the morphological importance of [ɛ], which is not a part of their vowel inventory, while speakers of EAS were quite accurate and had faster reaction times identifying these stimuli. Both studies emphasize the importance of laxing, although it is less clear what effect VH has to reinforce interpretations of plurality.

A couple additional perceptual studies have examined some aspects of regional variation. For example, in a study of the AS affricate, Regan (2020) uses matched guises to compare perceptions of [t͡ʃ] and [ʃ] in WAS. While the fricative variant was perceived overall as having less prestige than the NCPS variant, evaluations differed based on listeners' origins. Those participants who hailed from the city of Huelva were more likely to classify the regional [ʃ] production as having reduced status and cosmopolitan identity than those from the small town of Lepe. Regan argues that this points to an urban-rural divide in Andalusian perceptions, as seen in other studies of AS such as that by Villena-Ponsoda and Ávila-Muñoz (2014).

Some research has also examined the relative regional and national prestige of different varieties based on sociophonetic features. Milla-Muñoz (2020: 88) performs a perceptual study of rhoticism and deaffrication in Andalusia, finding that listeners are more likely to rate both regional variants as "friendly" and "honest," but that rhotacism, perhaps because of its greater salience, is rated as less "professional." Milla-Muñoz (2020:145) suggests that there is not in-group bias among Andalusians to rate their speech more positively, and that although certain adjectival descriptions of AS were positive, greater prestige is attributed to NCPS, especially by younger speakers.

Harjus (2017), on the other hand, argues for a more nuanced perceptual analysis of Andalusian attitudes, arguing that, when considering perceptions of *ceceo* and *seseo* by speakers from Jerez de la Frontera (WAS), participants acknowledge dialectal differences and use these measures in their evaluations. Even though they come from a region that uses greater *ceceo*, listeners rate speakers using *ceceo* who come from Seville (i.e., where *seseo* is dominant) as being inaccurate, not seeing it as part of that variety of Spanish. Harjus (2017:12) argues that speakers' linguistic awareness is based "on a center–periphery model," and that positive perception of

variants is tied to “political or cultural relevance and the larger population of the urban area,” even if variants are not part of the listener’s personal repertoire. In the case of speakers from Jerez de la Frontera, *seseo* is seen as “urban” and “educated,” while also being “artificial” and “forced,” showing that while it received positive rating, it is seen as foreign and is not readily accessible to most speakers for identity construction.

One of the key findings across these perceptual studies is a definitive acknowledgement of where contrast does and does not exist. Herrero de Haro (2017a, 2019) shows that speakers have a nuanced means of contrasting between vowels preceding different deleted consonants, showing that contrast is maintained in the wake of elision. Henriksen (2017), meanwhile, indicates that there are clear distinctions between vowel space boundaries for speakers of NCPS and EAS. While some of these results confirm production claims, they also go beyond that, providing more nuance regarding variation, and basing their results on speaker input.

3.2.3 Perception and political speech

While perceptual research in English sociolinguistics has been gaining popularity in the last decade, it has remained relatively rare in peninsular Spanish. However, particularly in the realm of political speech, it is a necessary next step to help us better understand third-wave stylistic meaning and use. Observations that arise from production can be tested, offering a more accurate vision of politicians’ available linguistic repertoire and explaining how certain speech decisions may be regarded by the community. If a speaker from a large urban area were to use a stigmatized

rural variant, perceptual study would help determine if it would be salient to urbanites and what sort of prestige or stigma the use might be seen to carry.

There are a couple examples of perceptual research being implemented in political speech. In their study of María Antonia Martínez, Hernández-Campoy and Cutillas-Espinosa (2013:90) trace "public opinion" through newspaper sources, comparing them with an analysis of the politician's use of regional variants, showing how production research correlates with her perceived identity in the press (for better or for worse) as being very "Murcian." Similarly, Podesva et al. (2015) study listeners' attitudinal responses (Figure 18) to variation in the released /t/ of 2000s era U.S. politicians, including former presidents George Bush and Barack Obama. Based in an analysis of production, they examine attitudinal perception of variants using adjectival continua (e.g., friendly-unfriendly, inarticulate-articulate). Listener perceptions were found to vary based on the position of released /t/ in a word, suggesting that variants possess a complex array of variables that contribute to attitudinal responses.

Adjective Scale	Politician	Released Guise Rating	Flapped Guise Rating	Significance
Articulate	John Edwards	3.29	2.81	$p = .042$; $F(1,313) = 4.15$
Articulate	Condoleezza Rice	4.2	3.5	$p = .005$*; $F(1,313) = 7.94$
Southern	John Edwards	5.26	5.81	$p = .002$*; $F(1,313) = 9.84$
Southern	Condoleezza Rice	1.55	2.07	$p = .02$; $F(1,313) = 5.39$
Accented	John Edwards	3.92	4.37	$p = .048$; $F(1,313) = 3.91$
Accented	Nancy Pelosi	2.34	1.39	$p = .004$*; $F(1,313) = 8.29$
Passionate	Barack Obama	3.45	4.38	$p < .001$*; $F(1,313) = 11.18$
Sincere	Nancy Pelosi	3.78	4.34	$p = .035$; $F(1,313) = 4.42$
Authoritative	John Edwards	3.79	3.22	$p = .038$; $F(1,313) = 4.28$

Figure 18: Perceptual rating of American politicians by /t/ release (Podesva et al., 2015: 78)

Finally, in an examination of public (rather than political) speech, Molina-Ortés (2020) considers four Andalusian tv presenters' Twitter presence to determine how presenters' identity work is interpreted by online audiences. Speakers that the author identifies as having "coherent" linguistic identities are perceived positively, whereas those that dialect-switch between Andalusian and normative Spanish receive critiques of being "bipolar" and even disloyal to their region. Across these cases, albeit through different methods, social and phonetic boundaries can be established that help to describe variation: not how researchers intuitively define regionalisms based on production trends, but rather how speakers themselves identify and evaluate differences. Notably, perception studies provide a more accurate and detailed description of variable meaning than linguist intuition alone can provide.

3.2.4 Takeaways from production and perception

Previous research detailed in 3.1 and 3.2 shows an unfortunate imbalance between production- and perception-based approaches. On the one hand, production research has historically shown no need to be based in or discuss perceptual results, presenting a picture of geographical and, occasionally, social variation in peninsular Spanish. On the other hand, perceptual studies require researchers to have a firm grasp of the community and local norms that they are studying, often implying a need for production data of the community under examination before moving into their focus on perception. This often means that while production rarely serves to further perceptual knowledge, perceptual research, either through an explicit production-focused first-step, or through a thorough

examination of previous literature, often contributes to both production and perceptual understanding.

Among the problems raised by this difference in approach is the fact that impressionistically and acoustically described differences in production studies are rarely confirmed as being perceptually salient, or socially-meaningful. For example, Morrison and Escudero (2007) describe differences in vowel length between peninsular and Peruvian Spanish that are statistically significant, without being able to indicate whether this consistent difference in milliseconds reaches the level of perceptibility for members of the communities in question. In order to determine if quantifiable phenomena reach the level of perception, or if phonetically measurable differences possess any degree of social meaning, auditive phonetics is necessary to offer a fuller picture of social and linguistic variation in language. By collecting listeners' perceptions and evaluations, the social valence and meaning of linguistic resources can be more accurately pinned down.

3.3 Research Goals

Based on production- and perception-based research in Andalusian Spanish, this dissertation aims to address several issues raised in this chapter, including the role of regional variants in style-shifting behavior, their social meaning in context, and the perception of these variants by listeners in the region. Hernández-Campoy and Cutillas-Espinosa's (2006) description of style as a multidimensional concept that requires analysis within the specific context where it was produced requires that approaches to variable use closely consider indexical meaning. By combining

examinations of perception and production related to regionalisms, using a methodology similar to Podesva et al. (2015), I hope to address this complexity in understanding stylistic meaning. Related to this combination of production and perceptual approaches, there are several gaps in the research that must be addressed.

First, while the field of sociolinguistic perception is still young, especially in the Spanish-speaking world, it offers considerable promise. Political speech, in particular, demonstrates fruitful contexts for closer examination. Podesva et al. (2015) and Hernández-Campoy (2013) both implement aspects of perceptual and attitudinal studies as a way of determining community values based on political speech. Although sociolinguistics may be able to point to released /t/ as indicative of articulateness, when placed into varying word positions or combined with other features (as in real speech), the complexity of variation quickly moves beyond our ability as linguists to predict evaluative decisions (Podesva et al., 2015). Instead, by tracking the responses of potential audiences, nuanced claims about style-shifting, identity performance, and audience reception can be made that are all part of the sociolinguistic speech situation.

Next, in order to determine the norms of the community, it must be listeners from within a given community that judge its speech evaluatively. Given their daily use of and exposure to the available linguistic resources that politicians may use and, importantly, as the intended audience of Andalusian politicians' messages, listeners from within the community will be best able to identify and distinguish social meaning. As Podesva et al. (2015) show for the released /t/, while this variant alternates in the production of politicians, perceptual data is needed to begin to unravel its role in identity construction and meaning creation.

Perceptual data is also fleeting: indexical meaning is changing all the time as part of a dynamic interplay of social and linguistic values in daily life (Eckert, 2008). If perceptual data is

not collected as close to the moment of production as possible, it becomes increasingly unreliable. Asking a listener about their attitudinal response to speech from fifty years past clearly would not represent the reaction of the then-intended audience. While it is less clear where that line is in more recent data, it must be preferrable to collect perceptual responses as close to the time of production as possible.

Finally, while perceptual research helps to fill gaps in our understanding about social meaning and attitudes toward variation, perceptual and production analyses should be combined whenever possible. Patterns gleaned from perceptual research can be used to confirm the claims made by production research (as discussed by Campbell-Kibler [2010: 385]). If a given researcher does not investigate both for a particular speech community, given the ephemeral nature of perception and the broad array of speech contexts yet to be studied, the likelihood of other linguists returning to confirm production results are low. Molina-Ortés (2020) and Hernández-Campoy and Cutillas-Espinosa (2013) both show how this can be done through less time-intensive means; rather than always requiring perceptual instruments and participants, publicly available data from social media and print media can be used to track general attitudinal responses.

Overall, through a combination of investigative methodologies targeting perception, community-level production, and individual stylistic variation, this dissertation examines linguistic norms in the community of politicians under examination, explores broader linguistic norms of the Andalusian community, and provides insight into possible ongoing processes of linguistic shift and change. Both through previous descriptions of production and perception data, as well as through a consideration of current patterns in the political community of practice in question, an understanding of speech norms is established that can then be confirmed through the attitudinal responses of Andalusian voters.

4 Stage 1: Community norms

This chapter presents an analysis of general norms in the speech of Andalusian and NCPS politicians, using a community-focused approach similar to those found in Labovian (1972) research, albeit with a variety of additional context-specific social variables that allow for greater insight into stylistic behavior and social variation. In order to examine so-called regional versus normative variation, it is necessary to understand the possible variation available for each of the ten regional phenomena under consideration.

First, when affricate variation is discussed in the peninsular context, it is often used to refer to deaffrication as [ʃ] (e.g., Herrero de Haro, 2017b; Samper-Padilla, 2011; Villena-Ponsoda, 2008), although there have also been cases of a voiced affricate [dʒ] in Murcia (Torrano-Moreno, 2017). For the most part, production of the fricative allophone correlates to working-class, male, elderly, and rural speech, although there are signs that it is disappearing due to stigma. No cases of the fricative were found in the 35 hours of audio examined in this dissertation. However, there were instances of a fronted fricative [t͡s] similar to the one described in Chile (Flores, 2018; Henríquez-Barahona & Fuentes Grandón, 2018), and examined acoustically by Pollock (2022), who found it to occur both in NCPS and EAS, and potentially result from analogy to the affrication of [t] to [t͡s] in contexts following an elided word-internal /s/ (e.g., Ruch, 2012; Del Saz, 2019). Thus, the contrast between the normative [t͡ʃ] and the potentially super-regional innovative variant [t͡s] were considered for this phenomenon.

Second, syllable-final /s/ can experience retention (i.e., [s]), aspiration (i.e., [h]), and elision (i.e., [Ø]). While the latter two phenomenon are most frequent in AS, they can also occur at lower rates in NCPS (Blanco, 2004; Gil-Peña, 2004; Ruiz-Martínez, 2003; Samper-Padilla, 2011). Social

descriptions in NCPS describe retention as more prestigious, being used by women and younger speakers, while male and working-class speakers favor [h] (Ruiz-Dominguez, 1997; Samper-Padilla, 2011). While there is little social description of /s/ specifically in AS, younger and more educated speakers are trending toward NCPS norms (Regan, 2017b; Torrano-Moreno, 2017; Valeš, 2013; Villena-Ponsoda, 2008, 2013), and comparable studies in the Canary Islands (Molina-Ortés, 2015), which sometimes trend with AS, suggest that elision may be losing prestige to aspiration and retention.

Third, this dissertation makes a novel distinction to allow for examination of a specific case of /s/ reduction that is particularly rare in NCPS: word-final onset /s/ elision, in which syllabic reinterpretation before a word beginning with a vowel causes a word-final /s/ to nonetheless be in initial position of a syllable. While onset /s/ elision in this context is rare in NCPS, it frequently occurs in AS, likely by analogy to coda /s/ elision (Pollock, 2023). While it likely shares a social profile with syllable-final /s/ reduction, its positioning as a more hyper-regional variant may associate it more with Andalusian identity than coda /s/ elision alone.

Fourth, the phenomenon of *seseo* (i.e., /θ/ → [s]) has been historically associated with AS, although its use in urban WAS spaces has decreased in recent years and rates of distinction (i.e., /θ/ → [θ] and /s/ → [s]) have been on the rise as NCPS norms have spread southward through education (Regan, 2017a; Villena-Ponsoda, 2008). There are signs that it continues to be employed in urban areas in WAS (e.g., Seville) and NAS (e.g., Córdoba), although elsewhere it has declined (Díaz-Campos & Pollock, submitted). *Seseo* is associated with speakers with moderate education, male speech, and older speakers, as opposed to distinction, which tends to be employed by college-educated speakers, women, and younger speakers.

Fifth, the related phenomenon of *ceceo* (i.e., /s/ → [θ]) is mainly found in southern parts of Andalusia, particularly in rural speech (e.g., Díaz-Campos & Pollock, submitted; Melguizo-Moreno, 2009). Speakers with minimal formal education, as well as elderly speakers and men, are most likely to use *ceceo*, (Ávila-Muñoz, 1994; Melguizo-Moreno, 2007a; Carbonero, 1985; García-Amaya, 2008). Overall, *ceceo* tends to be viewed as vernacular and lacking in urban associations, causing it to be used in more rural areas as a means of diverging from urbanites' identities (e.g., Dalbor, 1980; Melguizo-Moreno, 2007b).

The sixth phenomenon, perhaps one of the most wide-spread throughout the Iberian peninsula, is the reduction and elision of intervocalic /d/. While the normative production has been described as the dental approximate [ð̞] – with rare cases of the stop [d] often only found in utterance-initial and post-pausal contexts – elision in intervocalic position is not infrequent, with it being particularly common in the past participle suffix -ado (Gil-Peña, 2004; Ruiz-Martínez, 2003). In Madrid, elision is associated with "informal" speech, including at the end of interviews, in discussions of personal topics, as well as by male speakers, and young and elderly speakers (e.g., Gómez-Serrano, 1994; Molina-Martos & Paredes-García, 2014; Molina-Martos, 1998). In Andalusia, elision is favored by young and working-class speakers (e.g., Ruiz-Martínez, 2003; Samper-Padilla, 2011) and is gaining ground in EAS (Villena-Ponsoda & Moya-Corral, 2016).

Seventh, elision and variable production of /l/ receives relatively little discussion among analyses of liquids, in part because, unlike /r/ and /ɾ/, there is less chance for the overlapping of phonological categories based on occlusion count. Hernández-Campoy and Cutillas-Espinosa (2010, 2013) describe "consonant permutation," whereby /l/ and /ɾ/ alternate in speech, as a typical vernacular variant in Murcia. With that said, variable production of /l/ and /ɾ/ has also been

identified in AS, where it is most likely among working-class speakers who identify as socialists politically and are proud of their Andalusian identity (Ruiz-Peña, 2013).

Eighth, the tap /ɾ/ is frequently subjected to variable production and reduction processes across the Spanish-speaking world (Bradley, 2001; Díaz-Campos, 2008; etc.). Tap reduction has been found to be less frequent than trill reduction in high-frequency words in Andalusia (Zahler & Daidone, 2014), as well as by working-class socialists who associate with Andalusian identity (e.g., Ruiz-Peña, 2013).

Ninth, trill /r/ reduction and variable production has received considerable attention as well in Caribbean Spanish linguistics (Delgado-Díaz & Galarza, 2015; Pollock et al., in press). Older speakers and male speakers tend to reduce and elide the trill (Henriksen, 2014; Henriksen & Willis, 2010). Word-finally, /r/ deletion is more common among Murcian men (Hernández-Campoy & Cutillas-Espinosa, 2010).

Finally, the tenth phenomenon is the process of vowel laxing and harmony, found in EAS, that precedes elided /s/. Herrero de Haro (2017a) and Henriksen (2017) agree that the vowel system of this region contains three additional lax vowels to maintain contrast between plural and singular words (i.e., /æ ɛ ɔ/). Despite both generative and perceptual work related to the phenomenon, there is little information regarding social differences in production, suggesting that while this phenomenon may be categorical in certain varieties of EAS, further sociolinguistic analysis is merited to confirm this contrast.

Based on these ten regional phenomena, analysis of regional and normative productions by politicians from the EAS, NAS, WAS, and NCPS varieties are examined to determine how identity construction and style-shifting behavior are carried out. A more in-depth analysis of production differences is carried out in Section 4.2.4, which also discusses acoustic measures used to

distinguish between allophones. Over the course of this analysis, the chapter examines several questions based on theoretical and quantitative results, discussed in the following section.

4.1 Research questions and hypotheses

Based on descriptions of Andalusian Spanish by researchers including Samper-Padilla (2011), Villena-Ponsoda (2008, 2013), and Regan (2017a, 2019), the ten phenomena under consideration have received some attention both in AS and NCPS. Oftentimes, while regional variants are considered to occur with greater frequency in Andalusian than in the broader "standard" peninsular variety, there is also a tendency for the highest rate of use to reside among speakers from the working-class, those with minimal education, elderly speakers, and more generally those lacking social prestige. Other than for age (many politicians fall into the late-middle aged or elderly categories), these social categories tend not to apply to the typical politician, who tends to be middle- or upper-class, have higher education, and possess a great deal of social prestige. Nonetheless, Hernández-Campoy and Cutillas-Espinosa (2013) and Pollock and Wheeler (2022) have found that certain members of the political speech community produce speech high in vernacular productions.

This seeming incongruity leads to the first overarching question of this dissertation, which is addressed by the current chapter: namely, based on spectrographic and quantitative analysis, what social and linguistic variation actually exists within this community for the ten phenomena described in the previous section? Through an examination of the overall norms of a community of practice composed of politicians at the high regional and national levels, the current quantitative

analysis aims to document the major types of variation that occur and address this general question. Along the way, three additional questions are raised.

This includes, first, the matter of differences across social questions: What trends can be identified in variable use, based on social and linguistic divisions? The expectation, based on evidence from previous research into political speech (e.g., Hernández-Campoy & Jiménez-Cano, 2003; Schilling-Estes, 2013), is that politicians for the most part will follow the linguistic and social norms of their community, using the "automaticity" of less-conscious style based on the "script" established for the medium. While some style-shifting behavior is expected to break away from these norms, the broad approach of this first stage, which will not distinguish between nuanced differences in speech context, should provide a general overview of variation that mirrors that found in previous research.

The second question addresses novel social categories, such as those discussed in third-wave studies (e.g., burned-out burnouts; Eckert, 2000; smooth operators, Zhang, 2005): are there specific social categories relevant to the current context that can help define variation in greater detail than through macro-social factors? This can include factors like political parties, which previous research has found to be predictive of linguistic differences (e.g., Hall-Lew et al. 2017).

Lastly, the third question examines variation within specific words as a means of beginning to access stylistic variation: based on lexical analysis, which regional variables appear to demonstrate fruitful stylistic variation in select lexical items? This examination provides an idea of how macro- and contextually-relevant-micro-social variation occurs within the group, with the expectation that politicians may be more likely to group along the lines of political party than region, gender, or other context-nonspecific social groupings (e.g., Hall-Lew et al., 2012).

4.2 Methodology

Variationist sociolinguistics focuses on changes in the linguistic system of speakers, seeing this change not as random and chaotic but rather as measured and systematic. This study involves identifying systematic variation within a community and comparing mixed-effect logistic regression models to describe variation based on the predictiveness of social and linguistic factors. Given historical approaches to style as part of a variable context of sociolinguistic values and the quantitative approaches to its role in identity work in political discourse, a variationist methodology is necessary to analyze the behavior of this community of speakers.

Both social and linguistic factors for each variant are considered, following methodological frameworks of previous studies of public and political speech. This includes Flores' (2014, 2017, 2018) work with Chilean radio discourse, in which she finds social variation across speaker gender, age, and interlocutor gender, and Hernández-Campoy and Cutillas-Espinosa's (2010, 2013) series of articles about the hyper-vernacular speech of former Murcian president María Antonia Martínez, in which they compare her regional variant production with that of politicians and community members. As with these previous studies impressionistic coding will be used to identify phonetic and phonological differences – however, this will be supplemented with visual inspection of waveforms and spectrograms and reinforced statistically through analyses of acoustic measures to confirm impressionistic categorization quantitatively.

The remainder of this section on methodology presents an explanation of five central parts of the design of this experiment. First, the corpus of politicians is discussed in detail, explaining the social background and demographic breakdown of speakers. Next, two sections discuss the

dependent and independent variables under consideration in the statistical analysis. Following that, the fourth and final section covers the data collection and analysis process, describing how tokens were identified and coded, as well as presenting the results of an inter-rater reliability analysis, comparing the coding of myself and two native Andalusians with backgrounds in linguistics.

4.2.1 Corpus

The group of 32 speakers analyzed in this chapter was first assembled for analysis by Pollock (2022, 2023). Audio was collected from audiovisual news media recordings made publicly available online that were created between the years 2011 and 2019. In order to ensure the broadest availability of speech examples, members of the two largest political parties in Spain were included in the examination: the conservative *Partido Popular* 'People's Party' (PP), and the socialist *Partido Socialista Obrero Española* 'Spanish Socialist Worker's Party' (PSOE).

The 32 politicians under examination are evenly divided by city/region (i.e., 8 each from Córdoba/NAS, Málaga/EAS, Seville/WAS, and Madrid/NCPS), speaker gender (i.e., 16 males, 16 females), age (16 born pre-1965, 16 born post-1965), and political affiliation (i.e., 16 PP politicians, 16 PSOE politicians). Individuals were selected based on the criteria that they had been born no more than fifty miles from the main city of the region they represent, and that they held major positions in urban, regional, or national politics that afforded them a regular opportunity to speak in news media.

Three speech contexts were selected for each individual, based on Flores' (2014) findings that speaker and interlocutor gender both play roles in variation, including scripted speeches,

unscripted interviews with male interlocutors, and unscripted interviews with female interlocutors. Only 2 speakers were lacking one of the three speech contexts, resulting in a total of 94 interviews and speeches under examination. The complete representation of the corpus, including the name, gender, city, political party, speech context, age, video year, video duration, and broadcasting channel for each politician can be found in Appendix A: Table A.

4.2.2 Dependent Variables

The ten regional phenomena discussed in the introduction to this chapter serve as the dependent variables for analysis in this project. They have been drawn by studies that have made general characterizations of AS regional tendencies (e.g., Pollock, 2022, 2023; Pollock & Wheeler, 2022; Samper-Padilla, 2011), as well as from specific studies focusing on the variation found in individual phenomena, such as *seseo*, *ceceo*, and distinction (e.g., Melguizo-Moreno, 2010; Regan, 2017a/b). For each phenomenon, possible productions were identified in audio files, and grouped into the categories of "normative" productions and "regional" ones that tend more often to be specific to AS. These divisions include:

1. Affricate /t͡ʃ/ fronting to [t͡s]: Based on previous work in Chile (Flores, 2018) and exploratory examination in AS (Pollock, 2023), the innovative apico-alveolar affricate variant [t͡s] was considered a regional production, while the pre-palatal affricate variant [t͡ʃ] was described as normative. COG was used as a secondary acoustic distinguisher between the apico-alveolar affricate (i.e., avg 4700 Hz) and the pre-palatal one (i.e., avg. 3900 Hz) based on previous findings (e.g., Gordon et al., 2002).

2. Syllable-final /s/ reduction to [h] or [∅]: A highly salient AS feature (Gil-Peña, 2004; Ruiz-Dominguez, 1997; Samper-Padilla, 2011) as a key feature of Andalusian Spanish, productions indicating "aspiration," *ceceo*, and elision (i.e., [h], [θ], and [∅]) were defined as regional, while retention and cases of fortition (i.e., [s] and [z]) were coded as normative. As with the other sibilant phenomena, COG was used as a secondary measure to distinguish normative (i.e., avg. 3300 Hz) and regional (i.e., avg. 825 Hz).
3. Word-final onset /s/ reduction to [h] or [∅][9]: A frequent phenomenon identified in AS (Pollock, 2023) that occurs by analogy to coda /s/ elision. As with syllable-final /s/, productions of "aspiration," *ceceo*, and elision (i.e., [h], [θ], and [∅]) were defined as regional, while retention and fortition (i.e., [s] and [z]) were coded as normative. COG was also used to support differentiating normative (i.e., avg. 3700 Hz) and regional (i.e., avg. 550 Hz) productions.
4. *Seseo*, where /θ/ is produced as [s]: Associated in particular with WAS and Seville (Harjus, 2017), rates are falling across Andalusia in favor of distinction (Regan, 2017a; Villena-Ponsoda, 2008). For *seseo*, all cases where /θ/ was produced as [s] or underwent reduction processes experienced by [s] were treated as regional productions (i.e., [s], [h], and [∅]), while all instances of production as /θ/ were treated as normative. COG helped to confirm this differentiation between regional retention of [s] (i.e., avg. 4900 Hz), regional reduction to aspiration and elision (i.e., avg. 590 Hz), and normative production of the interdental fricative (i.e., avg. 3500 Hz).
5. *Ceceo*, where /s/ is produced as [θ]: Mainly associated with rural, elderly, and male speech, use has decreased heavily over time (Díaz-Campos & Pollock, submitted; García-Amaya,

[9] For example, when resyllabification would be expected to occur for a word-final, pre-vocalic /s/, e.g., *las otras* [la.so.tɾas] 'the others,' reduction would instead lead to production as [la.ho.tɾas] or [la.o.tɾas].

2008; Melguizo-Moreno, 2007a). For *ceceo*, all cases of /s/ being produced as an interdental fricative were treated as regional (i.e., [θ]), while all cases of production as [s] or reduction processes were, for this variable, treated as "normative" (i.e., [s], [h], and [∅]). Despite being regional syllable-final /s/ productions, aspiration and elision were not representative of *ceceo*. This was the only variable for which tokens were duplicated in the analysis, allowing a separate consideration for reduction and *ceceo* phenomena. COG supported the impressionistic examination, showing a difference between normative retention (i.e., avg. 4050 Hz), normative reduction (i.e., avg. 834 Hz), and regional *ceceo* (i.e., avg. 3000 Hz) productions.

6. Intervocalic /d/ deletion to [∅]: Following previous findings in Spain, reduction was expected to occur throughout the peninsula, but be most frequent in AS (Gil-Peña, 2004; Ruiz-Martínez, 2003; Samper-Padilla, 2011). The regional production was considered to be the elided variant (i.e., [∅]), whereas the normative one was treated as production of the dental approximate (i.e., [ð̞]). Intensity range[10] was used as an acoustic correlate to track reduction, measuring the maximum intensity in the preceding vowel, and subtracting the minimum intensity of the occlusion valley of /d/. The presence of a closure yields a much higher range (i.e., avg. 8.7 dB) than absence of the same (i.e., avg. 4.1 dB).
7. Reduction and variable production of /l/: This phenomenon has been suggested in previous research to correlate with Andalusian and Southern Spanish identity (Hernández-Campoy & Cutillas-Espinosa, 2010, 2013; Ruiz-Peña, 2013). Both instances of elision and variable production (i.e., [∅], [ɾ], and [r]) were treated as regional variants, while the lateral (i.e., [l]) was considered normative. Both intensity range and F3 range were used as acoustic

[10] This variable is defined by Torreira and Ernestus (2011: 342) as "maximum intensity velocity in CV transition," who use it to compare differences between voiceless Spanish and French stops.

measures to support impressionistic coding, taking the maximum and minimum values from across the segment to distinguish regional (i.e., avg. F3: 630 Hz; Intensity Range: 10.3 dB) and normative (i.e., avg. F3: 530 Hz; 7.1 dB) productions. These measures were most useful in distinguishing lateral and rhotic productions.

8. Reduction and variable production of /ɾ/: There are connections between these processes and Andalusian identity, although tap reduction may not be as common as trill reduction (Ruiz-Peña, 2013; Zahler & Daidone, 2014). As with the lateral, all instances of elision and variable production (i.e., [∅], [l], and [r]) were treated as regional, while production of the tap (i.e., [ɾ]) was seen as normative. F3 range was somewhat helpful in differentiating regional (i.e., avg. F3: 640 Hz) and normative (i.e., avg. F3: 605 Hz) variants.
9. Reduction and variable production of /r/: For these processes, there is evidence associating reduction with male and older speech (Henriksen, 2014; Henriksen & Willis, 2010; Hernández-Campoy & Cutillas-Espinosa, 2013; Zahler & Daidone, 2014). Elision and variable production (i.e., [∅], [l], and [ɾ]) were treated as regional, while production of the trill (i.e., [r]) was classified as normative. Intensity range and F3 range served as secondary measures to identify regional (i.e., avg. F3: 690 Hz; Intensity Range: 10.7 dB) and normative (i.e., avg. F3: 710 Hz; Intensity Range: 10.2 dB) variants.
10. Mid- and low-vowel laxing and harmony preceding elided /s/: Mainly associated with EAS, this phenomenon has not been associated with specific sociolinguistic variation (Henriksen, 2017; Herrero de Haro, 2017a). Rather than solely categorical, vowel height and backness were also treated as dependent variables to allow an examination of continuous variation among speakers using EAS, with lower values of F2 (i.e., backing) and higher values of F1 (i.e., lowering) correlating with regional “laxed” productions. In

order to create a categorical dependent variable for comparison with other phenomena, normalized F1 values above the mean production for each file were coded as “regional,” while values below the mean were coded as “normative” to capture quantitative differences equivalent to laxing.

Tokens were inspected in Praat based on their spectrogram and waveform. Acoustic cues were considered both during the initial coding and in the mixed-effects logistic and linear regression models. Each sound was listened to several times to help with impressionistic coding and to determine spectral properties. Dependent variables were identified as regional and normative groups, allowing for specific analysis of each sound, as well as a combined analysis of all phones, to distinguish between AS and NCPS norms through binary analysis. This description is summarized in Table 4.

#	Variable	Normative	Regional
1	Affricate fronting	/t͡ʃ/	[t͡s]
2	Syllable-final /s/ reduction	/s/	[h], [θ], [∅]
3	Word-final onset /s/ reduction	/s/	[h], [θ], [∅]
4	*Seseo*	/θ/	[s], [h], [∅]
5	*Ceceo*	/s/, [h], [∅]	[θ]
6	Intervocalic /d/ reduction	[ð̞]	[∅]
7	Lateral /l/ reduction & variable production	/l/	[ɾ], [r], [∅]
8	Tap /ɾ/ reduction & variable production	/ɾ/	[l], [r], [∅]
9	Trill /r/ reduction & variable production	/r/	[ɾ], [l], [∅]
10	Vowel laxing	-scaled F1	+scaled F1

Table 4: Dependent Stage 1 variables, including normative/regional divisions

4.2.3 Independent Variables

A total of twenty variables were analyzed across the ten phenomena, including two categorical random effects, seven categorical and five continuous linguistic variables, five categorical and two continuous social variables. The two random effects, “speaker” and “word,” were included to account for the base assumptions in mixed-effect models, which treat every individual token as coming from a different speaker and representing a different context. By treating these two variables as random effects, their influence on variation is controlled in the model, while their influence is reported separately from the fixed-effects that compose the central part of the model. While all seven social variables were considered in each model, the linguistic variables differed based on the dependent variable under consideration.

With respect to the categorical linguistic variables, three were considered in all models, and the other four were included to represent tendencies of specific dependent variables. Preceding phonological context and following phonological context were examined in each analysis, although specific features of these contexts differed based on the phenomenon (e.g., features related to backness were key in examining affricate fronting, while those for tongue height were more important in the consideration of vowel laxing). Lexical position was also included in every model to distinguish between word-final, word-medial, and word-initial productions. In addition to these, participial usage of intervocalic /d/ was coded to determine if the specific <-ado> and <-ido> suffix experienced greater elision (e.g., Pollock & Wheeler, 2022). The presence or absence of elision in a word for the examination of vowel harmony was coded (e.g., Henriksen, 2017). The orthographic representation was coded for the trill /r/ to distinguish between <rr> and phonological contexts described as requiring the trill (e.g., <lr>, <sr>, word-initial position; this tested the role of writing in pluralization). Finally, morphological function was coded for sibilant phenomena where /s/ represented plurality (testing the functional hypothesis).

Next, for the continuous linguistic variables, only one was examined in all models, while the remaining four were phenomenon-specific. Segment duration in milliseconds was included for each factor, with the expectation that it captured information about local speech rate, which has been associated with Labovian style. Center of gravity (COG) in Hz was analyzed for all five sibilant phenomena, including the affricate, syllable-final /s/ and word-final onset /s/ elision, *ceceo*, and *seseo*. Intensity range was included for the intervocalic /d/, the lateral, the trill, and the tap. F3 range was examined for all three liquids, given that F3 can hint at rhoticity and tongue shape, meaning that the range in difference from the proceeding to following sound was expected to be more drastic for rhotics, followed by liquids, and lowest for cases of elision. Additionally, scaled F2 was included for vowels to investigate the relationship between vowel height and backness in the process of vowel harmony.

Alongside these linguistic factors – many of which were included either as a sanity check to reinforce impressionistic coding and to confirm that these speakers followed linguistic norms expressed by previous research – were seven social variables. These served as the main focus of the analysis and were included in each of the mixed-effect logistic and linear regression models to determine how information related to the speaker, expected addressees, stylistic goals, and speech situation influenced production.

The first categorical social variable was politician gender. Both in the peninsular context and in sociolinguistic research more broadly (e.g., Labov, 1972; Chappell, 2016), female speech is often associated with normative and prestige variants. Particularly in AS, it was expected that women would be more likely to use forms indexing overt prestige for their direct constituents (i.e., AS regional norms), whereas men were expected to be more likely to use NCPS variants, as well as those variants with covert prestige in AS (e.g., *ceceo*).

Next, speech audience was a categorical variable that reflected the realities of the medium on which the speech or interview was recorded. News channels were coded as local, regional, national, and international given a composite of their location, the extent of their reach, and the topics they discussed. For example, an interview from a small rural broadcaster discussing local policy changes would be coded as "local," whereas one on a broadcaster in an urban area with local topics would be coded as "regional," one on a nationally-available channel was coded as national, and one associated with EU committees and global politics was coded as international. Although these files were made globally available online, speakers are most focused on expected listeners, described in Audience Design Theory as addressees and auditors (Bell, 1984). Regional variants were expected to be more frequent for local and regional audiences, while normative variants were expected for national and international audiences, as Cruz-Ortiz (2019) reported historically for AS-speaking politicians in Madrid.

Following that, political party was the third categorical social variable, divided between left-leaning PSOE politicians and right-leaning PP ones. Based on the findings of Hernández-Campoy and Cutillas-Espinosa (2013) with regard to María Antonia Martínez, Pollock and Wheeler (2022) for Susana Díaz, and Ruiz-Peña's (2013) claim that variable liquid production occurs most frequently among members of a community with left-leaning affiliations, members of the PSOE are expected to be more likely to use regional variants as a form of showing regional solidarity with voters. Normative variants, described as correlating with bourgeoise values, are expected to be used by PP politicians.

Speech context is used to refer to one aspect of the Labovian conceptualization of style, separating more "formal" scripted speech from less "formal" unscripted interviews with male and female interlocutors. Based on previous findings in public speech (e.g., Flores, 2014), speakers on

the radio can vary regional variant use in speech based not only on their own gender, but also on that of their interlocutor. Flores (2014, 2017) found that men used more vernacular variants in speech with other men as opposed to women, whereas "formal" scripted speech is often correlated with the most careful speech.

The final categorical social variable, politician city of origin, was used not only to separate out the NCPS variety of Madrid, but also to distinguish between NAS (Córdoba), WAS (Seville), and EAS (Málaga). Geographical trends between NCPS and AS were expected to be the strongest, as the two varieties differ across all the phenomena under consideration. However, phenomenon-specific trends were also expected to emerge, following previous studies (e.g., vowel harmony in EAS, *seseo* in WAS, reduced intervocalic /d/ and syllable-final /s/ elision in NAS, etc.).

The first continuous social variable, age, reflects both general sociolinguistic norms as well as specific peninsular realities surrounding education and political change. The death of Franco in 1975 served as a dividing point between a Spanish education system with considerable illiteracy, and a modernized system with increased access to higher education for women (García-Amaya, 2008). Politicians were selected based on their year of birth, divided between pre- and post-1965. It was expected that speakers over roughly 50 years of age, as of the time of recording, would be more likely to use regional variants as a result of reduced access to higher education and NCPS norms early in life. Additionally, given the general tendency for younger speakers to prefer normative variants (e.g., Regan, 2017a), this continuous variable is expected to reflect ongoing change.

The final social variable, video percent, was a continuous measure reflecting another aspect of traditional Labovian style, namely the attention paid to speech over the course of an interview. This measure takes the timestamp of a variant and calculates where it occurs over the course of an

interview (e.g., occurrence in the first second = 0.01%, occurrence in the final second = 99.9%). This ties into issues of speech formality (e.g., Fafulas, Díaz-Campos, & Gradoville, 2018), who find that speakers tend to start interviews by using prestigious or normative variants, but transition over time to more innovative, regional variants. The traditional Labovian approach would dictate that regionalisms should increase over time, although in the current case, with the matter of identity construction in play, it is less clear whether that should result from this case.

These independent variables, as well as the factors considered and the phenomena to which they were just applied, are presented in Table 5.

Variable Type: Variable	**Factors**	**Phenomena**
Random		
Speaker	32 individuals	All
Word	Unique words	All
Categorical linguistic		
Preceding	Height/backness	All
Following	Height/backness	All
Position	Word-initial/medial/final	All
Participial use	-ado, -ido, other	intervocalic /d/
Elision	Present, Absent	vowels
Orthography	<r>, <rr>	Trill
Plurality	plural, other	*ceceo*, coda /s/, word-final onset /s/
Continuous linguistic		
Segment duration	milliseconds	All
COG	Hertz	affricate, *seseo*, *ceceo*, syllable-final /s/, word-final onset /s/
Intensity Range	decibels	intervocalic /d/, lateral, trill, tap
F3 Range	Hertz	lateral, trill, tap
Scaled F2	Hertz	vowels
Categorical social		
Gender	male, female	All
Audience	local, regional, national, international	All
Party	PSOE (left), PP (right)	All
Context	Scripted speech, unscripted (female interloc.), unscripted (male interloc.)	All

City	Seville, Málaga, Córdoba, Madrid	All
Continuous social		
Age	years	All
Video percent	Timestamp in video/total seconds in video	All

Table 5: Independent Stage 1 variables, including factors and phenomena

4.2.4 Data Collection and Analysis

Based on the 94 audio files under consideration, two minutes of audio were completely coded for the presence of all ten regional phenomena using the Praat phonetic software program[11] (Boersma & Weenink, 2023). This coding methodology was based on the concept of the principle of accountability in sociolinguistics, which stipulates that the entirety of the envelop of variation should be defined and examined by a linguist as a means of documenting the overall frequency of a variant in comparison with other productions (Díaz-Campos, 2014). Only by fully coding all possible productions can a quantitative analysis make claims about the spread of a linguistic phenomenon in a speech community.

The two minutes examined in each audio file were selected in part as an initial step to orient the Lectal Focusing in Interaction (LFI) analysis in Stage 2. Once the first and midpoint minutes were coded, it became possible to determine the extent of variation across a broad swath of the interview, and allowed for identification of speakers with increased vernacular usage that could be selected for further investigation in Stage 2, when additional audio was coded.

[11] Note that vowels were only coded for audio files of speakers from Málaga to investigate the possibility and extent of laxing in EAS.

Consistency was key across all segments coded in the analysis. Based on important acoustic measures, including intensity range, COG, and formant values, intervals were created based on the same acoustic landmarks. The phenomena and their related acoustic measures are included in Table 6. In addition to this acoustic phonetic information, social information was collected by probing into each audio file, as gathered in Table A in Appendix A. This included the gender of the interlocutor, year in which the video was filmed, age of the speaker at the time of filming, duration of the video, and channel on which the video was posted.

		Acoustic Measures			
#	**Phenomenon**	**COG (Hz)**	**Intensity range (dB)**	**F3 range (Hz)**	**F1 & F2 (Hz)**
1	Affricate fronting (frication)	X			
2	Syllable-final /s/ reduction	X			
3	Word-final onset /s/ reduction	X			
4	*Seseo*	X			
5	*Ceceo*	X			
6	Intervocalic /d/ reduction		X		
7	Lateral /l/ reduction & variable production		X	X	
8	Tap /ɾ/ reduction & variable production		X	X	
9	Trill /r/ reduction & variable production		X	X	
10	Vowel laxing				X

Table 6: Acoustic measures considered for the stage 1 phenomena

In this acoustic analysis, identification for impressionistic coding relied on several visual measures that were reinforced later through quantitative analysis. The next ten figures provide example spectrograms of normative and regional productions for each of the phenomena. First, in reference to the affricate, Figure 19[12] shows the difference between the pre-palatal and alveolar productions. While both were composed of an occlusion and frication, the pre-palatal production

[12] For the spectrograms demonstrated in this section, normative productions were drawn from Madrid female political speech, while regional productions were collected from male and female politicians in Málaga.

[t͡ʃ] in mucho ‘a lot’ has a much lower COG than in the alveolar production [t͡s] taken from charanga ‘charanga (dance),’ which makes sense given the higher intensity and pitch of the [s]-like segment in the latter. The average COG of frication for the pre-palatal production was 3907Hz across all productions, while the average for the alveolar production was 4729Hz.

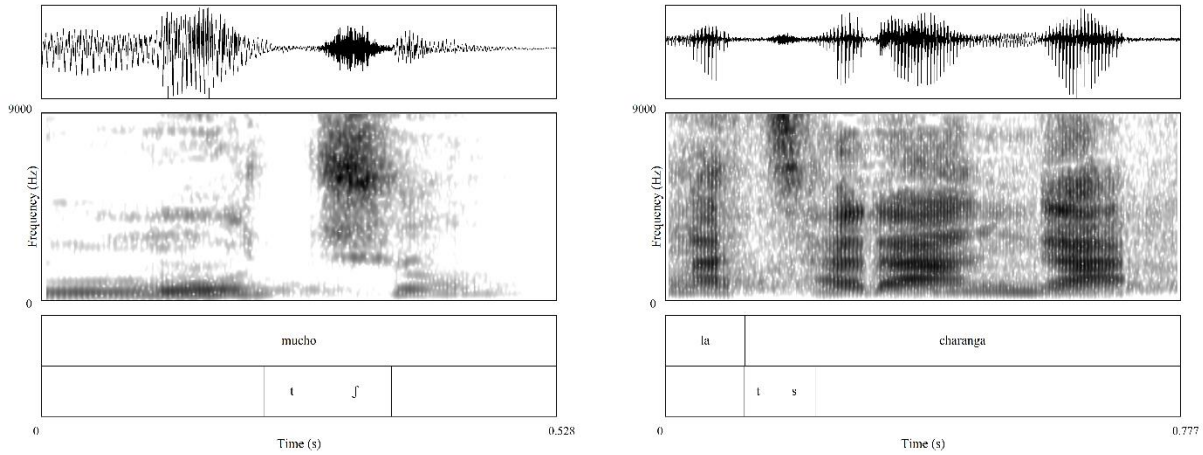

Figure 19: Normative (pre-palatal) and regional (alveolar) productions of the affricate in mucho *‘a lot’ and* charanga *‘charanga (dance).’*

Next, for syllable-final /s/, the presence of frication beginning around the range of 4500 Hz was treated as a marker of the production of the alveolar fricative.[13] Figure 20 shows normative and regional productions. There is a clear bunching of sound in the upper ranges of the spectrogram of *los bosques* ‘the forests’ in the former, while the latter shows only an occlusion for /p/ in *despacho* ‘office.’ The average normative COG was 3307 Hz, while the regional one was around 850Hz.

[13] Although rare, voicing also occurred in some tokens, particularly those produced by Madrid speakers. These tokens were grouped with [s] as cases of sibilant retention.

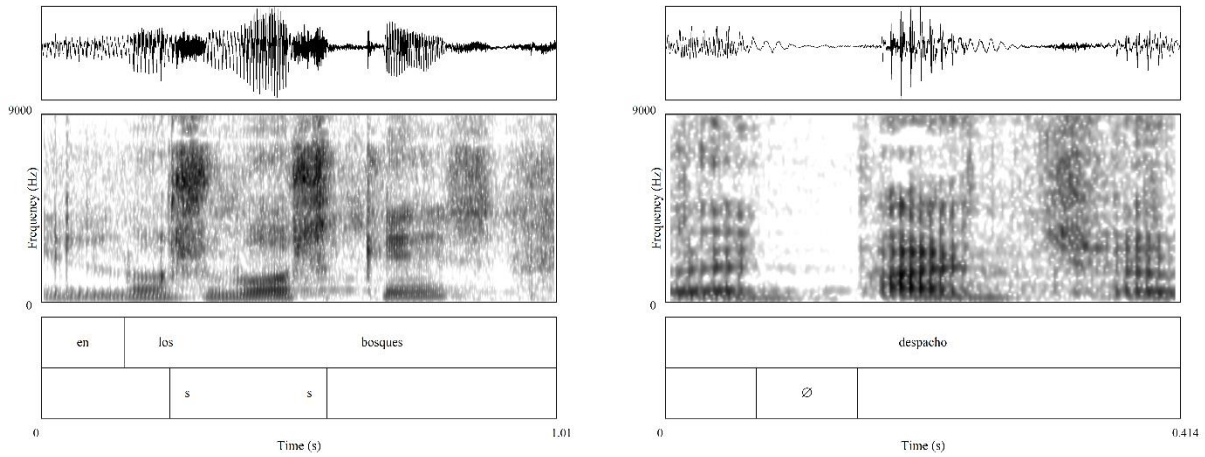

Figure 20: Normative and regional productions of coda /s/ in los bosques *'the forests' and* despacho *'office.'*

Following that, for word-final onset /s/ in pre-vocalic position, similar markers were examined as for syllable-final /s/. The normative and regional productions in Figure 21 provide a case of retention in *es algo* 'it's something' and elision in *dos actividades* 'two activities.' The average normative COG was 3713 Hz, while the regional one was around 550 Hz.

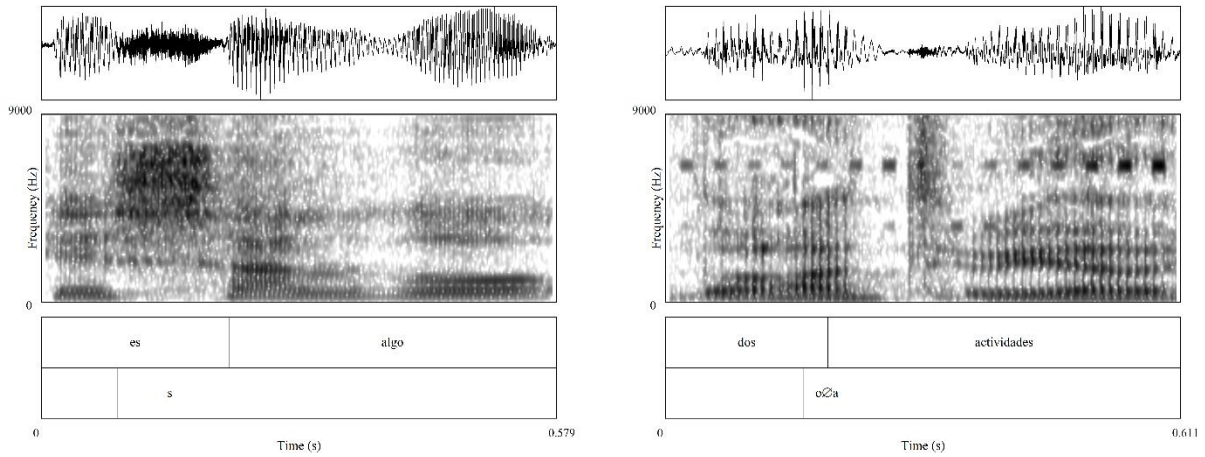

Figure 21: Normative and regional productions of word-final onset /s/ in es algo 'it's something' and dos actividades 'two activities.'

Using a similar set of acoustic markers, analysis of *seseo* made the distinction between the interdental fricative [θ] and the alveolar [s]. Figure 22 shows the difference between the interdental fricative in *capacidad* 'capacity' and the alveolar one in *[3sg] celebraron* 'celebrated.' Differences in COG were able to help distinguish production types: the average normative COG for the interdental fricative was 3511 Hz, while production as the regional [s] had an average COG of 4893 Hz.

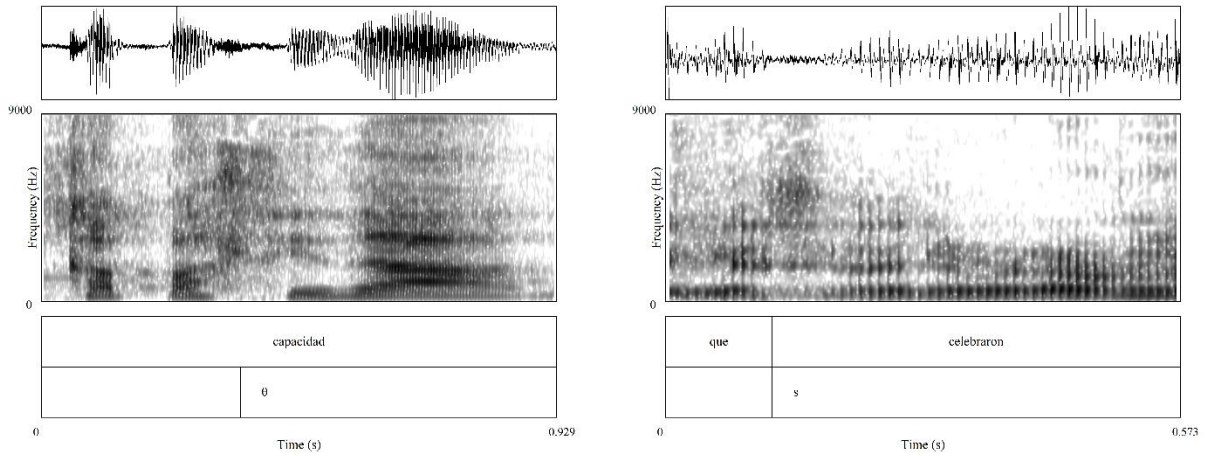

Figure 22: Normative and regional productions of seseo in capacidad 'capacity' and celebraron '[3sg] celebrated.'

The treatment of *ceceo* essentially inverted the consideration used for *seseo*, with [s] being treated as the normative production. As before, COG was useful in distinguishing between tokens. Figure 23 shows the difference between the alveolar production of *pasado* 'past' and the interdental production of *vicepresidente* 'vice president." The average normative COG for the alveolar fricative was 4052 Hz, while the regional [θ] had an average COG of 2386 Hz.

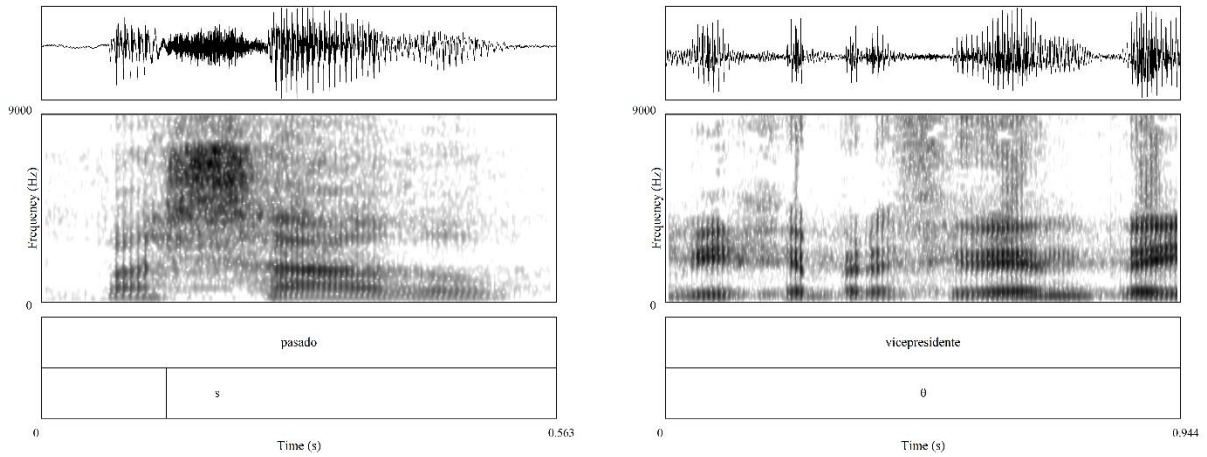

Figure 23: Normative and regional productions of ceceo in pasado 'past' and vicepresidente 'vicepresident.'

Next, for intervocalic /d/, intervals were created to bridge the intensity peak of the preceding vowel and the trough of the stop, allowing for a tracking of intensity range from the highest to lowest point. Greater intensity range correlated with an increased occlusion size, and a higher likelihood of normative production. Figure 24 demonstrates two spectrograms, contrasting the normative [ð̞] in *capacidad* 'capacity' with the elided [∅] in *abogado* 'lawyer.' The average range in intensity for retention was 8.7 dB, while, for elision, it was 4.1 dB.

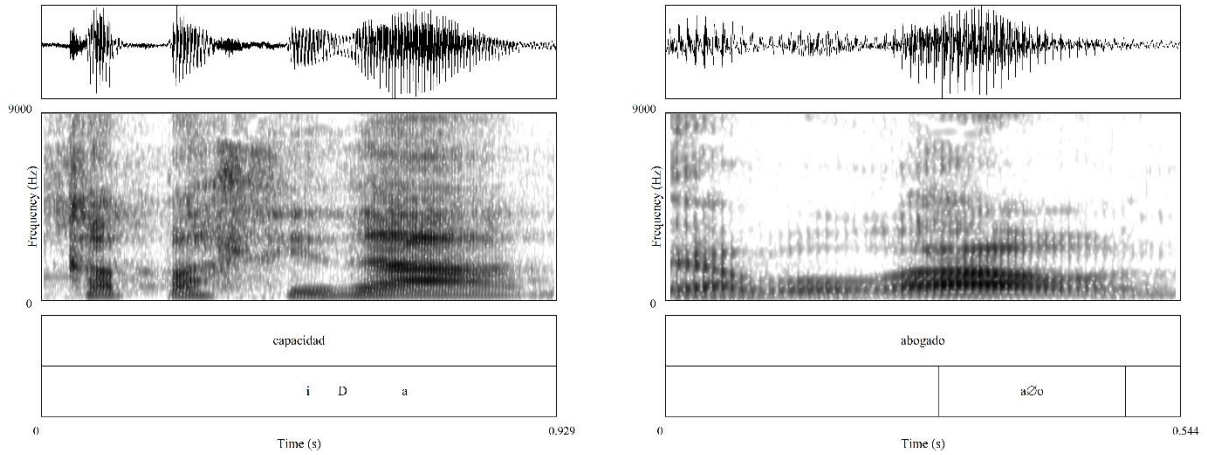

Figure 24: Normative and regional productions of intervocalic /d/ in capacidad 'capacity' and abogado 'lawyer.'

Liquids proved more difficult to classify with a single acoustic measure, although both the intensity range and F3 range provided a degree of distinguishing information. Figure 25 provides a normative production in *el cambio* 'the change' and a regional one in *de la noche* 'at night.' The average range in intensity for the normative production was 7.1 dB and the range in F3 was 529 Hz, while the regional production had an intensity range of 16.4 dB and an F3 range of 598 Hz.

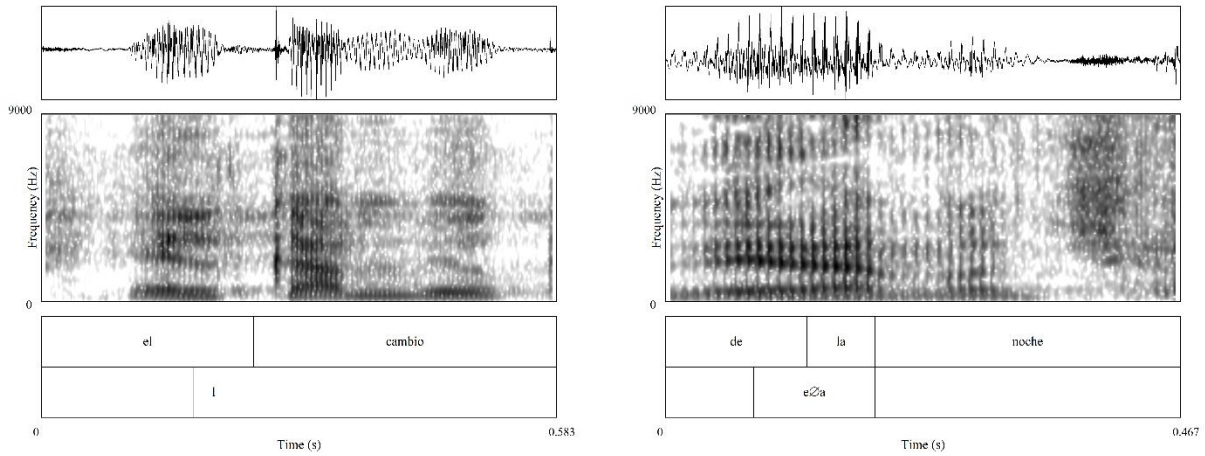

Figure 25: Normative and regional productions of the lateral in el cambio 'the change' and de la noche 'at night.'

Next, for the tap [ɾ], Figure 26 offers an example of a normative production in *nosotros* 'we' and a regional one in *mejorar* 'to improve.' The average range in intensity for the normative production was 9.8 dB and the range in F3 was 605 Hz, while production as [l] had an intensity range of 8.5 dB and an F3 range of 525 Hz and production as [r] had an intensity range of 12.1 (dB) and an F3 range of 639 Hz.

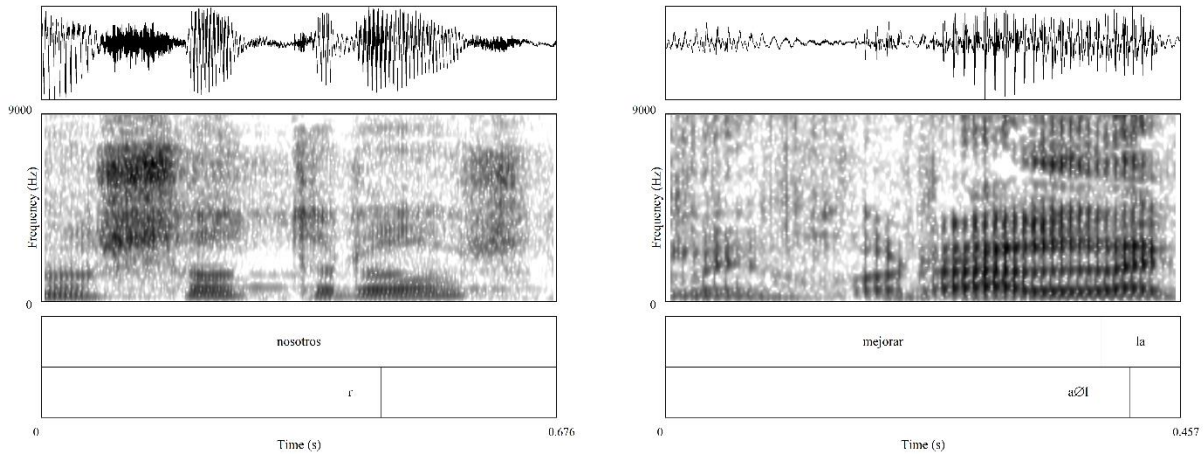

Figure 26: Normative and regional productions of the tap in nosotros 'we' and mejorar 'to improve.'

In the case of the final liquid, [r], Figure 27 gives examples of the normative production from *territorial* 'territorial' and the regional production from *irresponsabilidades* 'irresponsibilities.' The average range in intensity for the normative production of [r] was 10.2 dB and the range in F3 was 710 Hz, while regional productions had an intensity range of 11.3[14] (dB) and an F3 range of 575 Hz.

[14] There were cases where differences in acoustic measures may fall below the expected threshold of perceptibility. While one of these two measures supported and predicted the outcome of the acoustic coding, further perceptual research should be used to determine the extent to which these measures, as well as others like segment duration and occlusion number, are able to account for perceptible differences between allophones.

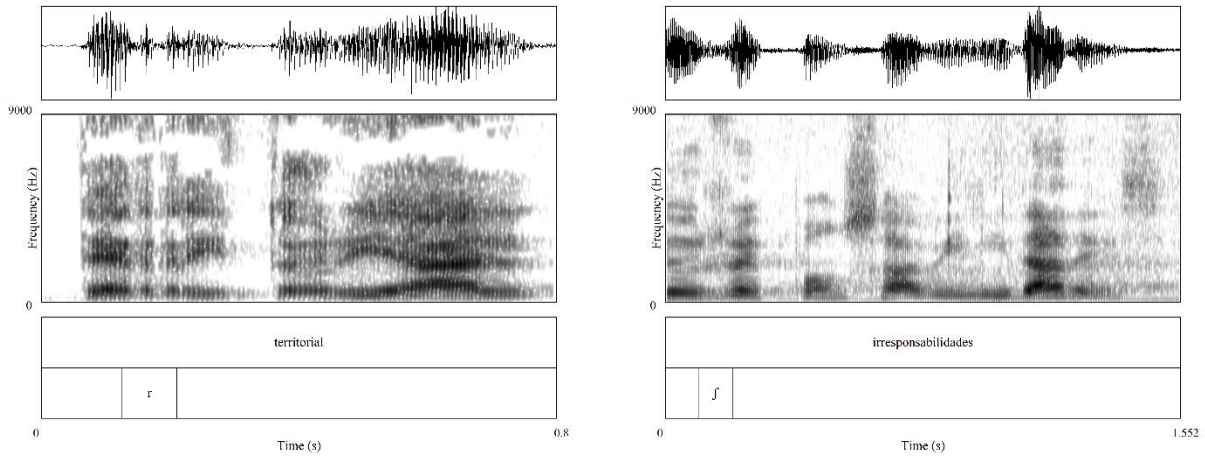

Figure 27: Normative and regional productions of the trill in territorial 'territorial' and irresponsabilidades 'irresponsibilities.'

Finally, the last phenomenon under consideration was that of vowel harmony in specifically eastern Andalusian Spanish. Figure 28 demonstrates two analyzed words – the first with the context of *abogado* 'lawyer' where normative production would be expected (i.e., no word-final /s/ elision), and the second in the context of *abogados* 'lawyers' where vowel lowering and laxing might be expected (i.e., a word with final /s/ elision) based on previous production research (e.g., Henriksen, 2017). Both normalized, scaled F1 and F2 values were collected in these two contexts to determine lowering and fronting tendencies. The differences in these values were relatively low: across all tokens, the average scaled F1 for vowels in words without final /s/ elision was 431 Hz and the scaled F2 was 1503 Hz, while the scaled F1 in cases of elision was 445 Hz and the scaled F2 was 1529 Hz.

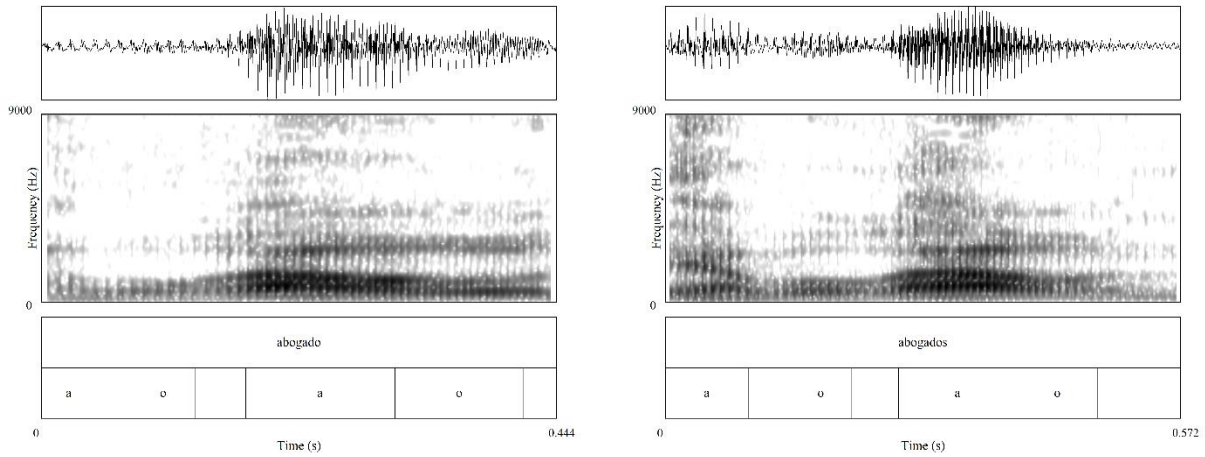

Figure 28: Normative and regional productions of vowels in abogado 'lawyer' and abogados 'lawyers.'

Impressionistic coding was supported not only by quantitative acoustic data and visual inspection in Praat, but also by a process of inter-rater reliability (IRR). Two native speakers of WAS from Seville, Spain, were provided with a guide explaining the navigation of the six tiers in Praat, as well as possible allophones identified by the author in the coding of stimuli (although they were also encouraged to identify and note down other allophones they heard). These instructions are replicated in Appendix C. Speakers, who both had a background in linguistics, coded a randomly-chosen selection of audio taken from the overall corpus (n=1343 tokens). Gradoville (2014: 96) and Brown (2008: 48) have used similar approaches, although limiting IRR analysis to 1% of tokens. Given the broad spread of phenomena under consideration, and in order to represent the responses of these two coders, roughly twice that percentage was coded. Overall, this represents 2.3% of the total amount of data coded for both Stage 1 (this chapter) and Stage 2 (Chapter 5) of the dissertation. Notably, the examination of vowel harmonic processes, which relied on continuous acoustic measures rather than impressionistic coding and was limited to only the EAS speakers, was not included in the IRR coding. Overall, the two native speakers had an

89.8% agreement with the author's coding, with agreement at its highest for sibilant phenomena, where the presence and absence of /s/ is extremely notable. Agreement did not fall below 80% for any single phenomenon. The results of IRR are demonstrated in Table 7, which includes the number of tokens considered for each phenomenon across Stages 1 and 2 of the analysis, the number of tokens used in IRR, and the agreement by phenomenon, where the Andalusian coders and the author identified tokens under the same allophonic category.

#	Phenomenon	Stage 1	Stage 2	Total Tokens	IRR %	IRR Tokens	Agreement
1	Affricate Fronting	3174	265	3439	2.0%	68	80.90%
2	Coda /s/ elision	5963	2652	8615	2.3%	198	97.50%
3	Word-final onset /s/ elision	1446	-	1446	2.9%	42	97.60%
4	*Seseo*	2669	965	3634	2.3%	82	92.70%
5	*Ceceo*	11145	4163	15308	2.3%	352	98.30%
6	Intervocalic /d/ elision	4253	1527	5780	2.2%	125	80.80%
7	Variable tap production	8050	3020	11070	2.3%	259	88.00%
8	Variable trill production	605	240	845	1.8%	15	80.00%
9	Variable lateral production	6199	2327	8526	2.4%	202	92.10%
	Total	43504	15346	58663	2.3%	1343	89.80%

Table 7: Interrater reliability using data from 2 native Seville phoneticians and linguists

Following successful coding and inter-rater reliability comparison, a script was created to collect the acoustic and linguistic coding data from the tiers in Praat. This included taking information from the first (word) tier, as well as each of the lower tiers (representing phenomena under examination). This script is provided in Appendix B, showing the specific information collected from the six coded tiers in Praat. By running this script, information about the file and speaker was collected alongside quantitative data about each segment, including formant values, duration, intensity, and COG. In addition, as every segment was identified by its underlying phoneme, allophonic production, and surrounding context (e.g., <s>oH#k, meaning underlying /s/,

produced as [h], in word final position #, preceded by /o/, and followed by /k/), this data was also collected to provide information on the phonetic environment and impressionistic production, based on analysis of the waveform and spectrogram.

Next, data were analyzed using a series of mixed-effects logistic and linear regressions in Rbrul[15] (Johnson, 2009), which measured the ability of both linguistic and extralinguistic factors to predict variation in the data. Each dependent variable was simplified into a regional versus normative dependent variable to fit the requirements of a categorical mixed-effects logistic regression. In order to create each model, a number of comparative variants were created, and predictive values related to AIC and log-likelihoods were compared to determine the most predictive and most simple model.

Overall, eleven mixed-effects models are presented in the following section to explore the data from two perspectives. The first ten models represent the best description of variation for each of the individual linguistic phenomena, determining how social and linguistic factors influence variable production. The final model combines all of these models, excluding linguistic factors (which tend to differ significantly across phenomena), and provides a holistic description of the impact of social variation on the production of regionalisms in AS, as compared to NCPS. The section concludes with a consideration of a small number of salient lexical items, comparing their manner of production based on the political party of speakers, to determine how individual words (e.g., Hall-Lew et al., 2012) can serve as touchstones for variation.

4.3 Results

[15] The first nine phenomena were considered using logistic regressions. However, as the dependent value for vowels was the scaled normalized F1, the dependent variable was continuous, and thus necessitated a linear regression.

In sum, a total of 44195 tokens were analyzed in this chapter. Those results are broken down by phenomenon, regional versus normative type, and specific variant in Table 8. For affricates, while the pre-palatal production was most common (72%), the fronted variant was identified as occurring in more than a quarter of cases (28%). In the case of both syllable-final and word-final onset /s/, elision was most frequent, with retention occurring second most frequently, followed by aspiration (which was much more frequent in word-final onset [27%] as opposed to syllable-final [6%] position). With respect to *seseo*, normative production as the interdental fricative /θ/ was by far the most common (89%), followed by production as [s], elision, and finally aspiration. There were very few cases of *ceceo* in the data (n=43, 0.4%), with the vast majority of tokens being either retained (53%) or elided (46%).

With respect to the intervocalic /d/, elision was slightly more frequent (52.5%) than retention (48.5%). For both the tap /ɾ/ and the lateral /l/, production as the underlying phoneme was by far most common, occurring in more than 92% of cases, and elision was next most common, at around 5%, in word-initial, medial, and final position for /l/ and in medial and final position for the tap. Variable production was the least frequent. This differed for the trill /r/, with the most frequent production being reduction to the tap (51%), followed by the normative trill (49%), although again there were very few cases of variable production (n=4). Notably, there was a higher rate of tap production in word-initial position, while normative production was more frequent in word-medial position.

Finally, for vowels, a division was made between tense and lax productions based on the normalized value acquired from each speaker's vowel space using a Lobanov equation (Lobanov, 1971). This value accounts for variation in each individual's articulatory tract, determining the

medial point and reporting numbers with negative and positive values in relation to that. Vowels came from words with an elided final /s/ (e.g., *estudiantes* [estudiante] ‘students’) and those with final vowels (e.g., *entiendo* [entiendo] ‘I understand’) to determine if vowel harmony occurs across words with elided /s/ in eastern Andalusian Spanish. Positive F1 values, representing lowering, were coded as the laxed variant, while negative F1 values, representing a normative or raised production were coded as the underlying phoneme. Laxing was most common for /a/, while laxed and tense productions of /o/ occurred almost equally, those for /e/ represented about a third of production, those for /i/ were negligible (n=14), and none were identified for /u/.

Phenomenon	Phone	N vs R	Variant	#	COG (Hz)	Intensity Range	%
Affricate	/t͡ʃ/	Normative	[t͡ʃ]	2296	3907		72.3%
		Regional	[t͡s]	878	4729		27.7%
			Total	3174			
Syllable-final /s/	/s./	Regional	[Ø]	3743	816		62.8%
		Normative	[s]	1836	3307		30.8%
		Regional	[h]	358	866		6.0%
		Regional	[θ]	26	901		0.4%
			Total	5963			
Word-final onset /s/	/s#V/	Regional	[Ø]	603	591		41.7%
		Normative	[s]	456	3713		31.5%
		Regional	[h]	384	445		26.6%
		Regional	[θ]	3	2308		0.2%
			Total	1446			
Seseo	/θ/	Regional	[Ø]	80	552		3.0%
		Regional	[h]	10	899		0.4%
		Regional	[s]	214	4893		8.0%
		Normative	[θ]	2365	3511		88.6%
			Total	2669			
Ceceo	/s/	Normative	[Ø]	5162	865		46.3%
		Normative	[s]	5940	4052		53.3%
		Regional	[θ]	43	2386		0.4%
			Total	11145			
Intervocalic /d/	/d/	Regional	[Ø]	2190		4.1	51.5%
		Normative	[ð̞]	2063		8.7	48.5%
			Total	4253			
Lateral	Word-initial /l/	Regional	[Ø]	35		6.2	0.6%
		Normative	[l]	2299		6.6	37.1%
		Regional	[ɾ]	9		6.6	0.1%
		Regional	[Ø]	119		9.0	1.9%

	Word-medial /l/	Normative	[l]	2110		6.2	34.0%
		Regional	[ɾ]	58		9.5	0.9%
		Regional	[r]	1		29.0	0.0%
	Word-final /l/	Regional	[∅]	140		12.6	2.3%
		Normative	[l]	1400		9.2	22.6%
		Regional	[ɾ]	28		11.5	0.5%
			Total	6199			
Tap /ɾ/	Word-medial /ɾ/	Regional	[∅]	360		10.5	4.5%
		Regional	[l]	20		7.8	0.2%
		Normative	[ɾ]	6013		9.5	74.7%
		Regional	[r]	40		10.2	0.5%
	Word-final /ɾ/	Regional	[∅]	139		10.2	1.7%
		Regional	[l]	10		9.8	0.1%
		Normative	[ɾ]	1431		10.8	17.8%
		Regional	[r]	37		14.1	0.5%
			Total	8050			
Trill /r/	Word-initial /r/	Regional	[∅]	1		15.3	0.2%
		Regional	[ɾ]	245		10.8	40.7%
		Normative	[r]	191		10.3	31.7%
	Word-medial /r/	Regional	[∅]	2		10.3	0.3%
		Regional	[ɾ]	60		10.2	10.0%
		Normative	[r]	103		10.1	17.1%
			Total	602			
					F1	F2	
Vowels	/a e i o u/	Normative	[a]	17	499	1427	2.5%
		Regional	[æ]	149	514	1441	21.6%
		Normative	[e]	125	419	1641	18.1%
		Regional	[ɛ]	60	431	1646	8.7%
		Normative	[i]	114	370	1803	16.5%
		Regional	[ɪ]	14	385	1849	2.0%
		Normative	[o]	102	435	1252	14.8%
		Regional	[ɔ]	84	444	1313	12.2%
		Normative	[u]	26	375	1259	3.8%
			Total	691			
Grand Total				44192			

Table 8: Findings for variant production across phenomena in Stage 1

In the analysis of the alveolar and pre-palatal sibilant affricates, a comparison was made between the /s/ productions recorded for *ceceo* and coda /s/ deletion, the fronted [t͡s] variant, and the prepalatal /t͡ʃ/ (Figure 29). The comparison shows a marked difference between productions. While the pre-palatal affricate has a peak around 4000 Hz, the fronted alveolar affricate peaks at 6000 Hz, and the retained /s/ has its peak at 5000 Hz. This helps to demonstrate clear acoustic

differences between the various productions, although notably there is variation, especially for [s] and [t͡s], which are produced across a wide range of COG values.

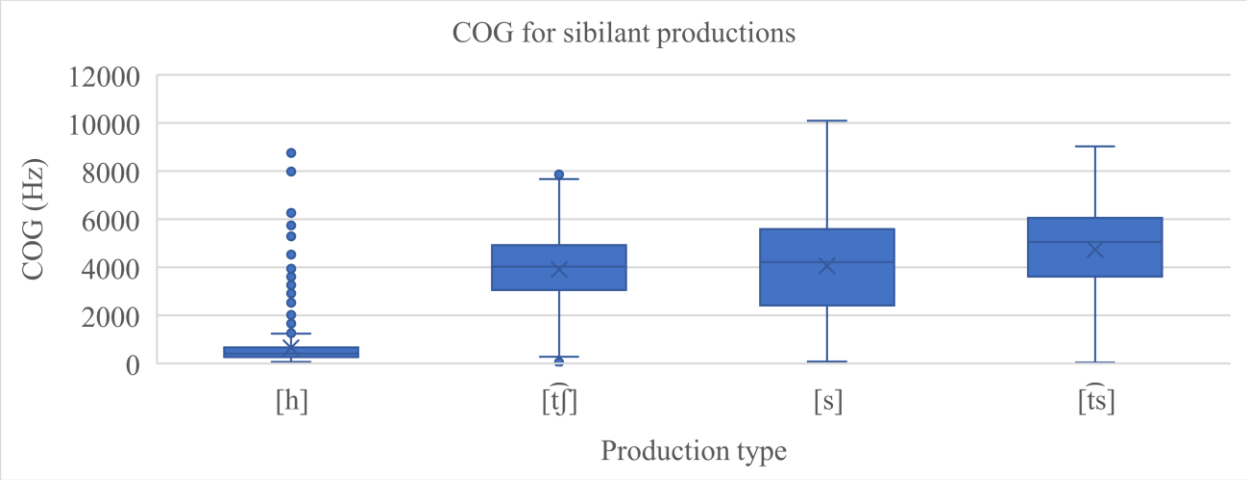

Figure 29: Frication COG (Hz) for variants of /s/ and /t͡ʃ/

Next, in a comparison of the main two *seseo* productions, there was a notable contrast based solely on COG for interdental fricative [θ] variants and the alveolar ones (Figure 30). Overall, the interdental fricative tended to be produced with a much lower COG, with almost half of all productions being under 3000 Hz, while the alveolar fricative (as in Figure 29) was produced most at 5000 Hz. The fact that COG for both variants has a long tail reflects previous discussions of *seseo* and *ceceo* (e.g., Dalbor, 1980) that reference medial-sounding productions like [θs] and [sθ], which exist in a medial space.

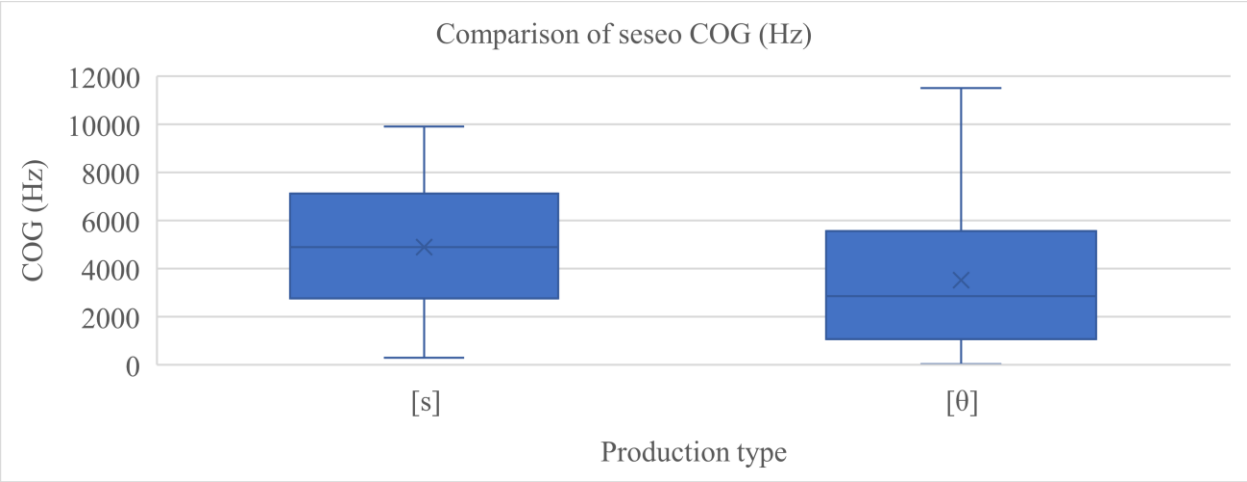

Figure 30: A comparison of COG (Hz) for seseo variants

Following that, the intensity range for intervocalic /d/ was compared in Praat, using decibels relative to $2x10^{-5}$ Pascal, from the preceding vowel peak to the occlusion valley (Torreira & Ernestus, 2011). In the case of a complete occlusion, intensity should drop in the direction of zero, whereas a fully-produced stop would have a closure dropping from around a sound pressure level (SPL) of 74 dB (i.e., the average dB of a verb in the corpus). Given that Spanish /d/ is typically described as a dental approximant [ð̞], retained variants were found to have a modal peak around 8 dB with a long right tail, whereas elided variants [∅] peaked at 2 dB and dropped off quickly (Figure 31).

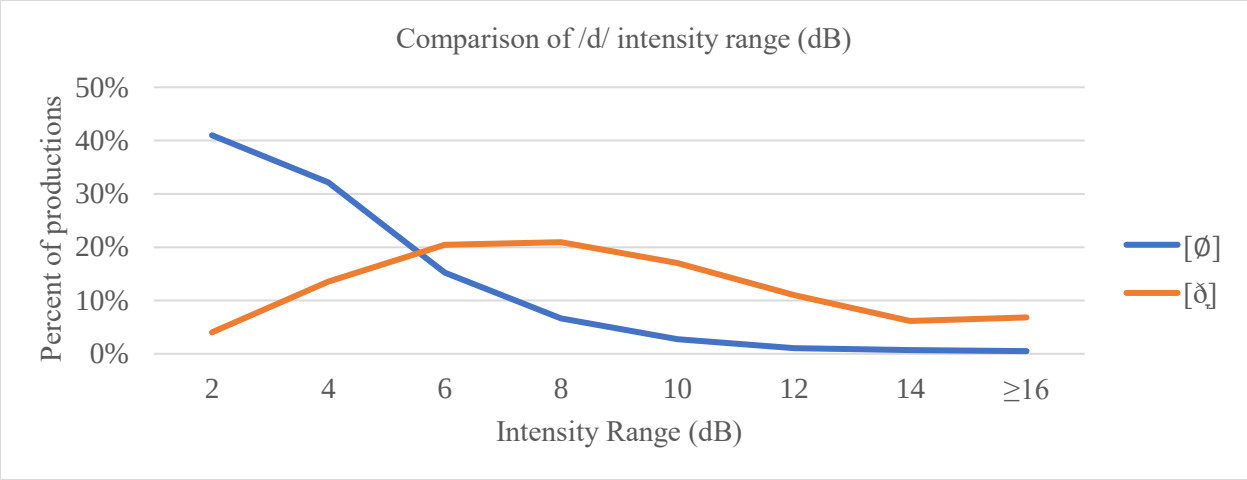

Figure 31: A comparison of the range of intensity for /d/ variants

Variation across the liquids was found to be best quantified using the intensity range measure to differentiate possible occlusions (Figure 32a). Range was tracked from the peak of the previous sound to the liquid valley, and was expected to differentiate between productions with occlusions, such as the tap or trill, and those without occlusions, such as the lateral or a fully elided production. Elision was most likely to have a range of 2 dB or below, while [l] had a peak at 4 dB, and the tap and trill peaked around 8 dB, seeming to bear out this expected articulatory difference. The range in F3 was also tracked as a means of examining tongue position. There was less differentiation here; while [l] in particular was much more likely to show under 500 Hz difference between the preceding sound and the lateral, the trill was most likely to have a broader range of F3 ranges.

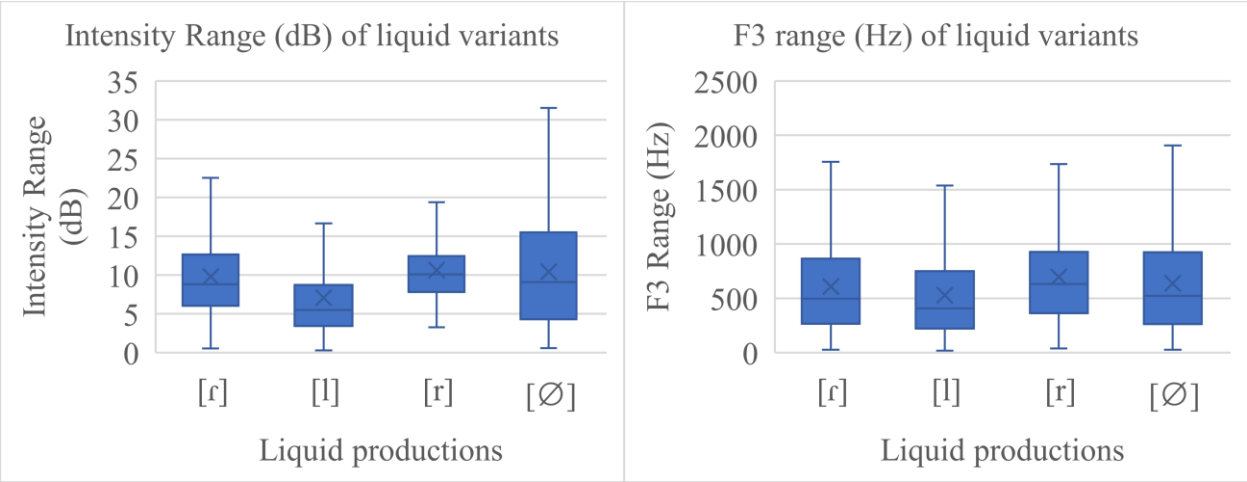

Figure 32a/b: Intensity Range (dB) and F3 range (Hz) for allophonic liquid productions

Finally, with respect to vocalic variation, the vowel space of EAS politicians is presented in Figure 33. Vowels were normalized using the Lobanov method (Lobanov, 1971), then scaled using formula (1) for F1 and (2) for F2 based on the formula from the R *vowels* package (Kendall & Thomas, 2010). A scatterplot of results is shown in Figure 33, with a triangle representing the average vocalic production for each group.

(1) F'1 = 250 + 500 (FN1 - FN1MIN) / (FN1MAX - FN1MIN)

(2) F'2 = 850 + 1400 (FN2 - FN2MIN) / (FN2MAX - FN2MIN)

Based on the results for F1, it can be seen that /a/ has the lowest production, while being produced with a long left tail. The peak for /o/ is the next lowest, and those for /e/ and /u/ are close (although /u/ has a larger left tail), with /i/ having the highest overall production (Figure 34a). Meanwhile, for F2, productions are slightly more discrete, with /o/ and /u/ having the most backed production, although /o/ has a long right tail. Next, /a/ is more medial, followed by /e/, and last of all by /i/, which has a longer leftward tail. Thus, we can see that while the vowel space does follow the divisions we would expect, there is considerable variability in positioning (even for these

normalized data) that could suggest differences between tense and lax productions (even if there is no clear bimodality that would strongly indicate a lax distinction).

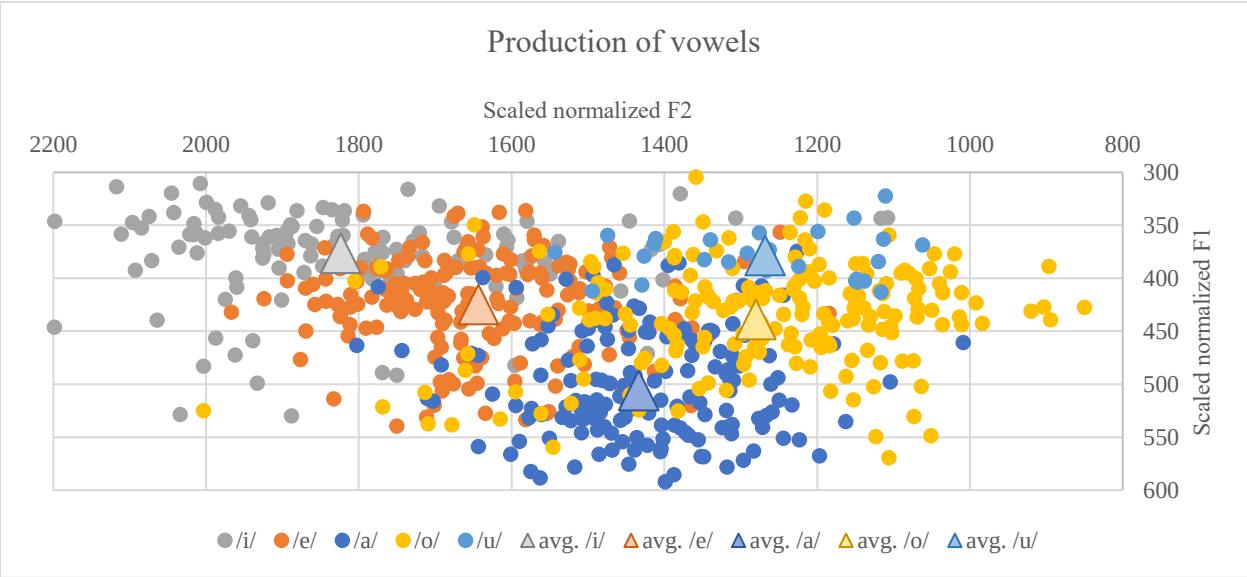

Figure 33: Variation across EAS politicians' vowel space

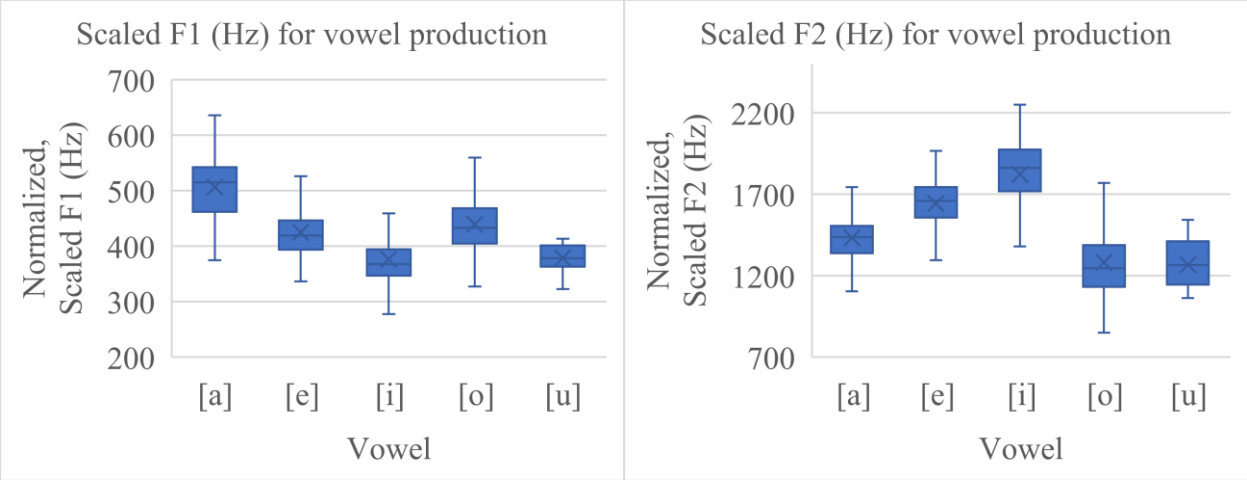

Figure 34a/b: Normalized scaled F1 (Hz) and F2 (Hz) values for EAS vowels

These data show that considerable variation exists across variants of the ten phenomena. Nonetheless, acoustic measures exist to help distinguish between impressionistic coding, demonstrating a consistency across the data that is reinforced by the IRR results. Following this descriptive presentation of acoustic differences, the next sections present the mixed-effect models designed to describe the influence of both social and linguistic variation on regional production in AS and NCPS.

When reading data from a statistical mixed-effects model used in variationist sociolinguistics, there are several types of information to key in on (Tagliamonte, 2006, 2011). The first of these are p-values, which indicate the possibility that a correlation could have resulted from random chance alone. In the social sciences, it is often considered best practices to treat p-values below 0.05 as "significant," meaning that it is 5% or less likely that the variable was only coincidentally a predictor of variation. The second type of information is represented by both the factor weight and log-odds, which represent the same information through two slightly different scales. For factor weight, an effect favors the application value when between 0.500 and 1.000 and disfavors the same when between 0.500 and 0.000. Log-odds, which is unbounded, shows factors that favor the application value as positive numbers, and those that disfavor the application value as negative ones. An increased distance from zero indicates an exponentially greater degree of effect. Finally, the range is included as a measure of the distance between the highest and lowest factor weight, providing a means of approximating the variables that have the greatest and least degrees of effect. Continuous variables, as well as linear regressions, do not include factor weights or range, instead providing only log-odds and p-values.

4.3.1 Affricate fronting

The first phenomenon under consideration is the fronting of the pre-palatal affricate [t͡ʃ] to the alveolar affricate [t͡s]. Figure 35 shows a breakdown by politician gender, political party, and word-position of the phenomenon. Female socialists were most likely to produce the fronted variant.

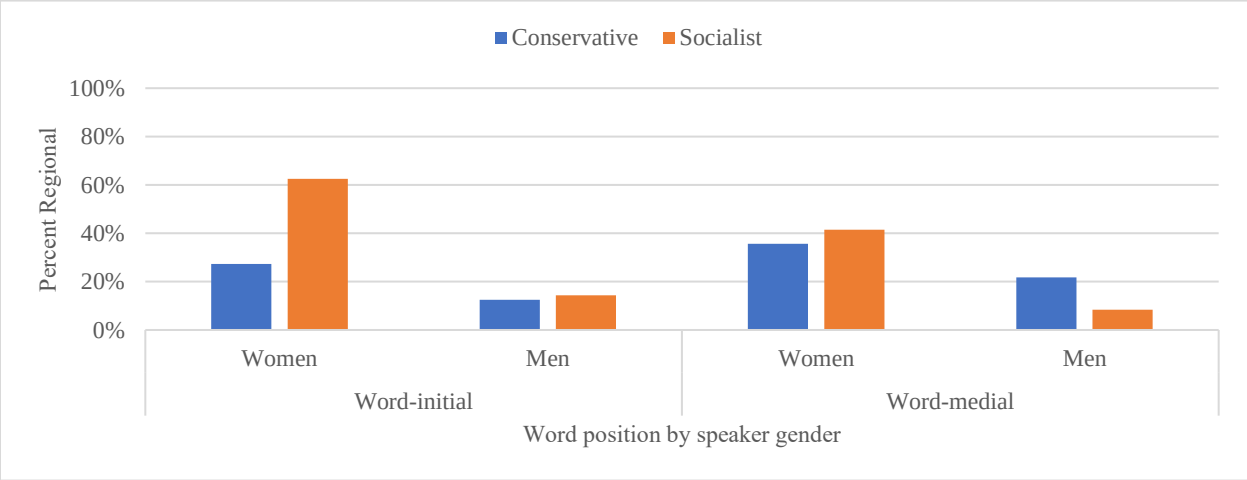

Figure 35: Regional production of the affricate by political party, speaker gender, and position in the word.

Next, Figure 36 shows the acoustic measure used for allophonic distinction. In this case, the fronted alveolar affricate was distinguished from the pre-palatal one by COG (Hz) of the frication portion of the segment. The alveolar production tended to have a higher value – however, differences in categories was reduced for men from Málaga and Seville and women from Madrid.

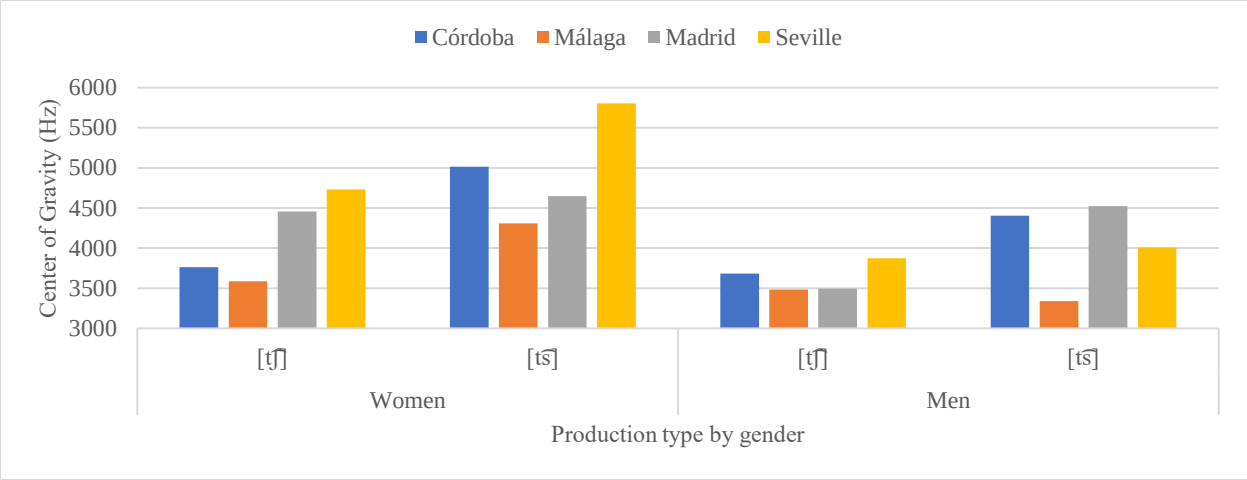

Figure 36: COG (Hz) of frication portion of the affricate by production type, city, and gender.

There were two interactions within the model identified that resulted from a lack of tokens in certain factors, including preceding context and word position. In order to prevent the reduced number of occurrences in certain cross-tabs from having an outsized weight on the model, and following the method described by Daleszynska (2015) to resolve factor interactions in mixed-effects logistic regressions, all 81 tokens of /t͡ʃ/ preceded by a medial vowel and all 42 tokens of word-initial /t͡ʃ/ were excluded from the analysis. The resulting mixed-effects logistic regression – which treated the speaker and word as a random effect and set the regional fronted production as the application value – identified seven variables that predict variation and is presented in Table 9.

Variable	factor	log-odds	tokens	% regional	factor weight
Preceding Category (p<0.001)					
	Alveolar				
	Consonant	0.735	206	51.0%	0.676
	Front vowel	0.037	1067	30.6%	0.509
	Back vowel	-0.772	1783	22.9%	0.316
	Range				*36.0*
Politician Gender (p=0.010)					
	Women	0.715	1566	38.1%	0.672

Men	-0.715	1490	16.4%	0.328
Range				*34.4*
Following Category (p<0.001)				
Medial vowel	0.520	1003	36.0%	0.627
Front vowel	0.083	454	29.5%	0.521
Back vowel	-0.602	1599	21.6%	0.354
Range				*27.3*
Audience (p=0.010)				
National	0.290	666	36.0%	0.572
International	0.265	505	27.5%	0.566
Local	-0.239	943	25.2%	0.441
Regional	-0.316	942	23.8%	0.422
Range				*15.0*
Context (p<0.001)				
Female interlocutor	0.221	933	30.9%	0.555
Scripted speech	0.068	1074	26.9%	0.517
Male interlocutor	-0.289	1049	25.2%	0.428
Range				*12.7*
Affricate Duration (p<0.001)				
continuous	log-odds			
+1	0.018			
Fricative COG (p<0.001)				
continuous	log-odds			
+1	0.001			
n=3056 df=15 Log-Likelihood=-1304 AIC=2638 R^2 Fixed=0.249 R^2 Total=0.552				

Table 9: Mixed effects logistic regression for affricate fronting with the fronted production as the application value, and Speaker and Word as random effects to account for individual variation

This model, which considered 3,056 tokens of /t͡ʃ/, has variables ordered based on the range in factor weight. The most predictive variable is the preceding category, with alveolar consonants and front vowels before the affricate favoring the fronted variant, and back vowels favoring the pre-palatal one, suggesting an effect of gestural positioning of the tongue in production. Second, for politician gender, women were considerably more likely to produce the fronted variant, using it about twice as much as their male peers. Third, the following phonetic category was predictive in describing variation, as both medial and front vowels favored the fronted variant, while back vowels favored the pre-palatal one. Fourth, the speech audience factored into variation. Both

national and international contexts favored use of the innovative fronted variant, whereas local and regional contexts disfavored it in favor of the pre-palatal variant. Fifth, for speech context, both scripted speeches and interviews with female interlocutors favored the alveolar variant, while unscripted interviews with male interlocutors had the greatest degree of pre-palatal variant usage.

The final two variables selected as predicting variation were continuous. First, affricate duration represents the length of the entire affricate segment, including both the fricative and stop portions. As duration increased, production of the fronted alveolar variant also increased, suggesting that faster speech rates may also play a role in innovative variant use. Finally, the COG of the fricative portion of the segment was the last selected variable. While the low log-odds signifies that the effect is slight, increasing COG values significantly favor increased use of the fronted variant, as shown descriptively in Figure 29, indicating that the difference in categories extends beyond the impressionistic alone.

4.3.2 Syllable-final /s/ aspiration and elision

For the next phenomenon, syllable-final /s/ aspiration and elision, another minor interaction between factors resulted from an imbalance in factor weighting in the preceding context. Following the description of balancing logistic regression models employed by Daleszynska (2015), all 3 tokens of /s/ with a preceding bilabial consonant were excluded from the analysis. Figure 37 shows a breakdown by politician gender, political party, and word-position of the phenomenon. Conservatives, particularly in word-medial cases, had greater levels of elision.

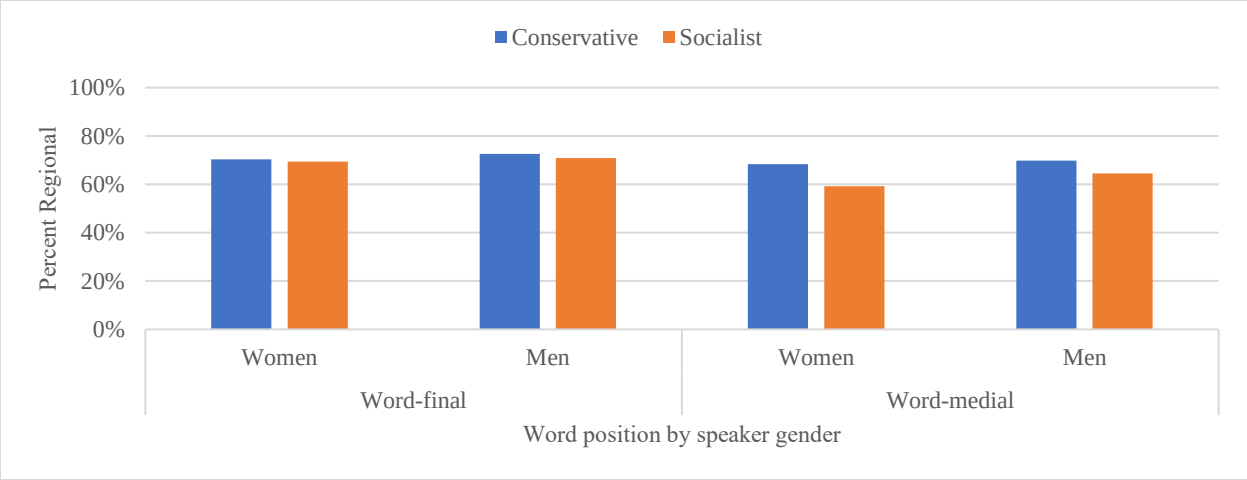

Figure 37: Regional production of syllable-final /s/ by political party, speaker gender, and position in the word.

Figure 38 shows the acoustic measure used for allophonic distinction. In this case, elision of syllable-final /s/ was distinguished from retention using COG (Hz). The alveolar production had by far the highest value, although it was particularly low in Málaga and among male speakers.

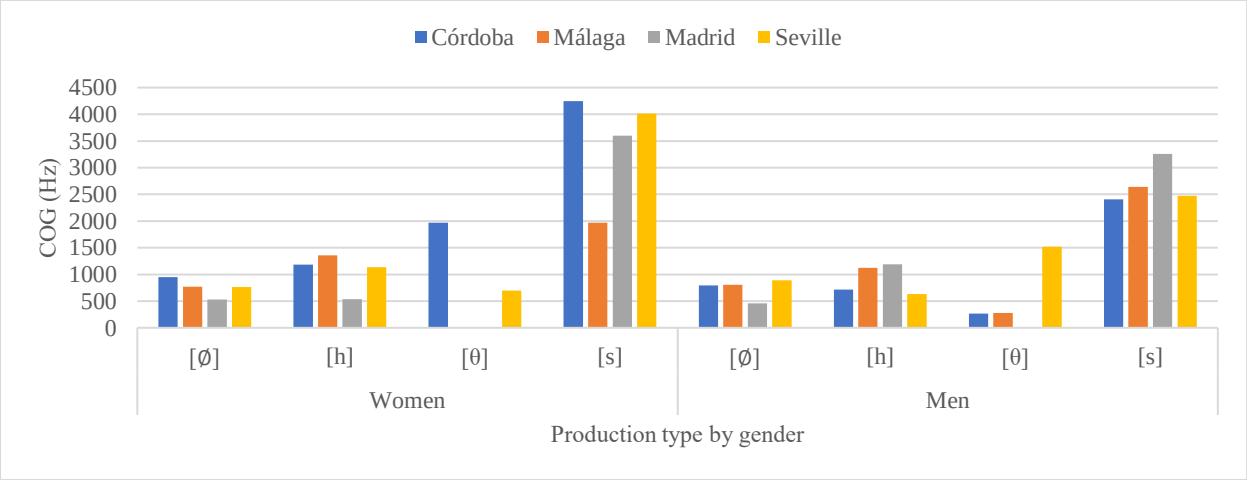

Figure 38: COG (Hz) of syllable-final /s/ by production type, city, and gender.

Following that, the mixed-effects logistic regression model, which set both aspiration and elision as innovative, regional variants, established six variables that significantly condition variation, as presented in Table 10. In this model, 5929 tokens ended up being considered. The first predictive variable in the model, politician city, reflects the strong regional tendencies related to syllable-final /s/ elision, with Seville, Córdoba and Málaga all favoring aspiration and elision, while Madrid strongly disfavors it. Second, following context shows a gestural effect, with pauses disfavoring the regional production in favor of retention. Third, for speech context, interviews with male interlocutors and scripted speeches favored regional productions, while interviews with female interlocutors favored retention.

Variable	factor	log-odds	tokens	% regional	factor weight
Politician City (p<0.001)					
	Seville	1.761	1631	91.80%	0.853
	Córdoba	1.712	1375	89.10%	0.847
	Málaga	0.984	1407	87.10%	0.728
	Madrid	-4.457	1516	9.80%	0.011
	Range				*84.2*
Following Context (p<0.001)					
	Labial	0.815	1354	75.70%	0.693
	Dorsal	0.606	903	73.80%	0.647
	Coronal	0.53	2761	68.90%	0.629
	Pause	-1.951	911	55.10%	0.124
	Range				*56.9*
Speech Context (p<0.001)					
	Male interlocutor	0.292	1946	70.10%	0.572
	Scripted Speech	0.23	1892	69.50%	0.557
	Female interlocutor	-0.522	2091	67.80%	0.372
	Range				*20.0*
Duration (ms; p<0.001)					
	continuous	log-odds			
	+1	0.009			
Style (Percent in video; p<0.001)					
	continuous	log-odds			
	+1	-1.12			
Center of Gravity (p<0.001)					
	continuous	log-odds			
	+1	-0.001			
n=5929 df=14 Log-Likelihood=-1113 AIC=2253 R^2 Fixed=0.720 R^2 Total=0.841					

Table 10: Mixed effects logistic regression for coda /s/ elision and aspiration, with non-sibilant productions (e.g., [h], [θ] and elision) as the application value, and Speaker and Word as random effects to account for individual variation

The final three variables are continuous. First, for segment duration, longer segments are more likely to favor aspiration and elision. Next, as the timestamp of variant productions increases, the regional variant is less likely to be produced, going against expectations regarding style as attention paid to speech and suggesting an influence of style-shifting behavior. Finally, for COG, as the Hz value increases in a segment, there is a reduced chance that the segment will be produced as a regional variant, given that sibilants tend to have much higher COGs than other segments, especially instances of [h] or elision.

4.3.3 Onset /s/ aspiration and deletion in word-final position

Following the syllable-final /s/, instances of word-final onset /s/ were examined to compare production tendencies between the two. These cases involved instances where resyllabification is expected in a word-final /s/ in pre-vocalic position (e.g., *las otras* [la.so.tɾas]), as these were seen as likely to be cases where weakening was more likely to be avoided, given the pseudo-onset position that these instances of /s/ take on. Figure 39 shows a breakdown by politician gender, political party, and morphological context for the phenomenon. Conservative women favored elision in plural and verbal contexts.

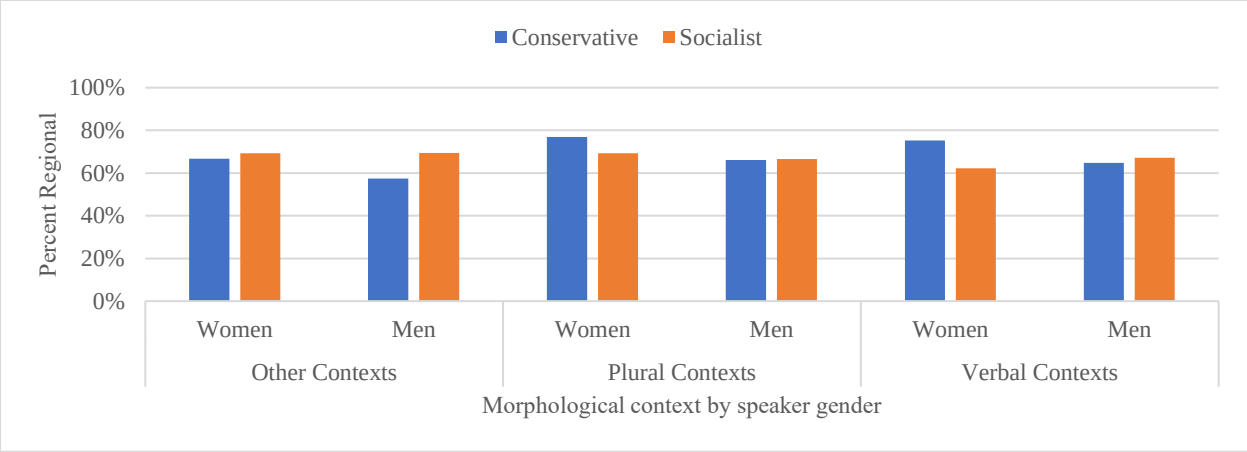

Figure 39: Regional production of word-final onset /s/ by political party, speaker gender, and morphological context.

Next, Figure 40 shows the acoustic measure used for allophonic distinction. In this case, elision of onset /s/ was distinguished from retention by COG (Hz). The alveolar production had the highest value.

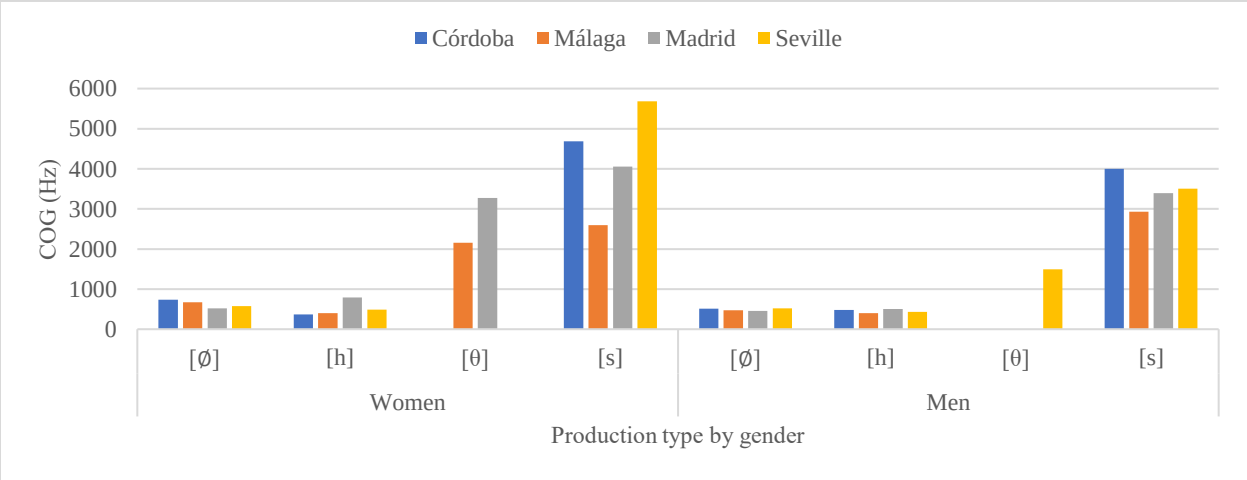

Figure 40: COG (Hz) of word-final onset /s/ by production type, city, and gender.

No interactions were found to result from cross-tab imbalances that required exclusion of data. The mixed-effects logistic regression model for this phenomenon pointed to five variables predictive of variation, presented in Table 11. In this analysis, a total of 1446 tokens were considered. The factor with the highest factor weight range is the city of the politician, as with syllable-final /s/. A similar hierarchy is shown here, although even more dramatic; AS varieties strongly favor aspiration and elision in this context about as often as in syllable-final position, while NCPS speakers elide and aspirate around a third as much in word-final onset position as in the syllable-final one. Second, speech context is also predictive here. Scripted speeches and interviews with female interlocutors favor regional production (a reverse from syllable-final /s/). Following speech context, the phonetic environment is represented by following context, with following dorsal sounds (including back vowels) being more likely to favor regional production, while coronal sounds (including front and central vowels) favoring retention.

Variable	factor	log-odds	tokens	% regional	factor weight
Politician City (p<0.001)					
	Seville	1.814	379	92.6%	0.86
	Málaga	1.417	368	90.8%	0.805
	Córdoba	1.058	336	86.9%	0.742
	Madrid	-4.289	363	3.6%	0.014
	Range				*84.6*
Speech Context (p=0.023)					
	Female interlocutor	0.644	526	69.2%	0.656
	Scripted speech	0.15	418	68.4%	0.537
	Male interlocutor	-0.794	502	67.7%	0.311
	Range				*34.5*
Following Context (p=0.002)					
	Dorsal	0.681	643	70.0%	0.664
	Coronal	-0.681	803	67.2%	0.336
	Range				*32.8*
Center of Gravity (p<0.001)					
	continuous	log-odds			
	+1	-0.003			
Politician Age (p=0.006)					
	continuous	log-odds			
	+1	-0.151			

n=1446 df=11 Log-Likelihood=-123 AIC=267 R^2 Fixed=0.922 R^2 Total=0.959

Table 11: Mixed effects logistic regression for onset /s/ elision and aspiration in word-final position, with non-sibilant productions (e.g., [h], [θ] and elision) as the application value, and Speaker and Word as random effects

Finally, the last two variables are continuous. For COG, as the value in Hz increases, the likelihood of regional production decreases, showing the reliability of this measure as a differentiating acoustic correlate. Lastly, for politician age, the older the speaker, the lower the chance that they will produce regional variants, reflecting a willingness for younger speakers to avoid normative productions, and potentially set aside traditional expected speech norms.

4.3.4 *Seseo*

The next sibilant phenomenon, *seseo*, yielded a minor interaction between factors related to the following context. As before, following Daleszynska (2015), all 35 tokens of /θ/ followed by a pause and all 10 tokens of /θ/ followed by a labial consonant were excluded from the model. Figure 41 shows a breakdown by politician gender, political party, and word-position of the phenomenon. *Seseo* was most common in word-final position, and was generally favored by female conservatives.

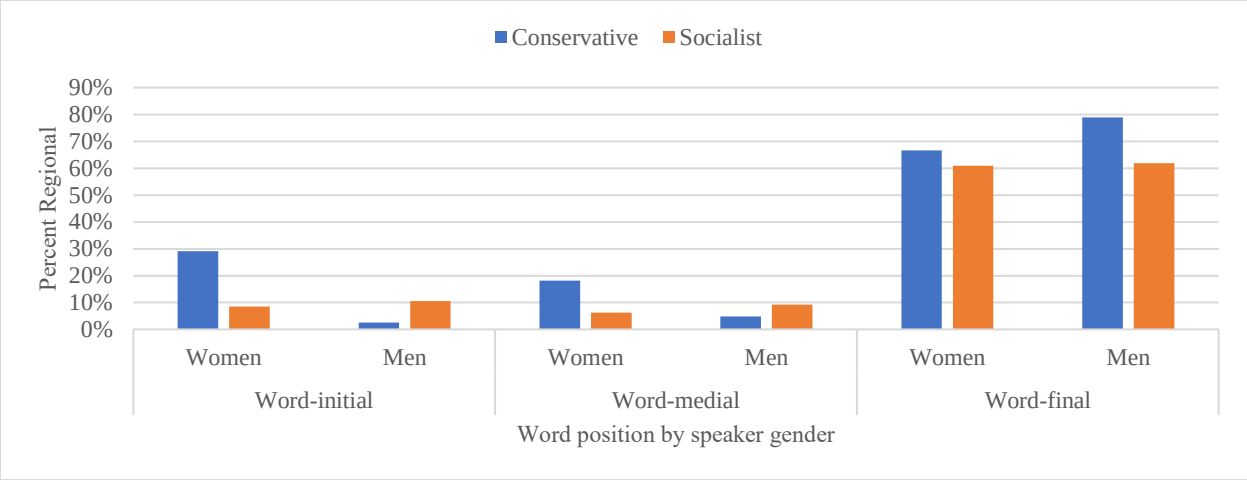

Figure 41: Regional production of seseo by political party, speaker gender, and position in the word.

Figure 42 shows the acoustic measure used for allophonic distinction. In this case, *seseo* was distinguished from [θ] by COG (Hz). The alveolar production had the highest value.

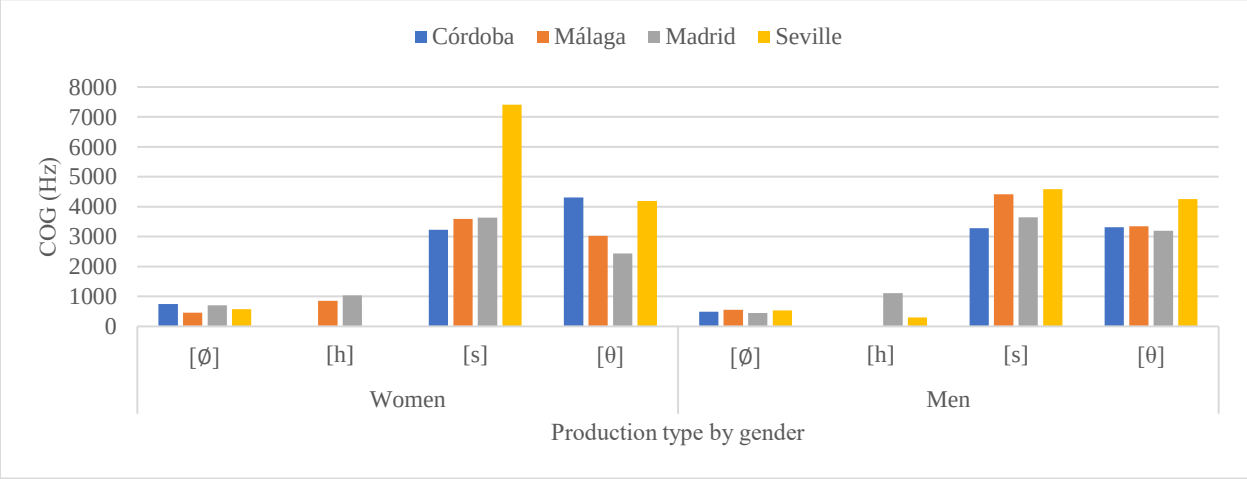

Figure 42: COG (Hz) of seseo by production type, city, and gender.

The resulting mixed-effects logistic regression model – which coded productions of /θ/ as retained [s] and aspirated or elided /s/ under the "regional" umbrella – determined that six variables conditioned variation, as presented in Table 12. Overall, 3174 tokens were considered in the analysis of *seseo*. The variable with the highest factor weight range, syllable position, was added to the independent variable list based on the differences noted between syllable-final and word-final onset /s/. While the cases are rare, it is much more likely for *seseo* to occur when /θ/ is in coda position. Second, for position in the word, *seseo* occurred most often in word-final position, with it being disfavored in both word-initial and -medial position. Third, for politician city, there is a considerable reduction in regional differences here. While all three AS cities favor *seseo*, even the city with the highest rates (Seville, 15%) is not so distance from the NCPS norm (Madrid, 5%), suggesting that the ongoing demerger in favor of distinction may well be playing a role in the speech of AS politicians. Fourth, for following context, dorsal sounds (including back vowels) favor *seseo* while coronal ones (including other vowels) disfavor them. Fifth, for the speech context, unscripted interviews with male and female interlocutors pattern similarly, favoring *seseo*, while it is disfavored in scripted speeches.

Variable	factor	log-odds	tokens	% regional	factor weight
Syllable Position (p<0.001)					
	Coda	2.272	53	79.20%	0.907
	Onset	-2.272	2571	9.10%	0.093
	Range				*81.4*
Position (p<0.001)					
	Word-final	1.573	39	69.20%	0.828
	Word-initial	-0.506	248	11.70%	0.376
	Word-medial	-1.067	2337	9.40%	0.256
	Range				*57.2*
Politician City (p=0.016)					
	Seville	0.732	687	14.80%	0.675
	Málaga	0.545	667	12.40%	0.633
	Córdoba	0.359	630	9.20%	0.589
	Madrid	-1.636	640	5.00%	0.163
	Range				*51.2*

Following Context (p=0.003)					
	Dorsal	0.377	308	20.80%	0.593
	Coronal	-0.377	2316	9.10%	0.407
	Range				*18.6*
Speech Context (p=0.017)					
	Male interlocutor	0.175	823	11.80%	0.544
	Female interlocutor	0.168	839	11.10%	0.542
	Scripted speech	-0.343	962	8.80%	0.415
	Range				*12.9*
Politician Age (p<0.001)					
	continuous	log-odds			
	+1	0.121			
n=3174 df=17 Log-Likelihood=-1356 AIC=2747 R^2 Fixed=0.246 R^2 Total=0.546					

Table 12: Mixed effects logistic regression for seseo, with expected sibilant-type productions of /θ/ (e.g., [h], [s] and elision) as the application value, and Speaker and Word as random effects

The final factor, politician age, is continuous. The older the politician, the greater the likelihood that they will produce *seseo*, following generational norms discussed in previous research surrounding the spread of NCPS norms through education among younger speakers.

4.3.5 *Ceceo*

The final sibilant phenomenon, *ceceo*, had no interactions that required exclusions from the model. Figure 43 shows a breakdown by politician gender, political party, and word-position of the phenomenon. *Ceceo* was most common in word-medial position, where it was favored by socialists – however, elsewhere it was more often used by conservatives.

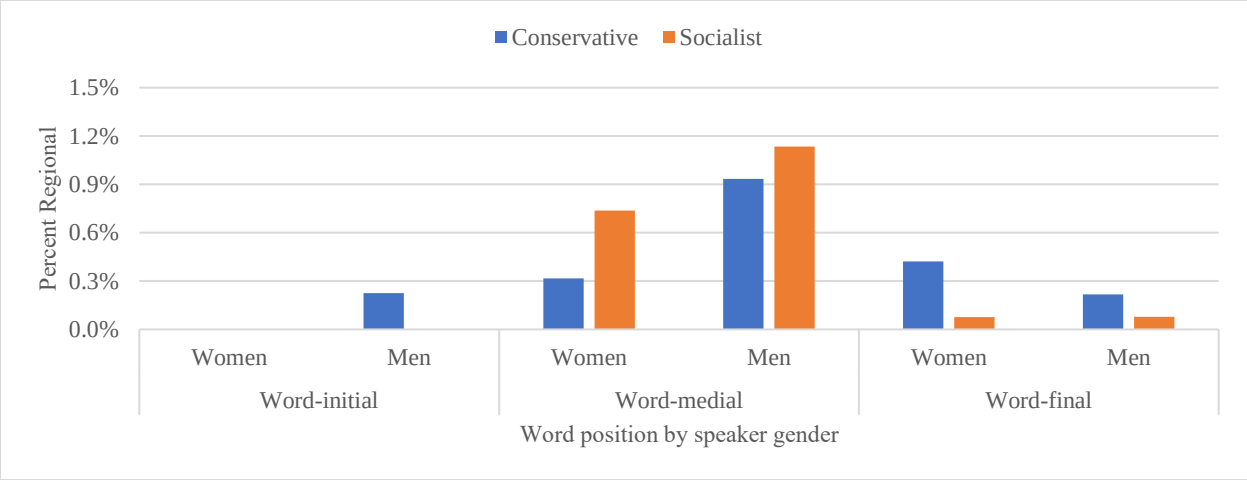

Figure 43: Regional production of ceceo by political party, speaker gender, and position in the word.

Next, Figure 44 shows the acoustic measure used for allophonic distinction. In this case, *ceceo* was distinguished from [s] by COG (Hz). The alveolar production had the highest value, particularly in Seville among women and in Córdoba and Madrid among men.

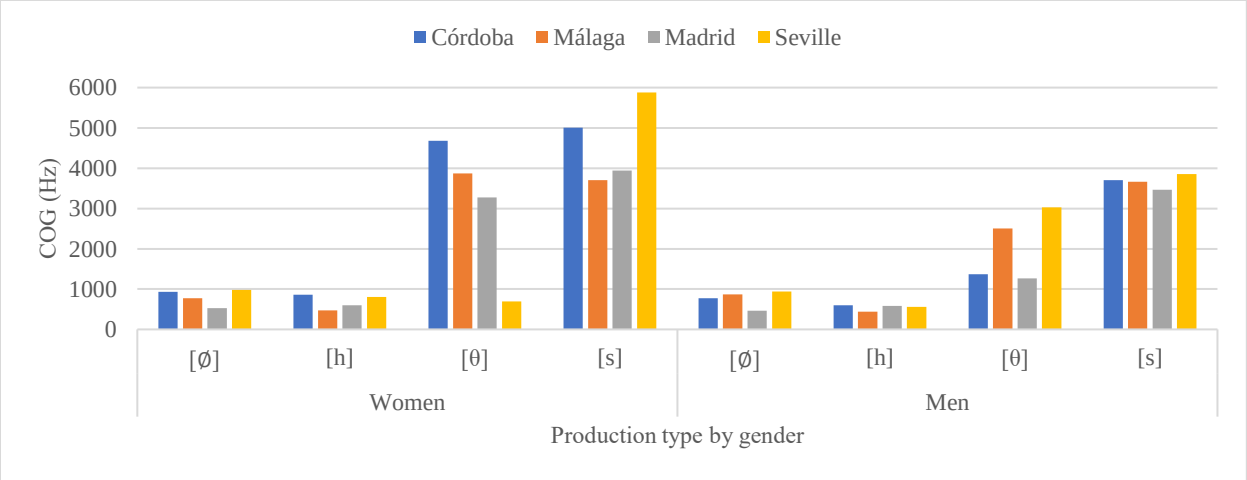

Figure 44: COG (Hz) of ceceo by production type, city, and gender

In the mixed-effects logistic regression model, which coded only productions of /s/ → [θ] as regional variants, five factors were predictive of variation, as seen in Table 13. It is important to note that, while 11,145 tokens of /s/ were analyzed in this model, only 43 of those were produced using *ceceo* as [θ] (i.e., 0.4%), driving home the extreme rarity of this phenomenon in political speech in current day AS. The model with the highest factor weight range is following context, with pauses disfavoring *ceceo*. Next, for position in the word, word-medial and word-final contexts favor *ceceo,* while word-initial ones disfavor it. Third, the morphosyntactic function of /s/ was also found to influence production of [θ], with both second-person use in verbs and plural use disfavoring production of *ceceo,* while all other contexts were more likely to favor it. Following morphosyntax, the politician's city did show a difference by region, with both EAS and WAS varieties in Seville and Málaga favoring *ceceo,* and NAS and NCPS in Córdoba and Madrid following previous dialectological descriptions and disfavoring it.

Variable	factor	log-odds	tokens	% regional	factor weight
Following Context (p<0.001)					
	Coronal	27.746	5534	0.7%	0.999
	Labial	26.323	1354	0.1%	0.999
	Dorsal	25.159	3346	0.1%	0.999
	Pause	-79.228	911	0	0.001
	Range				*99.8*
Position (p<0.001)					
	Word-medial	2.568	3935	0.8%	0.929
	Word-final	0.417	5398	0.2%	0.603
	Word-initial	-2.985	1812	0.1%	0.048
	Range				*88.1*
Morphosyntax (p<0.001)					
	Other	1.684	6412	0.5%	0.843
	Verbal	-0.693	1371	0.3%	0.333
	Plural	-0.991	3362	0.1%	0.271
	Range				*57.2*
City (p=0.005)					
	Seville	0.798	3085	0.6%	0.69
	Málaga	0.348	2616	0.6%	0.586
	Córdoba	-0.088	2603	0.3%	0.478
	Madrid	-1.058	2841	0.1%	0.258

Range		*43.2*
Center of Gravity (p<0.001)		
continuous	log-odds	
+1	-0.001	
n=11145 df=14 Log-Likelihood=-212 AIC=451 R^2 Fixed=0.985 R^2 Total=0.996		

Table 13: Mixed effects logistic regression for ceceo, with interdental production of /s/ (i.e., [θ]) as the application value, and Speaker and Word as random effects

The final significant variable, COG, reflects an acoustic trend. As COG increases, there is a reduced chance that *ceceo* will occur, reflecting the tendency for sibilants to have a higher COG value than the interdental fricative.

4.3.6 Intervocalic /d/ elision

Next, moving on to stop norms, elision of the intervocalic /d/ also yielded no factor interactions that required further cleaning of the data. Figure 45 shows a breakdown by politician gender, political party, and word-position of the phenomenon. Intervocalic /d/ elision was most frequent among conservatives, as well as in word-final, pre-vocalic position where it resyllabifies with the following word.

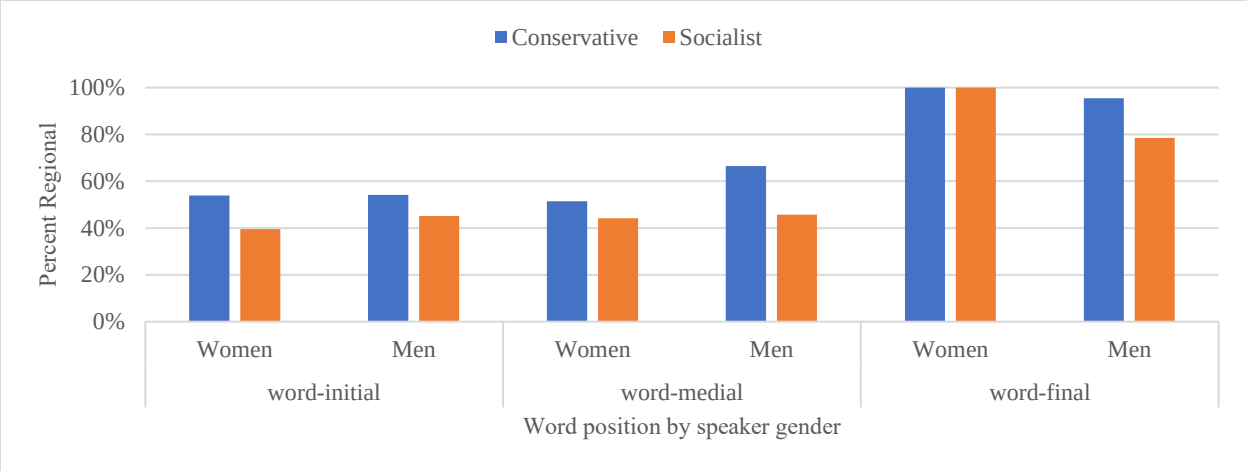

Figure 45: Regional production of intervocalic /d/ by political party, speaker gender, and position in the word.

Figure 46 shows the acoustic measure used for allophonic distinction. In this case, retained intervocalic /d/ was distinguished from elision by intensity range (dB). Occlusions were most marked for women, particularly those from Córdoba.

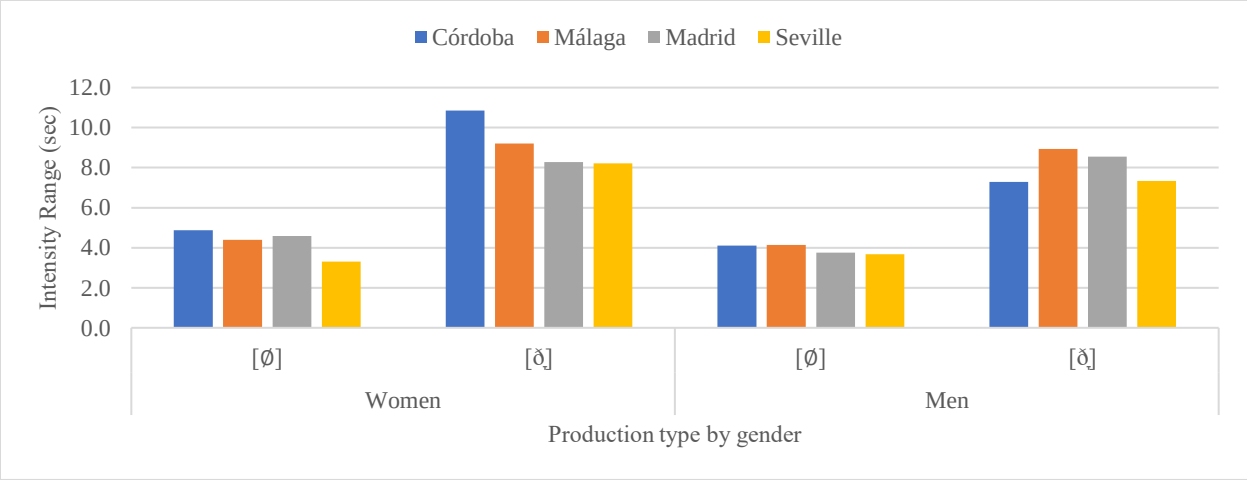

Figure 46: Intensity range (dB) of intervocalic /d/ by production type, city, and gender

In the mixed-effects logistic regression model, which treated elision as the regional variant, the seven factors predicting variation are shown in Table 14.

Variable	factor	log-odds	tokens	% regional	factor weight
Position (p<0.001)					
	Word-final	5.771	61	93.4%	0.997
	Word-medial	-2.845	2637	52.3%	0.055
	Word-initial	-2.926	1555	48.6%	0.051
	Range				*94.6*
Participle usage (p=0.002)					
	ado	0.478	483	54.5%	0.617
	other	0.154	3404	51.5%	0.538
	ido	-0.632	366	47.8%	0.347
	Range				*27.0*
Politician Political Party (p=0.017)					
	Conservative	0.304	2206	57.9%	0.575
	Socialist	-0.304	2047	44.6%	0.425
	Range				*15.0*
Preceding Context (p=0.003)					
	Low vowel	0.248	1519	54.4%	0.562
	High vowel	0.022	1154	51.6%	0.505
	Mid vowel	-0.27	1580	48.7%	0.433
	Range				*12.9*
Speech Context (p=0.038)					
	Female interlocutor	0.108	1415	57.5%	0.527
	Male interlocutor	0.057	1454	52.8%	0.514
	Scripted speech	-0.164	1384	44.1%	0.459
	Range				*6.8*
Intensity Range (p<0.001)					
	continuous	log-odds			
	+1	-0.468			
Duration (ms; p<0.001)					
	continuous	log-odds			
	+1	-0.028			
n=4253 df=14 Log-Likelihood=-1767 AIC=3562 R^2 Fixed=0.655 R^2 Total=0.727					

Table 14: Mixed effects logistic regression for intervocalic /d/ deletion with elision as the application value, and Speaker and Word as random effects

For intervocalic /d/ elision, 4253 tokens were considered. The factor with the greatest factor weight range is segment position in the word, with word-final productions of /d/ heavily favoring elision, while word-medial and -initial ones disfavor it. Second, for usage in participles, there was

a marked difference, as other studies in peninsular Spanish have observed, between production in -ado (which most strongly favored elision) and in other contexts, as compared to production in the participial ending -ido, which disfavored elision. Third, for political party, conservatives were markedly more likely to favor elision than socialists. Fourth, for preceding context, low and high vowels favored elision, while mid vowels disfavored. While this makes sense articulatorily for high vowels, it may be the case that -ado contexts increase the likelihood of elision following low vowels. Fifth, for speech context, unscripted interviews (especially with female interlocutors) favored elision while scripted speeches were more likely to show retention.

The final two variables were continuous. First, intensity range was shown to serve as a significant correlate to describe variable production – as the range increases (i.e., greater closures), regional variant use diminishes. Finally, for segment duration, longer segments are less likely to be produced with elision.

4.3.7 Lateral elision and variable production

The first liquid phenomenon, lateral elision and variable production[16], yielded a minor interaction between factors for position in the word. Following Daleszynska (2015), all 187 tokens of /l/ in word-final position preceded by a pause were excluded from the analysis. Additionally, in order to better understand the interaction related to coda /l/, the factor of syllable position was included in the consideration of variation. Figure 47 shows a breakdown by politician gender, political party,

[16] For all the liquid cases, possible productions included elision to [∅], as well as "variable production" as an allophone approximating another liquid (i.e., lateralization as [l], rhotacization as [ɾ], or production as the trill [r]).

and word-position of the phenomenon. Lateral elision and variable production was most frequent among socialist men and conservative women, and in word-final position.

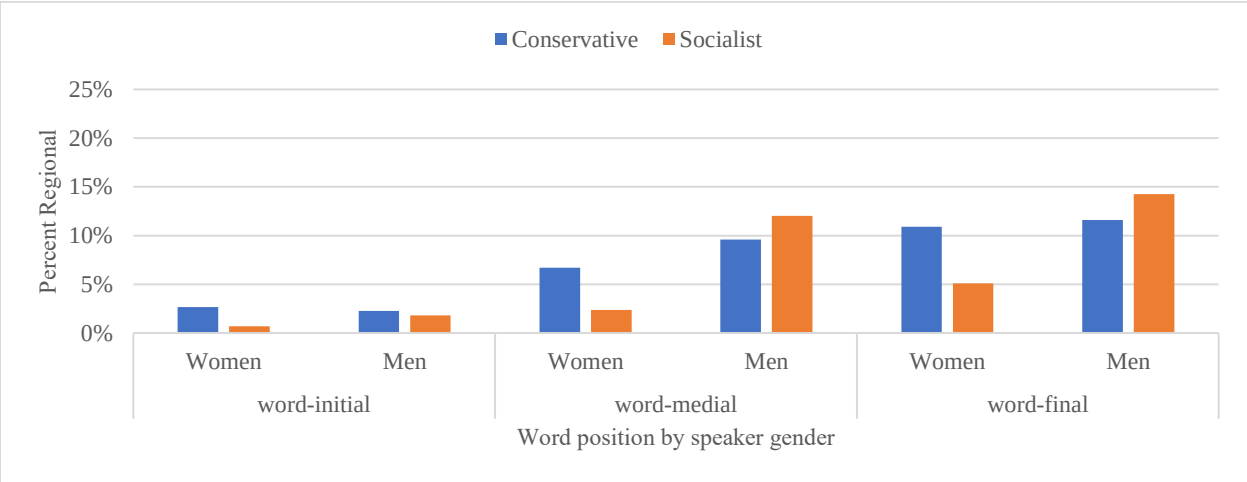

Figure 47: Regional production of the lateral by political party, speaker gender, and position in the word.

Next, Figure 48 shows the acoustic measure used for allophonic distinction. In this case, laterals were distinguished from taps, trills, and elision by F3 range. The range tended to be lowest for the lateral.

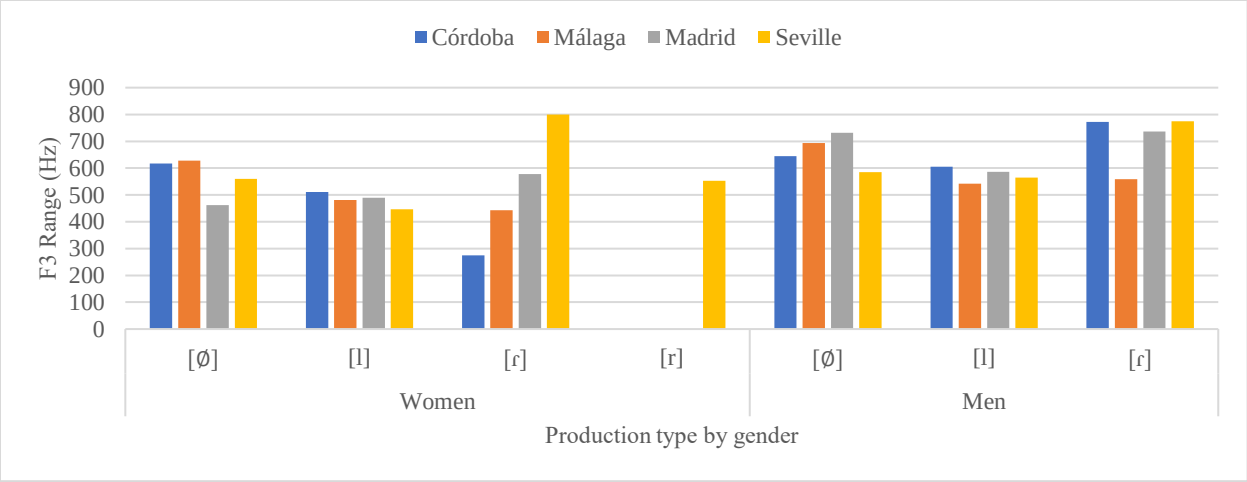

Figure 48: F3 (Hz) range of the lateral by production type, city, and gender

After these chances, a mixed-effects logistic regression model was developed that coded all productions of /l/ other than [l] as regional, yielding seven total variables that were found to condition variation that are provided in Table 15. A total of 6012 tokens were examined for the lateral. The highest factor weight range was associated with the preceding context, under which all contexts other than pauses were found to favor regional production. Following that, for the city, there was an interesting distinction drawn between EAS and NAS, on one side, and WAS and NCPS on the other. While Málaga and Córdoba favored regional production of the lateral, Madrid and Seville both disfavored it, following some suggestions in research on laterals that southeastern Spain has high rates of liquid permutation (Hernández-Campoy and Cutillas-Espinosa, 2010, 2013; Zahler and Daidone, 2014). Third, the variable of syllable position indicated a marked difference between /l/ in coda contexts, where it favors regional productions, and in onset contexts, where it disfavors it. The fourth variable, politician gender, reflects some community trends in previous dialectological research, showing that male politicians favor regionalisms while female politicians

disfavor them. Subsequently, for speech context, as with the intervocalic /d/, both unscripted interview contexts favored regional production, while scripted speech disfavored it.

Variable	factor	log-odds	tokens	% regional	factor weight
Preceding Context (p<0.001)					
	Labial	0.994	376	7.4%	0.730
	Dorsal	0.521	2481	7.1%	0.627
	Coronal	0.167	2962	5.4%	0.542
	Pause	-1.683	193	1.0%	0.157
	Range				*57.3*
City (p=0.014)					
	Málaga	0.727	1462	10.4%	0.674
	Córdoba	0.032	1399	6.1%	0.508
	Madrid	-0.313	1615	4.2%	0.422
	Seville	-0.446	1536	4.0%	0.390
	Range				*28.4*
Syllable Position (p<0.001)					
	Coda	0.572	1805	12.5%	0.639
	Onset	-0.572	4207	3.3%	0.361
	Range				*27.8*
Politician Gender (p=0.008)					
	Men	0.369	3168	7.8%	0.591
	Women	-0.369	2844	4.2%	0.409
	Range				*18.2*
Speech Context (p=0.009)					
	Male Interlocutor	0.145	2101	7.0%	0.536
	Female Interlocutor	0.129	2007	6.6%	0.532
	Scripted Speech	-0.274	1904	4.6%	0.432
	Range				*10.4*
F3 Range (p=0.032)					
	continuous	log-odds			
	+1	-0.001			
Intensity Range (p<0.001)					
	continuous	log-odds			
	+1	0.074			
n=6012 df=15 Log-Likelihood=-1158 AIC=2345 R^2 Fixed=0.204 R^2 Total=0.375					

Table 15: Mixed effects logistic regression for lateral elision and variable production with regionalisms (i.e., the tap, trill, and elision) as the application value, and Speaker and Word as random effects

The final two variables were continuous. First, for F3 range, there was a slight correlation between greater ranges in F3 and normative productions of [l]. Second, with respect to intensity range, as the range in intensity increased (i.e., closure size increased), regional productions were more likely, given the articulatory differences between taps/trills and laterals.

4.3.8 Tap elision and variable production

When analyzing the elision and variable production of the tap /ɾ/, no factor interactions were discovered that required elimination of tokens in either onset or coda position. Figure 49 shows a breakdown by politician gender, political party, and word-position of the phenomenon. Tap elision and variable production was most frequent among conservative women and socialist men, particularly in word-final position.

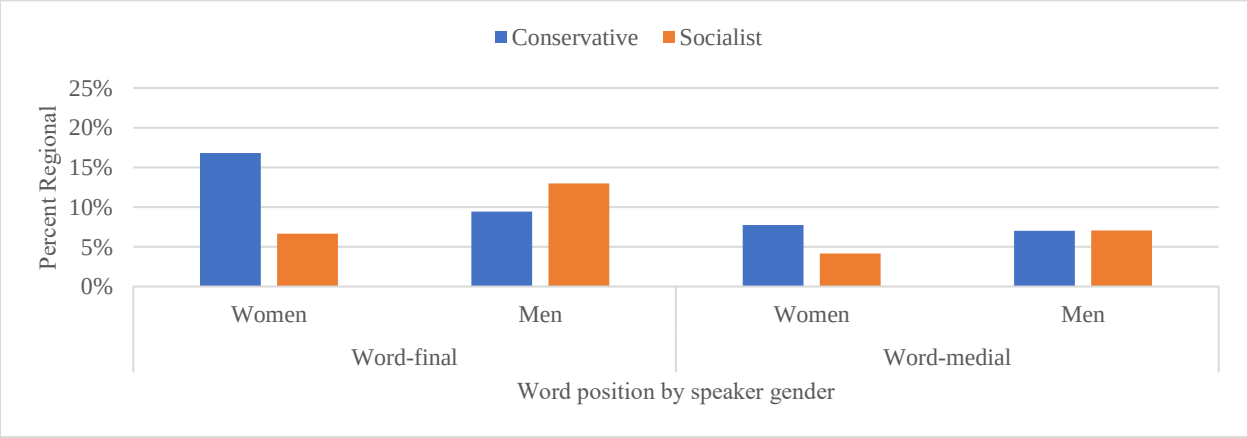

Figure 49: Regional production of the tap by political party, speaker gender, and position in the word.

Figure 50 shows the acoustic measure used for allophonic distinction. In this case, taps were distinguished from the lateral, trills, and elision by F3 range. The range trended higher for the trill and lower for the lateral overall.

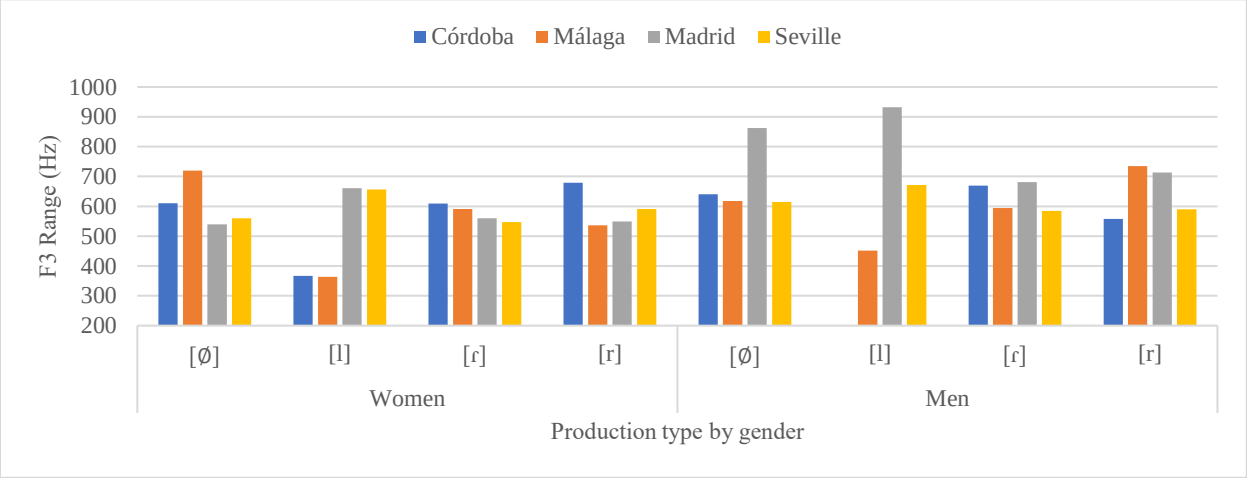

Figure 50: F3 (Hz) range of the tap by production type, city, and gender.

In the mixed-effects logistic regression model, which treated elision, trill, and lateral productions as regional variants, seven factors were found that significantly conditioned variation, as provided in Table 16. This analysis of the tap took into consideration 8050 total tokens. The variable with the highest factor weight range, following context, showed a distinction between pauses, which favored regional variants, and all other environments, which disfavored them. The next variable, city, showed a trend similar to the one determined for laterals, although now NCPS and NAS (rather than WAS and NCPS) group together, in opposition to WAS and EAS. While Madrid and Córdoba politicians were less likely to produce regional variants of the tap, politicians from Seville and Málaga were more likely to do so. Following that, for proceeding context, coronals (including central and front vowels) are more likely to favor regional production, whereas dorsals and labials

both disfavor it. Fourth, for political party, conservatives favored regionalisms, while socialists disfavored them. Fifth, for speech context, unscripted interviews favored regional productions while scripted speeches disfavored them.

Variable	factor	log-odds	tokens	% regional	factor weight
Following Context (p<0.001)					
	Pause	1.08	354	39.0%	0.746
	Labial	-0.132	422	19.2%	0.467
	Coronal	-0.391	3701	14.7%	0.404
	Dorsal	-0.557	3573	12.9%	0.364
	Range				*38.2*
City (p=0.016)					
	Málaga	0.336	1921	18.8%	0.583
	Seville	0.245	2081	18.5%	0.561
	Córdoba	-0.098	1976	13.4%	0.475
	Madrid	-0.483	2072	10.3%	0.382
	Range				*20.1*
Preceding Context (p<0.001)					
	Coronal	0.378	3236	20.0%	0.593
	Dorsal	-0.154	3853	12.3%	0.462
	Labial	-0.224	961	11.1%	0.444
	Range				*14.9*
Politician Political Party (p=0.018)					
	Conservative	0.226	4186	17.0%	0.556
	Socialist	-0.226	3864	13.3%	0.444
	Range				*11.2*
Speech Context (p<0.001)					
	Female interlocutor	0.175	2767	16.9%	0.544
	Male interlocutor	0.008	2628	15.4%	0.502
	Scripted speech	-0.183	2655	13.2%	0.454
	Range				*9.0*
Intensity Range (p=0.007)					
	continuous	log-odds			
	+1	-0.001			
F3 Range (p=0.027)					
	continuous	log-odds			
	+1	0.001			
n=8050 df=16 Log-Likelihood=-3101 AIC=6235 R^2 Fixed=0.078 R^2 Total=0.256					

Table 16: Mixed effects logistic regression for rhotic tap elision and variable production with regionalisms (i.e., the lateral, trill, and elision) as the application value, and Speaker and Word as random effects

The last two variables were continuous. First, for intensity range, as the range (i.e., closure size) increased, production of regional variables decreased, likely referencing the loss of the tap in favor of elision or lateral production. Finally, for the F3 range, as the range increased, regional production also increased, potentially indicating the larger range existent for laterals.

4.3.9 Trill elision and variable production

For the final liquid context, elision and variable production of the trill /r/, no factor interactions were discovered that required exclusion of tokens from the analysis. Figure 51 shows a breakdown by politician gender, political party, and word-position of the phenomenon. Trill elision and variable production was most frequent among conservative politicians, and particularly among men in word-initial position.

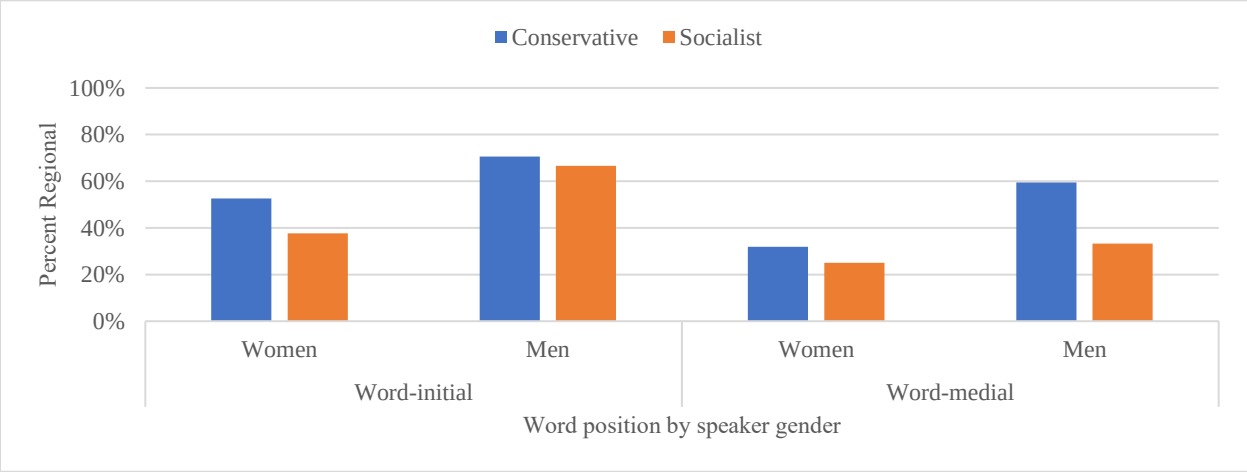

Figure 51: Regional production of the trill by political party, speaker gender, and position in the word.

Meanwhile, Figure 52 shows the acoustic measure used for allophonic distinction. In this case, trills were distinguished from taps and elision by F3 range. The range tended to be greater among men, but remained similar among women.

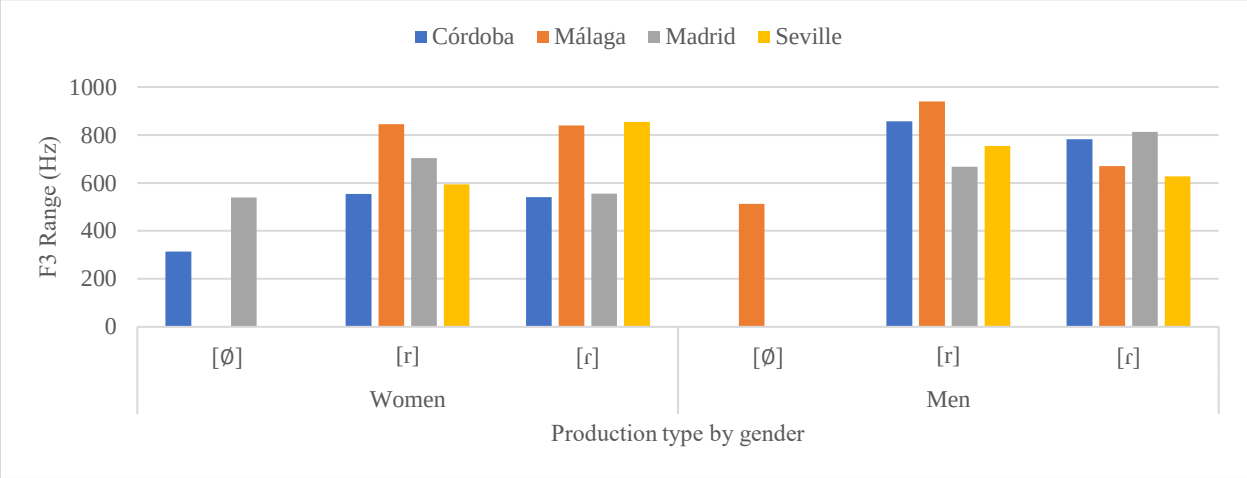

Figure 52: F3 (Hz) range of the trill by production type, city, and gender.

In the mixed-effects logistic regression model, which treated elision, tap, and lateral productions as regional variants, only two factors were found that significantly conditioned variation. These are shown in Table 17.

Variable	factor	log-odds	tokens	% regional	factor weight
Politician Gender (p=0.007)					
	Men	0.533	299	62.9%	0.63
	Women	-0.53	306	39.9%	0.37
	Range				*26.0*
Orthography (p<0.001)					
	<r>	0.52	447	57.0%	0.627
	<rr>	-0.52	158	34.8%	0.373
	Range				*25.4*
n=605 df=5 Log-Likelihood=-367 AIC=745 R^2 Fixed=0.108 R^2 Total=0.288					

Table 17: Mixed effects logistic regression for trill elision and variable production with regionalisms (i.e., the tap, lateral, and elision) as the application value, and Speaker and Word as random effects

In this analysis, a total of 605 tokens were considered – the trill being one of the most infrequent phones to occur in this analysis. Regional productions here included elision, as well as instances of lateralization to [l] and weakening to the tap [ɾ], and included both word initial and medial positions. Although both variables had similar factor ranges, gender was higher, with men being considerably more likely to produce regional variants than women. Following that, orthography was the next most predictive variable. Contexts without an explicit doubling of the grapheme (i.e., <rr>), which linguistic texts often refer to as requiring trill production, such as in word- and syllable-initial position (e.g., /#r/, /s.r/, /l.r/, /n.r/), favored production as a regional variant. This distinction reflects instances where the orthographic and phonemic boundaries do not align, as in word-medial position, <rr> corresponds to [r] and contrasts with <r>, which corresponds to [ɾ]. Instances of /r/ with the doubled grapheme that form contrastive minimal pairs with taps (e.g., *carro* ‘car’ vs. *caro* ‘expensive’) were much less likely to reduce or experience variable production.

4.3.10 Vowel raising and harmony

For the last phenomenon considered in this chapter, vowel raising and harmony in exclusively EAS, the scaled normalized F1 value was used as a continuous dependent variable for vowel

productions by Málaga politicians. No factor interactions were discovered that required exclusion of tokens from the analysis. Figure 53 shows a breakdown by politician gender, political party, and word-position of the phenomenon for F2. Vowel fronting did not differ markedly across phones, although word-initial /i/ and word-final /o/ showed slightly more fronted productions in words with elision.

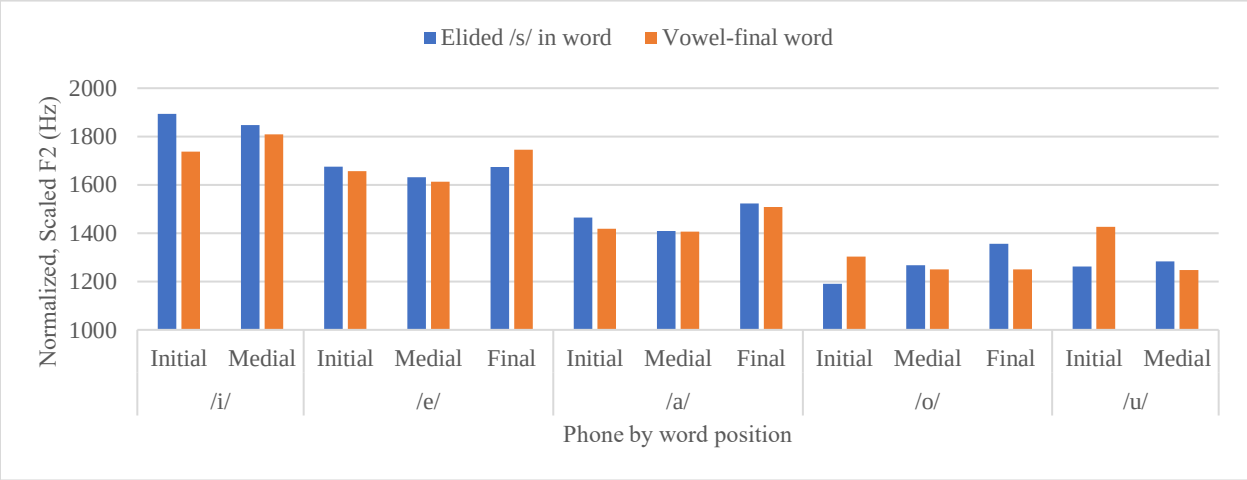

Figure 53: Vowel fronting (F2, Hz) by phone, context, and position in word.

Meanwhile, for vowel lowering in Figure 54, there was a greater tendency for all vowels to lower in elided contexts, although this was especially marked in word-initial position for /i/, /e/, and /a/, and in word-final position for /o/.

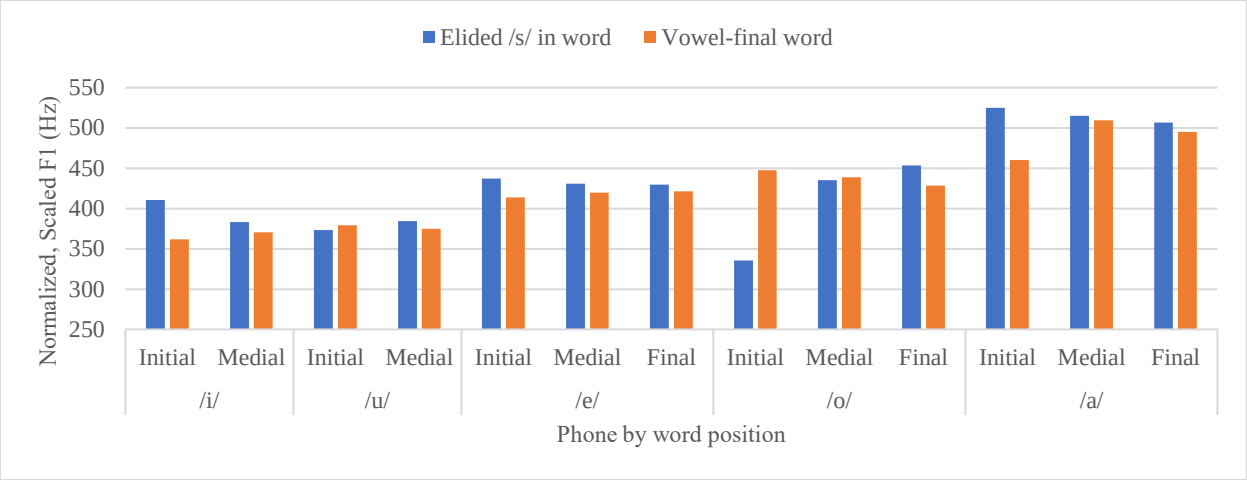

Figure 54: Vowel raising (F1, Hz) by phone, context, and position in word.

In the mixed-effects linear regression model, three factors were to significantly describe variation, given in Table 18. Given the high frequency of vowel occurrence, approximately 90 tokens were collected from each of the speakers, divided evenly between words with a final, elided morphological /s/ (e.g., *estudiantes* [estudiante] 'students') and those without the final /s/ underlyingly (e.g., *estudiante* [estudiante] 'student'). In sum, 691 vowels were analyzed in this model. As linear regressions do not provide factor ranges, the ordering of the three variables in the model is random, without indicating a meaningful hierarchy. One variable, the presence of elision, was used to investigate vowel harmony processes, as both words with word-final elided /s/ and words ending in an underlying vowel were examined equally. The significance of this variable indicates that, among these speakers, there was a slight lowering (i.e., avg. 15 Hz) among all vowels produced in words that had a final elided /s/, as compared to all vowels in words with an underlying final vowel. It is not clear if this falls within the realm of perceptibility.

Variable	factor	coefficient	tokens	mean F1 (Hz)
Elision Present (p=0.002)				
	Preceding elided /s/	5.811	317	445.1

No elision	-5.811	374	430.9
Vowel duration (ms; p=0.003)			
continuous	coef		
+1	230.88		
Scaled F2 (Hz; p=0.006)			
continuous	coef		
+1	0.030		
n=691 df=7 Log-Likelihood=-3693 AIC=7401 R^2 Fixed=0.028 R^2 Total=0.453			

Table 18: Mixed effects linear regression for vowel raising in EAS with normalized, scaled F1 height set as the continuous dependent variable, and Speaker and Phone as random effects

Another variable, vowel duration, suggests that as vowel length increases, vowel lowering is also more likely to occur, indicating that there may be an effect of speech rate and stress on vowel height. Speakers may lengthen their final utterance as a means of using duration for impact, which can in turn affect vowel height. Finally, the scaled value for F2 was also found to predict variation; as F2 rises (i.e., fronting), F1 also tends to rise (i.e., lowering). Notably, no social variation was found to occur for this variable, with male and female speakers from both political parties seeming to pattern similarly.

4.3.11 Regional trends

Narrowing our focus to just social variables, the next step in the analysis was to compile all phenomena into a single analysis. The significant social factors in each of the previous analyses are provided below, in Table 19. While city and speech context were predictive for six and seven of the phenomena respectively, the role of factors across individual variables drops off, with

gender descriptive for three phenomena, party and age conditioning two, and audience and percent in video each only significant in describing one phenomenon.

Phenomenon	**City**	**Speech Context**	**Gender**	**Party**	**Age**	**Audience**	**% in Video**
Affricate fronting		X	X			X	
Syllable-final /s/ aspiration & elision	X	X					X
Word-final onset /s/ aspiration & elision	X	X			X		
Seseo	X	X			X		
Ceceo	X						
Intervocalic /d/ elision		X		X			
Lateral elision & variable production	X	X	X				
Tap elision & variable production	X	X		X			
Trill elision & variable production			X				
Vowel laxing & harmony							

Table 19: Significant social factors across the ten phenomena

The resulting combined model considered how social variation predicted regional variant use, with all formerly established regional-normative distinctions in use. The speaker and the phenomenon were both treated as random effects to account for variation across individuals, as well as to differentiate between phenomenon-specific social tendencies. Four social factors were selected as predictive of regional variation across nine phenomena. Due to the high token to regional variant ratio of *ceceo*, that variant was not included in the analysis. The results for the mixed-effects logistic regression model are presented in Table 20, including results for phenomenon as a random effect to portray regional differences.

Variable	factor	logodds	tokens	% regional	factor weight
City (p<0.001)					
	Malaga	0.429	8377	39.1%	0.606
	Seville	0.386	8581	37.2%	0.595
	Cordoba	0.351	7684	36.2%	0.587
	Madrid	-1.166	8408	18.4%	0.238
	Range				*36.8*
Audience (p<0.001)					
	Local	0.427	12121	36.7%	0.605
	International	-0.069	5234	31.8%	0.483
	National	-0.129	6914	31.2%	0.468
	Regional	-0.23	8781	28.8%	0.443
	Range				*16.2*
Speech Context (p<0.001)					
	Female interlocutor	0.147	10948	34.2%	0.537
	Male interlocutor	0.096	11475	33.8%	0.524
	Scripted speech	-0.242	10627	29.8%	0.44
	Range				*9.7*
Politician Political Party (p=0.047)					
	Conservative	0.106	17260	34.5%	0.527
	Socialist	-0.106	15790	30.7%	0.473
	Range				*5.4*
Phenomenon (random effect)					
	Syllable-final /s/	1.622	5963	69.2%	0.835
	Word-final onset /s/	1.541	1446	68.5%	0.823
	Vowels	0.798	691	55.6%	0.689
	Intervocalic /d/	0.661	4253	51.5%	0.659
	Trill	0.471	605	51.2%	0.615
	Affricate	-0.325	3174	27.7%	0.419
	Tap	-1.141	8050	15.2%	0.242
	Seseo	-1.486	2669	11.4%	0.184
	Lateral	-2.131	6199	6.3%	0.106
n=33050 df=12 Log-Likelihood=-15190 AIC=30404 R^2 Fixed=0.080 R^2 Total=0.393					

Table 20: Mixed effects logistic regression for regional production across all 10 phenomena, with "regional" features as the application value[17], and Speaker and Phenomenon as random effects

Considering the nine phenomena, this model included 33050 tokens. The variable with the highest range in factor weight was city; AS varieties favored regionalisms, while the NCPS variety of Madrid did not. Following that, audience showed an interesting trend – local broadcasts favored

[17] Regional and normative productions are identified in Table 4.

regional variants, while all other contexts, including more broadly regional ones, disfavored them. Third, speech context was found to influence variation, with both types of unscripted interviews (i.e., with male and female interlocutors) favoring regional production and scripted speeches disfavoring them. The last variable, political party, showed a reversal of expectations set out by Hernández-Campoy and Cutillas-Espinosa (2010, 2013), among others in the realm of political speech. Overall, there was a greater tendency for conservative politicians to favor regional productions and socialist politicians to disfavor them.

4.3.12 Lexical Variation

Throughout the process of coding and analysis, a number of key terms were highlighted as being frequent in political speech, and therefore potential loci of variable production, in the vein of Hall-Lew et al. (2010, 2012). In particular seven high frequency, salient words were selected to examine the interaction of social factors in productions across several of the regional phenomena: Sánchez (referencing the PSOE Prime Minister, n=56)[18], *socialista(s)* 'socialist(s)' (referencing both the political party and the affiliation, n=135), militante(s)/militancia (referencing the fervor of the socialist constituency, n=23), Madrid/ *Madrileño* 'denizen of Madrid' (referencing the seat of government, n=54), *partido(s)* 'party(ies)' (referencing political parties by name, n=282), *político/a(s)* 'politics/political/politicians' (referencing specific parties or groups of individuals) and Andalucia/*Andaluz* 'denizen of Andalusia' (referencing the region, n=168).

[18] Numbers in this section represent all phenomena examined across all words, rather than a count of lexical items.

Sánchez: The first word under consideration, references the last name of the current Prime Minister of Spain, Pedro Sánchez. In sum, the word was examined a total of 56 times for both the fronted affricate (n=47) and *seseo* (n=9).[19] As represented in Table 21, regional differences were relatively few – however, at the level of political parties in NCPS, conservatives were much more likely to produce the fronted affricate than liberals, and women were more likely to front than men. Meanwhile, although political parties in AS produced similar rates of regional variants, male conservatives and female socialists favored fronting, showing a curiously inverted pattern.

Region/Social Variables	#	% regional
AS	**30**	**53%**
Conservative	15	53%
Women	9	44%
Men	6	67%
Socialist	15	53%
Women	8	75%
Men	7	29%
NCPS	**26**	**62%**
Conservative	13	85%
Women	8	100%
Men	5	60%
Socialist	13	38%
Women	5	60%
Men	8	25%
Total	**56**	**57%**

Table 21: Lexical productions of Sánchez

Socialismo/Socialista(s): Next, for words related to the socialist ideology, four phenomena were examined across 135 contexts, including syllable-final /s/ (n=43), the lateral (n=46), word-final onset /s/ (n=3), and *seseo* (n=43). The spread for these words are shown in Table 22. There is a marked regional difference, with speakers of AS referencing the term more frequently and using a

[19] Tokens for the *ceceo* context are not examined in this section, given the extremely low rate of regional variants used in the corpus.

much higher rate of AS regional variants (34%) as compared to NCPS speakers (4%). This is unsurprising, as unlike the fronted affricate in Sánchez, the phenomena considered herein are much more closely associated with southern Spanish. While the differences are slight, women are slightly more likely to use regional variants in AS, and socialist politicians use more regional variants than conservatives.

Region/Social Variables	#	% regional
AS	**85**	**34%**
Conservative	36	31%
Women	15	33%
Men	21	29%
Socialist	49	37%
Women	27	37%
Men	22	36%
NCPS	**50**	**4%**
Conservative	14	7%
Women	9	0%
Men	5	20%
Socialist	36	3%
Women	11	9%
Men	25	0%
Total	**135**	**23%**

Table 22: Lexical production of socialista/socialismo

Mi<u>l</u>itante(<u>s</u>)/Mi<u>l</u>itan<u>c</u>ia/Militar: Turning to a less-produced lexical item, examined in 23 contexts throughout this investigation, this word group is used most often to affectionately reference fervent left-wing supporters and is relegated exclusively to socialist politicians. A number of phenomena come into play here, including syllable-final /s/ (n=4), laterals (n=10), word-final onset /s/ (n=1), the tap (n=1), *seseo* (n=3), and EAS vowels (n=4). As shown by Table 23, this term is used most by women in AS, and while lateral productions and *seseo* never use regional variants, the remaining sibilant and vowel phenomena have high rates of regional production.

Region/Social Variables	#	% regional
AS	**16**	**31%**
Socialist	16	31%
Women	15	33%
Men	1	0%
NCPS	**7**	**0%**
Socialist	7	0%
Men	7	0%
Total	**23**	**22%**

Table 23: Lexical production of militante, militancia, militar

Madrid/Madrileño/a(s): Following that comes both the city name and referent to people from Madrid. The term is used considerably more in NCPS by Madrid politicians than by speakers of AS, for a sum total 54 occurrences in the data. Phenomena examined for these words include syllable-final /s/ (n=5), intervocalic /d/ (n=6), laterals (n=7), word-final onset /s/ (n=1), and the tap (n=35). As shown in Table 24, the highest rates of "regional" production of Madrid come, in fact, from the north. Likely in part due to the low lexical frequency in AS, no politicians produce it with regional variants. Even in NCPS, there is only a single socialist production of the item: it is conservatives, and particularly male conservatives, who lead the way in non-normative productions. This occurs through both the peninsular phenomenon of intervocalic /d/ elision (n=4), and the reduction of the tap (n=8), a phenomenon mainly only observed among these politicians in the production of these words.

Region/Social Variables	#	% regional
AS	**8**	**0%**
Conservative	5	0%
Women	1	0%
Men	4	0%
Socialist	3	0%
Women	3	0%
NCPS	**46**	**37%**
Conservative	45	36%
Women	32	28%
Men	13	54%

Socialist	1	100%
Men	1	100%
Total	**54**	**31%**

Table 24: Lexical productions of Madrid, Madrileño/a(s)

Partido/a(s): The next word is quite frequent across both regions under consideration, with 282 total occurrences across all the data considered in this chapter. Phenomena involved with production of *partido/a(s)* include syllable-final /s/ (n=15), intervocalic /d/ (n=130), the tap (n=128), and EAS vowels (n=9). As shown in Table 25, rates of non-normative variant use are comparable across regions and political parties. In AS, the largest divide is by gender, with male speakers tending to produce reduced /d/, /s/, and /ɾ/, as well as lowered vowels, most frequently. In NCPS, tap elision is nearly non-existent, whereas intervocalic /d/ elision averages above 60% across all social sub-groups (and approaches 100% in three). While the other gender-party intersections are all comparable in NCPS, the two male socialists have the lowest rate of intervocalic /d/ elision.

Region/Social Variables	#	% regional
AS	**175**	**39%**
Conservative	98	38%
Women	65	31%
Men	33	52%
Socialist	77	40%
Women	31	35%
Men	46	43%
NCPS	**107**	**37%**
Conservative	65	42%
Women	36	42%
Men	29	41%
Socialist	42	31%
Women	9	44%
Men	33	27%
Total	**282**	**38%**

Table 25: Lexical production of partido/a(s)

Po̲lítico/a(s̲): While not as frequent as *partido/a(s)*, this word was often used in conjunction with it, as well as to refer to other groups of politicians. The 115 tokens examined in these words came from four phenomena: syllable-final /s/ (n=17), the lateral (n=80), word-final onset /s/ (n=6), and EAS vowels (n=12). Table 26 shows variation in production for these words. While there is little lateral reduction in AS, all non-normative productions in NCPS come from variable production and elision of /l/ (i.e., [poi̯tico], [poɾitico]). On the other hand, sibilant reduction is highly frequent in AS. Female conservatives, in particular, have the highest regional production rate, dwarfing both male conservatives and all socialists in the region.

Region/Social Variables	#	% regional
AS	**90**	**32%**
Conservative	50	34%
Women	24	46%
Men	26	23%
Socialist	40	30%
Women	11	27%
Men	29	31%
NCPS	**25**	**8%**
Conservative	10	10%
Women	5	20%
Men	5	0%
Socialist	15	7%
Women	5	20%
Men	10	0%
Total	**115**	**27%**

Table 26: Lexical production of político/a(s)

Anda̲lu̲cia/Anda̲lu̲z̲(a): The final lemma under consideration relates to both the region of Andalusia and those who dwell therein. As with "Madrid," speakers from the area use the term much more than those from outside of it. A total of 168 tokens were drawn from these words, including the four phenomena related to syllable-final /s/ (n=8), the lateral (n=79), word-final onset /s/ (n=3), and *seseo* (n=78). As shown in Table 27, only conservative men from NCPS produced

the word at all, with 2 reduced lateral productions leading to their 25% rate of non-normative production. Among AS speakers, on the other hand, production was more frequent. Notably, while the political parties are comparable, men were less likely to use *seseo* and reduce sibilant phenomena than women. This is a phenomenon which will be discussed in greater depth in Stage 2 of the analysis, although broadly there seems to be a tendency for regional sibilant phenomena, particularly *seseo* and the fronted affricate, to be favored in female speech, hinting at the social prestige that these forms may possess.

Region/Social Variables	#	% regional
AS	**160**	**13%**
Conservative	81	10%
Women	12	17%
Men	69	9%
Socialist	79	15%
Women	73	16%
Men	6	0%
NCPS	**8**	**25%**
Conservative	8	25%
Men	8	25%
Total	**168**	**13%**

Table 27: Lexical production of Andalucia/Andaluz

4.4 Discussion

Based on the analysis, there is a clear divide for some phenomena between innovative variation attributed to Andalusian Spanish, while others offer good evidence to suggest that variation is less geographically restrained than has been previously described. Table 28 offers a breakdown for each phenomenon by city. For sibilant phenomena, while syllable-final and word-final onset /s/ reduction are both massively more common for all AS varieties, and *seseo* occurs around twice as

frequently in AS, affricate fronting is most common in NCPS and NAS. With respect to intervocalic /d/, elision is common across all varieties of Spanish, although it occurs the most among Madrid politicians. Among cases of liquid phenomena, EAS shows the highest rate of variable production and elision, followed by NAS for laterals and trills, and WAS for taps, with NCPS having near or the lowest rate. Finally, as vowels were only examined among EAS speakers, no comparisons can be made.

Phenomenon	**Córdoba**	**Málaga**	**Seville**	**Madrid**	**Total**
Affricate fronting	39%	18%	18%	37%	28%
Syllable-final /s/	89%	87%	92%	10%	69%
Word-final onset /s/	87%	91%	93%	4%	68%
Seseo	11%	13%	15%	6%	11%
Intervocalic /d/	46%	52%	49%	58%	51%
Variable /l/ production	7%	11%	4%	4%	6%
Variable tap production	13%	19%	19%	10%	15%
Variable trill production	50%	63%	43%	48%	51%
Vowels	n/a	56%	n/a	n/a	56%
Total	**36%**	**39%**	**37%**	**18%**	**33%**

Table 28: Production of AS variants by city

Overall, these results suggest that, while certain hyper-salient regional contexts for variation (e.g., production of /s/) differ considerably across speakers, there is less variation among politicians for others, especially phenomena like intervocalic /d/ elision, which has been previously described as more frequent in Andalusia (e.g., Samper-Padilla, 2011). In the sections that follow, we return to the research questions presented at the onset of this chapter, determining what takeaways this data has provided with respect to our understanding of social and linguistic variation in AS.

4.4.1 Variable use

The first research question pursued in this chapter involved patterns found in variable use, with respect to both social and linguistic variables. First and foremost, based on the description of acoustic correlates in Figures 29-34, as well as the inclusion of measures like COG, duration, and both intensity and F3 range in the mixed-effect models, the coding based on visual and auditory inspection of tokens correlates to the acoustic data. Differences exist between productions that permit analysis in categorical statistical models, as were presented here. Future research could also treat these continuous measures as dependent variables, following non-categorical approaches to variation such as those used by Díaz-Campos, Cole, and Pollock (2023) for the Caracas affricate.

In terms of linguistic factors, preceding and following context, COG, word position, syllable position, intensity and F3 range, morphological differences, duration, orthography, scaled F2, and the presence of elision were all found to condition variation for at least one phenomenon. Preceding and following context tended to reflect articulatory and gestural differences between variants, suggesting, for example, that affricates were more likely to be fronted when surrounded by segments produced near the front of the mouth. COG was an important measure for sibilants, reflecting acoustic differences described by Gordon et al. (2002) between alveolar and pre-palatal sibilants, as well as those expected between sibilants, [h], and the interdental fricative [θ]. Both word and syllable position tended to reflect tendencies for word-final and coda position to favor increased rates of vernacular productions, while word-initial and onset contexts favored normative productions, potentially resulting from differences in emphasis across these contexts. While intensity range distinguished between normative productions with occlusions (e.g., /d/, /ɾ/, and /r/) and regional instances of reduction, F3 range distinguished between laterals, which had greater

contrasts, and other liquids. Plural and second-person verbal contexts with morphosyntactic meaning disfavored *ceceo*, while for intervocalic /d/, the site of greatest regional production was the participial ending -ado, which others have described as a common context for elision among politicians (Ortiz-Cruz, 2019). For duration, regional productions of the affricate and syllable-final /s/ were more likely in longer segments, while they were less likely for intervocalic /d/. Finally, the orthography of <r> vs <rr> for trills influenced production, as did the height of vowels (i.e., F2), and contexts with or without preceding /s/ elision (e.g., Henriksen, 2017).

Meanwhile, for social factors, variation was conditioned by the speech context, audience, city, politician party, gender, age, and the percent in the video. With respect to speech context, interviews with female interlocutors tended to favor regional productions the most, followed by interviews with male interlocutors, while scripted speeches tended to disfavor them, although this ordering was reversed for syllable-final /s/ elision. On the one hand, this reflects Flores' (2014) findings in Chilean radio that speakers tend to alter their speech depending on the gender of their interlocutor. However, on the other hand, this finding also points to the common Spanish sociolinguistic trend favoring the use of Labovian style as "attention paid to speech," and casting speeches as differing in formality based on their scripted or unscripted nature (e.g., Hernández-Campoy & Cutillas-Espinosa, 2010; Labov, 1972; Ruch, 2012). Nonetheless, the fact that this hierarchy is consistent in the analysis of all regional productions calls back to the description of Cutillas-Espinosa & Hernández-Campoy (2007), who see style as a multidimensional entity that requires consideration in its original context. Politicians clearly reposition themselves linguistically based on their platform, a reality that also emerges through the audience of their discourse.

With respect to audience, there was a universal tendency for speech on local channels, directed at local audiences, to have the highest rate of regional variant usage, while speeches at the regional, national, and international levels were all more likely to contain normative variants. Interestingly, despite its predictive role in the overall model, audience was only found to be additionally predictive for affricate fronting, where its factors were reversed: the affricate was most likely to be fronted in discourse presented to national and international audiences, suggesting that this phenomenon has a different valence than many of the other AS-specific variants.

Following audience, the grouping for city often reflected the differences between AS and NCPS. Málaga, Seville, and Córdoba often favored regional variants and Madrid disfavored them, as was the case in the regional model. This was also the case for most of the sibilant phenomena—syllable-final /s/ elision, word-final onset /s/ elision, *seseo*, *ceceo*—while differing slightly for liquids, as EAS and NAS grouped together for lateral reduction and variable production, and EAS and WAS paired for tap reduction and variable production. Previous research has tended to associate EAS, in particular, with these liquid processes (Samper-Padilla, 2011; Ruiz-Peña, 2013; Zahler & Daidone, 2014).

Next, political party was the last variable to be selected in the overall regional model, with conservatives tending to produce more regional variants overall. While the overall difference between parties in the regional model was less than 4%, the differences were considerably higher in the individual models for intervocalic /d/ and the tap, suggesting a possible analysis of these differences. Schilling-Estes' (2013: 15) idea of automaticity of speech may play a role here, with many politicians deciding to follow general political speech "script" norms, while certain individuals (i.e., those associated with a conservative political party that, at the time of recording, had seen little success in Andalusia for decades) were more willing to deviate in order to make

themselves heard. Further, it may be the case that, as in the United States through the 2010s, there is a shift in political rhetoric and stance, beginning to appeal to more nationalistic ideals that would logically see conservative politicians appealing to regionalisms as a means of showing solidarity with "the common man." These possibilities, along with the political success the right-wing PP experienced in the 2018 elections, will be discussed in more depth in later chapters.

While politician gender was not included in the regional model, it was found to be predictive of variation for three individual models, including affricate fronting, the lateral, and the trill. While women favored the innovative fronted affricate, it was men who were more likely to use the reduced or variably produced lateral and trill. This suggests a difference not only in regional boundaries for the fronted affricate, but for prestige – it seems to experience mainly use by women, particularly those in or near NCPS. Meanwhile, the results for the liquids follow previous gendered findings, with other evidence from Andalusia suggesting that men favor reduced variants (Henriksen & Willis, 2010; Henriksen, 2014).

The final two variables were found to be significant predictors of variation for relatively few phenomena: age described variation for word-final onset /s/ and *seseo*, while the percent in the video (related to "attention paid to speech") was only predictive for syllable-final /s/. Interestingly, older speakers were less likely to reduce word-final onset /s/, while they were more likely to produce *seseo*. While the latter trend is not unexpected, given the correlation between older speakers, reduced access to nationalized education during the time of Franco, and the historical prestige of *seseo* in urban areas of WAS, as well as the tendency for younger speakers to move toward NCPS norms (e.g., Regan, 2017a; García-Amaya, 2008; Ruiz, 2017), the tendency for younger speakers to reduce word-final onset /s/ is more curious, and may result from a progressing change in which reduction in these contexts has become, by analogy, as frequent as in

syllable-final position. Finally, the variable of style-over-time, represented by the point in the speech when a given variant appears, was only selected as significant for word-final /s/ reduction. This defies expectations, with aspiration and elision being more likely earlier in the interview as other studies have found (e.g., Fafulas, Díaz-Campos, & Gradoville, 2018). This point is returned to in the next chapter with respect to individual changes in variation.

Thus, this chapter has presented a number of fruitful contexts in which variation is possible, demonstrating a rich tapestry of regional and normative variant production across both AS and NCPS political speech. While there are several clear differences, the tendencies with respect to social factors like political party, gender, and city are more complex and interconnected than previous studies of political speech have been able to show.

4.4.2 Contextual micro-social categories

The next research question aimed to move away from macro-social categories in an attempt to determine how third-wave categorizations could be applied to this data, in the same way that Eckert (2000) defined a sub-category of "burned-out" burnouts who were most extreme in their phonetic tendencies, as compared to "regular" burnouts, or Zhang (2005) defined the "smooth operator" based on their social backgrounds and phonetic performance.

In the current context, the overall regional analysis showed a tendency for conservatives to use increased rates of regional variants. When comparing regional production by political party, politician gender, and speech context, the highest rates of regional production are consistently in interviews with female interlocutors, particularly with female conservatives in all three AS

regions. Upon scrutiny, not even the majority of phenomena in these groups are produced with the highest rates of regional variants across the corpus: instead, there is consistency across the entire array. This suggests that a specific female conservative identity exists among speakers of AS, in which they use high rates of regional variants with female interviewers, a slightly lower rate with male interviewers, and around a 10% drop in scripted contexts. Female Andalusian socialists, on the other hand, tend to differ in their production norms, being some of the only groups to have comparable (i.e., Málaga) or even higher rates (i.e., Seville, Córdoba) of regional productions in scripted speech than in interviews, although their rates do not ever reach the same heights as conservative women.

These results follow Hernández-Campoy and Cutillas-Espinosa (2010, 2013) and Pollock and Wheeler (2022) in presenting a gendered political identity that goes beyond specific macro-social groupings. Both sets of women in Andalusia have found alternative and distinct means of integrating regional variation into their speech. Conservative women are more likely to emphasize Andalusian solidarity and belonging in unscripted interviews with other women, where they can use regionalisms as a means of expressing sincerity and a connection to the community (e.g., Sharma, 2018; Hernández-Campoy & Cutillas-Espinosa, 2013). Socialist women, on the other hand, are more likely to use locally-prestigious regional variants in ostensibly more "formal" scripted speeches, potentially as a means of emphasizing their sincerity directly to their voters, related to their addresses directly to the "militantes" who support them.

These differences in political party and the ways that regional variation is used as a means of expressing one's stance toward Andalusia, voters, and opponents, is discussed further in Chapter 5 and Chapter 6 through the lens of qualitative analysis and perception.

4.4.3 Lexical variation

Finally, the last research question opens the door to style-shifting analysis, albeit from a more Labovian community-wide perspective, asking how lexical items reflected variable use as a means of distinguishing individuals' identity goals. Seven lexical items were selected that were common in the data and related to discussions of specifically political matters.

The lexical analysis provided insight into certain group differences across lexical items. For example, there is a general tendency for socialist politicians to use words like *militance* and *socialism* at greater rates, while conservatives speak more about *political parties* and *Madrid*. Male conservatives are more likely to produce reduced variants when referencing *Madrid* and *socialism*, whereas socialist women are most likely to produce innovative variants when discussing *militancy* in their voting base. Female politicians, overall, were most likely to produce a fronted alveolar affricate [t͡s] when referencing the prime minister – although conservative male AS politicians also manifested this tendency.

Sharma's (2018) description of a correlation between regionalisms and a 'real me' personae among public speakers may help explain this behavior, with speakers using higher rates of innovative, non-normative variants as a means of drawing explicit attention to their speech behavior and the sincerity of their claims. Socialist references to *militantes*, and conservative productions of reduced *Madrid* may help emphasize their positive stance toward and affiliation with political ideology.

4.4.4 Final takeaways

Given the main goal of this chapter, which was to address the social and linguistic variation that exist across members of the political community of practice represented by this corpus, this closing section describes major takeaways from the quantitative analysis. Some sounds analyzed in this chapter, such as syllable-final /s/, word-final onset /s/, *seseo* and *ceceo* have followed previous expectations from the literature, with innovative productions being associated broadly with AS norms. Notably, while syllable-final and word-final onset /s/ were separated in this chapter to distinguish whether resyllabification in onset position led to higher rates of retention, the differences were not marked between these two contexts. Elision is frequent in Andalusian Spanish even in pre-vocalic position.

Some sounds, including laterals and taps, have had innovative productions associated with a sub-section of AS while still generally following expectations from previous literature. Other sounds have been less consistently associated with regional speech; for example, intervocalic /d/ elision might be better termed an "informal" and conservative phenomenon. Meanwhile, affricate fronting seems to possess a degree of overt prestige, occurring more frequently in female speech, with female interlocutors, and in national and international contexts. Regional reduced production of the trill, which is rare overall, is associated with male speech, rather than a specific region, while vowel laxing was only considered in EAS, and is not governed by social variation at all.

When combined with the group trends and lexical results, a picture forms of speakers, for the most part, following the script of political speech. Examining these dozens of audio files, we see that context is of the utmost importance. Hall-Lew et al. (2012) narrowed their focus by considering a single day's vote in U.S. Congress to examine variation in production of the second

vowel in Iraq, finding that political party became an important factor in distinguishing between productions. Hernández-Campoy and Cutillas-Espinosa (2010) and Pollock and Wheeler (2022) instead chose to examine a broad variety of production contexts while focusing on the tendencies of a single individual. While these results provide a good idea of how variation manifests at the national level in peninsular politics, it can only form a part of the overall picture.

In order to find where speakers deviate from the political script, it is necessary to zoom in further and examine individuals. However, in order to grasp the uniqueness of individual deviations, it is first necessary to understand the variation and norms existent in a community, as this chapter has done. While style-shifting is omnipresent in speech, with everything from topic to audience to individual identity goals influencing speech, some speakers choose to deviate more dramatically than others. Chapter 5 delves into a couple such individuals. Additionally, while we now understand variation among politicians, and how groups move outside of those norms, another vital piece of the puzzle is saliency and social meaning as perceived by the community. Chapter 6 examines how a group of 75 listeners from Seville evaluate audio taken from this corpus, assigning attitudinal values to regional and normative productions to determine the extent to which difference is observed, and the reactions that these differences receive in an urban space in Andalusia. Taking these results together, Chapter 7 assembles the pieces of the jigsaw puzzle, looking beyond any one part of this dissertation to understand how the community, the individual, and the listener fit together to give us a complete idea of variation in AS political speech.

5 Stage 2: Style-shifting and Lectal Focusing in Interaction

Based on the quantitative results in Chapter 4, the current chapter presses forward with an examination of linguistic variation at the individual level, tracking how style-shifting is conducted from moment to moment within speech. The basic idea is founded on the concept of indexicality (Silverstone, 2003) and identity construction (Le Page & Tabouret-Keller, 1985:5), which describe social meaning invested in language as serving a role in how speakers and interlocutors behave and interrelate. In particular, the current analysis relies on a methodology described by Sharma and Rampton (2015) that tracks Lectal Focusing in Interaction (LFI), which they see as a bridge between broad group analysis and individual behavior. LFI analysis tracks style shifting over time within speech, providing a combination of quantitative and qualitative methods to describe changing linguistic behavior and indexical meaning, which has been previously described as flexible and transitory (e.g., Johnstone & Kiesling, 2008; Eckert, 2008).

In their initial study proving the concept of LFI, Sharma and Rampton (2015) examine how members of a Punjabi community in London move between Standard British English, Vernacular London English, and Indian English. The authors discuss the typical quantitative approach to variation in sociolinguistic research, which cannot distinguish between the speech of younger and older male members of this community. Referencing the work of Bell, Coupland, and other second- and third-wave sociolinguistic studies, the authors argue that social meaning is more complex than macro-social factors alone can acknowledge, especially in the understanding of stylistic behavior.

Speaker agency, which is an important part of a modern understanding of sociolinguistic behavior (e.g., Ahearn, 2001), requires insight into the multidimensional reality of style (Cutillas-

Espinosa & Hernández-Campoy, 2007). Rather than accepting the community-level reality that young and old British Asian men do not differ significantly, Sharma and Rampton (2015: 10) look at the complexity of speech in the moment to see how similar production results are yielded from differing real-time tendencies. Variants can operate at differing orders of indexicality, they argue, based on the individuals under examination, meaning that old versus young male behavior may appear similar but be based on different reasons for speech differences. Thus, LFI analysis provides proportions of dialectal productions (i.e., in this case, the two varieties of British English and Indian English).

Sharma and Rampton reveal patterns of variable use that differ across age groups. While older speakers spent much of their lives in a hostile British context, younger speakers experienced a more welcoming context, and had fewer ties to India than their elders. Older speakers shift in a more nuanced way between ethnolinguistic traits (i.e., Indian English) and class ones (i.e., Vernacular London English), showing a high range of variation, whereas younger speakers have much more limited lectal focusing, with fewer ethnolinguistic traits emerging specifically in discourse work. This suggests a shift in meaning of Indian English traits from markers (i.e., possessing social meaning) among older speakers to indicators (i.e., devoid of that meaning) among younger speakers. As Sharma and Rampton (2015: 26) argue, the ethnolect becomes more accepted and less politicized in British life.

In a second study examining LFI analysis, Sharma (2018) goes on to apply the methodology to Fareed Zakaria, a CNN host in the public sphere. She compares variation between Indian English and American English. Here, she argues that there are clear moments where American English variants dip in a way that signifies something important about discourse (see Figure 4 in Chapter 2 for the LFI visualization). These moments, which she describes as

opportunities for the speaker to turn "off other styles in order to display the 'real me,'" or a moment when speakers allow their main style or lect to come forward, emphasizing their real feelings on a topic (Sharma, 2018: 24). When using stance markers showing his sincerity (i.e., "okay," "look"), or to use ironic humor, Zakaria also uses the highest rates of Indian English.

Sharma also references the effect that this behavior has in the realm of pragmatics on social face, noting that moving to one's primary lect can appear to be a type of divergence from the norms of the interlocutor under Communication Accommodation Theory (Giles et al., 1991). However, this behavior allows for a degree of sincerity and closeness that it insinuates reduced social distance, leading to a type of "divergence as solidarity" (Sharma, 2018: 26). Overall, these two studies show that LFI can be examined quite fruitfully not only in speech communities, but also among individuals in the public sphere, both of which serve as the goal of the current chapter.

This chapter is organized in the following manner. The first section discusses investigative goals, including questions that the LFI analysis aims to address and hypotheses for how variation may occur based on previous findings. The second section introduces the methodology of the analysis, including a description of the regional phenomena under consideration, the speakers selected from the corpus in Chapter 4, a description of the data coding process, and a rundown of the LFI approach. The third section introduces the results, giving a brief descriptive overview of variation in the data, and then looks at two groups of speakers selected for further examination following different selection criteria. Finally, the fourth section discusses the findings of the chapter, returning to the research questions and previous literature to examine style-shifting patterns in light of other third-wave sociolinguistic research.

5.1 Research questions and hypotheses

In southern Spanish politics, there has been a tendency for politicians to converge with Madrid norms when speaking to northern audiences. Hernández-Campoy and Cutillas-Espinosa (2010) find this tendency for politicians in Murcia, and Cruz-Ortiz's (2019) examination of political speech over decades shows a similar convergence in Andalusia. However, as both texts note, there is the possibility for individual variation. The former authors key in on María Martínez's hyper-vernacular tendencies, while Cruz-Ortiz finds at the speaker level that some individuals show high rates of Andalusian variant production that deviate from overall norms of political speech. These deviations from the "script" of the medium suggest a process of identity construction, by which some individuals strike out on their own, using high rates of regional variants in the process of identity construction. This begs the question; to what extent can this behavior be tracked using LFI analysis, and are Sharma's 'real me' moments also identifiable in political speech as a way of distinguishing politicians in a more nuanced fashion? The 'real me' moment can be used to show increased sincerity, sarcasm, and reduced social distance, potentially demonstrating moments where the speech norms identified in Chapter 4 fall away.

This reality leads to the second overarching question of the dissertation, which the current chapter aims to address: which regional variants are used fruitfully in style-shifting behavior, and how does use differ across communities of practice and individuals? By looking at moment-to-moment variation, patterns of variation can be identified – as Sharma and Rampton (2015) do – while individual behavior and topical shifts can also be examined in closer detail as Sharma (2018) does. All told, by also examining the tapestry of regional variants, we will see how variable usage of Andalusian phenomena patterns in speech, and how their meaning may differ based on

production. In the process of answering this question, the LFI analysis will address three related matters.

The first of these involves identifying trends of style-shifting behavior based on regional variants: which regional variants demonstrate stylistic variation and how? As Sharma and Rampton (2015) show, while variation may appear comparable from an overarching quantitative perspective, looking at individual behavior can result in marked differences. In the same way, while the quantitative results have shown broad patterns for each phenomenon, LFI analysis will help to show if they are all employed in the same way and in the same contexts. While some features of Andalusian Spanish may be produced regularly across speech, others may serve discursive functions, such as in 'real me' moments, or through the consistent navigation of social pressures in speech.

The second question addresses individual differences, asking: what patterns of stylistic variation exist across individuals and how do they both compare and differ? Individual behavior has played a major role in explaining variation in political speech in Spain (e.g., Hernández-Campoy & Cutillas-Espinosa, 2010, 2013; Cruz-Ortiz, 2019; Pollock & Wheeler, 2022), stemming from differences in agentive decisions and identity construction goals, as speakers navigate indexicality in the political sphere. At the level of LFI, this chapter aims to determine how the usage of Andalusian variants differs across individuals, and how speakers stand in their patterns of speech (e.g., hypervernacularity: Hernández-Campoy & Cutillas-Espinosa, 2010; MUSE versus AAL norms: Holliday, 2017; contextual differences: Kirkham & Moore, 2016).

Finally, the last question aims to connect the quantitative findings from Chapter 4 with the qualitative-quantitative mix yielded from LFI analysis, asking: what role do social categories and community-specific social divisions play in explaining style-shifting and identity work? In the

process of examining correlations and differences for the previous research question, commonalities may begin to emerge that group speakers based on their background, be it by factors such as gender, politically-affiliation, or region, as Hall-Lew et al. (2010) find for speakers from the U.S. Congress, or Hall-Lew et al. (2017) determine for Scottish parliamentarians.

Through these three questions, this chapter establishes how individuals use variants to construct public identities, and how that behavior maps onto the larger community of peninsular political speech. While previous research in English has shown that political speech can be classified by political parties (e.g., Hall-Lew et al. 2010, 2012; Hall-Lew et al. 2017), results from this project suggest that such classification may not hold true outside of specific moments and environments (e.g., voting for war, among politicians in a minority party). Speakers share overarching norms from their political communities of practice, but their goals are diverse, the expectations of their constituents are varied, and their contexts are ever changing. Thus, individuals likely have the ability to deviate from the automaticity of public speech. The LFI analysis in the current chapter provides unique access to these deviations.

5.2 Methodology

Third-wave sociolinguistics focuses on conceptualizations of speech style that go beyond macro-social factors, or even small-group tendencies alone, looking also at how individuals navigate specific contexts, and differ across specific environments and with differing agentive goals. This study took the results from Chapter 4 and identified 10 speeches and interviews out of the 94 that merited further consideration. These samples were taken from 10 of the 32 speakers, accounting

for gender and region to provide as balanced a view of style-shifting as possible. Rather than identifying independent variables, rates of style shifting are mapped out on a graph over time, breaking interviews and speeches into manageable units with sufficient instances of regional and normative productions to provide a nuanced depiction of variation. The resulting analysis offers insight into both broad patterns around style-shifting (e.g., Sharma & Rampton, 2015), as well as permitting a focus on individual moments where regional variant production rises. Uniquely, this LFI analytical approach determines how the topic and discussion itself influence use.

The remainder of this section is sub-divided into three parts. First, phenomena are described as dependent variables that contribute to the composite of regional variation. Next, the process is described for selecting speakers from the corpus in Chapter 4. Third, the process of data collection is described, following the same method as used in Chapter 4 to code audio. Finally, the methodology for LFI is described in detail, adapting the process used by Sharma and Rampton (2015) to allow for visualizations of real-time variable use.

5.2.1 Regional phenomena

As in Chapter 4, regional and normative features were used to distinguish Andalusian from Northern Central Peninsular Spanish (NCPS). Of the ten phenomena previously discussed, three changes were made regarding phenomena representation in the style-shifting portion of the analysis. First, vowel laxing, which did not reveal any social correlations and was confined solely to Eastern Andalusian Spanish (EAS) was not included here to avoid influencing comparability across Andalusian speakers (although that offers an opportunity for future examination). Next, as

the phenomenon of word-final onset /s/ reduction performed similarly to syllable-final /s/ reduction, the two categories were elided into the descriptor of syllable-final /s/ reduction. Finally, although coded in the analysis, only 10 tokens of 4163 included instances of *ceceo*. In order to prevent rare regional production of this process from unduly diluting regional production measures, *ceceo* was also excluded from the final analysis.

The remaining seven phenomena were considered in the LFI analysis, based on their association with Andalusian Spanish (AS) regional expectations (e.g., Pollock, 2022, 2023; Pollock & Wheeler, 2022; Samper-Padilla, 2011; Melguizo-Moreno, 2010; Regan, 2017a/b). For each phenomenon, possible productions were coded, following the same schema as in Chapter 4, based on "normative" productions associated with NCPS and "regional" ones expected for AS. These phenomena include:

1. **Affricate /t͡ʃ/ fronting to [t͡s]:** This variation seems common across the peninsula. The normative production was defined as the alveolar affricate [t͡ʃ], while the regional one was described as the fronted [t͡s].
2. **Syllable-final /s/ reduction to [h] or [∅]:** A very salient feature of AS (e.g., Gil-Peña, 2004; Ruiz-Dominguez, 1997; Samper-Padilla, 2011). Aspiration, *ceceo*, and elision (i.e., [h], [θ], and [∅]) were regional, retention and fortition (i.e., [s] and [z]) were normative.
3. ***Seseo*, where /θ/ is produced as [s]:** Associated with WAS and Seville (Harjus, 2017), although rates are falling across Andalusia in favor of distinction (Regan, 2017a; Villena-Ponsoda, 2008). Production as [s], [h], [θ], and [∅] were regional productions expected from /s/, while production as [θ] were normative.

4. **Intervocalic /d/ deletion to [∅]:** Reduction occurs throughout the peninsula, but is associated in the literature with AS (Gil-Peña, 2004; Ruiz-Martínez, 2003; Samper-Padilla, 2011). Elision to [∅] was regional, while the dental approximate [ð̞] was normative.
5. **Reduction and variable production of /l/:** Correlates with Andalusian and Southern Spanish identity (Hernández-Campoy & Cutillas-Espinosa, 2010, 2013; Ruiz-Peña, 2013). Elision to [∅] and variable production as [ɾ] and [r] were regional, while lateral production as [l] was normative.
6. **Reduction and variable production of /ɾ/:** Part of EAS, although tap reduction may be less common than trill reduction (Ruiz-Peña, 2013; Zahler & Daidone, 2014). Elision to [∅] and variable production as [l] and [r] were regional, while the tap [ɾ] was normative. Productions included instances of the tap in onset and coda position word-initially, medially, and finally.
7. **Reduction and variable production of /r/:** Associated with male and older speech (Henriksen, 2014; Henriksen & Willis, 2010; Hernández-Campoy & Cutillas-Espinosa, 2013; Zahler & Daidone, 2014). Elision to [∅] and variable production as [l] and [ɾ] were regional, while production of the trill [r] was normative.

Tokens were subjected to the same process as in Chapter 4, being inspected in Praat based on acoustic correlates in the spectrogram and waveform. Cues including the center of gravity (COG, Hz), duration (ms), intensity range (dB), and formant range (Hz) were considered during coding and data cleaning to ensure that tokens followed the same trends as previously identified for the quantitative analysis of community norms. Sounds were heard several times to aid impressionistic coding and determine spectral properties. After this analysis, dependent variables

were placed into the aforementioned regional and normative groups. This permitted the categorical analysis of lects, for which LFI analysis is perfectly suited (Sharma & Rampton, 2015). Normative and regional productions for this chapter are distinguished in Table 29.

#	Variable	Normative	Regional
1	Affricate fronting	/t͡ʃ/	[t͡s]
2	Syllable-final /s/ reduction	/s/	[h], [θ], [∅]
3	*Seseo*	/θ/	[s], [h], [∅]
4	Intervocalic /d/ reduction	[ð̞]	[∅]
5	Lateral /l/ reduction & variable production	/l/	[ɾ], [r], [∅]
6	Tap /ɾ/ reduction & variable production	/ɾ/	[l], [r], [∅]
7	Trill /r/ reduction & variable production	/r/	[ɾ], [l], [∅]

Table 29: Dependent Stage 2 variables, including normative/regional divisions

5.2.1 Corpus

The data analyzed and described in Chapter 4 were used as a means of reexamining the 32 speakers in the corpus, previously assembled and analyzed by Pollock (2022, 2023). These 94 audio files came from news media and were made publicly available online between 2011 and 2019. Appendix A: Table A provides a full breakdown of these files, which are evenly balanced by politician gender, age, political affiliation, city, and speech context. This included 16 male and 16 female politicians, half born before 1965, half after 1965, with 8 from Córdoba, Málaga, Seville, and Madrid. Half were from the left-leaning political party Spanish Socialist Worker's Party (PSOE), and half were from the right-leaning People's Party (PP). Three contexts were included for each speaker: interviews with male interlocutors, with female interlocutors, and scripted speeches. Politicians were born less than fifty miles from their city, and held comparable positions

in urban, regional, and national politics, which gave them regular chances to speak with news media.

From this cohort of speakers, a series of t-tests were performed to examine the differences in regional production among speakers in the coded audio. Stage 1 examined both the first minute of speech and a minute at the midpoint of each speech to examine the oft-described "formality" effects of style common in Spanish sociolinguistics. Taking this distinction in speech utterance, a t-test was run for each speaker that compared average regional productions of each phenomenon at the beginning and midpoint of speech. The results are provided in Table 30, showing a spread of variance from 2.14 to 12.36. Individuals were selected across the spread of variance (and are presented in the results in this order) to determine the extent to which increased regional production affected style-shifting over time. Individuals were picked based on social categories: two speakers were picked from every city, alternating based on political affiliation and gender, except in Seville – the focus of the perceptual examination – from which four speakers were selected.

Speaker	**Variance**	**Speaker**	**Variance**
1	2.14	32	5.06
7	2.91	27	5.06
16	2.97	31	5.07
15	2.99	23	5.18
8	3.03	11	5.42
4	3.06	2	5.44
21	3.15	14	5.61
5	3.19	22	5.75
20	3.22	24	6.16
10	3.48	25	6.24
30	3.50	17	7.22
12	3.51	3	7.30
9	4.11	29	7.64
13	4.68	19	7.67
26	4.89	18	9.47
6	4.95	28	12.36
Total Average			**5.08**

Table 30: T-test for politicians comparing beginning and midpoint

Following the identification of these speakers, a second t-test was run to separate audio files based on the context: scripted speeches, interviews with women, and interviews with men. For the three speakers with the lowest variance (i.e., 7, 4, and 5), the context with lowest variance was chosen; for those three with the highest variance (i.e., 22, 19, 18), the highest variance was chosen; and for the remaining four speakers, the context with mid-level variance was selected. The exact social information about these ten speakers is provided in Table 31, including their political party, city of origin, gender, speech context, year of speech, and duration of the audio. It is curious to note that socialists had the lowest rates of variance over time, while conservatives had the highest rates, both within the table and across the entire corpus: across all 32 speakers, 10/16 speakers with the lowest variance were socialist, while 10/16 of those with the highest variance were conservative.

#	Party	City	Gender	Context	Year	Age	Length
7	Socialist	Seville	Female	Speech	2015	41	24 min.
4	Socialist	Madrid	Female	Male interlocutor	2013	53	12 min.
5	Socialist	Seville	Male	Female interlocutor	2014	59	10 min.
26	Socialist	Malaga	Male	Speech	2017	43	14 min.
32	Conservative	Malaga	Female	Female interlocutor	2016	43	8 min.
11	Socialist	Cordoba	Female	Male interlocutor	2018	61	12 min.
14	Conservative	Madrid	Male	Speech	2018	54	31 min.
22	Conservative	Cordoba	Male	Male interlocutor	2015	40	19 min.
19	Conservative	Seville	Female	Female interlocutor	2014	53	4 min.
18	Conservative	Seville	Male	Male interlocutor	2017	60	47 min.

Table 31: Audio files for Stage 2 by speaker social background

5.2.3 Data Collection

As with the community analysis in Chapter 4, the contexts where any of the seven final regional phenomena could appear were coded using the Praat phonetic software (Boersma & Weenink, 2023). This coding methodology observed the sociolinguistics Principle of Accountability, requiring examination of the whole envelop of variation to establish community norms in a balanced way (Díaz-Campos, 2014).

Consistency was important across tokens, although acoustic measures would not be implemented in the LFI analysis. Nonetheless, those measures found to significantly describe variation of the phenomena in the previous chapter, including intensity range (dB), COG (Hz), and differences in formant values (Hz) were used to identify and distinguish between categorical coding of tokens as regional or normative. This provided consistency across coding of not only the 76 minutes of speech analyzed here, but also of the 188 minutes previously examined in Chapter 4, providing comparability across both stages of the dissertation. The same Praat script was used to collect information from the textgrid files as in the previous chapter (shown in Appendix B), although all information besides the speakers' identification and categorical production type were left for future consideration, as they could not be employed in the LFI analysis.

In addition to acoustic correlates, it is important to note that the inter-rater reliability testing described in Chapter 4 also applies to these data. Two native speakers of Western Andalusian Spanish (WAS) from Seville, Spain, coded stimuli following the instructions provided in Appendix C. These raters both had linguistics backgrounds, and coded a randomly-chosen selection of audio from the corpus (n=1343 tokens). Previous studies have performed IRR on around 1% of tokens (e.g., Gradoville, 2014: 96; Brown, 2008: 48), but given the range of social factors, phenomena, and geographical areas under consideration here, 2.3% of the total data across both stages was considered. The native speakers had around 90% agreement with the author's

coding, with agreement at its highest for sibilant phenomena, where the presence and absence of /s/ is extremely marked. Agreement did not fall below 80% for any phenomenon. These tendencies are provided in Table 7, reproduced below.

#	Phenomenon	Stage 1	Stage 2	Total Tokens	IRR %	IRR Tokens	Agreement
1	Affricate Fronting	3174	265	3439	2.0%	68	80.90%
2	Coda /s/ elision	5963	2652	8615	2.3%	198	97.50%
3	Word-final onset /s/ elision	1446	-	1446	2.9%	42	97.60%
4	*Seseo*	2669	965	3634	2.3%	82	92.70%
5	*Ceceo*	11145	4163	15308	2.3%	352	98.30%
6	Intervocalic /d/ elision	4253	1527	5780	2.2%	125	80.80%
7	Variable tap production	8050	3020	11070	2.3%	259	88.00%
8	Variable trill production	605	240	845	1.8%	15	80.00%
9	Variable lateral production	6199	2327	8526	2.4%	202	92.10%
	Total	43504	15346	58663	2.3%	1343	89.80%

Table 7: Interrater reliability using data from 2 native Seville phoneticians and linguists

5.2.4 Lectal Focusing in Interaction

Following data coding, the results were subjected to the tracking methodology for LFI described by Sharma and Rampton (2015), in which the metric reflects proportional change in style over time. While these authors focus only on direct interactions with interlocutors, political speech differs from regular speech in that scripted delivery of discourse is a much more frequent means of interacting with constituents, and still showed considerable variation in certain contexts. As such, to explore variation across the three speech contexts included in the corpus, these differences were balanced across the overall analysis.

When calculating lectal variation, Sharma and Rampton (2015:12) emphasize the importance of selecting an appropriate unit of measurement, from tokens, to utterances, to topics.

Their method, in particular, requires a balance that provides sufficient tokens to prevent skew resulting from low token numbers, while at the same time having sufficient blocks to provide more nuanced insight into moment-by-moment shifting patterns. In the case of the current analysis, a unit of ten seconds was selected, providing a robust sample across the eight phenomena while still providing insight into momentary differences.

Variation was collapsed into the categories of regional and normative productions and mapped out over time. As Sharma and Rampton (2015: 12) suggest, variation by AS phenomenon is also represented in the analysis, providing what they term "a more data-driven approach, acknowledging that each variable may have a distinct indexical field." While longer than the turn-constructional units Sharma and Rampton employ, twenty-second blocks were used to ensure that sufficient instances of individual phenomena were represented.

The final metric weights all tokens equally in each 10-second block, such that contexts that occur more frequently (e.g., syllable-final /s/) have a greater weight than those that may only occur a few times across a speech context (e.g., affricate production). Considering a usage-based framework of language (e.g., Bybee, 2007), as well as a sociolinguist conception of speech markers (e.g., Labov, 1972), it is expected that tokens occurring more frequently have a greater contribution to perception than those that occur more rarely. The end result being, in the case of 10 cases of 100% regional production of syllable-final /s/ and 10 cases of 50% regional production of the lateral /l/ in a single block, that the overall "regional" production for that time would be 75% (i.e., 15/20 regional productions). This approach avoids the possibility of skewing feature weightings, which Sharma and Rampton strongly suggest avoiding as a best practice in LFI consideration.

In this way, the LFI analysis offered an in-depth analysis of peninsular political speech patterns, allowing for further qualitative analysis of moments of high regional production (as Sharma [2018] shows in depth for Fareed Zakaria).

5.3 Results

Across the ten audio files considered in the current chapter, 76 minutes of interlocutor-free speech were analyzed for the eight total phenomena, yielding 15346 tokens. On average, this represents eight minutes of speech per politician, although in the case of shorter audio files, speech was coded in its entirety. Tokens of *ceceo* were excluded, showing fewer than one regional production per 400 cases, suggesting a minimal degree of use in political speech. Additionally, in order to account for moments when interlocutors spoke, politicians paused, or simply few contexts for variation were produced, time blocks with fewer than three tokens were excluded from analysis. As Sharma and Rampton (2015) reference, it is important to avoid fluctuation caused by low token counts, to get an accurate representation of regional variation over time.

All told, 11149 tokens remained in the LFI analysis. This represents slightly more than 1100 tokens per speaker, and roughly 160 tokens of each phenomenon per speaker (although as can be seen in Table 7 in the previous section, affricates, trills, and *ceceo* were much less represented than taps, laterals, and coda /s/).

5.3.1 Susana Díaz (07FSSD)

The speaker with the lowest degree of variance was the former President of Andalusia (2013-2018), Susana Díaz. The identifying code (i.e., 07FSSD) provides information about the audio file coded: Speaker 7 is a 41 year old Female Seville Socialist, in an audio file with scripteD speech. As further background, Díaz was elected to presidency in 2015 (taking on the role previously following the resignation of the former president), and lost in the 2019 elections to a member of the PP. A total of 1796 tokens were collected for this speaker from her inauguration speech, which lasted 23 minutes. From that audio, 12.2 minutes were coded, representing 53% of the overall file.

For each comparison of Lectal Focusing in Interaction across the speakers moving forward, a composite view of variation will first be presented, as in Figure 55 for Díaz. Peaks were selected based on their relative deviation from the mean, with those reaching over 50% regional production and/or being near valleys indicating high normative production being favored. These moments are discussed in greater depth to determine topical shifts in the audio when regionalisms reach their summit for a speaker. For Díaz, the speech she presents is an inaugural address, following her election in 2015. The very first moment sees the highest rate of regionalisms, as she addresses those who came out for her speech:

(1) *Buenos días. Muchísimas gracias a cuantos se han desplazados para acompañarme en un día tan importante.*

Good morning. Many thanks to all of those who have travelled to join me on such an important day.

After thanking individuals in her government and her *militantes*, that is, the voters who supported her, she turns toward the future, discussing the contributions her government will make to Andalusia. Then, leading up to the peak at 130 seconds, she turns her discussion to the suffering of working-class mothers and pensioners trying to provide for their children and grandchildren:

(2) *Mi cabeza está donde estan las madres preocupadas, porque no pueden garantizar dar las tres comidas al día a sus hijos.*

My mind is with the worried mothers who can't give their children three meals a day.

She describes various contexts of suffering and worry in the region, and the hope that her party represents. She then turns to a discussion of voter wishes and desires, including sensibility and honesty, and her desire, at 390 seconds, to be the president of all Andalusians, including:

(3) *De aquellos Andaluces que votaron a otras fuerzas políticas. Y de muchos que, por distintos motivos, ni siquiera acudieron a votar en la jornada del veintidos de marzo.*

Of those Andalusians that voted for other political parties, and of those many who, for various reasons, didn't vote in the March 22 elections.

Her rhetoric then turns to a discussion of a need for a united front to address these concerns in Andalusia, and the clear consensus among voters in electing her that they have a concern for the general welfare. The concern of politics should be to address the concerns of citizens, she begins, nearing the 530 second mark:

(4) *Para ninguna otra cosa estamos aquí, y nunca debemos olvidarlo. Quiero que sepan los andaluces y las andaluzas que así quiero actuar siempre.*

We are here for no other reasons, and we should never forget it. I want Andalusians to know that I always want to act in this way.

The final highlighted peak, at 560 seconds, furthers this same point. After saying that her only legacy is her word, she continues by saying that:

(5) *Solo tengo un patrimonio, que es mi palabra y la fuerza de los valores que represento y con los que me identifico. Para algunos será un escaso patrimonio, pero para mi es un patrimonio inmenso y grande.*

I only have one legacy, and that is my word, and the force of the values that I represent and which which I identify. For some, that is a small legacy, but for me, it is massive.

In each of these cases, the discourse turns less charged, describing lists of ideas and enlarging on those concepts she has just introduced. At 150 seconds, a peak is followed by a slight slip of the tongue[20], which causes Díaz's speech rate to slow and her regional productions to drop off abruptly. In terms of Sharma's 'real-me,' however, every moment serves as an instance for Díaz to lower the veil of political speech and embrace sincerity, from thanking those who attended, to emphasizing the struggling individuals she hopes her presidency will help, to emphasizing her desire to be accepted by all Andalusians, to finally emphasizing the importance of respecting the political process and carrying through on her word. In each of these moments, the emphasis of her

[20] *Y que ven amenazada no solo perder su vivienda sino vend- per- perder un proyecto de vida suyo…* 'And who are threatened not only with the loss of their home, but also the selling- lose- losing of their life's work.'

Andalusian identity serves as a way to reinforce her point and, for a moment, to set aside the political script of the prestigious, televised inauguration.

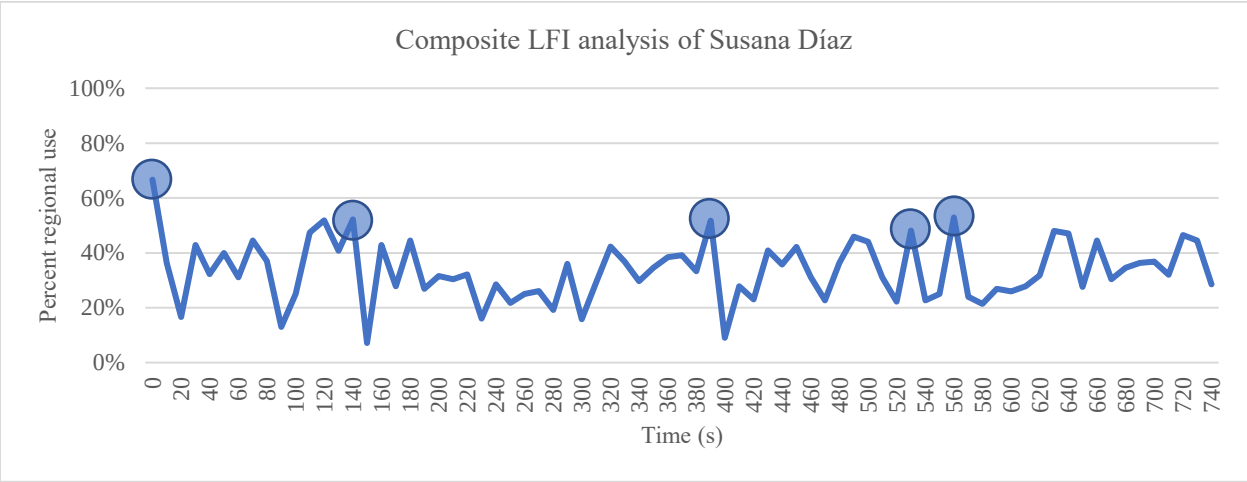

Figure 55: Composite LFI analysis of Susana Díaz in ten-second intervals

Following the consideration of composite variation and topical shifts, the LFI analysis for each speaker takes a look at the differences across regional phenomena in the moments when regional productions are at their peak, as represented by the line graph in Figure 56 for Díaz. This graph shows each of the phenomena in question as they rise and fall over the course of the interview, represented by the percentage of regional production per "block" of time. Highlighted in blue rectangles are the moments when local peaks occur in the combined view, allowing us to see which phenomena are most influential in these "real me" moments.

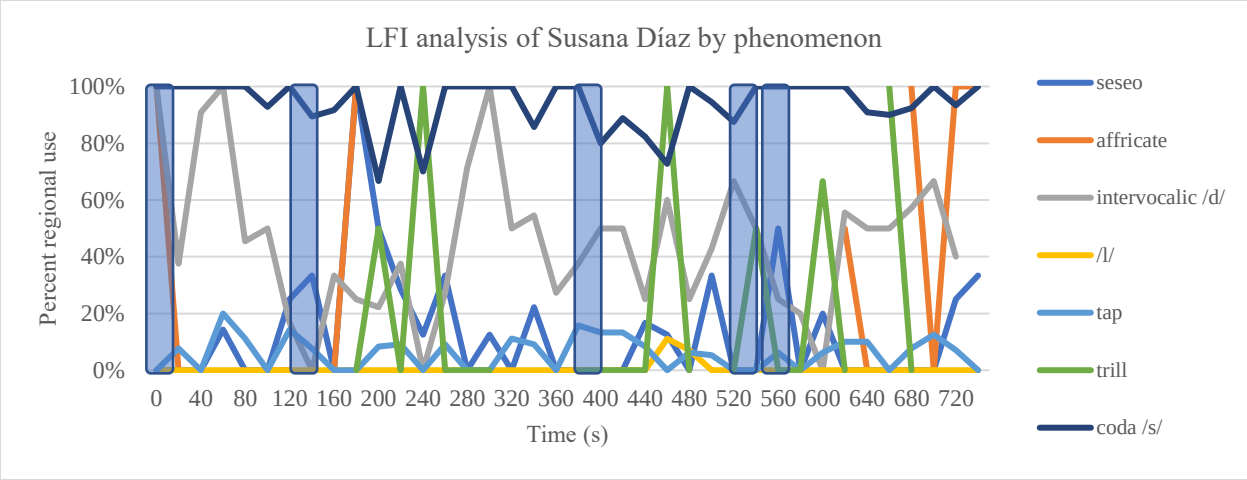

Figure 56: LFI analysis by phenomenon for Susana Díaz in twenty-second intervals

The start of the interview includes intervocalic /d/ elision, affricate fronting, and coda /s/ reduction at their peaks, with both the first and the last being typical of Andalusian political speech (Cruz-Ortiz, 2019) and AS. Then, when discussing working mothers at 130 seconds, *seseo* and tap reduction are both near a peak, while intervocalic /d/ elision and coda /s/ reduction fall. Next, at 390 seconds, when expressing a desire to be everyone' president, it is tap reduction and intervocalic /d/ elision that contribute to regional perception, along with coda /s/ reduction. Finally, in the last two peaks at 530 and 560, when describing the importance of serving the people and following through on her word, there is a spike in reduced trill use, *seseo*, high coda /s/ reduction, and falling intervocalic /d/ elision.

Several regional features work in concert to emphasize regional speech, especially *seseo*, while coda /s/ elision is so common that it has little effect on norms. Already, this first examination of LFI shows possible differences in relative effect on social meaning based on frequency of use and regularity of regional productions. While intervocalic /d/ and *seseo* are common enough in Díaz's speech to affect regional association, liquid phenomena and the affricate have low regional

rates and low frequency, while coda /s/ is so often reduced as to not be able to convey much meaning. Moving through the other nine speakers, it will be important to note if these same patterns and trends differ by social factor, or if WAS and EAS have specific differences in these patterns.

5.3.2 Elena Valenciano (04FMSM)

The next speaker, Elena Valenciano, was a longtime member of the European (1999-2008, 2014-2019) and Spanish Parliament (2008-2014). Speaker 4 is a 53 year old Female Madrid Socialist, and this audio represented an interview with a Male interlocutor. She has been a feminist activist and member of EU committees on human rights. A total of 1248 tokens were collected for this speaker from an interview with a male politician on a morning talk show (i.e., Las Mañanas Cuatro) that had 7.5 minutes of politician speech. The entire interview was coded (i.e., 100% of the file).

In her interview, Valenciano has a much lower composite than Díaz (Figure 57) – unsurprising, given her NCPS background. While coda /s/ reduction is much less frequent for these speakers, those other variables classified as “regional” that also experience some working-class and rural production in NCPS continue to serve a stylistic function for politicians with access to these norms. This interview with Elena Valencian focuses on national finance, including social security and the ongoing financial crisis in Spain (stemming in part from the 2008 recession). She discusses the move of middle-class earners into the working-class, the state of the national debt, and methods of securing credit for small businesses. The interviewer presses her to commit to a

bill addressing the debt, and after repeatedly arguing that she would prefer to find another way, she says dismissively at 210 seconds:

(6) *Con la ley esta de seguridad ciudadana, gracias a eso, los van a risar la cabeza y encima vais a tener que aplaudir con las orejas.*

Thanks to this law for citizen security, they're going to laugh their heads off, and on top of that, you are going to have to applaud with your ears.

The topic continues to focus on the citizen security law, which Valenciano describes as stemming from a university protest preventing former Deputy Prime Minister Alfredo Rubalcaba from speaking, which ended in police intervention. Valenciano argues that the new law is going further than necessary, but at 260 seconds, she says the following about Ana Botella, the mayor of Madrid at the time:

(7) *Es que creo que a Botella hace esto porque efectivamente se está jugando en este momento la candidature al Ayuntamiento de Madrid en el seno del PP, ¿no?*

I think Botella is doing this because the candidacy for the Madrid City Council is being played out at the moment inside the PP, right?

After a lengthy question from the interviewer, asking about upcoming budget cuts and their effects on the economy, Valenciano becomes heated, arguing that education is meant to be open to all, not just those who can pay, and that:

(8) *Los recortes en la educación producen peores resultados de educación. Si hay menos profesores y más alumnos, si hay menos becas… [será peor].*

Budget cuts produce worse educational outcomes. If there are fewer professors, more students, fewer scholarships… [it will be worse for education].

In each of these cases, there is a moment where Valenciano uses lectal focusing as a means of doing what Hernández-Campoy and Cutillas-Espinosa (2013) discuss, employing rural and working-class variants as a way of pushing forward a socialist ideology and identity. In the first case, combating what she sees as a governmental overreach, she makes the type of sarcastic humor with lectal focusing that Sharma (2018) describes for Zakaria, an instance of divergence as solidarity, mocking the legislation proposed by her opponents. Then, in the second case, she takes the interviewer into her confidence, speculating as to the reason for the citizen security law, and its process within the opposing party. Finally, for the last peak, in a discussion of equity and universal access to education, her anger at the idea that budget cuts might benefit education combines with her rural presentation as a means of showing her true, strong feelings on the subject.

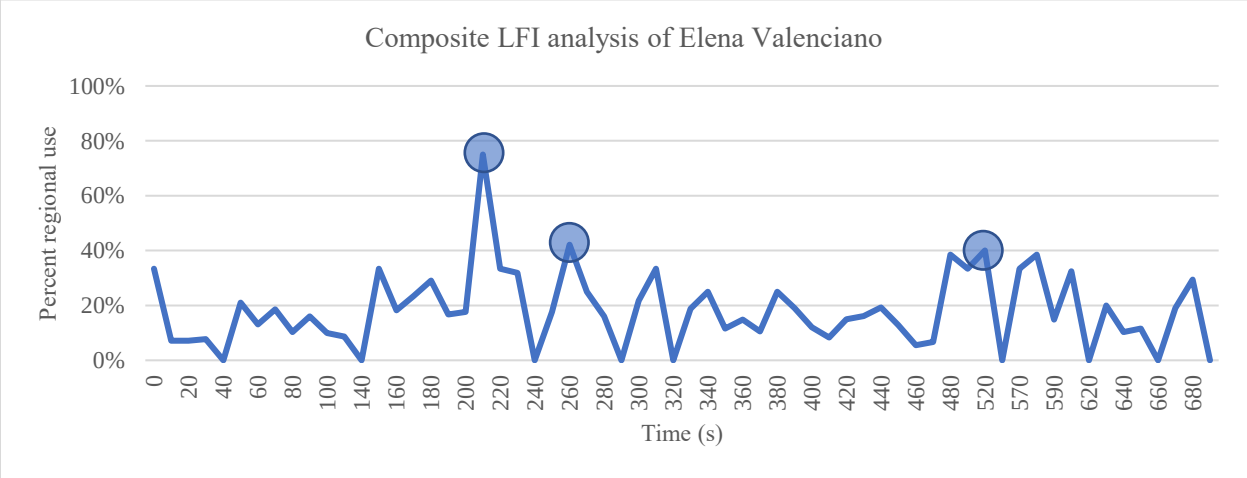

Figure 57: Composite LFI analysis of Elena Valenciano in ten-second intervals

Next, we examine these three moments in terms of the presence of specific phenomena (Figure 58). When joking about the severity of the citizen security law (210 seconds), intervocalic /d/ is at its peak, and tap and lateral reduction are both raised. Next, when describing the rationale behind Botella's actions (260 seconds), coda /s/ elision, intervocalic /d/ elision and all three liquid phenomena play a role in variation. Finally, when addressing cuts in the education budget (520 seconds), these same phenomena come to the fore. While coda /s/ elision is infrequent, that makes its appearance all the more salient for Madrid speakers, while intervocalic /d/ is frequently elided.

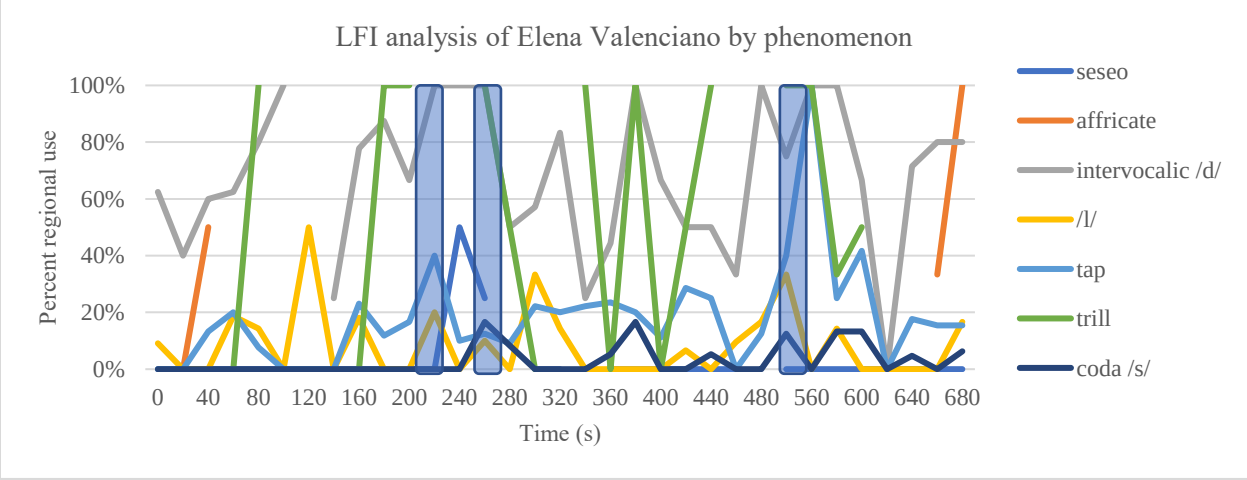

Figure 58: LFI analysis by phenomenon for Elena Valenciano in twenty-second intervals

Unsurprisingly, there are some marked differences between the AS and NCPS norms in these first two speakers' lectal focusing behavior. Intervocalic /d/ elision is used to stylistic effect by both speakers, suggesting it may be something of a performance marker (following Cruz Ortiz, 2019, and others' suggestions that it occurs frequently in political speech), while *seseo* is less

common in the north and coda /s/ is too frequent to vary meaningfully in the north. Díaz, as a female Seville politician, uses little liquid variation, but this phenomenon is more accessible to Valenciano in Madrid.

5.3.3 José Antonio Pérez Tapias (05MSSF)

Next is José Antonio Pérez Tapias, a university professor in Philosophy at the University of Granada, who served in the lower House from 2006 to 2011. Speaker 5 is a 59 year old Male Seville Socialist, and this audio represented an interview with a Female interlocutor. Despite stepping away from politics in 2011, he returned in 2014, vying for the position of general secretary and leader of the national PSOE. Although he failed to win, he published a novel critical of social democracy in 2017, then left the PSOE and founded his own party, "Socialism and Republic," in 2018. The morning interview lasted 6 minutes; 100% of the audio was coded, yielding 606 tokens.

In this interview, Pérez Tapias discusses the PSOE leadership elections in 2014, presenting rhetoric that foreshadows his later break from the party, arguing that the leadership was unable to make good on the integration of diverse voices as it had promised, and that campaign promises were bring broken. The average regional composite is much closer to 50% than we have seen for the two female speakers, as shown in Figure 59. After initial complaints about the direction of the party in the wake of the direction, the interviewer asks if Pérez Tapias had spoken with Pedro Sánchez, the winner of the PSOE elections (and now-Prime Minister of Spain) after the elections. Leading up to the 60 second mark, he confides that:

(9) *Bueno, tuvimos una primera reunion hace dos semanas o cosa así, una toma de contacto para en fin declarar nuestras respectivas inteciones de seguir hablando y concretar esa propuesta, pero luego ya no se produjo ninguna reunión más.*

Well, we had a first meeting two weeks ago or so, a moment of contact to finally declare our respective intentions to continue talking and formalize the proposal, but then no more meetings took place.

The drop in regional use from 60 to 70 seconds follows the narrative structure. Communicating the existence of this meeting, which is publicly expressed for the first time in this interview, provides Pérez Tapias the opportunity to emphasize the goodwill with which he treated the chance to meet with PSOE leadership. The lack of results from the meeting parallels the return to more normative variants, along with the more formal expression regarding the lack of a further meeting (rather than explicitly referencing or commenting on Sánchez's lack of continued communication). The conversation then continues, with the interviewer asking if Pérez Tapias feels that Sánchez tricked them with rhetoric of "integration," which Pérez Tapias takes a bit of pushing to directly admit, and an interviewer comments that "it's clear that you're a philosopher" based on his language. Pérez Tapias goes on to argue that representation in congressional committees and selections by Sánchez's government doesn't reflect the proportions laid out in the recent elections. Discussion then turns to Sánchez's governments stance against working with right-leaning parties like the PP, asking if this is part of the integration that Pérez Tapias wants, but he rejects that at 290 seconds, saying:

(10) *Estoy de acuerdo con esa decision, en ese sentido, vamos a ver, hay que distinguir entre hacer pactos y hacer una gran coalición, que es otro tipo de pacto.*

I actually agree with that decision, in the sense that, let's see, you have to distinguish between making pacts and making a grand coalition, which is another type of pact.

This claim seems like a reversal of his previous wish for integration, forcing Pérez Tapias to explain the difference between making individual agreements with other political parties (through the "integration" discussed previously), and creating a coalition government that is forced to rely on parties with different political goals that may prevent legislation from passing. Finally, as the interview begins to wrap up, the interviewer references the suspension of elections that has been rumored for the following year. Pérez Tapias notes leading up to 580 seconds that they seem to have been postponed indefinitely, but that a federal committee is talking about it, that the need for the democratic process is vital, and that:

(11) *Por lo tanto esa convocatoria y la participación ciudadana deberia cumplirse.*

Therefore this call and citizen participation should be fulfilled.

In this moment, the politician's desire for public elections (in spite of the current decisions of the party) places Pérez Tapias at odds with the party leadership and the General Secretary of the party, who makes the final decision about primary election timing. This peak in regional production comes at a moment when he is arguing the case for representation in government, an issue that pairs with his broader complaints about social democracy.

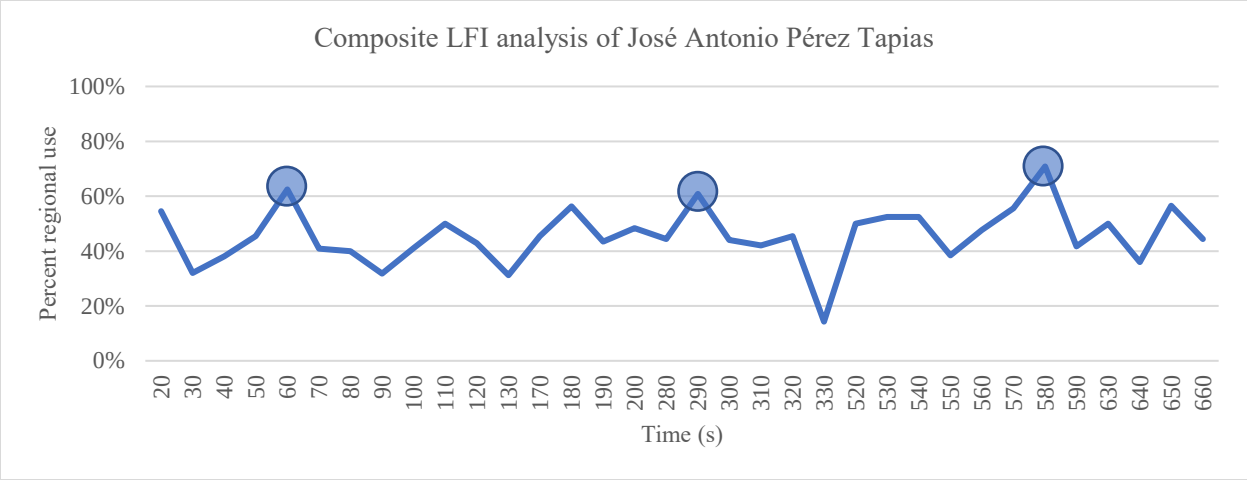

Figure 59: Composite LFI analysis of José Antonio Pérez Tapias in ten-second intervals

Looking at regional phenomena, *seseo*, intervocalic /d/ elision, coda /s/ elision, and tap phenomena are the most frequent to emerge at regional peaks (Figure 60). First, at the 60 second point when discussing his most recent contact with Pedro Sánchez, local peaks occur for regional production of the two aforementioned sibilant phenomena, as well as intervocalic /d/ and the tap. Next, at 290 seconds, falling rates of *seseo*, coda /s/ elision, and lateral reduction pair with tap and intervocalic /d/ peaks contribute to the rise in vernacular variant use when distinguishing between integrating voter desires into government and developing a grand coalition. Finally, at 580 seconds, as he discusses the need for a date for primary elections, the tap, lateral, intervocalic /d/, coda /s/, and *seseo* all experience high rates of regional production.

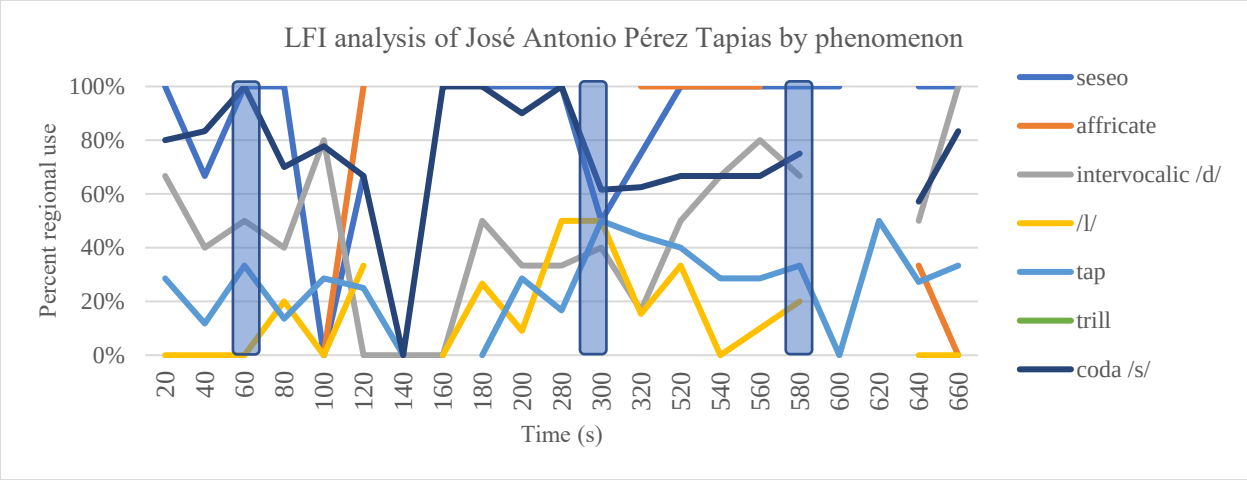

Figure 60: LFI analysis by phenomenon for José Antonio Pérez Tapias in twenty-second intervals

Coda /s/ elision is more frequent than for Valenciano from NCPS, but less categorical than it appears for Díaz. Meanwhile, tap lateral reduction and variable production is more common than seen for the female speakers. Few cases of trills occur in this audio sample, and the fronted affricate appears variably but infrequently. Pérez Tapias also has a relatively low rate of intervocalic /d/ elision compared to the former two speakers, with several drops below 20% regional production. Although his use of *seseo* and coda /s/ reduction is more reminiscent of Díaz's, it differs markedly from the first two female politicians.

5.3.4 Pepe Bernal (26MLSD)

The fourth speaker analyzed is José (Pepe) Bernal Gutiérrez, a university professor of History and Geography, who served as general secretary of the PSOE in Marbella, Málaga, in 2008, and took

on the role of mayor of Marbella from 2015 to 2017, until he stepped down after a vote of no confidence. Speaker 26 is a 43 year old Male Málaga Socialist, and this audio represents a scripteD speech. The speech lasted 14 minutes, of which, 7.3 minutes (i.e., 52.5%) were coded that included 1143 tokens for consideration.

In this interview, following the vote of no confidence in his leadership as mayor, Pepe Bernal reviews the contributions that his coalition government made in Marbella over the previous two years, including economic recovery, and his pride in the steps that his government was able to make (Figure 61). There are four clear peaks in the speech, where regional features reach or slightly top 60%. First, he discusses the "government of change" he assembled from four different factions, and at the 70 second mark, emphasizes the need for diverse opinions and mutual respect:

(12) *Porque eso fue lo que expresaron los ciudadanos en las urnas y fuimos capaces de hacerlo y lo conseguimos. Dijimos y hemos demostrado que hay otra forma de gobernar.*
Because that was what citizens asked for at the polls, and we were able to do it, and we achieved it. We said and demonstrated that there is another way to govern.

In this moment, he defends the legacy of his government, arguing that they carried out the goals set out for them by voters, crafting a coalition that demonstrated a new way to lead the city. He continues, arguing that his government focused on the general interest over the individual, saying at 180 second that:

(13) *No todos los anteriores pueden decir lo mismo. Y los datos avalan lo que os digo.*
Not all previous ones can say the same. And the data supports what I'm saying.

Again, Bernal shifts from describing the needs of the town, or the general outcomes of governance by his coalition, to a face threat toward previous governments in Marbella, and an implicit insult toward those who chose to end his tenure as mayor prematurely. The government was working so well: for what reasons would someone choose to end this, if not as a means to return to the way things were? Following that, he begins to describe in detail the projects that his government accomplished throughout the city, as well as those that are incomplete, saying at 300 seconds that:

(14) *Y además vamos a presentar a poner sobre la mesa muchos proyectos para ejecutar en los próximos meses. Algunos están los presupuestos y otros están en esas propuestas del remanente de más de diez millones de euros para los próximos meses.*

We're also going to advance many projects to carry out in the coming months. Some are in the budgets, while others are in proposals for the remainder of more than ten million euros for the coming months.

There is a curious shift here, moving from mainly normative variants as he describes completed projects at breakneck pace, to the sudden regional peak as he talks about the future. He conveys the overall message that he loves his city and will support the next government, while abstaining from pointing fingers at those involved in his ouster. This contributes to his image as a citizen of Marbella aiming to cause real change and development in his town, especially given the central placement of public works and development in this speech. Finally, moving toward the end of the segment, Bernal discusses the recent integration of workers and rural citizens in the city council,

a frequently unfulfilled promise in the area, including reinstating 35-hour work weeks, and (430 sec) putting up job boards:

(15) *...con procesos selectivos para que los ciudadanos puedan tener un empleo en el Ayuntamiento sin rodeo, por la puerta principal y en igualdad de condiciones y de oportunidades.*

With selective processes so that citizens can find work in the city council without detour, coming through the main door, on equal terms and with the same opportunities.

The regional peak here, amidst a list of other contributions, both suggests the importance of this achievement to Bernal, while also potentially referencing the same working-class values emphasized by Hernández-Campoy and Cutillas-Espinosa (2013) for María Martínez. In part, Bernal is directing these comments to the rural and working-class individuals his government supported.

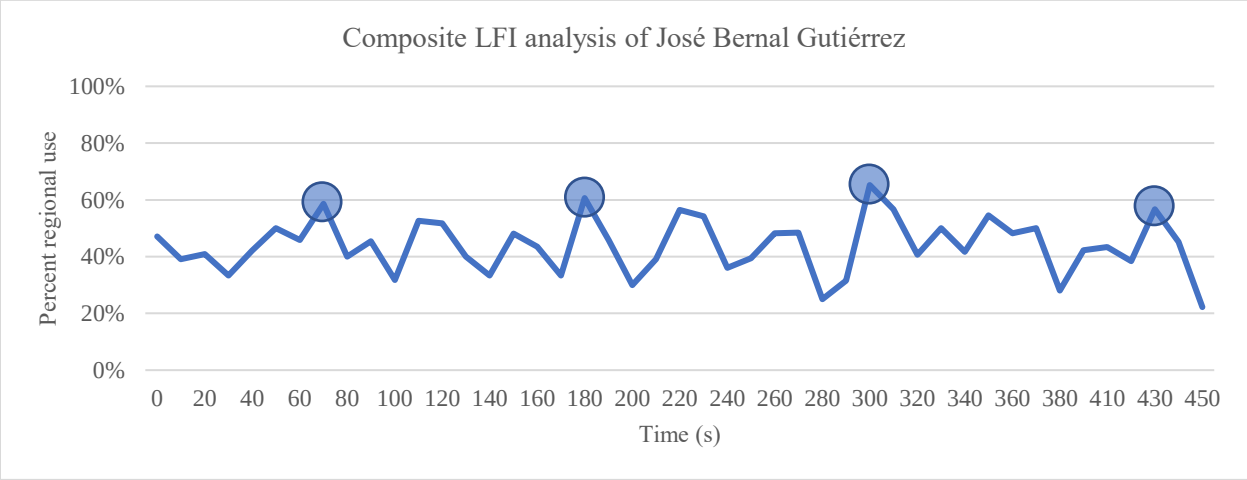

Figure 61: Composite LFI analysis of José Bernal in ten-second intervals

Looking at these four moments based on specific phenomena, we see a couple of differences here compared to the NCPS and WAS speakers examined so far (Figure 62). Regional tap, trill, and lateral phenomena are frequent (in fact, this speaker has among the highest rates of specifically variable production – while other speakers like Pérez Tapias elide liquids, Bernal often interchanges them in this speech). Intervocalic /d/ elision and *seseo* are infrequent compared to others. Additionally, what affricates occur in this segment are produced as normative alveolar allophones. For the peaks specifically, it is the combination of coda /s/, variable liquid production, and intervocalic /d/ that are the most important regional features, providing a very different profile from speakers in the east.

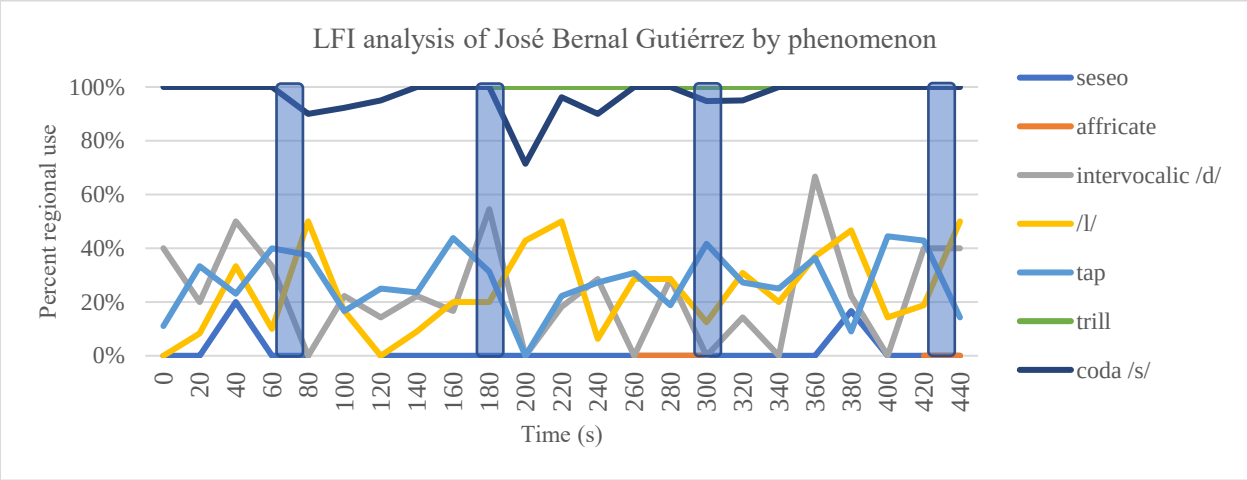

Figure 62: LFI analysis by phenomenon for José Bernal in twenty-second intervals

In comparison to NCPS and WAS, this EAS speaker uses regional liquid productions at a much higher rate, alongside moderate to low intervocalic /d/ elision and coda /s/ elision. The low rates of *seseo* found in the east by previous studies are borne out in Bernal's speech, and tied to

findings from Chapter 4, the affricate seems to have a combined gendered and regional social value that leads fronting to be infrequent in this case. When LFI occurs, Bernal uses it as a way to emphasize the claims he's making about the success of his government in a context where his credibility is considerably reduced (i.e., having received a motion of no confidence in his leadership), showing the most important ideology, public developments, and opportunities made. This emphasizes, as well, a difference between interviews and public speeches (i.e., the loss of the direct "interaction" reference by LFI). Rather than decreasing social distance with a single individual, public speeches like this have a massive audience, contributing to a levelling effect on regional variation (save in specific moments), which is less impactful in interviews.

5.3.5 Margarita del Cid (32FLCF)

The next speaker is Margarita del Cid, who served as a councilwoman and spokesperson for the PP in Málaga since before 2016, as well as serving as a vice-president (2019-2021) in the city and most recently as mayor of Torremolinos, Málaga, since 2021. Speaker 32 is a 43 year old Female Málaga Conservative, and this audio presents an interview with a female interlocutor on a local television station. The conversation lasted 6 minutes, allowing the entire audio to be coded, resulting in 684 tokens for LFI consideration.

The interview discusses the ongoing Fair of St. Michael, including its organization, Del Cid's favorite parts, and reminiscing on previous fairs in past years (Figure 63). Leading up to the first of four regional variant spikes approaching 60%, at the 40 second mark, the interviewer thanks

the politician for being available for the interview, given the busyness of the week, to which she responds:

(16) *Bueno, pues, muy buenas noches y en primer lugar, agradecerte la invitación.*

Well, good evening, and in the first place, thank you for the invitation.

This is one of the few cases where a speaker's first introductory words show high rates of regional features – typically, Labovian-focused sociolinguistic literature would expect the earliest speech to also be the most careful and formal. This quickly becomes the case after this introduction, with the regional rate dropping as she speaks about those who have contributed to the fair. The interviewer then asks about her clothing choices for the fair, specifically if she had worn a fair costume yet. Approaching the next peak at 150 seconds, Del Cid responds that she had selected an older skirt that she had worn previously, and:

(17) *Había hecho algunos retoquitos y eso ha servido pero bueno eso es una anécdota. Yo creo que es importante ir, participar.*

I had done some tweaks and that has helped, but in any case, that was an anecdote. I think it is important to go, participate.

In this moment, the politician shows an increased level of familiarity with the interviewer, going into detail about her clothing choices. She then seems to realize she may have gone too far into detail and reorients the conversation with a self-aware discourse move (i.e., "but, in any case, that was an anecdote") before returning to the main message of attending the fair. She moves on,

arguing that it is important to support the community by attending, not taking on a major role but simply being a part of the experience (210 sec), saying that:

(18) *Yo intento hacerlo así. Mis concejales que han estado también viviendo, pues, todos estos momentos lo han hecho así.*

I try to do that. My councilors, who have also been experiencing all these moments, have done this as well.

In this moment, Del Cid connects her suggested approach to the festival with the behavior of herself and her councilors in the town hall, emphasizing the supporting role of herself and her peers in the organization of the ongoing festivities. Next, approaching the 290 second point, Del Cid is asked her favorite part of the fair – the nighttime or daytime aspects. She discusses the pilgrimage (*la romería*) and its history and meaning, then concludes with an optimistic summation, saying that whether listeners prefer the nighttime or daytime fair, the important thing:

(19) *Es que todo tenga su momento. Poder disfrutarlo con las personas que te rodean y vivirlo bueno pues de una forma intensa pero también sana y respetando nuestas tradiciones.*

Is that everything has its moment. Being able to enjoy it with the people around you and experience it, well, both intensely and in a healthy way, respecting our traditions.

This message seems especially directed toward younger viewers and tourists in the town. She encourages safe drinking and celebrating as the fair gets underway, emphasizing the importance – above pointing out any single "best" part of the fair – of enjoying the experience as part of a

Torremolinos tradition, rather than an excuse for dangerous behavior. As the interview moves toward its conclusion, Del Cid speaks up the festival, referencing the pleasant weather, the nice time of year at the start of fall. She then makes a call to the people of Málaga and surrounding areas to visit the fair, at the 370 second mark, saying that in the meantime:

(20) *Hay que seguir trabajando hay que trabajar para que tengamos la major feria que se merecen nuestros vecinos y también las personas que nos visitan.*

We have to keep working, we have to work so that we can have the best fair that our neighbors and visitors deserve.

This interview, taking place in the midst of the set-up for the festival, comes across in part as a means of advertising the event. Del Cid makes comments appreciative of those working to set up the fair, while also emphasizing the fun of the event and the need for safety in celebrating. The 'real me' is distant in these instances of regional production. Rather than increased sincerity, especially with the latter two peaks, we see a sort of regional persona put on in the same way Schilling-Estes (1998) describes in Ocracoke. Rather than being a way of reducing social distance with the interviewer, it comes across as folksy, indexing appreciativeness and good values. The context of the interview likely contributes to that – rather than arguing a political point, the goal is to hear about the fair and to draw in tourists and locals based on a discussion of the event's history.

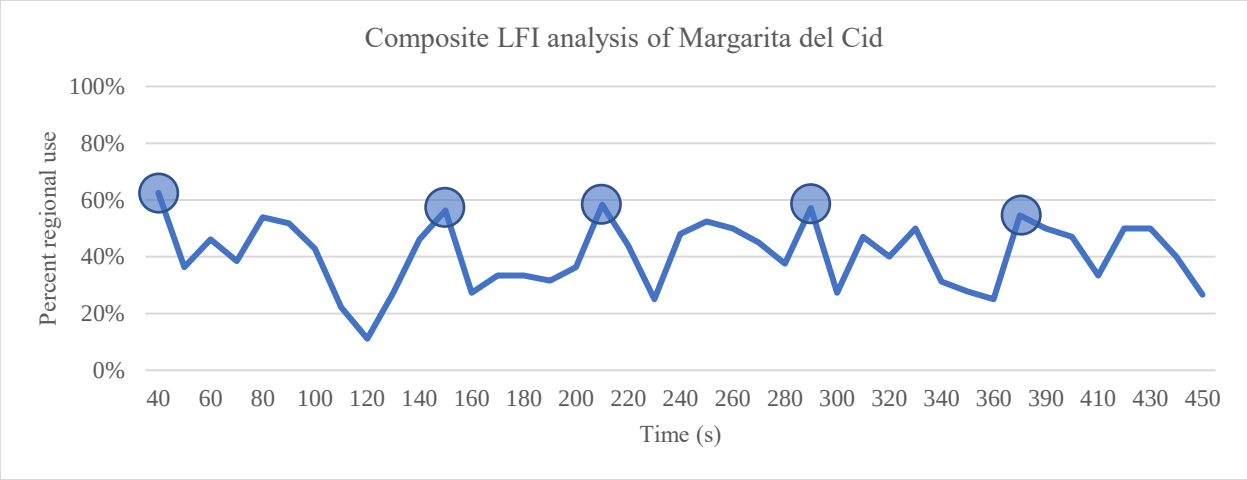

Figure 63: Composite LFI analysis of Margarita del Cid in ten-second intervals

With respect to regional phenomena, we can see that coda /s/ is high overall, while tap elision occurs during the first and final peak, remaining low in the other three moments (Figure 64). Regional intervocalic /d/ and affricate fronting are also frequent in these moments, while reduced laterals, trills, and *seseo* are infrequent. Interestingly, this seems to share patterns with both the Seville and Madrid speakers, having high rates of intervocalic /d/ and medial rates for the tap as in NCPS, but also high coda /s/ elision and occasional affricate fronting, more like the WAS speakers. There is relatively little similarity between Del Cid and the other EAS speaker (Bernal), except for the rates of tap elision, showing a marked difference between a speaker from a more urban space (Del Cid, associated with Málaga) and one from a smaller town (Bernal, from Marbella).

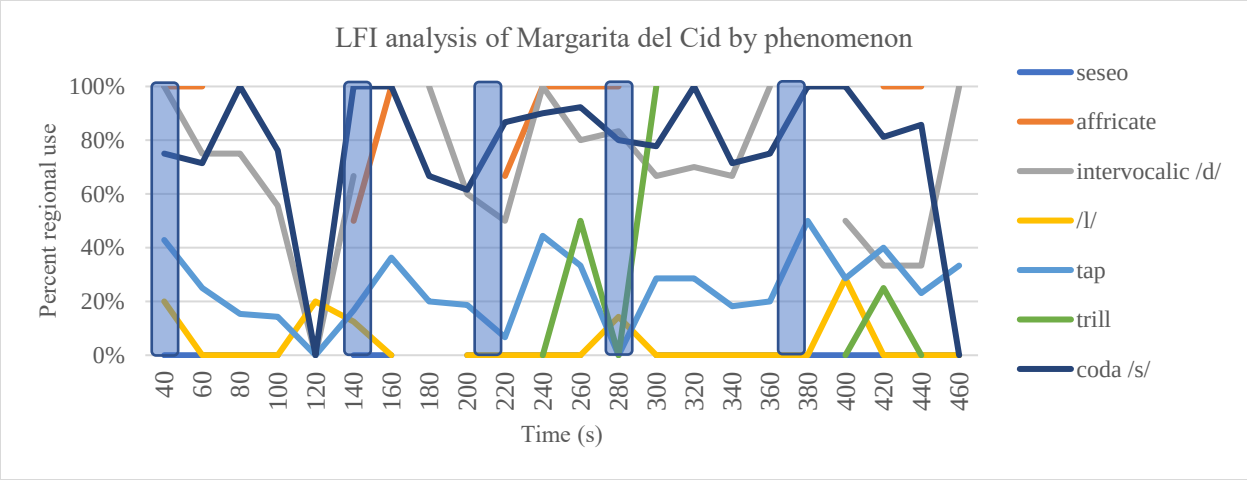

Figure 64: LFI analysis by phenomenon for Margarita del Cid in twenty-second intervals

5.3.6 Carmen Calvo (11FCSM)

The sixth speaker is María del Carmen Calvo Poyato, who has a long history of political service, including as parliamentarian (2000-2004), congresswoman (2004-2011), minister (2018-2021), and Deputy Prime Minister (2018-2021), while also working as a Professor of Constitutional Law at the University of Córdoba. Speaker 11 is a 61 year old Female Córdoba Socialist, while this audio involves an interview with a Male interlocutor on a major news network. The conversation lasted 10.2 minutes, of which 7.3 minutes were analyzed (i.e., 71.6% of the interview), yielding 1037 tokens.

This interview focuses on a recent humanitarian crisis, in which over 600 refugees travelled across the Mediterranean and were taken in by Spain after the Italian Interior Minister refused to accept them (Figure 65). The first speaker of a Northern Andalusian Spanish (NAS) variety, Calvo has a considerably lower overall rate of regional production than Seville and Málaga speakers,

with only two peaks reaching 60% regional variant production. She discusses the necessity to not just provide charity to those whose homes have been torn apart by civil war (220 sec), but also:

(21) *Ayudando al despegue de esos países donde la gente puede estar en su país. Esto es muy importante. Y luego no perdamos de vista que España ha cumplido casi literalmente la cuota de refugiado que nos correspondía*

Helping the ascent of those countries where people can be in their country. This is very important. And let us not lose sight of the fact that Spain has almost literally complied with our refugee quota.

This comment serves both to indicate a need for the EU to support nations before refugee status becomes a necessity, as well as introducing a slightly awkward topic – namely the fact that Spain had taken more refugees than many of its peers, and was forced to take in those aboard the boat in question when Italy refused to send aid. Both of these topics require the sincerity of regional variants, suggesting a policy need on the part of the EU and gingerly disapproving of other EU nations' refugee numbers. Next, leading into the 460 second point, the topic turns to other current events, including the state of Spain's ecology, including a comment by a minister that "Spain has returned," at which point the interviewer asks where Spain had gone. She responds:

(22) *Yo les comprendo muy bien a los periodistas porque hay que hacer su trabajo y además lo hacen al milímetro de las palabras que utilizamos. Bueno, somos un gobierno distinto. ¿Está claro, no? Un gobierno legal y legítimo.*

> I understand journalists very well, because they have to do their jobs, which they do while scrutinizing our words. Well, we are a different government. That's clear, isn't it? A legal, and legitimate government.

In this moment, Calvo seems to be presenting a bit of humor about the detail with which journalists look at political speech. However, it is particularly interesting here to see how the regional percentage falls off following this moment in the LFI analysis. The interviewer responds to the comment in (22) with the question "And the previous government wasn't?", which leads Calvo to affirm that it was as well, but that the new PSOE government is different. This suggests that her attempt to reduce social distance with the interviewer was unsuccessful, and an attempt at levity is not repeated.

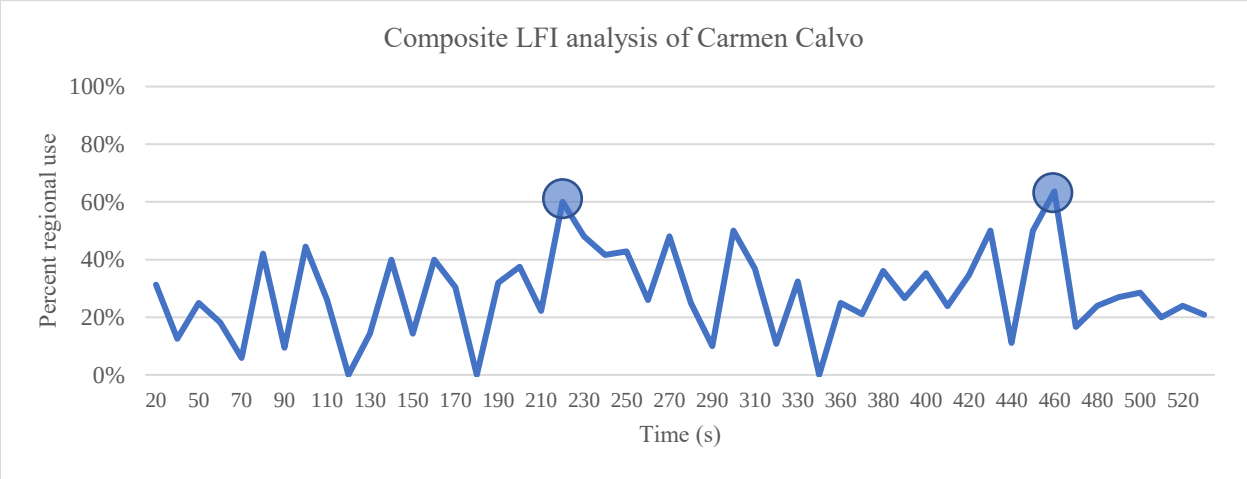

Figure 65: Composite LFI analysis of Carmen Calvo in ten-second intervals

According to the representation of regional variation by phenomenon in Figure 66, there are a number of regional features present in Calvo's speech. Interestingly, coda /s/ retention is

particularly high for an Andalusian speaker, while tap and trill reduction is comparable to EAS speech, and intervocalic /d/ elision is common. When broaching the topic of Italy's refugee responsibilities (220 seconds), coda /s/ reduction is at its highest rate in the interview, intervocalic /d/ elision is present, and liquid phenomena are at their most normative. Meanwhile, when joking about the scrutiny of journalists in the context of the recent government change in Spain, Calvo has some of her highest rates of *seseo* (for one of the few times in the speech), tap reduction, and intervocalic /d/ elision, as well as high coda /s/ reduction. This attempt at solidarity is even more notable as it is not an instance of divergence as in Zakaria's case (Sharma, 2018), but rather convergence, as her interviewer is a native speaker of the EAS variety.

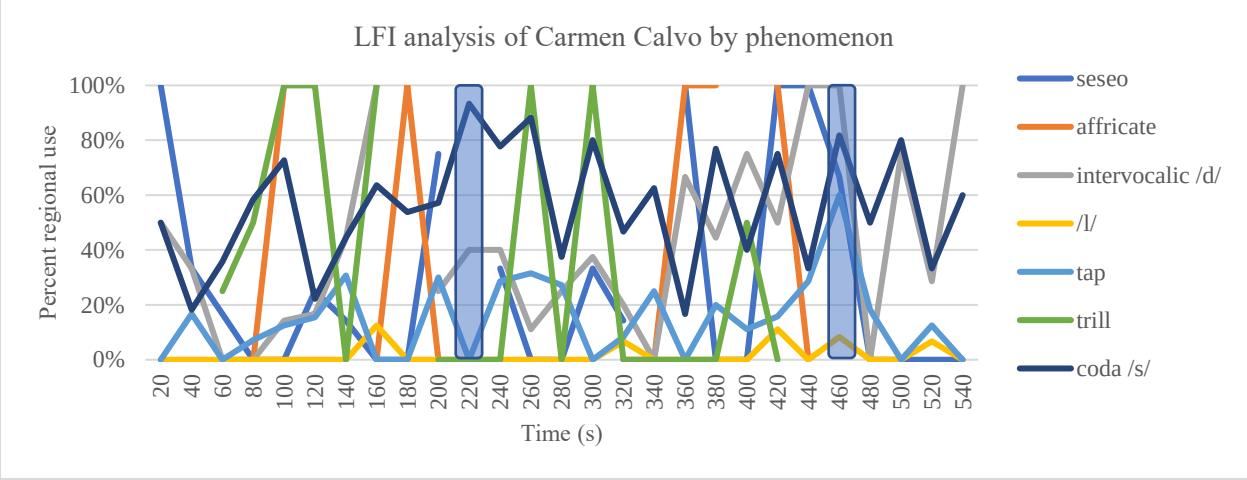

Figure 66: LFI analysis by phenomenon for Carmen Calvo in twenty-second intervals

With respect to the presence of liquid phenomena, Calvo is perhaps closest to the WAS speakers examined thus far, with regular instances of reduction. There are some marked differences between the AS and NCPS norms in these first two speakers' lectal focusing behavior. Intervocalic /d/ elision is used to stylistic effect by both speakers, while *seseo* is less common in

the north and coda /s/ is too frequent to vary meaningfully in the north. Díaz, as a female Seville politician, uses little liquid variation, which is more accessible to Valenciano in Madrid. With respect to intervocalic /d/, her rates of elision are high, comparable to those of other female socialists in Seville and Madrid. Finally, for sibilant phenomena, coda /s/ elision is curiously infrequent. Samper-Padilla (2011) describes NAS as having less reduction than more southern parts of Andalusia, which may help to explain this variable behavior. There are also clear instances of affricate variation, although there are fewer possible contexts where variable use can occur.

5.3.7 Ángel Garrido (14MMCD)

Following that is Ángel Garrido, a veteran councilor of Madrid (1999-2015) who held various high positions in the city and took over the presidency of the community of Madrid (2018-2019) after a scandal leading to the resignation of his predecessor. Speaker 14 is a 54 year old Male Madrid Conservative engaged in a scripteD speech in the current audio. He stepped down from the PP in 2018, joining *Ciudadanos*, a formerly leftist party that began to swing center-left in the same year. This speech lasted 31 minutes, of which 9.3 minutes were analyzed (i.e., 30.2% of the speech), resulting in 1666 tokens for LFI analysis.

This speech was presented by Garrido at the *Forum Europa*, an International European Economy Forum, focusing on the role the PP plays in Madrid's governance and going on to answer questions alongside the Spanish Senate President (Figure 67). He discusses a recent debate between members of the major parties, and emphasizes the 100 public work projects the PP will

be soon implementing. He begins with an early regional peak (20 sec), thanking the various politicians who have come to his talk, with an address to:

(23) *Presidente del Senado y presentado en este caso – de lujo para mí – Pío García Escudero.*

The Senate President, and presented in this case – a luxury for me – Pío García Escudero.

Moments later, as he lists other dignitaries and politicians, he pauses to say that they are "on their way or have already arrived… ah, on their way," suggesting that the audience may be relatively bare at the start of the proceedings. His wry reference to the fact that Pío García Escudero is present serves as a brief moment of relief and sincerity, especially given the topic of the next regional peak. As Garrido continues, he describes his path into politics, identifying himself as a politician "of Pío," having met and learned from him early in his career, before becoming a councilor in the Madrid town hall. Now, he says (210 sec), Pío supported him in assuming the role of President of Madrid following his predecessor's resignation, and that:

(24) *Tengo el honor de que la primera persona que confió en mi, que fue Pío García Escudero, pues tambien me presente como president de la comunidad de Madrid. Por tanto de verdad especial ilusión que me ha hecho Pío y tus generosas palabras que te vuelvo a agradecer.*

I have the honor that the first person to trust me, Pío García Escudero, also introduced me as president of the community of Madrid. Therefore, Pío has done me a great honor, and I again thank you for your generous words.

As a Madrid conservative, it is unsurprising that Garrido's regional production rate is comparable to that of Valenciano, another NCPS speaker. However, it is noteworthy that once again in reference to Pío García Escudero, Garrido's regionalisms spike, emphasizing the sincerity of his appreciation and excitement to be speaking to García. The final regional spike occurs (360 sec) as he begins to list several major accomplishments during his tenure, including:

(25) *...Por poner solo tres o cuatro ejemplos. Son en todo caso datos y hechos que creo que confirman que las políticas sociales no solo no son patrimonio de la izquierda sino que además… quien de verdad las platica, las amplia, y las mejora es el centro derecha…*
To give just three or four examples. In any case, they are data and facts I think confirm that social policies are not only the heritage of the left, but also… those who really talk about, broaden and improve them are those on the center right [like me].

Here, as with Bernal, the reference to social systems and individuals in need of social support may lead to a shift towards a more rural vernacular. This is, after all, a right-leaning politician who explicitly conveys the message that the right is responsible for issuing important social policies, and these lines serve as a conclusion to a lengthier discussion of policy changes that Garrido and his party have been responsible for. However, as Garrido's average rate of regionalisms hovers around 30%, and he comes from Madrid, these are relatively small bumps in comparison with Andalusian speech norms, which may result from his instead targeting rural NCPS speakers.

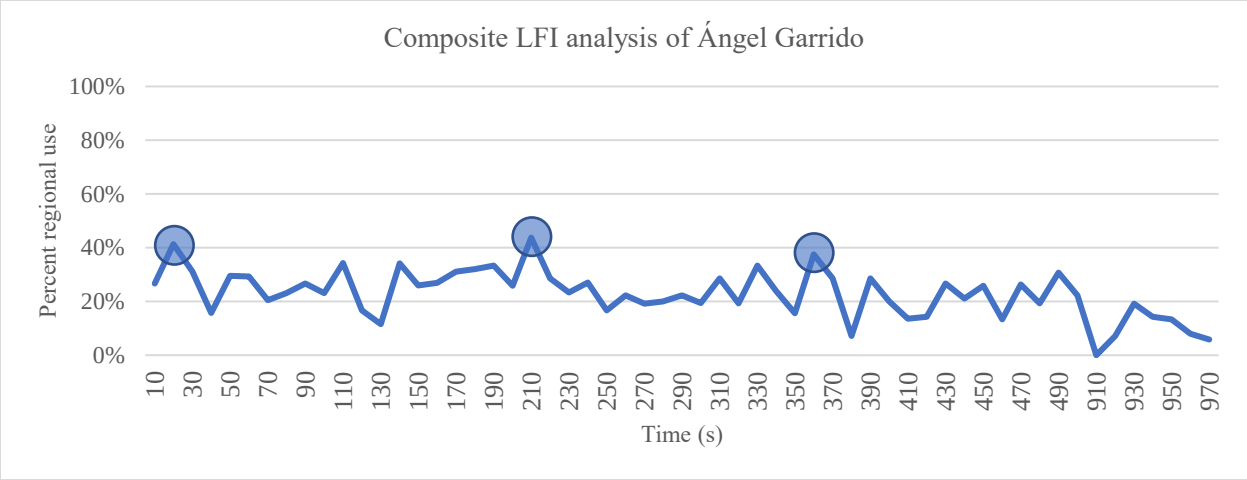

Figure 67: Composite LFI analysis of Ángel Garrido in ten-second intervals

Considering these three moments based on regional variant use (Figure 68), we see that tap and lateral reduction are both relatively high, while intervocalic /d/ elision tends to remain below 60%, and coda /s/ elision, while present, remains low. This patterning appears most similar to Valenciano, the other Madrid politician, although Garrido elides less intervocalic /d/ and has higher rates of liquid elision. When addressing Pío García initially (20 sec), thanking him later on (210 sec), and speaking to social services (360 sec), the same variables are in play each time – coda /s/ hovers near 20%, intervocalic /d/ is above 50%, and regional liquid productions are common (with the regional lateral least common, followed by the tap, followed by the trill).

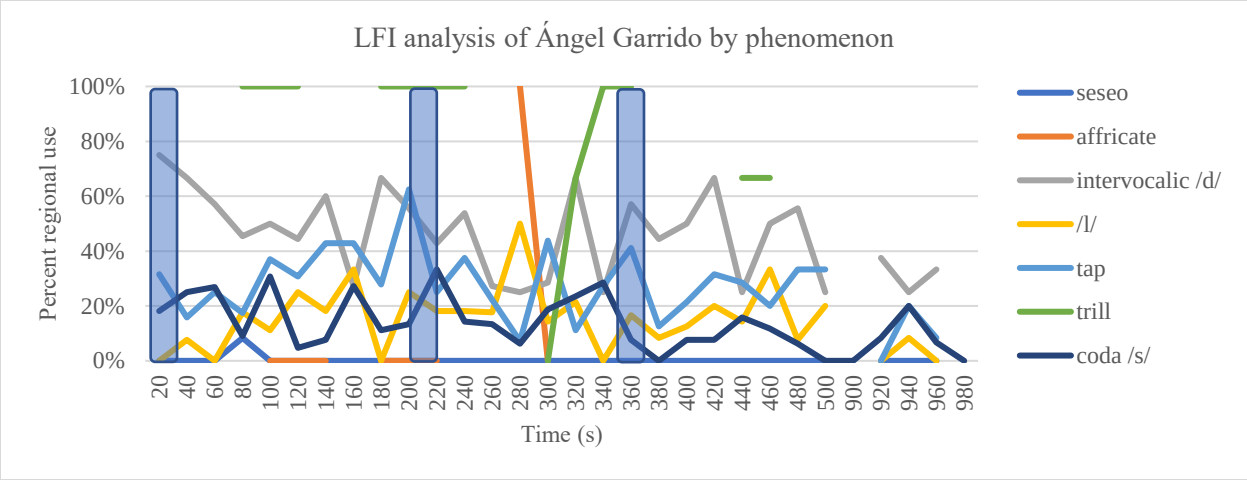

Figure 68: LFI analysis by phenomenon for Ángel Garrido in twenty-second intervals

5.3.8 Adolfo Molina (22MCCM)

Adolfo Molina Rascón, the current president of the Córdoba PP, was Secretary General of the party (2012-2017), a regional congressman (2012-2022), and president of the local Cabra PP (2005-2017). Speaker 22 is a 40 year old Male Córdoba Conservative engaged in an interview with a Male interlocutor in the current audio. The interview, which was broadcast on a local Córdoba station, lasts for 18 minutes. Eight minutes were coded for the LFI analysis (i.e., 44.1%), which considers 1209 tokens.

Molina speaks in this interview about evergreen topics, including ways to get into politics, the importance of democracy, and the desire among voters for change in the upcoming 2015 elections (Figure 69) The interviewer references recent cases of corruption, saying that voters have more and more begun to see politicians as the problem rather than the solution, and if that is placing

democracy at risk. Molina responds that some individuals are taking advantage of the situation to cry corruption in every situation (160 sec), but that:

(26) *Yo creo que el sistema democrático es el mejor qe podemos tener. En el case de nuestro país es el que más progreso ha traído a España y el que más ha cambiado nuestro país en estos años de democracia.*

I believe that the democratic system is the best we can have. In the case of our country, these years of democracy have brought the most progress and change to Spain.

In this moment of seriousness, Molina emphasizes the contributions of a democratic system, and the need for its maintenance, emphasizing his own belief in the importance of maintaining it. The interviewer goes on to reference recent cries of corruption from all corners of the government, and Molina agrees that people are tired, but that there is a need for a restructuring to take the focus away from such claims, and to reorient voters' attention toward meaningful project and social improvement. This change is needed not just in the news media (250 sec):

(27) *Pero efectivamente creo que ahora mismo se tiene también que producir un cambio en los partidos. Un cambio de relacionarnos con el ciudadano.*

But indeed, I believe that right now a change is necessary in the parties – a change in relating to the citizen.

In this conversation about accessibility of politicians and ways to engage the public, Molina's regional usage rises, providing a convergence not only to regional viewers, but also his

interlocutor. Finally, the last peak occurs near the end of the segment when the topic turns to a recent push on the part of the PP to allow rural mayors from simultaneously serving in Andalusian parliament, as members of city councils in urban areas (who are largely left-leaning) have no such limits. The interviewer notes that this has been an issue for decades, since the PSOE came into power, and asks why a new outcome is now expected, to which Molina responds:

(28) *Yo lo que veo es que después de 33 años hay cosas que no han cambiado en Andalucia, que siguen siendo desgraciadamente… pues, el paro, por ejemplo. Diez puntos por encima de la media nacional.*

After 33 years, I see things that haven't changed in Andalusia, which unfortunately continues to be… well, unemployment, for example. 10 points above the national average.

It is curious that, similar to Margarita Del Cid, this moment of regional usage does not correspond as neatly to moments of sincerity and reducing social distance with the interlocutor. Instead, Molina offers up points on the PP platform, from the importance of democracy, to a need to reconnect with citizens, to the need for change in the nation. Whereas the behavior comes off as more naturalistic for some of these socialist speakers, Molina shows how regionalisms can also come about at highly performative moments – rather than showing a 'real me,' these moments emphasize a united message offered on behalf of his party.

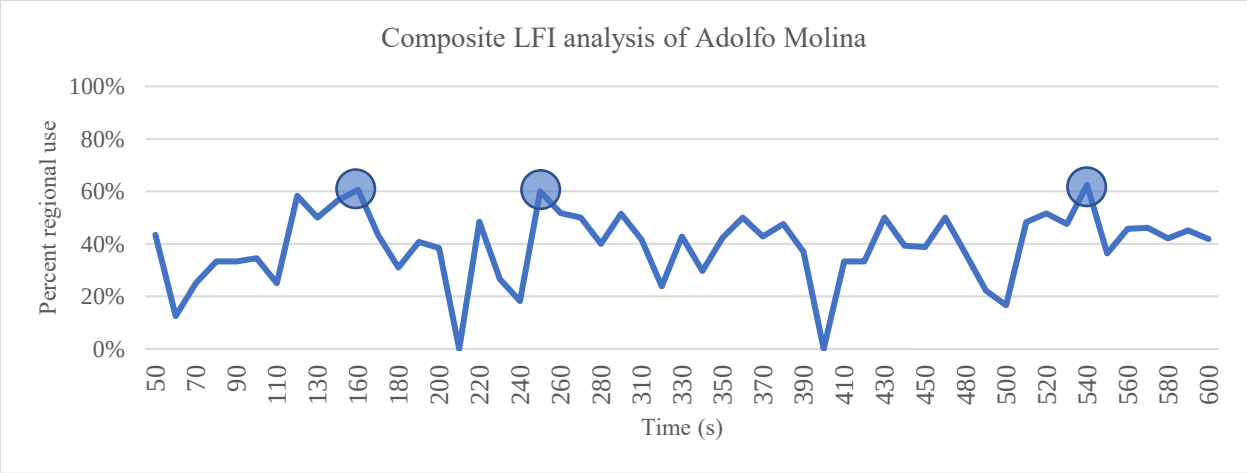

Figure 69: Composite LFI analysis of Adolfo Molina in ten-second intervals

Molina has a curious breakdown by regional variants, most similar to José Bernal from EAS (Figure 70). Like Bernal, there are moderate occurrences of liquid variation, and near categorical elision of coda /s/. However, in addition, there are high rates of intervocalic /d/ elision more reminiscent of Elena Valenciano and Susana Díaz. In the discussion of the importance of the democratic system at 160 seconds, high coda /s/ elision, moderate intervocalic /d/ elision, and low rates of tap and lateral reduction contribute to the regionalism peak. Then, at 250 seconds when discussing a need to better connect with voters, and at 540 seconds when discussing the need for change, those same regional features emerge at similar rates.

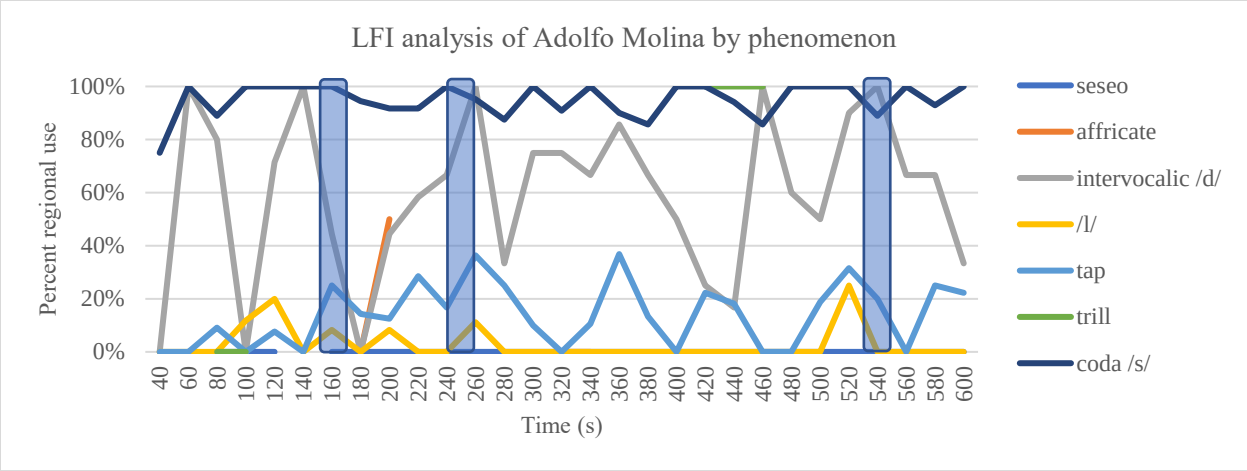

Figure 70: LFI analysis by phenomenon for Adolfo Molina in twenty-second intervals

It should be noted how different the two NAS speakers are. Carmen Calvo uses trill reduction and affricate fronting regularly, has generally lower rates of tap reduction, occasional uses of *seseo*, and medial rates of coda /s/ elision. Molina, on the other hand, has something approaching bands of variation – regular coda /s/ and intervocalic /d/ elision, moderate tap and lateral reduction, and negligible affricate fronting, trill reduction, and *seseo*. This emphasizes not only variation by political affiliation and gender, but also at the individual level.

5.3.9 Teresa Jiménez-Becerril (19FSCF)

The penultimate speaker, Teresa Jiménez-Becerril, was a longtime member of the EU Parliament (2009-2019), served in congress (2019-2021), and has served as an assistant to the Spanish Ombudsman to congress since 2021. Speaker 19 is a 53 year old Female Seville Conservative, who speaks with a female interviewer in the audio under examination. The interview took place on a local station, lasting four minutes, which were entirely coded, yielding 457 tokens.

This interview memorializes a very personal topic for Jiménez-Becerril, taking place fourteen years after the assassination of her brother and sister-in-law, Alberto and Ascension, by the terror group ETA (Figure 71). The interview has five notable peaks of regional variants, one of which tops 70% and another nearly reaching 80% regional production. After speaking briefly about Alberto's place in her memories, the interviewer affirms that "Seville won't forget them," to which Jiménez-Becerril responds at the 40 second mark:

(29) *No, Sevilla no se olvida y estoy segura que esta noche bueno que esta tarde a las cinco en la categdral habrá mucha gente.*

No, Seville won't forget, and I'm sure that tonight, there will, well, that this afternoon at five o'clock in the cathedral, there will be a lot of people gathered.

Her response reaffirms the interviewers claim, and references the mourners who will gather, before turning to describe other events, as well as a denunciation of the crimes ETA has committed, as her regional use falls. This moment of closeness, mirroring the words of the interviewer, allows for greater closeness between the two as they reflect on the weight of the day. As she describes the crimes of ETA and the need to pay, Jiménez-Becerril argues that their words and presence is necessary to remind the terror group of the crimes they have committed (80 sec), rather than allow them to be forgotten, at which point:

(30) *Ellos pueden igualar a víctimas, verdugos, poner a todos iguales, y al final no, nosotros temenos que estar aquí en la Calle Don Remondo…*

They can make the victims and the executioners equal, and in the end no, we have to be here on Don Remondo street [where her brother and sister-in-law lost their lives]…

There is a marked shift between her discussion of ETA at 50 and 60 seconds, and the return to the memories of the victims, and the stand she argues should be taken against the terror group. The interviewer asks what steps ETA needs to take to permit a ceasefire, causing Jiménez-Becerril to complain about the current state of ETA in the Basque Country (110 sec), where they aim to:

(31) *… completamente eliminar a los que no piensan como ellos o sea es la falta de libertad. Ellos no van a estar contentos hasta que el país vasco sea suyo.*

Completely eliminate those who do not think like them, that is, it is a lack of freedom. They will not be happy until the Basque Country is theirs.

In this moment, the politician's speech references those victims, and the future goals of the group, taking control of the territory. Again, as the frame of her reference shifts, there is again another drop in regional variation, with the topic turning to the demands, rather than apologies, that ETA has directed toward them, and the lack of patience she feels toward them (140 sec), saying:

(32) *… que no les debemos nada a ellos, los que nos deben son ellos a nosotros.*

… that we don't owe them anything, they are the ones that owe us.

Again, as the topic shifts from the frame of ETA and their crimes to the needs of the people of Spain, the regional usage again increases, reinforcing her words and tying her to the Andalusian

community as one of many who have been affected by the terror group, demanding they be held accountable. The final peak comes in the closing moments of the interview, as the interviewer thanks Jiménez-Becerril for her time (240 sec), and she returns the gratitude by replying:

(33) *Muchísimas gracias a vosotros porque sois vosotros con vuestras altavoces y vuestras cámaras lo que de verdad también frena allí es que esta gente se salgan con la suya.*
Thank you very much because it is you with your loudspeakers and your cameras that really also stops these people from getting away with it.

In this final segment before the video closes, Jiménez-Becerril also connects the interviewer and news media with the crusade against ETA. Just as she has addressed Seville, the victims, Basque citizens who don't agree with the terror group, and Andalusians in general, lectal focusing allows her to show not only her convergence with the Seville interviewer, but also set up a dichotomy between the terror group and those who oppose it.

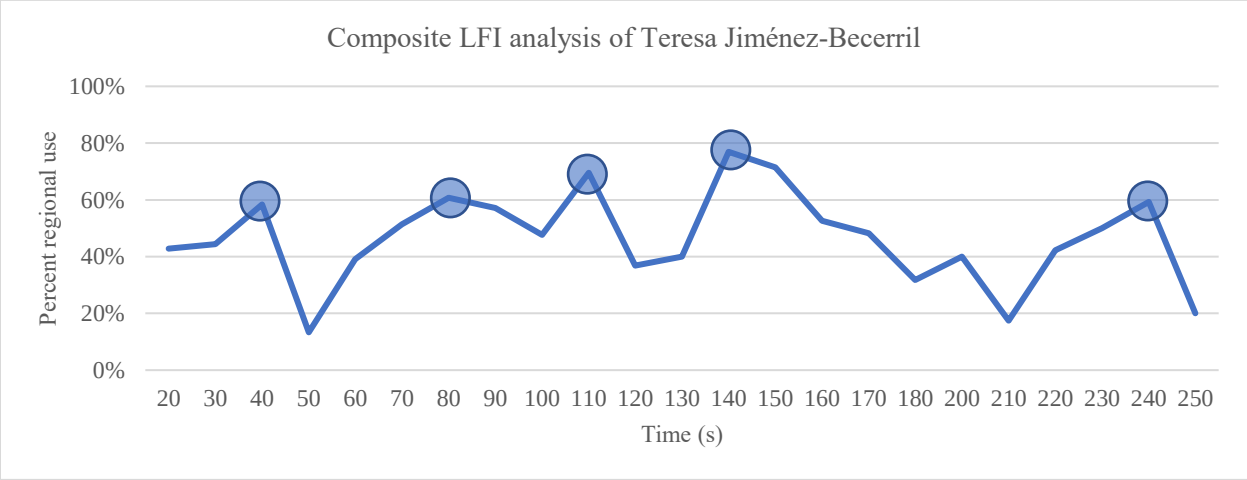

Figure 71: Composite LFI analysis of Teresa Jiménez-Becerril in ten-second intervals

The patterns we see from Jiménez-Becerril resemble those of other Andalusian speakers (Figure 72), perhaps most of all Adolfo Molina from Córdoba, with high rates of coda /s/ and intervocalic /d/ elision, and medial rates of liquid phenomena (other than the tap). However, with relation to her use of *seseo*, she most closely resembles Susana Díaz and José Antonio Pérez Tapias, with regular spikes throughout her speech. Across the overall regional spikes, however, *seseo* tends to be low or not present, while it is tap reduction, intervocalic /d/, and coda /s/ elision that most contributes to regional feature peaks in speech.

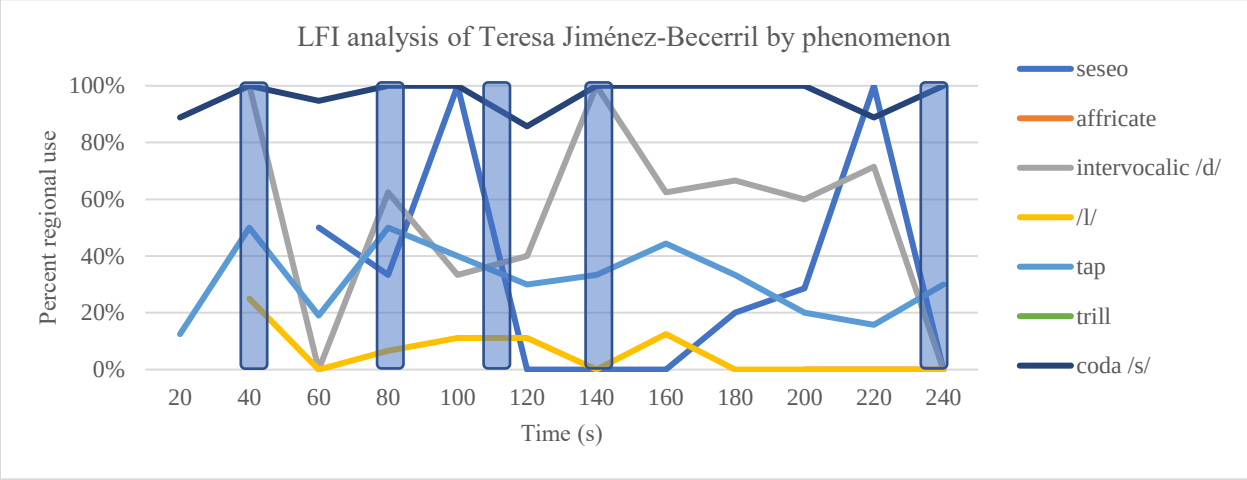

Figure 72: LFI analysis by phenomenon for Teresa Jiménez-Becerril in twenty-second intervals

5.3.10 Juan Zoido (18MSCM)

The final speaker is Juan Ignacio Zoido, a former mayor of Seville (2011-2015), congressman (2016-2019) and Minister of the Interior (2016-2018), and member of the EU parliament since

2019. Speaker 18 is a 60 year old Male Seville Conservative, who speaks with a Male interviewer in the current audio. The interview took place on a national station and lasted for 20 minutes, of which 8.6 minutes were coded (43.4%). All told, 1303 tokens were produced for analysis in the LFI analysis.

In this interview, Zoido discusses various matters related to interior affairs, including recent actions of the civil guard and investigations into corruption (Figure 73). In the first moments, the interviewer identifies a long history speaking with Zoido, and expresses his surprise that the politician would make his way to Minister of the Interior, asking if he had expected it, or if he has taken it on with resignation. Zoido responds (40 sec), saying that:

(34) *Tampoco. Pero así es la vida… [y] lo llevo con responsabilidad.*

Me either. But that's life, and I take it on responsibly.

As the interlocutor here is an NCPS speaker from Madrid, the 'real me' that emerges here at the start of the interview seems to suggest that a form of "divergence as solidarity" is ongoing here, with Zoido comfortably entering the conversation with a longtime colleague with less formality or normative speech. The conversation turns to Zoido's responsibilities in the government, including a discussion of the Civil Guard's agenda, including their work with drug trafficking, which is very broad. The interviewer interrupts momentarily, saying "We'll see, but you…" before being cut off as Zoido begins to speak (540 sec):

(35) *Yo les decía– Yo les decía que ellos hacen unas– unas investigaciones se la encargan los los fiscales y los jueces y ellos pues despúes emiten un parecer.*

I told them– I told them that they carry out some– some investigations entrusted to them by the prosecutors and the judges, and then they issue an opinion.

At this moment, the rushed rate of speech and repeated words, coming as Zoido interrupts the interviewer to continue speaking (for almost another minute), suggests both a desire to finish his train of thought and an apologetic approach. Regional variation comes into play almost as a way to continue speaking while attempting to dismiss the impoliteness toward his interlocutor – as Sharma (2018) discusses, these are moments when speakers are suggesting their true stances and reducing social distance with their interlocutor, making it also an effective way to step in with a clearly intended purpose. Near the end of the coded segment, the interviewer and speaker go back and forth. The interviewer asks if something the current PP had done would have been done previously. Zoido references the government of José María Aznar (1996-2004). The interviewer comments on the high quality of that specific government. Zoido defends the present one, insisting it is of similar quality. The interviewer doubles down, insisting that certain individuals in the Aznar government were particularly skilled. Zoido insists that there are such individuals in the current government. After all this back and forth, the original topic is nearly lost, and the interviewer apologizes for the distraction, letting Zoido get back to answering the original question (1580 sec):

(36) *No, no, sí, yo, lo que, lo que le quería decir queeee no existe es el fuego amgio. Yo yo pienso que hoy por hoy estamos más centrados en otro, en otro serie de temas.*

No, no, yes, I, what, what I wanted to tell you that doesn't exist is friendly fire. I think that today we are more focused on other, on other issues.

Both seeming a bit rattled from the interviewer's questions, and trying to be as political as possible in the face of a question critiquing the current government, regional features return. However, it is also contextually relevant to mention that, amidst the back and forth, Zoido insists that he is from the province of Seville and is "not accustomed to what happens in Madrid." In this way, as with Margarita Del Cid and Adolfo Molina, Zoido's behavior here seems almost to play into rural stereotypes, falling back on higher rates of regionalism as a means of playing up his role as an outsider, unfamiliar with the big city, a seeming throughline in the speech of some southern conservatives.

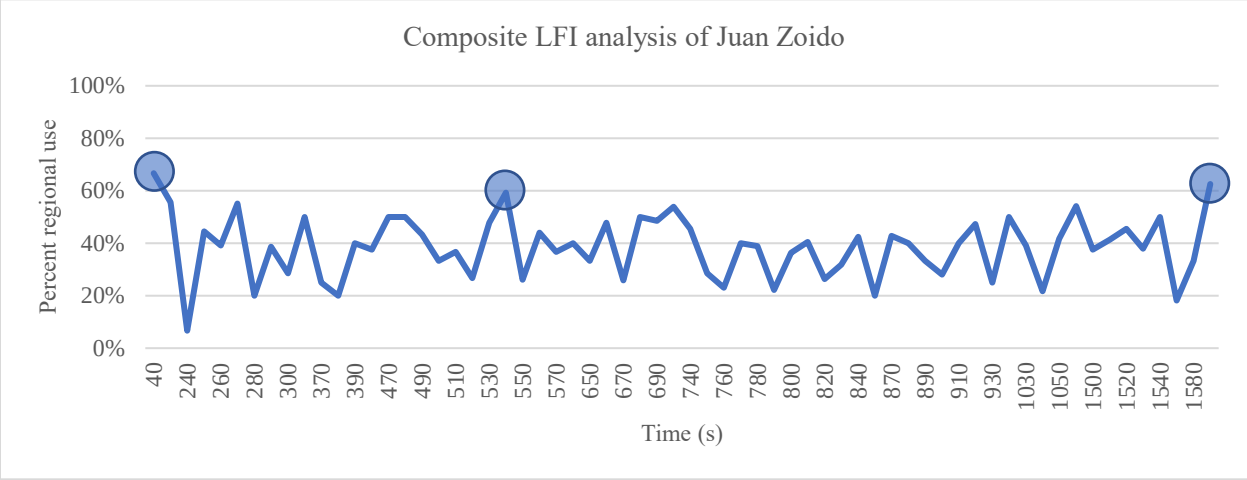

Figure 73: Composite LFI analysis of Juan Zoido in ten-second intervals

The last step is to examine the regional phenomena that emerge across these peaks (Figure 74). In the introductory segment (40 sec), coda /s/ is the only Andalusian feature at play. Meanwhile, mid-way through the segment as Zoido interrupts the interviewer (540 sec), there is nearly categorical coda /s/ elision, as well as medial rates of intervocalic /d/ and trill reduction, and lower rates of tap elision. In this way, Zoido resembles other Seville politicians like Díaz and

Pérez Tapias, with high rates of regional coda /s/, middle-high levels of intervocalic /d/ and liquid reduction, and low rates of affricate fronting like other male politicians. However, Zoido *is* the only Seville politician to show nearly no *seseo* use at all, fitting him in with politicians from the rest of Andalusia and Madrid. This could result from the venue and context of speech (i.e., Madrid, with a Madrid interlocutor), as well as from the national position he holds, given the lack of prestige *seseo* enjoys in the capital.

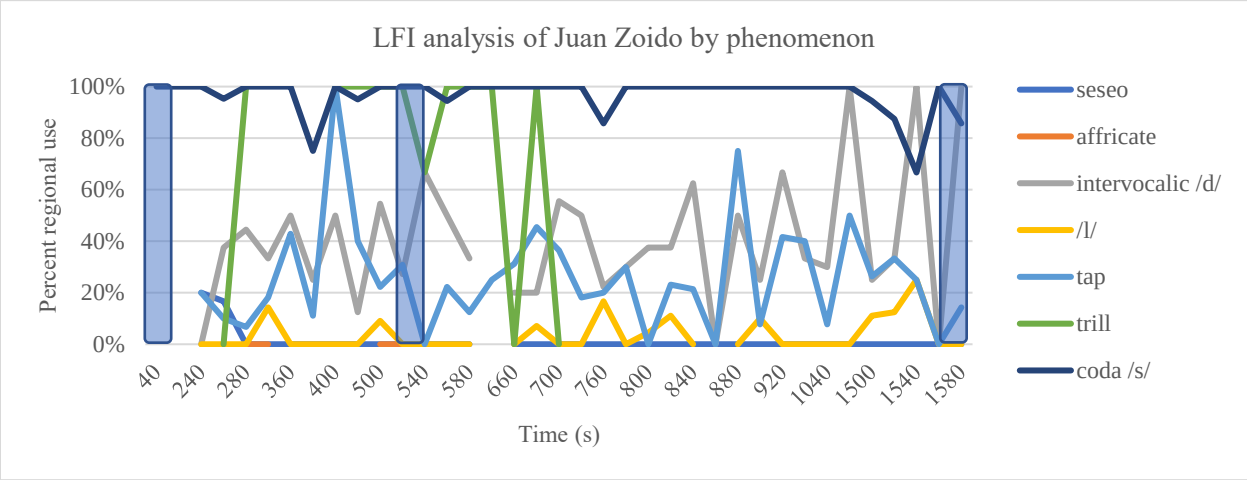

Figure 74: LFI analysis by phenomenon for Juan Zoido in twenty-second intervals

5.3.11 Comparison of speaker results

Based on the results from the analysis of LFI by phenomenon, a final accounting of variation was made across the speakers, presented in Table 32. This presents the speakers as ordered in the previous ten sections, including their number, variety, and name, as well as a composite percent of regional speech across the entire coded segment. Following that, phenomena are identified

based on their relative proportion in peaks discussed in the LFI analysis – those that have a rate at or above 60% at moments of regional production peaks are marked with "X," while those with a rate between 20-59% are coded with "--." In this way, three categories of use were created, differentiating between use greater than chance (60-100%), use that could be classified as occasional or chance (20-59%), and infrequent to nonexistent use (0-19%).

				Phenomena @ Peaks*						
#	**Variety**	**Name**	**% Regional**	**/t͡ʃ/**	**/s/**	***seseo***	**/d/**	**/ɾ/**	**/r/**	**/l/**
7	WAS	Susana Díaz	34%		X	--	--			
4	NCPS	Elena Valenciano	19%				X	--	X	--
5	WAS	José Antonio Pérez Tapias	46%		X	X	--	--		--
26	EAS	Pepe Bernal	45%		X		--	--	X	--
32	EAS	Margarita del Cid	42%	X	X		X	--		
11	NAS	Carmen Calvo	30%		X	X	X	--		
14	NCPS	Ángel Garrido	23%		--		--	--	X	
22	NAS	Adolfo Molina	22%		X		X	--		
19	WAS	Teresa Jiménez-Becerril	47%		X	--	X	--		
18	WAS	Juan Zoido	38%		X		X		X	

*Where X indicates values of 60-100% and -- indicates values of 20-59%.

Table 32: Regional variant tendencies from the LFI analysis by speaker

This chapter also examines the moments of those peaks in the text of the audio transcripts to determine what speakers are doing when regional features increase. Table 33 provides a visualization of these moments, breaking them down into topical categories. These post-hoc categories are broken down by greeting and gratitude (e.g., thank you for having me), opinions, confiding about heretofore unknown facts, expressing pride in one's party's accomplishments, verbal stumbles, promises made to listeners, references to the working class, and moments of humor or sarcasm. As Eckert (2008) and Cutillas-Espinosa and Hernández-Campoy (2007) describe, contexts where style occurs do not follow specific pre-existing categorizations – instead, they must be teased out through actual usage, as a way to represent the multidimensionality of

indexical meaning across a variety of speech situations. Thus, these categories, drawn from the contexts identified during regional peaks in the results section, provide insight into political speech norms among Andalusian and Madrid speakers.

#	Region	Greeting/ Gratitude	Opinion	Confiding	Pride	Verbal Stumble	Promise	Working Class	Humor
7	WAS	X					XX	XX	
4	NCPS		XX						X
5	WAS		XX	X					
26	EAS		X	X			X	X	
32	EAS	X	X	XX	X				
11	NAS		X						X
14	NCPS	XX			X				
22	NAS		XX			X			
19	WAS	XX	XX	X					
18	WAS			X		XX			

Table 33: Topics at regional production peaks by speaker in the LFI analysis

All told, LFI analysis presents an interesting visual display of variation that offers a way to pinpoint moments in speech worthy of further examination. The next section returns to these results, integrating previous research and theory into the discussion to determine what novel contributions this semi-qualitative analysis adds to the dissertation project, and our understanding of third-wave variation in Andalusian Spanish politics.

5.4 Discussion

Based on the LFI analysis, it can be seen that certain phenomena (e.g., affricates and trills) have a much lesser impact on moment-to-moment variation than others (e.g., coda /s/ or intervocalic /d/), based merely on the frequency of contexts where they can be used. Regional norms also play an

undoubtable role here, with NCPS speakers differing markedly from WAS and EAS speakers. There are also interesting qualitative differences that emerge with respect to the contexts where variation is most common, using regional variation not only as a means of emphasizing moments of sincerity, but also as part of the momentary construction of a rural or friendly persona. Analysis of community evaluations in Chapter 6 will serve to address the effectiveness of using regional variants in this way.

As Sharma and Rampton (2015) argue, LFI analysis also provides an interesting avenue for distinguishing between social groups that, at the surface level, do not appear to have any statistically meaningful differences in production. The problem with analyses of lectal focusing through traditional logistical means alone is the type of composite data that analysis requires. By eliding differences between phenomena, individuals, and regionalisms into a single dependent variable at the community level, shades of differences are collapsed, and transitory moments of identity construction and stance-taking are lost (Johnstone & Kiesling, 2008; Eckert, 2008). While, on its own, consideration of lectal focusing cannot paint a complete picture of linguistic variation, and it does not provide as robust a quantitative analysis as regression models, when the two approaches are paired, a better understanding of speaker behavior across the community and individual face work within it come to the fore.

Given how much variation Chapter 4 identifies across phenomena, with regional feature production remaining relatively stable in some cases by social group, and varying markedly for others, LFI analysis provides some much needed insight into style-shifting. In the sections that follow, the research questions that began this question make their return, allowing for a discussion of patterns of variation at the level of phenomena, individuals, and social groups.

5.4.1 Regional variant breakdown

The first issue to address involves stylistic variation at the level of regional phenomena, identifying how variation takes place. Sharma (2018) focuses particularly on regional variation peaks as moments in which speakers reduce social distance with their interlocutor, using Indian English norms in a form of 'divergence as solidarity.' However, for the politicians in the LFI analyses in this chapter, speaking at times with interlocutors from their own region and at times with those from other parts of Spain, divergence is not a requirement in these 'real me' moments.

When speakers are interviewed on local television, for example, their audience is very familiar with the regional linguistic features they employ. Rather than showing accommodation with one's audience, the 'real me' moment in these cases is speaker-led, reflecting something about the individual, independent of their hearers. Sharma (2018:25) describes this as "biographical indexicality," where it is a speaker's individual background rather than the norms of a community that dictate their usage. "An upper-class person may use their first-learned upper-class style to enact exactly the same frank stances that a working-class speaker achieves with their vernacular," Sharma (2018:25) writes. This goes toward explaining the dearth of AS phenomena in the speech of these individuals, such as deaffrication of [t͡ʃ] to [ʃ], *ceceo* and *heheo*, all of which are associated with rural and elderly speech, which these upper- and middle-class mainly urban individuals would not have had direct contact with. While speakers, like Hernández-Campoy and Cutillas-Espinosa's (2013) María Martínez, may emphasize certain forms in their environment during childhood to develop a hyper-vernacular pattern of speech, overtly stigmatized forms are much less likely to appear in political speech. Instead, the proportions of regional variants vary based on the perceived

persona with which they are associated (e.g., using more intervocalic /d/ to sound "political," Cruz-Ortiz [2019] finds it to be).

Looking at speech norms during peaks in regional variation, we return to Table 32 in section 5.3.11, with the breakdown by speaker. All Andalusian speakers had high rates of coda /s/ elision throughout their speech, but especially at regional peaks. All politicians had at least some presence of intervocalic /d/ reduction, although for at least one speaker from each region, those rates remained on average below 60%, regardless of speaker gender or political party (although more conservatives did elide more). Following that, medial levels of tap elision and reduction were common, although two Seville politicians did not have high rates of these. Also, unfortunately, the binary treatment of the regional variable meant that a speaker like Pepe Bernal, who had much higher rates of variable liquid production as compared to elision, was not identified as having markedly different behavior than his peers, despite nuanced differences in social meaning associated with variable liquid production versus elision. The remaining phenomena were less common, with trill reduction and *seseo* occurring in the speech of four speakers at regional peaks. While trill reduction was regionally spread out (and most common in NCPS), *seseo* was produced at high rates during regional speech by WAS and NAS speakers. Finally, the lateral was reduced by one EAS, WAS, and NCPS speaker, and the affricate was fronted by one EAS speaker, suggesting that these variants rely on individual differences rather than regional or political ones.

Thus, while there is the potential for all of these variants to play a role in regional identity construction, their rates differ by speaker region and other social factors. This is hardly unexpected – it would be strange for a Madrid politician like Ángel Garrido to use anywhere near as many "Andalusian" variants as a WAS speaker like Susana Díaz. As referenced in the results section, the social meaning differs greatly across these two individuals. While Díaz is approximating AS

norms, Garrido's productions are clearly not "Andalusian," but rather employing vernacular variants used on the national stage associated with working-class and rural speech. Thus, Garrido uses little of the salient AS coda /s/ elision, and none of the southern-marked *seseo*, while Díaz uses both. In this way, it is important to examine not only the overall Andalusian composite, but also to distinguish between individual phenomena as part of the whole picture of variable use.

5.4.2 Style-shifting patterns

The next research question guiding this chapter involved identifying style-shifting patterns across individuals to determine behavioral norms. Politicians have been explored in peninsular Spanish previously based on their agentive decisions, distinguishing themselves as individuals based on their profiles of regional feature use (e.g., Hernández-Campoy & Cutillas-Espinosa, 2010, 2013; Cruz-Ortiz, 2019; Pollock & Wheeler, 2022). Thus, this question aims to address the gap between the community-based analysis in Chapter 4 and the present individual-based LFI analysis, seeing how certain individuals' behavior can map onto one of several shared patterns of regional expression to better understand speaker-available norms and identity-related behavior.

While the same seven regional phenomena are presented in Table 32 for all ten speakers, and many of them overlap with respect to any one or two variant-types, it is curious to note that the overall patterns serve as a linguistic fingerprint distinguishing each speaker. No two individuals share the exact same pattern of production. Only two speakers approximate each other: Carmen Calvo and Teresa Jiménez-Becerril, the former from Córdoba, the latter Sevilla. Both women reduce coda /s/, the tap, and intervocalic /d/, and use *seseo*. Based on these seven phenomena and

three levels of regional production (i.e., full "X", medial "--," and none ""), there are 2187 possible patterns. Even presenting them as a binary (i.e., present vs. absent), there are 128. And this is further complicated by the fact that certain infrequent contexts like the affricate or trill simply may appear more in certain discussions (e.g., laws for *coches* 'cars' or acts of *terrorismo* 'terrorism'), and be wholly absent from others, making it impossible to consider some variants in the context of regional behavior. Thus, matching exact patterns through LFI analysis seems unlikely.

However, as referenced throughout the presentation of results, trends do emerge for individuals and small groups of phenomena across speakers. I begin with the sibilants. The most obvious of these involves coda /s/: NCPS speakers have low rates, while most AS speakers have near categorical usage. However, the female speakers from Córdoba and Málaga show that there is also an available reduction pattern in AS with closer to 70-80% reduction, which differs markedly from their WAS and male AS peers. Next, the affricate is relatively rare; many speakers use little of the fronted production – especially men – while AS liberal women tend to show variation in production. This makes Margarita del Cid stand out, as a conservative woman who nonetheless has almost categorically fronted productions, especially at regional peaks. Finally, for *seseo*, NCPS and EAS speakers use almost only distinction, while most WAS and NAS men avoid it —however, Pérez Tapias groups with female politicians in these regions.

Next, for the intervocalic /d/, there are essentially two patterns of use. Four speakers make extensive use of it, often approaching or at 100% regional production across blocks, including both Córdoba politicians, and women from NCPS and EAS. Meanwhile, the rest of speakers have a medial use, centering around 50%. No speakers routinely retain /d/.

Finally, the liquid phenomena mainly see a reduced use, tending to be medial, low, or non-existent. Women, for the most part, show low rates of regional tap and lateral production, although

Margarita Del Cid and Jiménez-Becerril have medial tap production rates. Meanwhile, men produce more medial rates of taps and low-to-medial rates of laterals. Trills are harder to track as, like affricates, they are infrequent, making occurrence in a given block less frequent, and regional rates more subject to fluctuation. However, men in general tend to have closer to 100% production, while female speakers tend to have more variation between 0% and 100%.

Another topic for consideration is the representation of the LFI composite. As shown in Chapter 2, Figure 4, Sharma (2018) finds a swing of up to 80% between Indian and American English features in Fareed Zakaria's speech. Over the course of the 105 utterance units (her temporal measurement), there are six moments when the speaker approaches and tops 80% Indian English variants. Across these politicians, there were few individuals who had such a drastic spread between regional and normative peninsular Spanish variants. Only Valenciano, Jiménez-Becerril, and Zoido (all Seville politicians) closely approached or passed the 80% mark a single time. Looking back to the regional percentage breakdown in Table 32, there seem to be three groupings by city: Madrid politicians had the lowest rate of regional productions, followed by Córdoba speakers, while Seville and Málaga speakers were comparable. This reflects the spectrum of regional production Samper-Padilla (2011) discusses, moving from most normative in Madrid down to most vernacular in the southern parts of Andalusia.

Looking at the LFI analysis composites, there are also different patterns by individual: some speakers have much greater variation from moment-to-moment (e.g., Díaz, Zoido), others have broad curves of variation (e.g., Tapias, Jiménez-Becerril), and some spike as much as 25% between blocks (e.g., Valenciano, Calvo). While this requires further examination and comparison with other speakers in the future, additional social variables involving the speech context, such as the gender of the interlocutor, venue of speech, expected audience, topic, or speech rate likely all

play a role in these differences at the discourse level, alongside individual agency and differences. Even across these ten speakers, there is simply too much potential for variation to account for all pattern differences.

Based on these results, this chapter demonstrates how speakers differ across speech patterns, providing a degree of visualization that could be fruitfully applied to other descriptions of political speech. For example, determining exactly how hypervernacularity (Hernández-Campoy & Cutillas-Espinosa, 2010), MUSE and AAL norms of coronal stop deletion (Holliday, 2017), and contextual differences (Kirkham & Moore, 2016) come into play in identity work can identify important moments in discourse deserving of more attention.

5.4.3 Social variables & style

Based on the phenomena- and individual-level findings presented in the first two sections, the final research question examines the role of social categories and divisions in the explanation of identity work. The breakdown of topics provided in Table 33 offers an idea of some differences in the way that the politicians navigate regional production. By political party, both sides use lectal focusing as a means of advancing their opinion and confiding in the interlocutor with frequency. Conservative politicians are more likely to use more regional features at the start of an interview and to express gratitude to interlocutors (e.g., thank you for having me). The only verbal stumbles that used increased rates of regional productions were produced by conservative male speakers of NAS and WAS, while only socialists made promises, referenced the working class, and

(specifically female socialists) used humor in moments of lectal focusing. Finally, pride in one's party and accomplishments occurred on both sides of the aisle, although more often among men.

As referenced in the results section, there also seemed a greater likelihood for conservative politicians to use regional variants in more intentional contexts. These included topics related to their political platforms and party talking points, referencing a need for change (Molina), indexing friendliness in relation to a town fair (Del Cid), and referencing one's own provincial background and lack of familiarity with "the way things work" in Madrid (Zoido). Meanwhile, the socialists examined tended to use regional variants in a way that furthered their arguments without being interpreted quite so cynically. These politicians reference working-class Andalusians (Díaz, Bernal), planned government works (Bernal), representative democracy (Tapias), and education funding needs (Valenciano), in addition to making jokes (Valenciano, Calvo).

Previous work has tied political party to identity work. At the phonetic level, Hall-Lew et al. (2017) find that Scottish parliamentarians differ based on their political background, with National Party vowel spaces resembling northern norms and Labor Party ones being closer to London expectations. These differences, they argue, reflect a tendency for the party arguing in favor of secession from the UK to diverge from British norms. Hernández-Campoy and Cutillas-Espinosa (2013) argue that Martínez's linguistic performance serves as a way to connect with working-class norms. Although regional production was found to be comparable across the two parties in Chapter 4, as Sharma and Rampton (2015) argue for London Punjabi communities, the LFI analysis shows differences at the moment-to-moment level based on topics, individuals, and social factors. Even when overall production tendencies are similar, underlying phenomenological differences contribute to differences in social meaning that will emerge through perceptual analysis.

5.4.4 Final takeaways

This chapter demonstrates how Lectal Focusing in Interaction can be analyzed across speakers to compare style-shifting tendencies at the individual level in order to better understand speech behavior. Fruitful shifting exists across all phenomena discussed in the chapter, including at the level of regional features, individual behavior, and social groupings. However, painting variation with a two-toned brush (i.e., regional vs. normative) is not sufficient to understand this behavior. Region, speaker gender and political party intersect with individual goals to present a complex web of variation. This analysis also shows how pairing qualitative analysis of topical differences with moment-to-moment variation and the community-level analysis in Chapter 4 presents a clearer picture of variation. Patterns emerge in behavior that groups speakers together, meaning that interpretation requires knowledge about the evaluative meaning of regional phenomena. The results from this chapter, in the context of community norms from Chapter 4 and perceptual responses to variation by Seville listeners in Chapter 6, will be discussed in greater depth in Chapter 7 as well, contextualizing each chapter's findings through the sum total of acoustic and auditive information assembled about political speech in Andalusia.

6 Stage 3: Andalusian perception

Perceptual analysis provides important insight into the way that members of a community evaluate and respond to speech stimulus. As Chapter 3 discusses in depth, there is a broad array of research on production phenomena in Peninsular Spanish sociolinguistics (and, of course, beyond), but the amount of perceptual work is still relatively sparse. Nonetheless, published studies in this realm provide important answers that production research on its own is often unable to address.

Several approaches to perception have provided insight into Andalusian Spanish (AS). There has been work examining differences in vowel height and laxing tendencies to determine whether contrast is maintained across contexts with elided coda /s/, /θ/, and /r/ (e.g., Henriksen, 2017; Herrero de Haro, 2017a, 2018, 2019). This project provides a key piece of insight into formant differences in production studies, determining that while marked differences in production exist, they also reflect real phonological boundaries possessed by speakers of Eastern Andalusian Spanish (EAS). Other work has used matched-guise tasks (originally described by Lambert et al. [1960]) to determine the prestige and attitudinal responses of AS speakers toward deaffrication of /t͡ʃ/ (Regan, 2020; Milla-Muñoz, 2020) and rhoticism (Milla-Muñoz, 2020). They find that deaffrication is seen as less urban, but more friendly, and that rhotacism is seen as less professional, while the Northern Central Peninsular variety is growing in prestige among younger speakers. Perceptual work has also been used to determine regional differences, and track how rural speakers account for local, regional, and national norms when evaluating speech from outside of their area (e.g., Harjus, 2017).

The current chapter adopts a political speech methodology similar to one used by Podesva, Reynolds, Callier, and Baptiste (2015) as a means of using rigorous production analysis (already

described in Chapters 4 and 5) to develop an attitudinal perceptual instrument. These authors examine the English released /t/, a linguistic feature that has previously been associated with intellect, as a means of examining digitally-manipulated stimuli from six mid-2000s era U.S. politicians. First, in their production study, the authors examine an hour of speech for each politician, collecting over five thousand tokens and reinforcing impressionistic coding with spectrographic cues such as following aspiration and aperiodicity, and present findings from regression analyses.

After their production analysis, they create a perceptual study using matched-guise tasks, in which listeners compared tokens with and without released /t/, created through digital splicing. The authors provided several adjective scales with which participants rated stimuli, including friendly/unfriendly, southern/not southern, sincere/insincere, intelligent/not intelligent. The resulting survey was distributed via a web platform, and politicians were identified by name, title, political party, and photo. Responses were found to differ based on the actual identity of the speaker, as well as the frequency with which individuals used the released /t/ in speech. Several aspects of the perceptual instrument distributed in this chapter rely on details from this article, including ideas for the adjectival scales, online distribution, and identification of politician information as a means to investigate bias toward members of certain political parties or even toward well-known individuals.

Following the perceptual methodology employed by Podesva et al. (2015), this chapter focuses on the attitudinal analysis of regional variants by listeners from western Andalusian Spanish (WAS) collected in the city of Seville. The first section begins by describing the key questions raised in this stage of the research, as well as hypotheses based on previous research and chapters within this dissertation. Following that, the next main section describes the methodology

used in examining WAS perception, including a description of the audio collection, the design of the perceptual instrument, the participants who took part in the study, and the means through which the data was collected and analyzed. The third main section presents the results of the three parts of the perceptual instrument, including stimuli provided blind to listeners, stimuli identified as coming from specific politicians, and stimuli comparing two politicians. That section closes with a comparison of the results across all three parts of the instrument. The final sub-section offers a discussion of the results in light of previous research and other chapters of the dissertation, returning to the research questions before concluding with main takeaways regarding regional Andalusian perception.

6.1 Research questions and hypotheses

Based on previous perceptual work in southern peninsular Spanish, attitudinal, geographical, and acoustic differences exist across some of the phenomena under examination in the current dissertation. There seems to be an ongoing trend, in which younger speakers and individuals with access to higher education are coming to favor certain features of the normative NCPS variety (e.g., Regan 2017b), while at the same time local distinctions may be giving way to more unified regional standards across WAS and EAS (Villena-Ponsoda, 2008, 2013).

This causes the valence of regional variation to face a number of perceptual outcomes – potential loss of prestige in favor of the spreading NCPS variety, or, alternately, region-specific prestige growing as a means of indexing belonging to Andalusia, in opposition to the north (e.g., a parallel process to the rural-urban divide discussed by Melguizo-Moreno [2010]). As discussed

in Chapter 3 with respect to regional variation, many of these AS phenomena are also associated with working-class, minimally-educated, elderly speakers who lack social prestige. Given the relatively widespread nature in political speech of some phenomena found in Chapter 4's analysis of production (e.g., word-final /s/, intervocalic /d/ reduction), in contrast with the much more infrequent nature of others (e.g., *ceceo*, liquid reduction and variable production, etc.), it seems likely that some phenomena will be less socially salient than others in the perceptual evaluations.

The differences in production lead us to the third overarching question presented in this dissertation, which the current chapter aims to address: namely, what attitudes do regional and normative variants garner among members of the AS community, and how does this influence their evaluation of political speech and identity construction in stimuli? The perceptual instrument is sub-divided into three parts to better address this question. At the same time, this division will allow three additional questions to be considered.

The first of these questions addressed in this chapter deals generally with evaluations of speech produced by other members of the community. The instrument does not initially specify the provenance of stimuli, allowing listeners to think of the audio as simply coming from AS and NCPS. In this way, it aims to avoid the concerns expressed by Podesva et al. (2015) initially, allowing speakers to express their general opinions about variation without being biased by partisan affiliations or opinions about individual politicians. This should provide insight into AS listeners' perception of normative variants in opposition to regionally-prestigious ones.

The next question, drawing on part two of the instrument, examines political speech and how individuals in the public sphere are evaluated with respect to language. In this second part, listeners receive an image of each speaker, as well as their name, and political information. By examining these results, and contrasting them with findings from the previous part, it is possible

to determine how perceptions differ between “community members” and politicians. As Hernández-Campoy and Jiménez-Cano (2003) describe, politicians use similar sociolinguistic variation to what is found in their communities, even though their speech is held to different standards of evaluation than non-politicians. Thus, while these stimuli come from the same individuals, it is likely that politicians will receive more negative attitudes for using regional variation than members of the community.

Finally, the last question is drawn from the final part of the instrument, in which listeners compare a regional and normative production of a phenomenon by two separate politicians without knowing any social information about these individuals. Based on their differentiation of these two speakers, how successful is the identity work of individual politicians based on community evaluations? In this analysis, the question comes to be whether politicians who produce high rates of regionalisms are evaluated more positively than those who produce normative variants, considering the interaction with underlying social factors, like gender and age.

These three questions should provide insight not only into the norms of evaluation, but also allow Chapter 7 to offer some insight into the results from the previous two analytical stages. By understanding how speech is evaluated, as Podesva et al. (2015) show in comparison to their own production analysis, better explanation of production behavior becomes possible. While others examining political speech elsewhere in Spain (e.g., Hernández-Campoy & Cutillas-Espinosa, 2010, 2013) have found correlations between regionalisms and working-class speech, perceptual results from members of the AS community will help to determine if listeners feel similarly.

6.2 Methodology

This study draws on work in English political speech (e.g., Podesva et al., 2015), perceptual research (e.g., Campbell-Kibler, 2009, 2010), as well as attitudinal studies in other varieties of Spanish (e.g., Dominican Republic: Willis & Díaz-Campos, 2021; Peninsular and American varieties: Díaz-Campos & Navarro-Galisteo, 2009; Díaz-Campos & Killam, 2012) to develop an instrument that both reflects social and political realities of the peninsular context. The goal of the resulting instrument is to determine the evaluative norms of speech communities, like the one in Seville, where voters reside. This allows insight into the sort of identity that is being portrayed, based on the "script" of the political medium of speech (Hernández-Campoy & Cutillas-Espinosa, 2007), as well as the performative deviations from that script that contribute to identity construction through stance accumulation (e.g., Hernández-Campoy & Cutillas-Espinosa, 2010, 2013; Pollock & Wheeler, 2022).

In order to avoid the difficulties of manipulating audio while maintaining a natural sound (discussed by Pollock [2019]), audio was drawn directly from the corpus analyzed in Chapter 4 and presented to listeners without a carrier phrase. Instead of using actors in a matched-guise methodology, this chapter adopted the verbal guise (Cooper, 1975) methodology employed by Willis and Díaz-Campos (2021), in which listeners receive stimuli taken from real speech scenarios.

This section continues with the presentation of four key components in the development of the perceptual methodology. The first section discusses the corpus of stimuli chosen for analysis in this study, including a description of the way that social factors were balanced across stimuli, and what distractors were added. The following section details the layout of the perceptual instrument itself, providing insight into the goals of the three main parts, as well as the background

information collected from participants. The third section discusses the listeners who took part in the project, including a summary of their background information, and how data was collected. Finally, the fourth section details the statistical analysis applied to the data.

6.2.1 Corpus

Speeches and interviews analyzed in Chapter 4 were used to find regional and normative stimuli for this instrument. This corpus includes 94 speech files, representing 35 hours of audio from 32 politicians evenly divided by gender, age, city (i.e., Córdoba, Málaga, Seville, and Madrid), and political party (i.e., the left-leaning Spanish Socialist Worker's Party (PSOE) and the right-leaning People's Party (PP)). Each speaker was represented by files from three different speech contexts, including scripted speeches, interviews with female interlocutors, and interviews with male interlocutors. Using the coding methodology identified in Chapter 4, regional AS variants, often containing reduction, were distinguished from normative NCPS variants.

One of the key goals in selecting stimuli to use in the perceptual instrument was to balance as many social factors as possible to help control for differences across individuals and groups. This involved selecting tokens produced by both speakers of AS and NCPS for both regional and normative variants to reduce the effect of other aspects of regional accent (e.g., F0 and intonation, etc.). Equal numbers of male and female speakers, conservatives and socialists, and at least one representative from each city were selected to represent the phenomena. This breakdown is shown in Table 34, in which the first full row for regional and for normative phenomena was included in part one of the instrument, the second row in part two, and the third row for part three.

Type	affricate fronting	*ceceo*	syllable-final /s/ reduction	intervocalic /d/ elision	lateral reduct.	*seseo*	tap reduct.	trill reduct.
Regional	FLC	FLS	FLC	FLC	MLS	MLS	MLC	MLS
	FSS	MSS	MSS	MSC	MSS	FSC	FSC	FSC
	MCC	MCC	FCS	MCS	FCC	FCC	FCS	MCS
Normative	MMC	FMS	MMC	FMS	FMS	MMS	MMC	FMC
	FSS	FSC	MSC	FSC	MSC	MSS	MSS	FSS
	MCS	MCC	FCS	MCS	FCC	FCC	FCS	MCC

The three letter codes represent: 1 = gender (Male/Female), 2 = city (máLaga, Seville, Córdoba, Madrid), and 3 = party (Socialist, Conservative)

Table 34: Social factors selected across the perceptual instrument

In order to draw attention away from the overall goal of the study, two types of distractors were also included. One sort, drawing on regionally-important lexical items, included politicians' productions of the items *Madrid* and *Sevilla*. The other included several stimuli unrelated to AS and NCPS variable production, presenting stimuli taken from the speech of a Catalonian politician, Carles Puigdemont. This offered less-familiar phenomena for listeners to classify.

All told, eight of the ten phenomena described in the production analysis of Chapter 4 were included in the perceptual instrument. *Ceceo*, while rare, was included due to the likely high salience it was expected to have. Meanwhile, given the lack of social factors conditioning variation of vowels, and the strong association between EAS and vowel laxing and harmony, these were not included for the WAS listeners to evaluate. Finally, given its similarity to syllable-final /s/ reduction, and to simplify the sibilant token variety offered to participants, word-final onset /s/ reduction was also excluded from consideration in the instrument. The eight phenomena participants heard included affricate /t͡ʃ/ fronting, intervocalic /d/ elision, syllable-final /s/ reduction, *ceceo*, *seseo*, and the elision and variable production of the liquids /l/, /ɾ/, and /r/.

6.2.2 Perceptual Instrument

A survey was designed on the online platform Qualtrics and distributed to participants. First, informed consent was obtained from all listeners. Following that, a background questionnaire was provided to participants, gathering information about their city of origin, gender, age, time spent living in Seville, level of education, university major (if applicable), political affiliation, and participation in the political process. A brief practice section with distractors was provided, which allowed listeners to adjust their headphones to a proper volume and familiarize themselves with the slider system and adjectival classification used in the rest of the survey.

Following that, the main body of the instrument was divided into three parts. In Part 1, participants were told only that they will hear speech from peninsular Spanish speakers and rated 20 words for five adjectives (1) that were used in all three parts of the instrument. Listeners also identified where they thought each speaker was from using a heat map.

(1) Adjectives used for rating tokens, scaling from 0 to 100.

1. *Urbano* --> *Rural* ‘urban/rural’
2. *Andaluz* --> *No Andaluz* ‘Andalusian/Not Andalusian’
3. *Culto* --> *Inculto* ‘educated/uneducated’
4. *Simpático* --> *Antipático* ‘likeable/not likeable’
5. *Liberal* --> *Conservador* ‘left-leaning/right-leaning’

Next, in Part 2, participants were told explicitly that the 20 stimuli were taken from politicians. They saw an image of each speaker (Figure 75), as well as the individual's name and most recent political office held, although their political party was not identified.

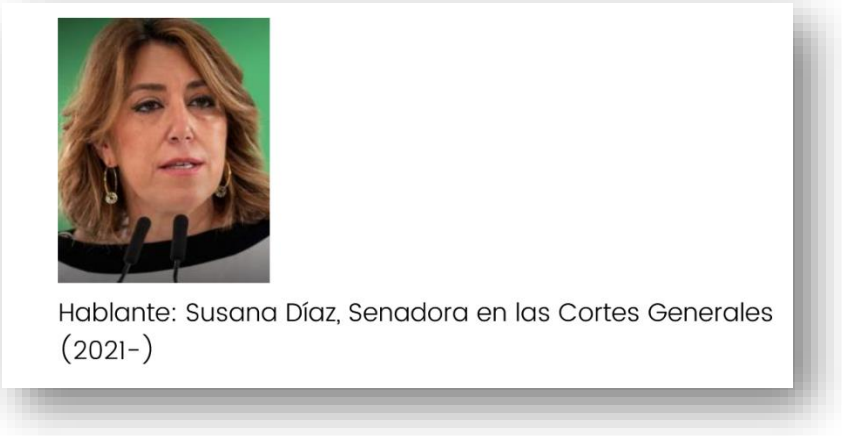

Figure 75: Example from Part 2: 'Speaker: Susana Díaz, Senator in the General Courts'

In addition to rating each token on the adjective scales, they were asked to again identify the origin of the politician, and now, additionally, to state the political party they thought they belonged to. Finally, in the third part of the questionnaire, participants received pairs of words, one produced using an NCPS production, one with an Andalusian production, and answered an AB perceptual task identifying tokens that best embodied each of the five adjectives. No information about the two speakers was provided. At the end, after identifying the political party and region of the speakers, they were also asked to identify one speaker who they would vote for if given the chance. An example of the sliders in Qualtrics is given in Figure 76.

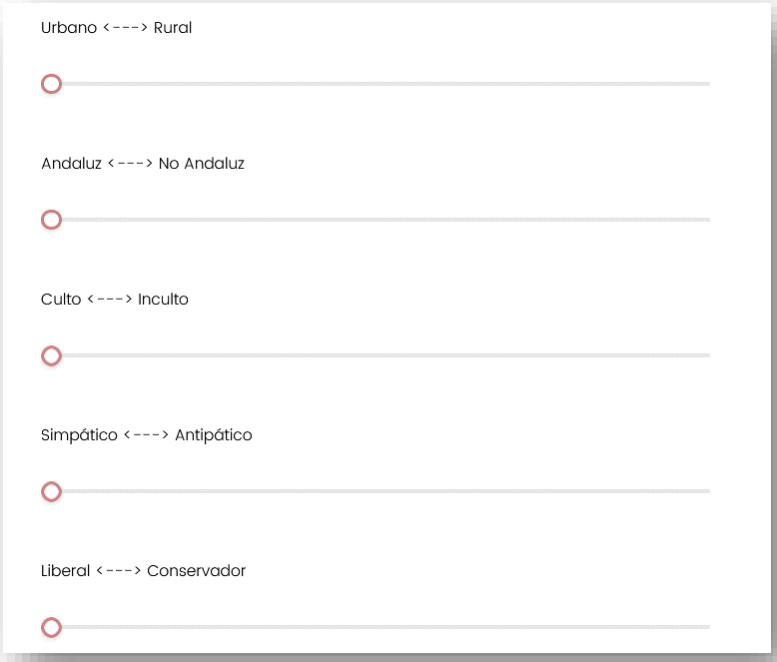

Figure 76: Adjective identification sliders in the Qualtrics perceptual instrument

For each part of the instrument, eight pairs of stimuli were selected to represent phenomena, produced by a politician from Madrid and another from a variety of AS. Thus, 4 distractors and 16 regional variants were presented per instrument part, yielding a total of 54 non-distractors heard and evaluated by listeners. In Part 1, the 16 target stimuli were ranked for the five adjectives and received a regional identification on a heat map (i.e., 16x6=96 responses per participant). In Part 2, the 16 stimuli were evaluated for the five adjectives, for regional identification, and for political party identification (i.e., 16x7=112 responses per participant). Finally, for Part 3, the 8 pairs of stimuli were compared for the five adjectives, for regional and party identification of both participants, and for "selected candidate" (i.e., 8x10=80 responses per participant). In sum, this resulted in 288 responses from each listener in the main three parts of the instrument. Tokens used in each part of the online survey are presented in Table 35.

#	Normative/ Regional	Phenomenon	Token	Phonetic Production	English
1.1	Normative	Affricate	*dieciocho*	[di͡eθiot͡ʃo]	‘eighteen’
1.2	Regional	*Ceceo*	*un sitio*	[unθiti͡o]	‘a place’
1.3	Normative	Syllable-final /s/	*buenas*	[buenas]	‘good day’
1.4	Regional	Intervocalic /d/	*nada*	[naØa]	‘nothing’
1.5	Normative	Lateral /l/	*cualquier*	[ku͡alki͡eɾ]	‘any’
1.6	Regional	*Seseo*	*solicitar*	[soli͡sitaɾ]	‘request’
1.7	Normative	Tap	*lideresa*	[liðeɾesa]	‘leader (f.)’
1.8	Regional	Trill	*reconocemos*	[ɹekonoθemos]	‘we recognize’
1.11	Regional	Affricate	*mucho*	[mut͡so]	‘a lot’
1.12	Normative	*Ceceo*	*presionando*	[pɾesionando]	‘pressing’
1.13	Regional	Syllable-final /s/	*intentamos*	[intentamoØ]	‘we try’
1.14	Normative	Intervocalic /d/	*alentada*	[alentaða]	‘encouraged’
1.15	Regional	Lateral /l/	*finalmente*	[finaɾmente]	‘finally’
1.16	Normative	*Seseo*	*emoción*	[emoθi͡on]	‘emotion’
1.17	Regional	Tap	*parlamento*	[paØlamento]	‘parliament’
1.18	Normative	Trill	*carro*	[caro]	‘car’
2.1	Regional	Affricate	*satisfecha*	[satisfet͡sa]	‘satisfied’
2.2	Normative	*Ceceo*	*posicion*	[posiθi͡on]	‘position’
2.3	Regional	Syllable-final /s/	*podemos*	[podemoØ]	‘we can’
2.4	Normative	Intervocalic /d/	*pedido*	[peðiðo]	‘requested’
2.5	Regional	Lateral /l/	*socialista*	[soθi͡aØi͡sta]	‘socialist’
2.6	Normative	*Seseo*	*generación*	[xeneraθi͡on]	‘generation’
2.7	Regional	Tap	*pero*	[peØo]	‘but’
2.8	Normative	Trill	*ocurre*	[okure]	‘happens’
2.11	Normative	Affricate	*dicho*	[dit͡ʃo]	‘said’
2.12	Regional	*Ceceo*	*pues*	[pu͡eθ]	‘well’
2.13	Normative	Syllable-final /s/	*consejeros*	[konsexeɾos]	‘counselors’
2.14	Regional	Intervocalic /d/	*visitados*	[visitaØos]	‘visited’
2.15	Normative	Lateral /l/	*siglo*	[siglo]	‘century’
2.16	Regional	*Seseo*	*celebrando*	[selebɾando]	‘celebrating’
2.17	Normative	Tap	*congreso*	[kongɾeso]	‘congress’
2.18	Regional	Trill	*realiza*	[ɹaliθa]	‘accomplishes’
3.1.1	Normative	Affricate	*dicho*	[dit͡ʃo]	‘said’
3.1.2	Regional		*marcha*	[mart͡sa]	‘progress’
3.2.1	Regional	*Ceceo*	*diecisiete*	[di͡eθiθi͡ete]	‘seventeen’
3.2.2	Normative		*resumen*	[resumen]	‘summary’
3.3.1	Normative	Syllable-final /s/	*buenas*	[bu͡enas]	‘good day’
3.3.2	Regional		*estar*	[eØtaɾ]	‘to be’
3.4.1	Regional	Intervocalic /d/	*presidente*	[presi͡Øente]	‘president’
3.4.2	Normative		*podido*	[poð͡iðo]	‘can (particip.)’
3.5.1.	Normative	Lateral /l/	*familia*	[famili͡a]	‘family’
3.5.2	Regional		*personales*	[pɾsonaØes]	‘personal’
3.6.1	Regional	*Seseo*	*modernización*	[modeɾnisasi͡on]	‘modernization’
3.6.2	Normative		*roce*	[roθe]	‘touch’
3.7.1	Normative	Tap	*septiembre*	[septi͡embɾe]	‘September’
3.7.2	Regional		*gobierno*	[gobi͡eØno]	‘government’

3.8.1	Regional	Trill	*barrio*	[ba.r̪i̯o]	‘neighborhood’
3.8.2	Normative		*terrible*	[te.r̪ible]	‘terrible’

Table 35: Lexical items in the perceptual instrument

These sub-divisions of the instrument will each address a research question. Part 1 establishes a baseline before introducing the bias inevitable by revealing the identity of politicians (as Podesva et al. [2015] discuss), showing how listeners evaluate variation outside of the political context. Part 2 makes clear the nature and backgrounds of the speakers, causing listeners to interpret speakers as public figures and introduce a degree of personal biases based on their political affiliations. Contrast between the first two parts should show how these factors prejudice evaluation. Additionally, listeners will be primed in Part 2 to associate regional variants with politicians from the AS variety (i.e., all regional variants will be produced by AS politicians). Finally, in Part 3, speakers are not identified, and regional productions are randomized across NCPS and AS speakers. Identification of speakers’ regions and political parties will show assumptions based on political preferences, as well as variants associated with AS.

Overall, this instrument demonstrates which variants are salient in AS, to what extent they carry social meaning, and the influence of political affiliation on that meaning. The instrument, including the original Spanish text and a translation, is provided in Appendix E.

6.2.3 Survey participants & data collection

The online survey in the Qualtrics platform was first distributed among a small group of phoneticians and graduate linguists at a large public midwestern U.S. university for pilot testing. Feedback was offered on the quality of the audios, the idiomatic Spanish phrasing used in some of

the background questions, and the overall layout of the instrument. Changes were made to address these concerns and improve the layout of the survey, resulting in the final product described in the previous section. Additional redundancy was also created for the audio files, after pilot testers discovered that some of the Qualtrics-hosted stimuli would occasionally fail to load. The final resulting survey was determined to have resolved these audio problems. Institutional Review Board approval was also obtained for this instrument (Appendix D).

In November of 2022, the survey was both presented on-location to students, faculty, and community members at the University of Seville and links were posted around campus inviting digital participation. Those individuals who took part in the experiment in-person received a small recompense of seven euros for their participation, made possible through a generous Grant-In-Aid from the Indiana University College of Arts and Sciences. In the on-location site, a large classroom provided through the generosity of the Department of English Philology at the University of Seville, tablets with local version of the survey were arranged at a large conference table, where participants who scheduled time in the room over the course of one week were able to set up and take part. This helped to offer more control over the audio quality of the survey, as headphones were provided. Online participants were also strongly encouraged to use headphones.

A total of 77 participants took part in the survey. There was an average survey time of 49 minutes, and a median time of 45 minutes. Since one individual selected “other” for gender identification, their responses were excluded from consideration to balance the groups, leaving 76 individuals in the study. Given the 288 responses per listener, this yielded 21,888 responses across all three parts of the survey, but the quantitative analysis presented in this chapter focuses on the 15,200 adjective-based responses.

The overall breakdown for these 76 participants is as follows. There were 25 male and 51 female listeners. The vast majority (n=64) were from Seville itself, while a smaller subset were studying or working in the city but originally came from east Andalusia (n=4), and other parts of west Andalusia (n=8). Most respondents were university-aged, either between 18-20 years old (n=39), or 21-30 (n=31). The remainder were older professors and community members (aged 40-59, n=6). A total of 66 participants were in-person on tablets taking the survey, while 10 took part virtually. Fifty-six listeners identified as having at least some familiarity with the field of linguistics, while the remaining 20 did not. The majority of respondents' highest level of education was a high school diploma (n=48), while lesser numbers had an AA (n=3), a BA (n=20), or a PhD (n=5). With respect to political affiliation, the majority of respondents identified as "liberal" (n=32), followed by "slightly liberal" (n=15), "very liberal" (n=11), and "politically moderate" (n=11). Only 7 participants identified as "slightly conservative" (n=4) or "conservative" (n=3), and no respondent described themselves as "very conservative."[21] The vast majority identified as being affiliated with major left-leaning political parties (n=31, PSOE; n=15, *Podemos*; n=3, *Ciudadanos*; n=2, Adelante Andalusia; n=1, PACMA), while a smaller number were associated with right-leaning parties (n=12, PP; n=6, *Vox*), and a few were apolitical (n=6). Finally, with respect to political information, most respondents identified their main source of political information as television and radio (n=38), with a smaller amount getting their information from both these sources and social media (n=22), still fewer only getting their information from social media (n=12), and very few saying they sought out no political information (n=4).

[21] Interestingly, potentially as a means of trying to (re)contextualize their rating, two separate participants came up to me after completing their survey, saying that they had identified themselves as "conservative," but that the scale should truly have been between "socialist" and "liberal," as the Spanish right is much further left than the U.S. one. While I have chosen to describe the division as socialist/conservative in this dissertation, it is curious the negative associations identified with American conservatism.

This information paints a picture of the relatively homogenous cohort one might expect to find at a large institution of higher learning, with a majority of female undergraduate students to the mid-left of the political spectrum. As the survey was advertised heavily in the department of English Philology, students with linguistics classes and studies were more likely to take part. The offer of financial remuneration for in-person participation was a large draw.

6.2.4 Data analysis

Following data collection in Spain, the responses from the 76 listeners were divided up by adjectives for the three parts of the survey, resulting in fifteen mixed-effect regression models. The participant and the politician were both treated as random effects to account for multiple responses produced by each participant, as well as occasional cases of politicians producing multiple words within a single section of the perceptual instrument. For the first two parts of the survey, when participants provided ratings from 0 to 100 for each of the five adjectives, five linear regressions were used in each part that treated individual adjective evaluations as a continuous variable, determining differences across responses. Meanwhile, in the third part of the instrument, when participants selected a preferred token as more “adjective” (e.g., more Andalusian), five logistic regressions were used to compare binary preferences between tokens.

Background questionnaire information about participants, including the differences described in the previous section, were treated as independent variables. In addition, specific variables related to region and political party identification differed across all three parts of the

analysis. Independent variables, including the listener variables included in every analysis and part-specific variables, are shown in Table 36.

Listener Variables		**Factors**		
Gender	Male	Female		
Phenomenon	Sibilant	Stop	Liquid	
Political leaning	-3 (very liberal)	-2 (liberal)	-1 (slightly liberal)	
	0 (politically moderate)	1 (slightly conservative)	2 (conservative)	
Linguistics Background	Yes	No		
Education	1 (High School)	2 (AA)	3 (BA)	
	4 (MA)	5 (PhD)		
Region	Seville	WAS	EAS	
Age	Continuous			
Part 1 Variables		**Factors**		
Region ID	By City			
Actual Politician Region	Madrid	Andalusia		
Politician gender	Male	Female		
Token regionalism	Regional	Normative		
Part 2 Variables		**Factors**		
Region ID	By City			
Actual Politician Region	Madrid	Andalusia		
Politician gender	Male	Female		
Token regionalism	Regional	Normative		
Party ID	PSOE	PP	*Podemos/ Ciudadanos*	*Vox*
Part 3 Variables		**Factors**		
Normative Production				
Politician Gender	Male	Female		
Politician Party ID	PSOE (L)	PP (R)	*Podemos* (L)	*Ciudadanos /Vox* (R)
Politician Region ID	By City			
Regional Production				
Politician Gender	Male	Female		
Politician Party ID	PSOE (L)	PP (R)	*Podemos* (L)	*Ciudadanos /Vox* (R)
Politician Region ID	By City			
Preferred Speaker (who would you vote for?)	Normative	Regional		
Comparison Variables		**Factors**		
Urban->Rural	Continuous			
Andalusian-> Non-Andalusian	Continuous			
Educated->Uneducated	Continuous			

Likeable->Unlikeable	Continuous
Liberal->Conservative	Continuous

Table 36: Independent Stage 3 variables in the analysis of perception

6.3 Results

This first subsection provides a descriptive overview of the main results from the instrument, as well as an idea of evaluations for each of the phenomena under examination. The remaining subsections discuss the mixed-effect regression models by adjective for each part of the survey.

In Part 1 of the instrument, 6080 continuous responses were provided by the 76 listeners across all five adjectives. Figure 77 offers a distribution of evaluations for regional and normative stimuli based on each of the three categories of phenomena under consideration, in which listeners received minimal information about the audio they were hearing. With respect to the urban-rural spectrum, normative and regional liquids and intervocalic /d/ were both seen as more urban and comparable in productions, while regional productions of sibilants received a higher rural identification. For Andalusian-Non-Andalusian, the normative productions of all three phenomena classes were seen as more normative, although intervocalic /d/ had the narrowest spread between regional and normative productions. This reflected production findings that intervocalic /d/ elision seems to be widespread throughout NCPS and AS. With respect to educated-uneducated, normative liquids and regional sibilants and intervocalic /d/ were seen as more uneducated. For likeability, normative variants were seen across the board as less likeable. Finally, for political affiliation, normative intervocalic /d/ was seen as more conservative, while political affiliation for sibilants and liquids was comparable across productions.

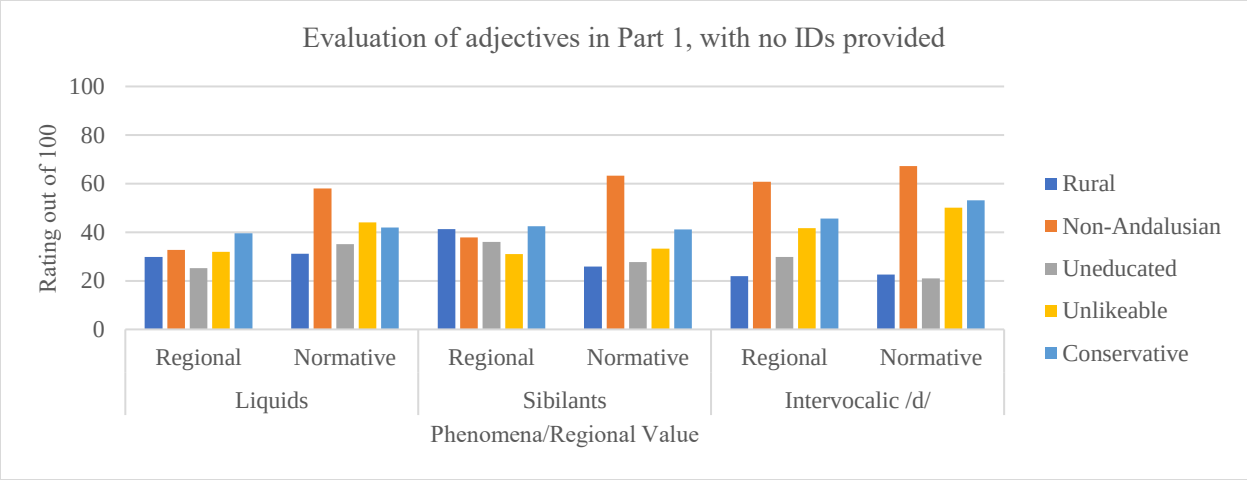

Figure 77: Adjective coding for Part 1 of the perceptual instrument

As in the previous portion, part 2 of the instrument yielded 6080 continuous responses from the listeners across the adjective continua. Figure 78 demonstrates this distribution for the task in which listeners received direct information about the identity of each speaker. In the contrast between urban-rural, unlike in the previous part, all three regional phenomena were seen as markedly more rural than their normative counterparts when produced by politicians. With regard to Andalusian-Non-Andalusian, the divide was again strong, with normative productions receiving a considerable evaluation as non-Andalusian across the board. All regional productions, including for intervocalic /d/, were seen as more Andalusian. For degree of education, the trend has actually reversed altogether – regional variants are now seen as less educated than normative ones for all phenomena. In terms of likeability, the gap narrowed – while normative variants are seen as less likeable, the difference is less marked, and both regional and normative sibilant phenomena are generally less well liked as a whole. Finally, for political affiliation, political differences for liquids and sibilants have become more marked, with normative variants being described as more

conservative across the board. This is curious, given the findings in Chapter 4 that conservatives, overall, produce more regional variants than socialists.

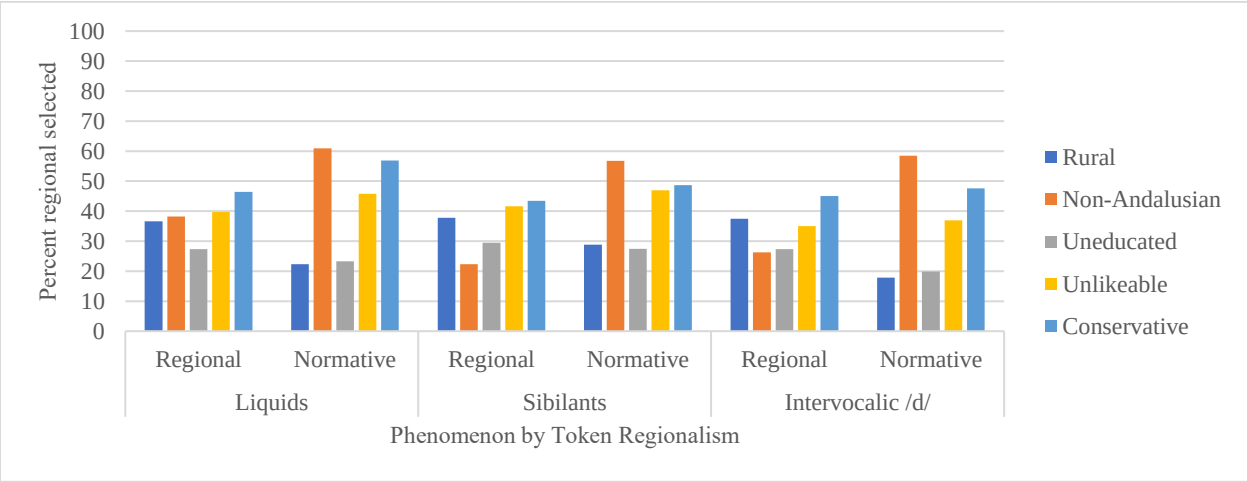

Figure 78: Adjective coding for Part 2 of the survey instrument

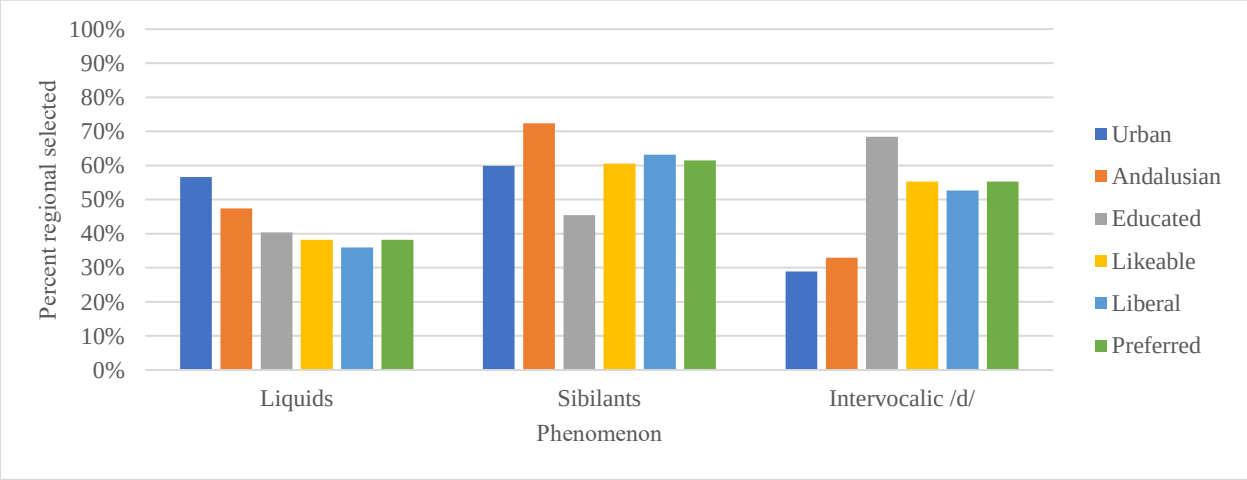

Figure 79: Regional tokens selected for adjectives by phenomenon

Finally, for part 3, responses came in the form of comparisons rather than continuous ratings, meaning that only about half as many responses were provided (as listeners also identified

a preferred politician who they would “vote” for if given the choice.) In sum, 3648 responses from listeners were collected from this section of the instrument. Figure 79 presents the percent of responses for each adjective that selected the regional production over the normative one as being more “adjective” (e.g., more Andalusian). With respect to the adjective urban, the regional production was selected around chance for liquids, more frequently for sibilants, and much less than chance for intervocalic /d/. Meanwhile, for identification as Andalusian, liquids were again selected around chance, while regional sibilants were much more likely to be identified as Andalusian, and regional intervocalic /d/ was rarely coded as Andalusian. For perception as educated, regional liquids were selected slightly less than chance as uneducated, while elided intervocalic /d/ was perceived as more educated, and sibilants were evaluated around chance. For likeability, regional liquids were selected less often, while regional sibilants were selected more often, and intervocalic /d/ was around chance. For political affiliation, regional liquids were seen as less liberal, while regional sibilants were described as more so, and intervocalic /d/ was around chance. Finally, for the “preferred” category added in part 3, listeners were slightly less likely to vote for speakers who used regional liquid productions, while they were slightly more likely to vote for those using regional sibilant productions and were around chance for intervocalic /d/.

While the results for part 3 of the survey provided comparative information, data from both part 1 and part 2 included ratings from zero to one hundred, and as such allow for comparative descriptions. These comparisons were made separately for each of the phenomena under examination in the survey, and are provided here. First, for the alveolar and pre-palatal allophones of the Spanish affricate, Figure 80 shows a breakdown of evaluations by the gender of the participant, the type of production, and the adjective. Among male listeners, the regional

production is seen as more Andalusian and socialist, while among female listeners, it is evaluated as more Andalusian, likeable, and socialist.

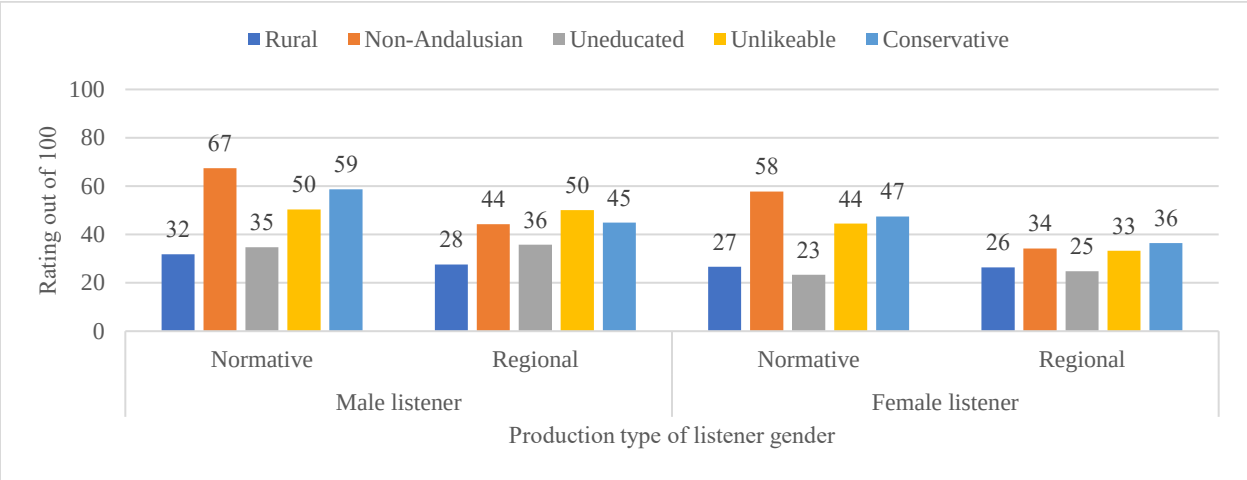

Figure 80: Affricate evaluation by listener gender, production type, and adjective.

Figure 81 provides a breakdown between the first part of the survey, where respondents were asked to simply evaluate "peninsular Spanish," and second part of the survey, where they found out that these were politicians and received some identifying information about the speakers. This figure breaks down the evaluations by survey part, production type, and adjective to show how evaluations differ for what are ostensibly members of their broader speech community, and for public speakers, specifically for the affricate. Both normative and regional productions were judged to be more rural and more unlikeable when known to have been produced by politicians. However, normative productions were seen as slightly less educated sounding when produced by politicians, and regional ones were slightly more educated. Additionally, normative productions were seen as more conservative, while regional productions were seen as more socialist and much more Andalusian, when produced by politicians as opposed to community members.

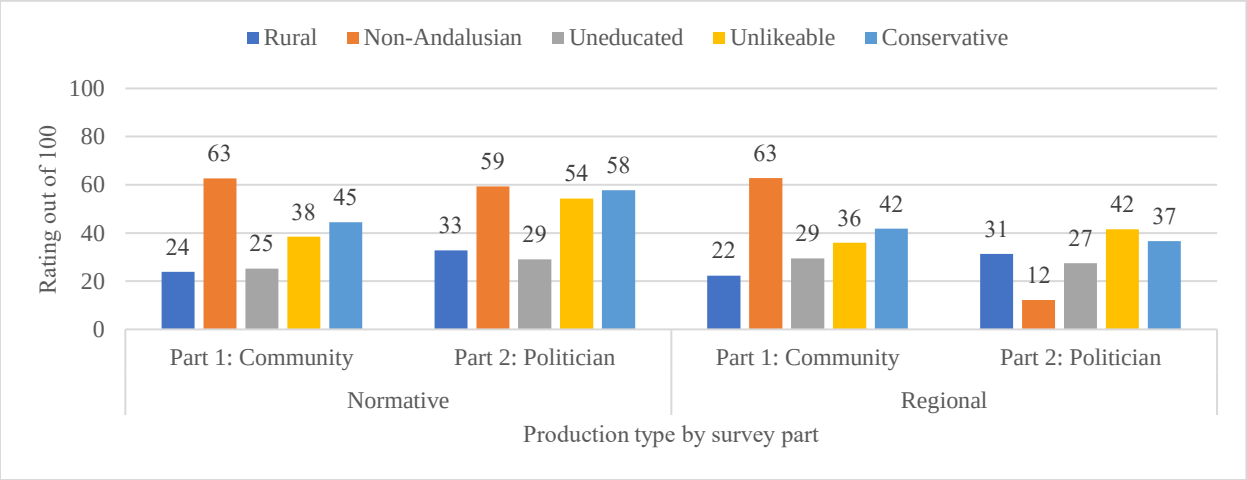

Figure 81: Affricate evaluation by survey section, production type, and adjective.

Next, for the retention and elision of coda /s/, Figure 82 shows a breakdown of evaluations by the gender of the participant, the type of production, and the adjective. Male listeners see the regional production as much more Andalusian and somewhat more socialist, unlikeable, and rural. Female listeners describe elision as more Andalusian, unlikeable, and rural.

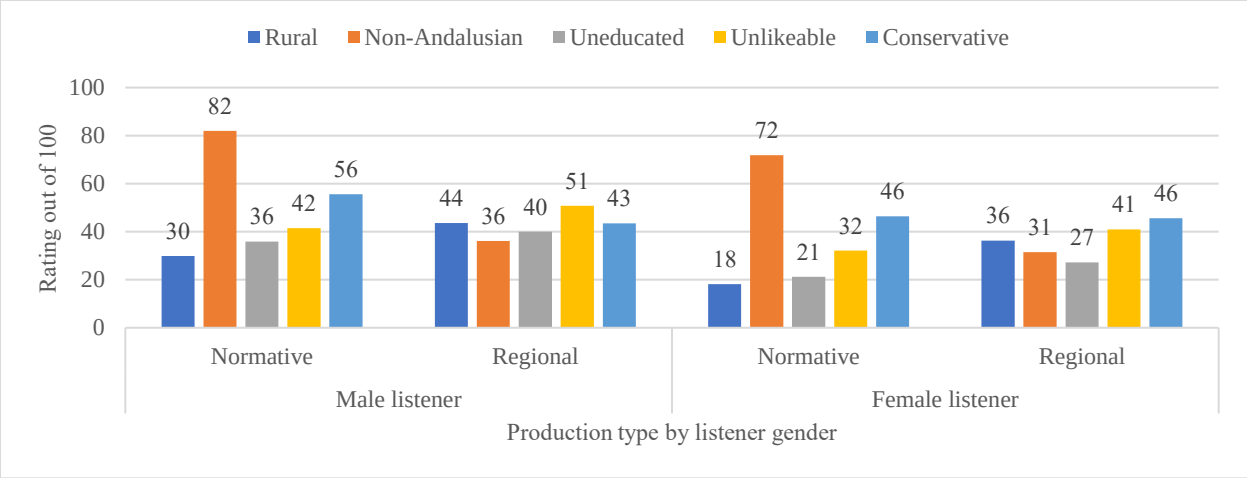

Figure 82: Coda /s/ evaluation by listener gender, production type, and adjective.

Figure 83 divides the results between the first and second part of the survey, production type, and adjective to show the evaluations of public speakers in contrast with community members for coda /s/. Both normative and regional productions were classified as more unlikeable and conservative when known to have been produced by politicians. However, results for the other three adjectives differed. Normative tokens produced by politicians were seen as less rural, more Andalusian, and more educated, while regional tokens known to come from politicians were seen as more rural, Andalusian, and uneducated as compared to speech seen as coming from members of their community.

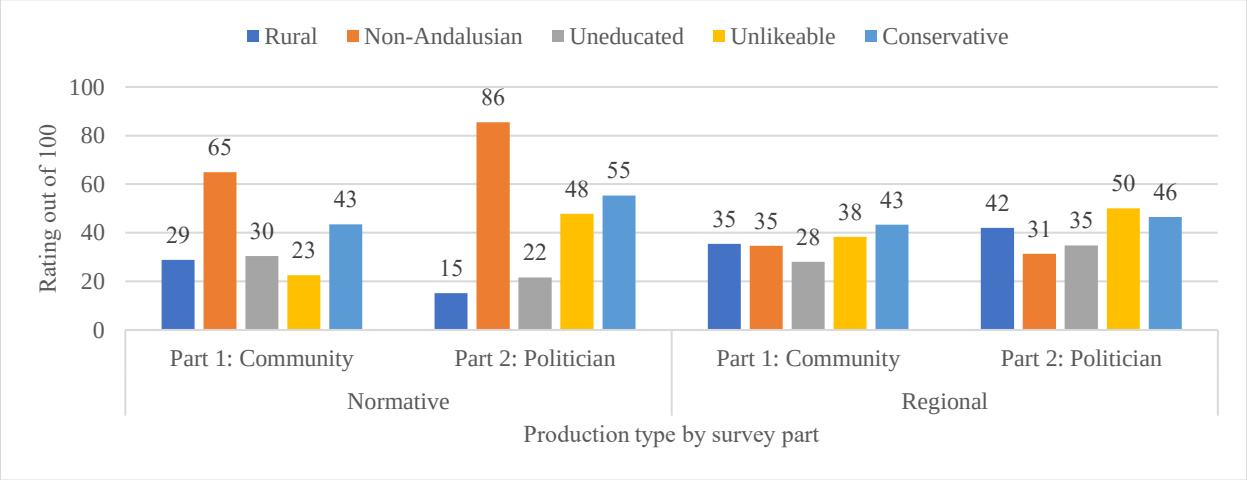

Figure 83: Coda /s/ evaluation by survey section, production type, and adjective.

Following that, for the production of /θ/ as [θ] or as *seseo* [s], Figure 84 shows a breakdown of evaluations by the gender of the participant, the type of production, and the adjective. Both male and female participants evaluated regional *seseo* as more Andalusian, educated, likeable, and conservative. However, men described it as more rural, while women did not.

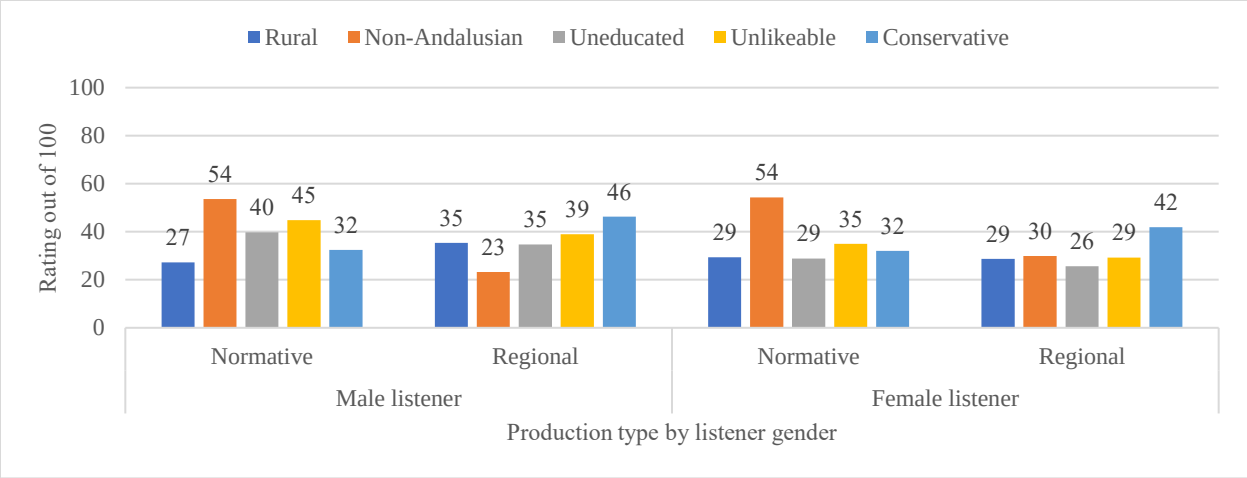

Figure 84: Seseo evaluation by listener gender, production type, and adjective.

Figure 85 provides insight into differences by survey part, production type, and adjective for *seseo*. Both normative and regional productions were classified as more Andalusian and unlikeable when known to have been produced by politicians. However, normative tokens produced by politicians were seen as more rural, uneducated, and socialist than those produced by community members, while regional tokens were see as less rural, more educated, and more conservative.

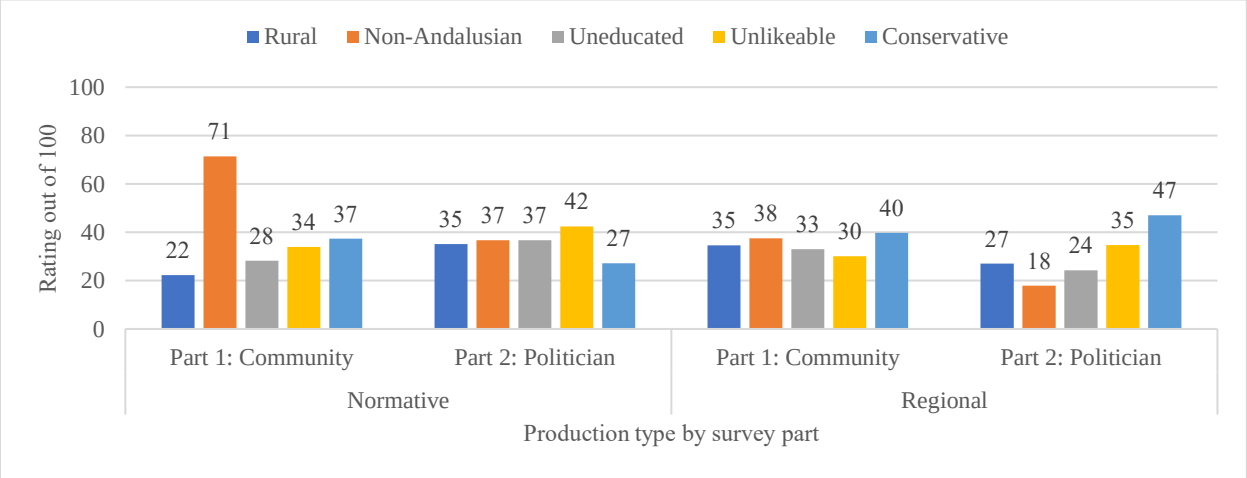

Figure 85: Seseo evaluation by survey section, production type, and adjective.

For the production of /s/ as [s] or as ceceo [θ], Figure 86 shows a breakdown of evaluations by the gender of the participant, the type of production, and the adjective. Both groups identified *ceceo* as more rural, Andalusian, uneducated, likeable, and slightly more tied to socialist identity than distinction.

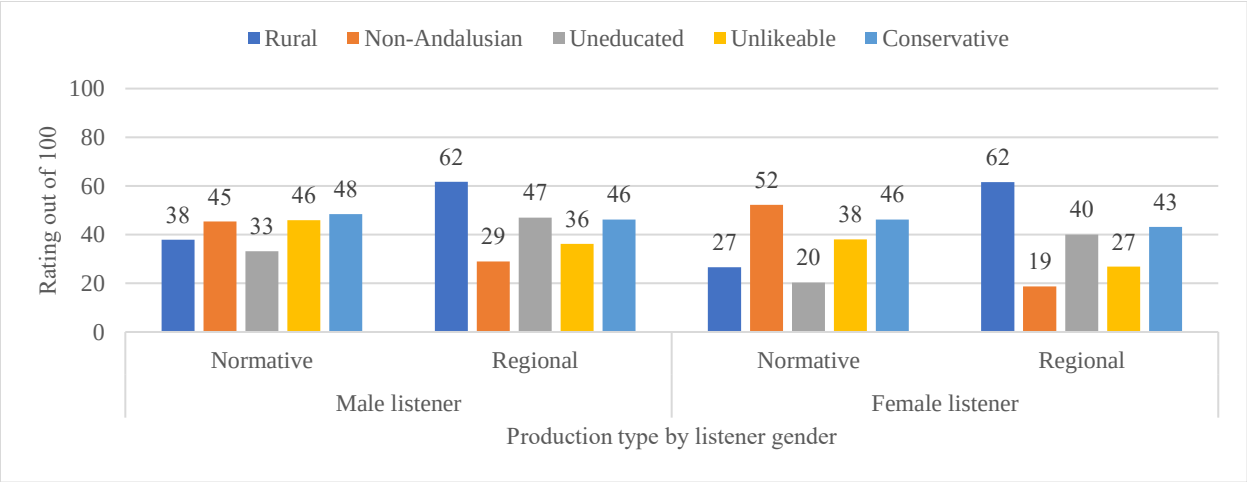

Figure 86: Ceceo evaluation by listener gender, production type, and adjective.

Figure 87 provides insight into differences by survey part, production type, and adjective for *ceceo*. Both normative and regional productions were classified as more educated and unlikeable when known to have been produced by politicians. Normative tokens were evaluated as slightly more rural, Andalusian, and conservative when produced by politicians compared to community speakers. Regional tokens of *ceceo*, in turn, were seen as less regional and Andalusian when produced by politicians. There was not a clear shift in political valence for regional productions across survey parts.

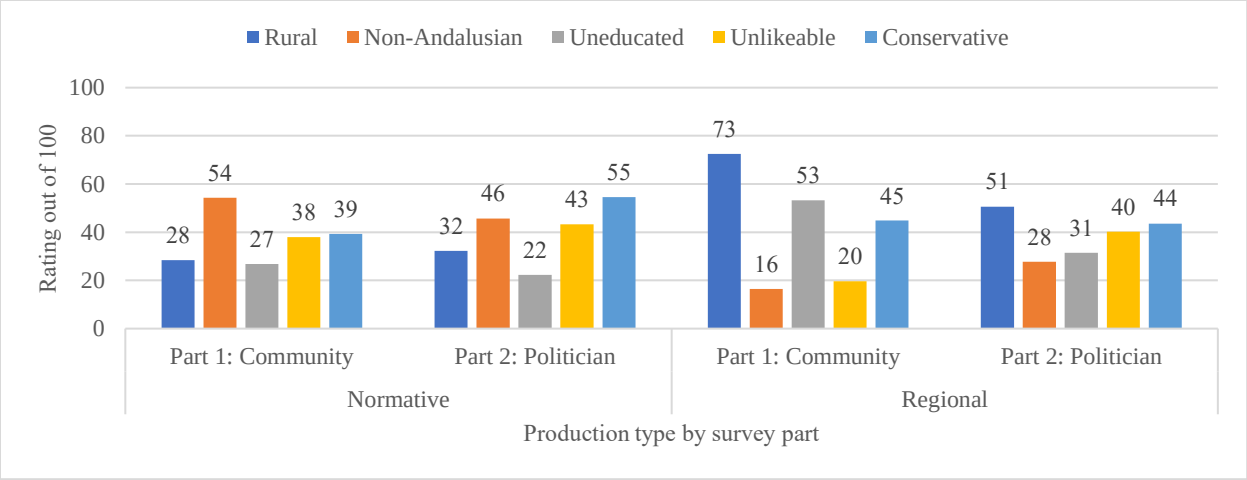

Figure 87: Ceceo evaluation by survey section, production type, and adjective.

Next, for the production and elision of intervocalic /d/, Figure 88 shows a breakdown of evaluations by the gender of the participant, the type of production, and the adjective. Both men and women evaluated intervocalic /d/ elision as sounding more rural, Andalusian, uneducated, likeable, and socialist than retention.

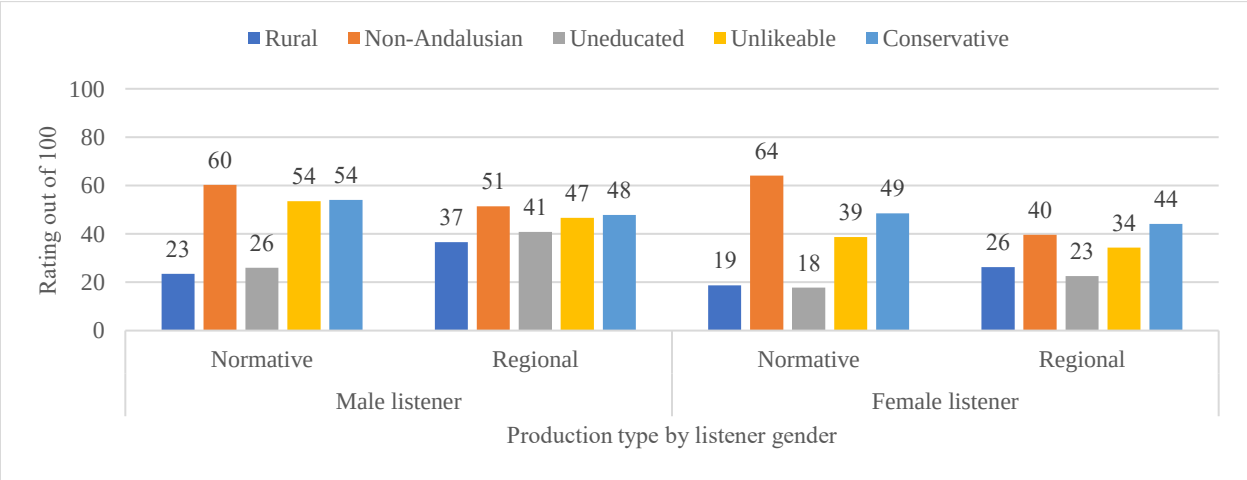

Figure 88: Intervocalic /d/ evaluation by listener gender, production type, and adjective.

Figure 89 provides insight into differences by survey part, production type, and adjective for the intervocalic /d/. Both normative and regional productions were classified as more Andalusian, likeable, and socialist when known to have been produced by politicians. The regional production was seen as more rural, while the normative one was perceived as slightly less rural, when produced by a known politicians over a community member. The reverse in likeability compared with other productions may support Cruz-Ortiz's (2019) claim, as well as that borne out in Stage 1 and 2 of the current dissertation, that intervocalic /d/ elision is an identity marker for political speech.

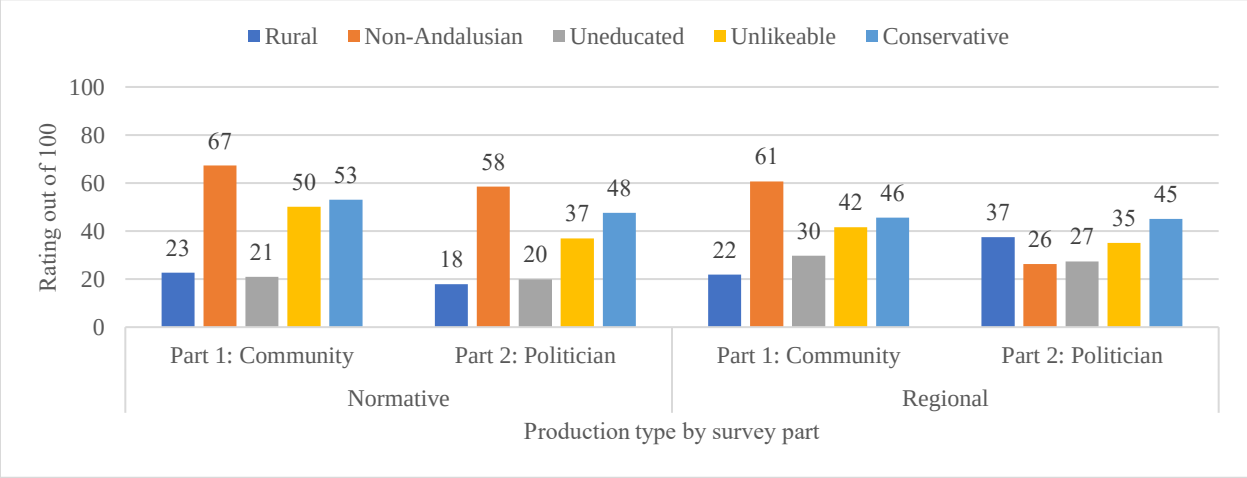

Figure 89: Intervocalic /d/ evaluation by survey section, production type, and adjective.

Following that, for the production and elision of /l/, Figure 90 shows a breakdown of evaluations by the gender of the participant, the type of production, and the adjective. Both men and women categorized lateral elision as more rural, Andalusian, and likeable than retention. Women were slightly more likely to classify elision as uneducated sounding, and neither group associated it strongly with a political ideology.

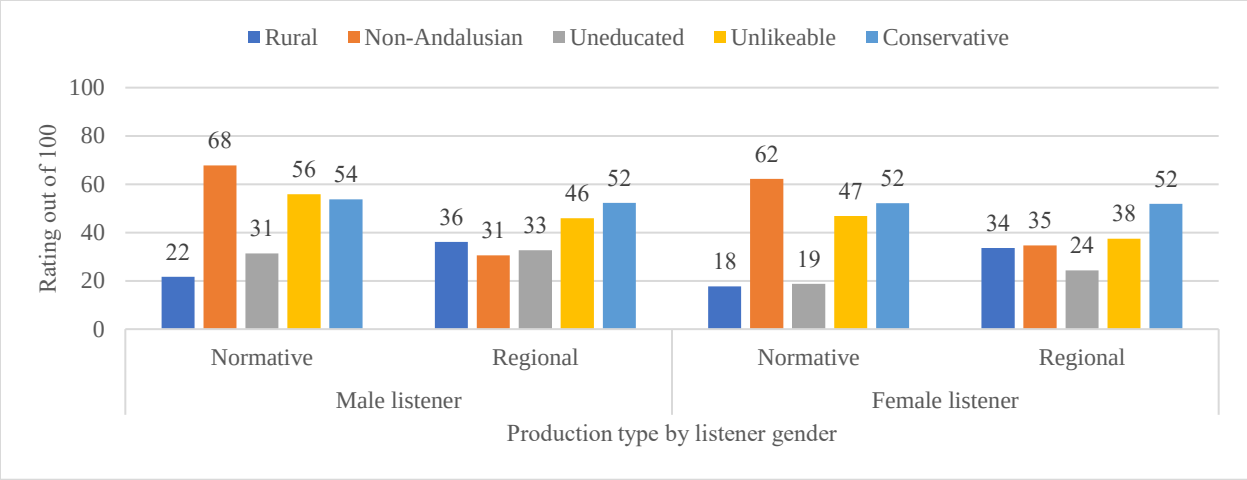

Figure 90: Lateral evaluation by listener gender, production type, and adjective.

Figure 91 offers insight into differences by survey part, production type, and adjective for the lateral. Both normative and regional productions were evaluated as more rural and Andalusian when known to have been produced by politicians. The normative production was seen as more likeable and more conservative when produced by politicians, while both of these categories remained unchanged for regional productions. Degree of education did not vary markedly for either production type.

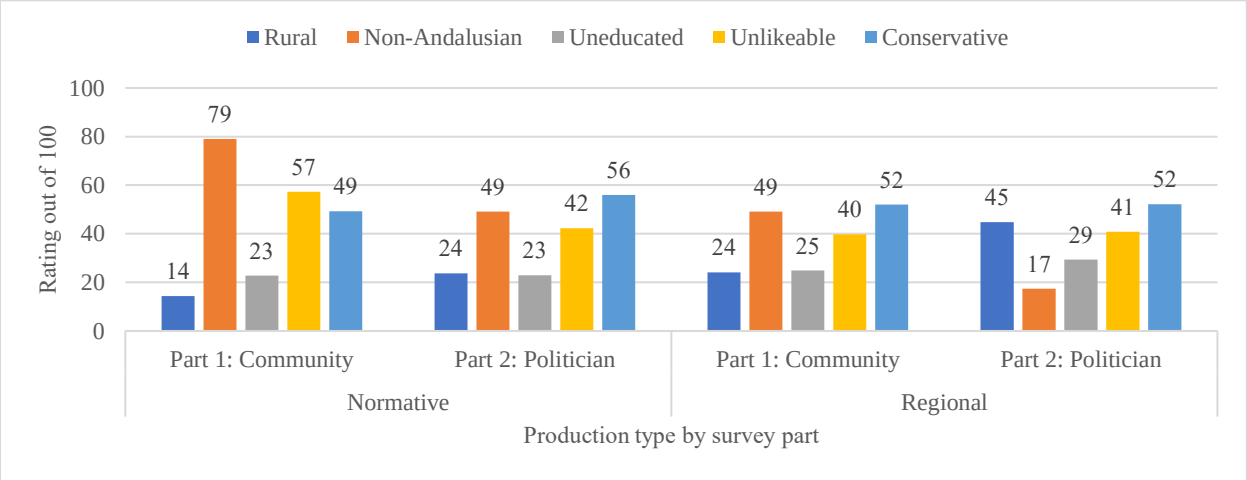

Figure 91: Lateral evaluation by survey section, production type, and adjective.

For the production and elision of /ɾ/, Figure 92 shows a breakdown of evaluations by the gender of the participant, the type of production, and the adjective. Participants of both genders perceived tap elision as more rural, Andalusian, likeable, and socialist. Interestingly, men were slightly more likely to describe elision as educated sounding as well.

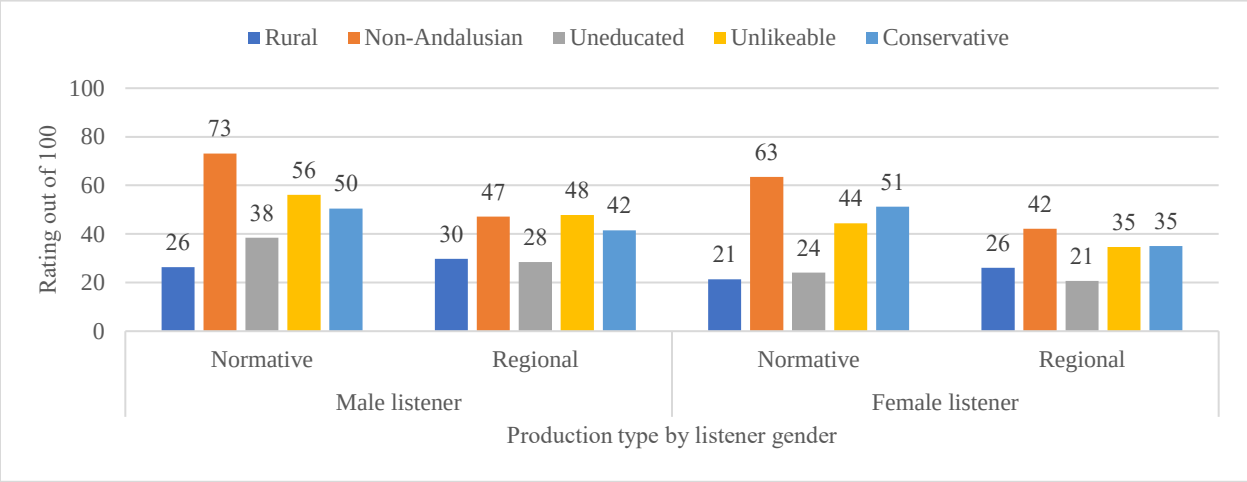

Figure 92: Tap evaluation by listener gender, production type, and adjective.

Figure 93 offers insight into differences by survey part, production type, and adjective for the tap. Both normative and regional productions were evaluated as less Andalusian and more unlikeable and conservative when known to have been produced by politicians. The normative production was seen as less rural and more educated in political speech, while the regional production remained unchanged regardless of the survey part.

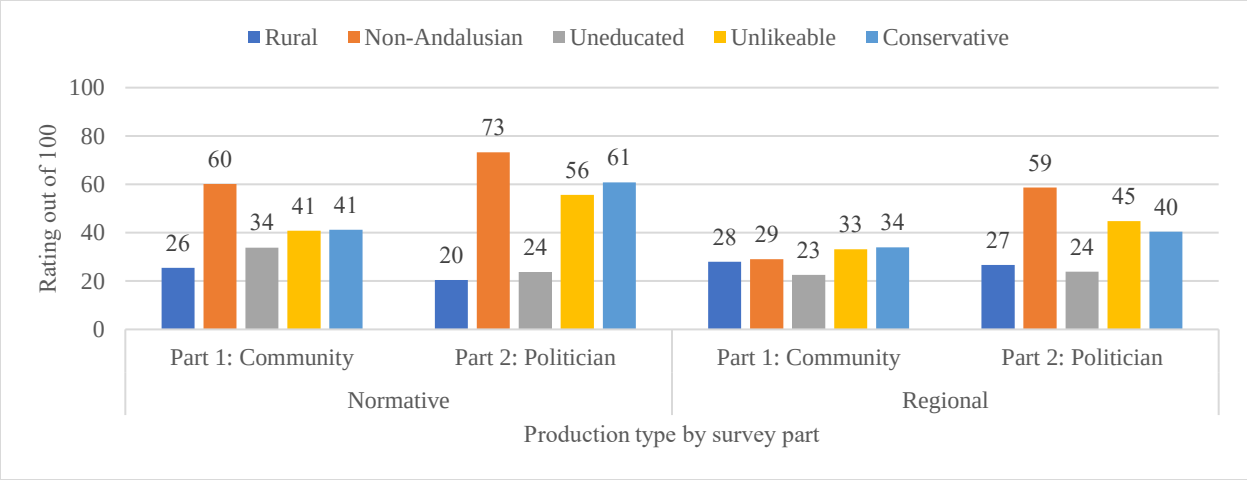

Figure 93: Tap evaluation by survey section, production type, and adjective.

Lastly, for the production and elision of /r/, Figure 94 shows a breakdown of evaluations by the gender of the participant, the type of production, and the adjective. Results for the trill were somewhat more uniform than for other phenomena. Men and women both described regional production as more Andalusian, educated, and likeable. However, men additionally identified regional production as more tied to socialist identity and rural, while women were slightly more likely to see the regional production as less rural.

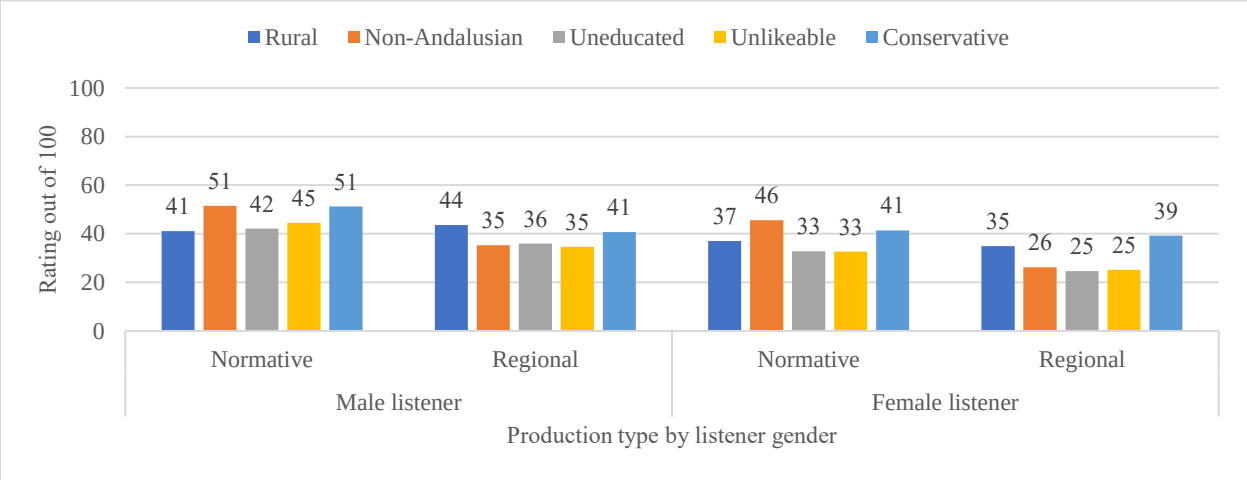

Figure 94: Trill evaluation by listener gender, production type, and adjective.

Figure 95 offers insight into differences by survey part, production type, and adjective for the trill. Both normative and regional productions were evaluated as less Andalusian, slightly more unlikeable, and more conservative when known to have been produced by politicians. Meanwhile, the normative production was described as much less rural and more educated when participants were describing political speech, while these adjectives remained unchanging for the regional production.

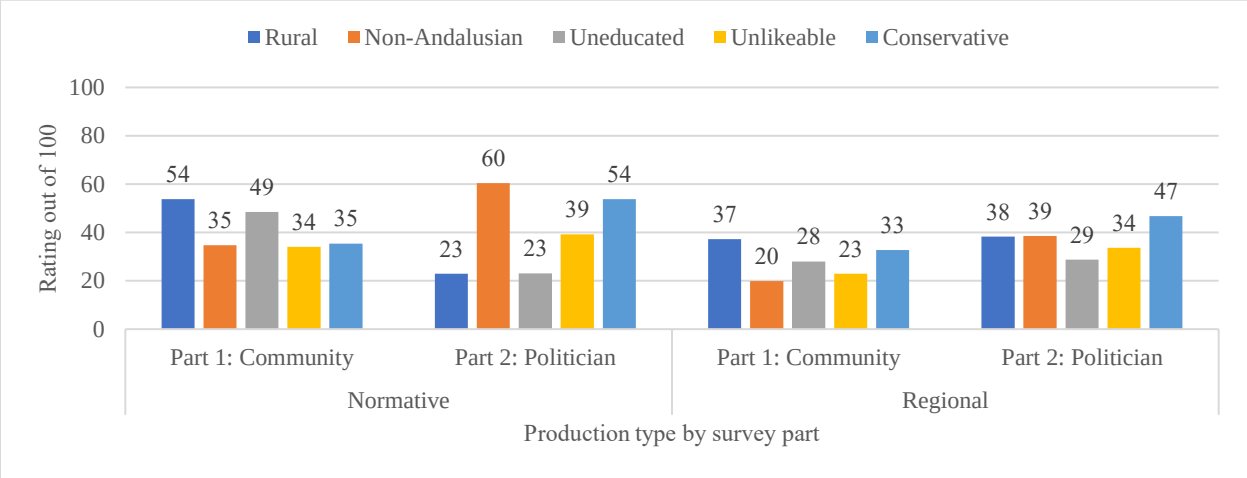

Figure 95: Trill evaluation by survey section, production type, and adjective.

In sum, 15,808 responses were analyzed across the three parts of the instrument, totaling 208 responses per informant. The following four sections break down these results using statistical models for each of the three parts of the instrument, then comparing those results comprehensively.

6.3.1 Perceptions of community speech

Part one presented listeners with eight regional and eight normative stimuli described as coming from peninsular Spanish speakers and asked them to evaluate these tokens along five adjectival spectra from one to 100.

Variable	factor	coefficient	tokens	mean
Region ID (p<0.001)				
	EAS	9	63	46.2
	Murcia	8.2	23	45.1
	WAS	5.53	366	42.0
	Don't know	2.147	140	35.7
	Canary Islands	1.55	25	33.5
	Other country	1.142	48	32.2
	Extremadura	-0.771	45	31.1
	Castilla-La Mancha	-4.483	53	25.2
	Catalonian	-10.307	51	19.8
	North Central Spain	-12.008	402	18.0
n=1216 df=13 Log-Likelihood=-5648 AIC=11322 R^2 Fixed=0.069 R^2 Total=0.416				

Table 37: Mixed effects linear regression for classification of tokens as urbano, 'urban,' in part 1 of the questionnaire, with Listener and Politician as random effects

The first adjective pair under consideration is urbano-rural 'urban-rural,' where ratings above 50 indicate an increased rural value, and those below fifty point to a greater urban meaning. The mixed-effects linear regression, which treated the listener and politician as random effects, identified one variable that conditioned variation, listener identification of token region (Table 37). Tokens perceived as Andalusian, Murcian, Canarian, or coming from another country were described as more rural, whereas those described as coming from northern provinces like Catalonia and North-Central Spain were identified as more urban.

While not a significant factor, Figure 96 distinguishes regional identification by the actual region of politicians to show differences in urban-rural identification. In Part 1, NCPS politicians

only produced normative variants, and differed markedly based on their perceived urbanity. Madrid politicians identified as being from the south of Spain were rated as much more rural than those perceived as coming from the north. This trend was less pronounced among Andalusian politicians regardless of the token they produced, with only a slightly increased tendency for regional variants to be coded as more rural. Instead, for the most part, this adjective seems to have relied on regional identification based on various acoustic correlates (whether accurate or not) as a means of showing perspectives toward differing varieties of Spanish.

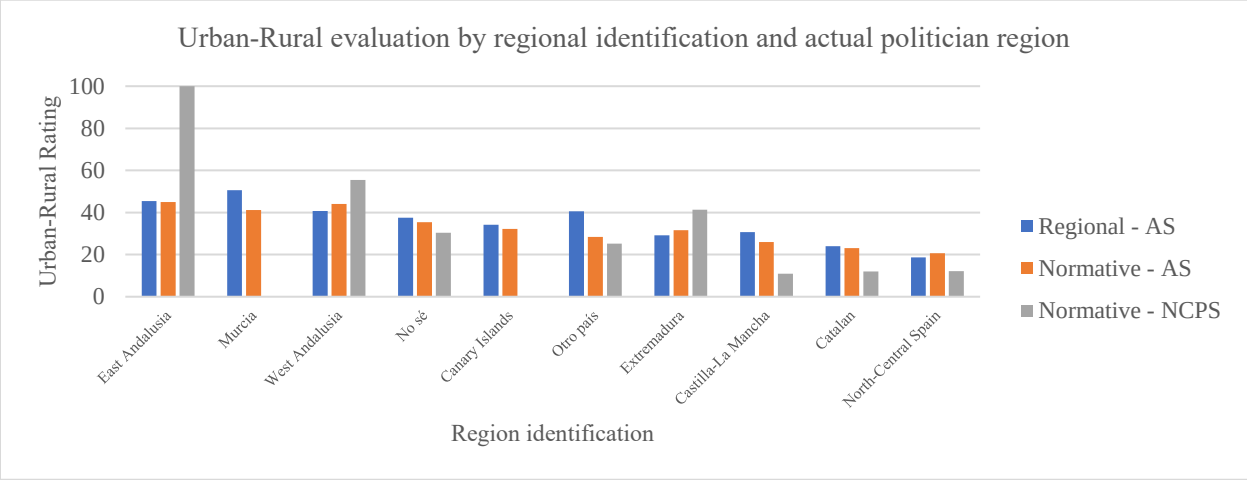

Figure 96: Evaluation of urban-rural adjectives by region ID and real politician region

The next adjective pair under consideration is andaluz-no andaluz 'Andalusian-Non-Andalusian,' where ratings above 50 correlate with Andalusian identity and those below it suggest association with a different variety. In this linear regression, three main effects and two interactions were identified as conditioning variation in the model (Table 38). As with the previous adjective pair, region identification was an important factor in describing variation, with northern regions receiving a much higher rating as not belonging to Andalusian Spanish, and only Extremadura,

EAS, and WAS being described as Andalusian. Following that, the factor of Politician Real Region reinforced these findings, with tokens from NCPS speakers being rated as much more non-Andalusian than those from AS speakers. Next, token regionalism was also descriptive of differences for Andalusian: normative productions were rated as non-Andalusian, whereas regional productions were rated more highly as Andalusian. Two additional variables, while not significant on their own (p>0.05, included in italics in the table), were included in significant interactions with token regionalism. In the interaction with political leaning, regional tokens were rated as more non-Andalusian by conservative listeners, while normative tokens were rated as more Andalusian. Then, in the interaction with listener gender, Men were most likely to code normative productions as non-Andalusian and regional productions as Andalusian, while women made less sharp a distinction between the two categories.

Variable	**factor**	**coefficient**	**tokens**	**mean**
Region ID (p<0.001)				
	North Central Spain	23.174	402	80.1
	Catalonian	20.524	51	78.6
	Other country	20.304	48	74.5
	Canary Islands	19.454	25	73.7
	Castilla-La Mancha	2.341	53	58.0
	Don't know	2.202	140	53.9
	Murcia	0.757	23	53.6
	Extremadura	-11.19	45	43.1
	EAS	-37.87	63	16.7
	WAS	-39.7	366	12.8
Politician Real Region (p=0.026)				
	Madrid	2.55	152	75.2
	Andalusia	-2.55	1064	46.7
Token Regionalism (p=0.002)				
	Normative	1.351	608	61.8
	Regional	-1.351	608	38.8
Listener Gender (p=0.488)				
	Men	*1.064*	*400*	*53.8*
	Women	*-1.064*	*816*	*48.6*
Political Leaning (p=0.610)				

	continuous	*coef*		
	+1	*-0.58*		
TokenRegionalism:Political Leaning (p=0.021)				
	factor:continuous	coef		
	Regional:+1	1.165		
	Normative:+1	-1.165		
TokenRegionalism:ListenerGender (p=0.004)				
	factor:factor	coef	tokens	mean
	Normative:Men	1.967	200	68.6
	Normative:Women	1.967	408	58.5
	Regional:Women	-1.967	408	38.7
	Regional:Men	-1.967	200	39.0
n=1216 df=19 Log-Likelihood=-5551 AIC=11140 R^2 Fixed=0.579 R^2 Total=0.664				

Table 38: Mixed effects linear regression for classification of tokens as andaluz, 'Andalusian,' in part 1 of the questionnaire, with Listener and Politician as random effects

The following adjective pair, *culta-inculta* 'educated-uneducated,' was oriented such that ratings above 50 correlated with uneducated classification and those below 50 were instead considered educated. This linear regression included three significant main effects to describe variation (Table 39). Region identification was selected, this time with southern identification correlating with slightly more uneducated speech, while northern identification correlated with more educated speech, although the overall spread was less drastic than for previous adjectives. Next, for listener gender, men rated tokens overall as less educated than women. Finally, for political leaning, more conservative individuals were more likely to rate tokens as less educated.

Variable	**factor**	**coefficient**	**tokens**	**mean**
Region ID (p<0.001)				
	EAS	7.298	63	37.0
	Extremadura	6.557	45	35.9
	Don't know	1.145	140	35.9
	WAS	0.723	366	33.4
	Murcia	0.664	23	33.3
	Other country	0.59	48	29.4
	Castilla-La Mancha	-0.97	53	28.2

	Catalonian	-4.436	51	26.8
	Canary Islands	-4.457	25	26.4
	North-Central Spain	-7.113	402	25.0
Listener Gender (p<0.001)				
	Men	5.599	400	38.0
	Women	-5.599	816	26.7
Political Leaning (p=0.007)				
	continuous	coef		
	+1	3.317		
n=1216 df=15 Log-Likelihood=-5418 AIC=10866 R^2 Fixed=0.090 R^2 Total=0.427				

Table 39: Mixed effects linear regression for classification of tokens as culto, 'educated,' in part 1 of the questionnaire, with Listener and Politician as random effects

While no interactions were significant in this model, Figure 97 offers some insight into the differences in coding based on the token regionalism and the phenomenon type. Male listeners rate all categories as more uneducated than female listeners, although both groups are more likely to perceive AS politicians' productions of normative liquid tokens as particularly uneducated. For both sibilants and intervocalic /d/, regional productions are seen as the least educated when produced by AS politicians.

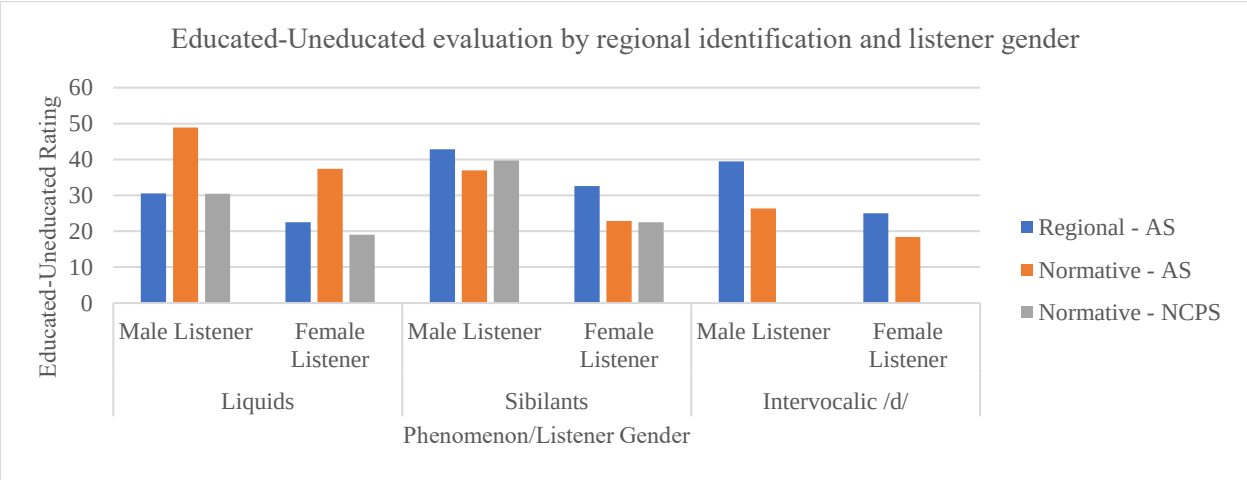

Figure 97: Evaluation of educated-uneducated adjectives by region ID and listener gender

The next adjective pair, *simpatico-antipático* 'likeable-unlikeable,' was organized so that evaluations above 50 correlate with the unlikeable description and those below are likeable. This linear regression included two significant main effects; region identification and listener gender (Table 40). In the case of the former, Andalusian varieties patterned together as being more likeable, while northern varieties were most unlikeable, with Catalonian Spanish breaking above 50 on average. Meanwhile, in the case of the latter, male listeners were more likely to rate tokens as unlikeable than female ones.

Variable	**factor**	**coefficient**	**tokens**	**mean**
Region ID (p<0.001)				
	Catalonian	11.767	51	53.0
	North Central Spain	8.744	402	46.1
	Murcia	7.249	23	42.2
	Don't know	2.369	140	40.3
	Extremadura	-2.27	45	34.5
	Castilla-La Mancha	-2.604	53	31.5
	Canary Islands	-5.088	25	30.5
	EAS	-5.562	63	27.4
	WAS	-6.206	366	25.4
	Other country	-8.399	48	20.0
Listener Gender (p<0.001)				
	Men	4.198	400	42.1
	Women	-4.198	816	33.1
n=1216 df=14 Log-Likelihood=-5536.8 AIC=11102 R^2 Fixed=0.093 R^2 Total=0.384				

Table 40: Mixed effects linear regression for classification of tokens as simpático, 'likeable', in part 1 of the questionnaire, with Listener and Politician as random effects

In order to better exemplify the relationship between listener gender and regional productions, Figure 98 offers a visualization of the data. While male listeners do rate all sub-categories as more uneducated than female listeners, this difference is more severe for intervocalic /d/ and normative productions of liquids by AS politicians. Interestingly, NCPS politicians'

productions of normative liquids are rated as sounding uneducated by both gendered groups, as are normative productions of intervocalic /d/ when rated by male listeners. Across all groups, regional productions by AS politicians are described as most likeable.

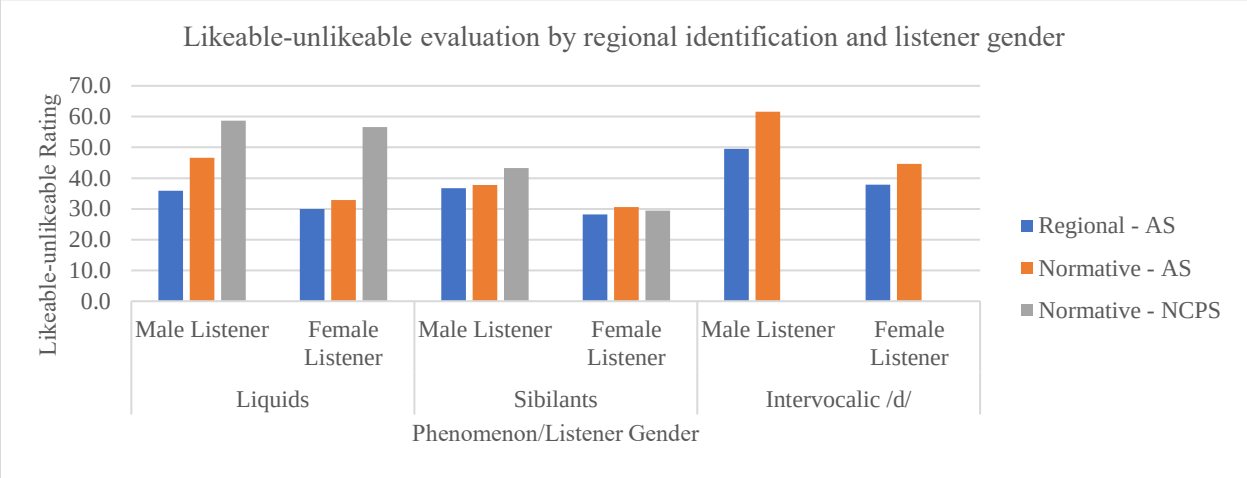

Figure 98: Evaluation of likeable-unlikeable adjectives by region ID and listener gender

Finally, the last adjective pair in Part 1 compared evaluations of *liberal-conservativo* 'liberal-conservative,' with ratings above 50 indicating conservative identification and those below 50 as liberal. This linear regression included two significant main effects and three interactions, with four additional non-significant main effects included in interactions (provided in italics in Table 41). The first fixed effect, region identification, showed that southern varieties of Spanish tended to be rated as more liberal than northern ones, although there is a degree of separation between WAS and EAS, with the former being perceived as more liberal. The next fixed-effect, phone, distinguishes between stop, sibilant, and liquid phenomena, with both productions of intervocalic /d/ receiving higher ratings as conservative than sibilants and liquids.

Of the four non-significant factors in the table, token regionalism forms an interaction with the remaining three. The first interaction, with listener gender, shows that men are most likely to rate regional tokens as more conservative, while women are most likely to rate them as liberal. The second interaction, with politician gender, shows that regional tokens produced by male politicians are also rated as most conservative, while those produced by female tokens are rated as most liberal. This suggests a degree of parallelism between speaker and listener that merits further discussion. Finally, the third interaction is with political leaning, in which listeners who identify as more conservative are more likely to rate regional tokens as more conservative and normative tokens as less conservative. This seems to equate to what was seen for Andalusian-Non-Andalusian, potentially suggesting a process by which conservatives are identifying regional speech as conservative, while also reducing the distinction between normative and regional production.

Variable	**factor**	**coefficient**	**tokens**	**mean**
Region ID (p<0.001)				
	North Central Spain	5.867	402	49.6
	Catalonian	4.317	51	46.5
	Extremadura	4.29	45	45.8
	Murcia	1.477	23	42.8
	EAS	0.777	63	41.8
	Don't know	-1.998	140	39.9
	Castilla-La Mancha	-2.206	53	37.6
	Canary Islands	-2.658	25	37.5
	WAS	-3.889	366	36.2
	Other country	-5.977	48	36.0
Phone (p=0.020)				
	Stop Phenomena	5.158	152	49.4
	Sibilant Phenomena	-2.261	608	41.8
	Liquid Phenomena	-2.898	456	40.7
Politician Gender (p=0.130)				
	Men	*2.291*	*380*	*45.7*
	Women	*-2.291*	*836*	*40.8*
Listener Gender (p=0.575)				
	Men	*0.873*	*400*	*43.7*
	Women	*-0.873*	*816*	*41.7*

Token Regionalism (p=0.500)				
	Normative	*4.645*	*608*	*43.0*
	Regional	*-4.645*	*608*	*41.7*
Political Leaning (p=0.610)				
	continuous	*coef*		
	+1	*-0.948*		
TokenRegionalism:ListenerGender (p=0.040)				
	Regional:Men	1.402	200	45.0
	Normative:Women	1.402	408	43.3
	Normative:Men	-1.402	200	42.3
	Regional:Women	-1.402	408	40.2
TokenRegionalism:PoliticianGender (p=0.001)				
	Regional:Men	3.298	152	48.4
	Normative:Men	3.298	228	43.9
	Normative:Women	-3.298	380	42.4
	Regional:Women	-3.298	456	39.5
TokenRegionalism:Political Leaning (p=0.006)				
	factor:continuous	coef		
	Regional:+1	1.395		
	Normative:+1	-1.395		
n=1216 df=22 Log-Likelihood=-5563 AIC=11169 R^2 Fixed=0.071 R^2 Total=0.265				

Table 41: Mixed effects linear regression for classification of tokens as Liberal in part 1 of the questionnaire, with Listener and Politician as random effects

6.3.2 Perceptions of political speech

The next part of the survey offered participants a photo, as well as the name and most recent position for each politician, although it left out the party affiliation by individuals. Despite having considerable identifying information, only about half of all PSOE politicians were identified as belonging to that party, with nearly 25% identified as belonging to the PP, and the rest balanced between two of the smaller centrist and left-leaning parties (i.e., *Ciudadanos* ‘citizens’ and *Podemos* ‘we can’) and a newer, vocal hard-right party (i.e., *Vox*). Meanwhile, under 50% of PP politicians were identified as belonging to that party, with a high number identified as being in the other four possible parties. This breakdown is provided in Figure 99. These results suggest that,

while there was only partial success in associating stimuli with accurate parties, listeners nonetheless began to pay more attention to party politics and affiliations when participating in this second part of the survey.

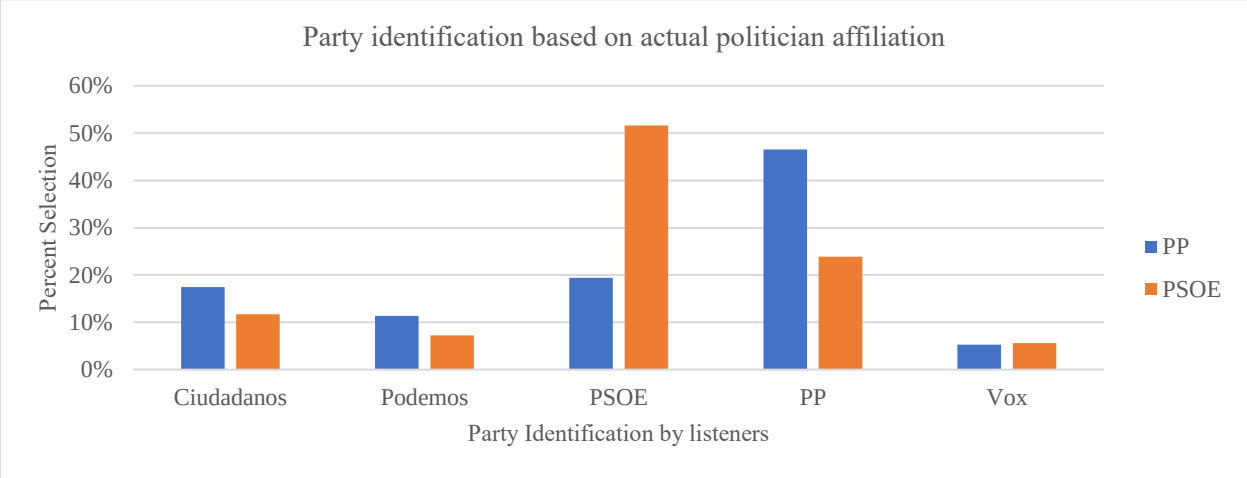

Figure 99: Listener identification of stimuli political parties compared with politicians' actual political affiliation (orange vs. blue)

The first logistic regression for part two looks at identification as urban-rural. Three main effects and one interaction were found to influence continuous ratings for this adjective pair, with one additional non-significant variable included in the interaction (Table 42). The first variable, region identification, shows a divide with AS and southern varieties being seen as more rural, while northern varieties are described as more urban. Next, for phenomenon, regional and normative sibilant phenomena are perceived as more rural than those with liquid and stop phenomena. Token regionalism is not significant but is included in italics for the interaction. Finally, the last main effect is political leaning, a continuous variable showing that listeners who identify as more conservative are more likely to rate all tokens as rural-sounding than liberals. The

interaction between token regionalism and political leaning shows that more conservative listeners tend to classify regional productions as having a greater rural valence and normative ones as more urban-sounding.

Variable	**factor**	**coefficient**	**tokens**	**mean**
Region ID (p<0.001)				
	Murcia	12.746	12	45.6
	Canary Islands	9.625	8	42.8
	Andalusia	4.296	532	39.4
	Extremadura	-3.679	29	36.4
	Other	-3.927	249	31.4
	Catalonia	-7.445	25	28.4
	Northern Central Spain	-11.617	361	17.9
Phenomenon (p=0.007)				
	Sibilant	5.149	608	33.3
	Liquid	-1.637	456	29.5
	Stop	-3.512	152	27.6
Token Regionalism (p=0.180)				
	Normative	*4.645*	*608*	*43.0*
	Regional	*-4.645*	*608*	*41.7*
Political Leaning (p=0.015)				
	continuous	coef		
	+1	3.483		
TokenRegionalism:Political Leaning (p<0.001)				
	factor:continuous	coef		
	Regional:+1	2.813		
	Normative:+1	-2.813		
n=1216 df=15 Log-Likelihood=-5607 AIC=11245 R^2 Fixed=0.130 R^2 Total=0.420				

Table 42: Mixed effects linear regression for classification of tokens as Urban in part 2 of the questionnaire, with Listener and Politician as random effects

The next model depicts the Andalusian-Non-Andalusian division, with four variables being selected as conditioning variation, as shown in Table 43. Listeners who coded stimuli as coming from northern Spain were more likely to code tokens as Andalusian, while only those stimuli coded as coming from AS were coded below 50 (i.e., as being Andalusian). This sharp contrast seems to suggest a stricter evaluation for politicians than community members. The next variable,

politicians' actual region, demonstrates that listeners were very accurate in identifying the origin of stimuli; while not quite as low an overall score as for region ID, actual Andalusian politicians were considerably more likely to be identified as Andalusian than those from Madrid. The third variable, listener gender, indicates that men were overall less likely to classify tokens as Andalusian than women by a small margin. Finally, token regionalism is the final variable selected in the model, with the indication that normative variants were relatively more likely to be seen as non-Andalusian, while regional ones had a high correlation with identification as Andalusian.

Variable	**factor**	**coefficient**	**tokens**	**mean**
Region ID (p<0.001)				
	Catalonia	26.078	25	88.3
	Canary Islands	23.973	8	79.4
	Northern Central Spain	14.794	361	78.1
	Extremadura	-4.976	29	53.3
	Murcia	-7.781	12	53.7
	Other	-8.215	249	51.5
	Andalusia	-43.873	532	13.2
Politician Region Real (p=0.018)				
	Madrid	4.215	228	72.7
	Andalusia	-4.215	988	37.0
Listener Gender (p=0.031)				
	Men	3.269	400	46.0
	Women	-3.269	816	42.5
Token Regionalism (p=0.008)				
	Normative	3.626	608	58.5
	Regional	-3.626	608	28.8
n=1216 df=13 Log-Likelihood=-5502 AIC=11029 R^2 Fixed=0.599 R^2 Total=0.690				

Table 43: Mixed effects linear regression for classification of tokens as Andalusian in part 2 of the questionnaire, with Listener and Politician as random effects

Although this interaction was not significant in the model, Figure 100 offers a visualization of the interaction between listener gender and token regionalism by phenomenon. Certain productions, including syllable-final /s/, the affricate, and the lateral show a clear distinction between regional and normative variants across listener genders. However, in cases like

intervocalic /d/ or the trill, male listeners identify a considerably reduced Andalusian distinction while the distinction for women is considerably greater. In fact, the trend overall seems to be that women are more likely than their male counterparts to rate regional tokens as more Andalusian, and normative tokens as less Andalusian. This trend is also identified in part 1 and follows other suggestions in descriptions of gendered identity performance and awareness of stigma among female speakers that is much more nuanced than that of men (e.g., Chappell, 2016; García, 2019).

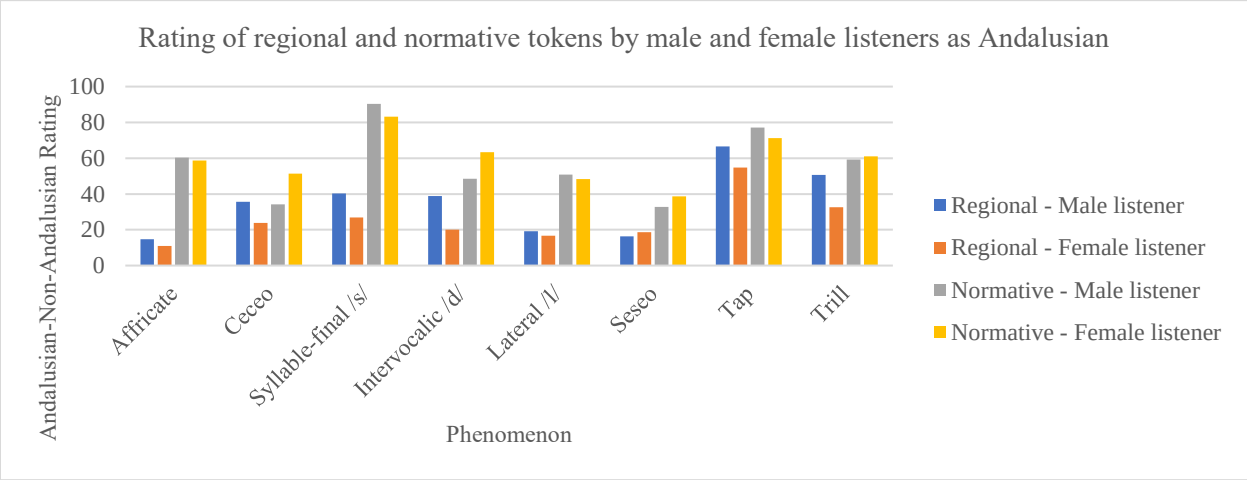

Figure 100: Listener identification of regional and normative tokens by politician gender and phenomenon.

The next adjective comparison is for educated-uneducated, for which three main effects and two interactions were chosen to describe variation, laid out in Table 44. The first variable was the listener gender, with men being more likely overall to describe tokens as uneducated compared to women. Next, for phenomenon, sibilant stimuli were more likely to be coded as uneducated compared to liquids and stops, a trend also identified for the urban-rural model. Following that, for region identification, there is a curious grouping here – while central and parts of southern

Spain were rated as most uneducated, Andalusian Spanish was medial, being more educated than these but less educated than Catalonian and NCPS. This suggests the same sort of negative self-perception of one's own variety that Hernández-Campoy and Cutillas-Espinosa (2010) discuss for Murcian. Next, three nonsignificant main-effect variables are included in italics due to their use in interactions: politician gender, token regionalism, and listener political leaning.

Following these fixed-effects, there were two interactions with token regionalism and the remaining two variables. First, with political leaning, regional productions were rated as more uneducated by more conservative listeners, while regional productions were rated as more educated. Next, for politician gender, male voices producing regional productions were ranked as the most uneducated, while male voices producing normative productions were ranked as the most educated, with female voices occupying the middle space. This is curious, based on the opposite results for listener gender discussed for the Andalusian adjective evaluation: while female participants are more severe in their ranking, here male politicians receive the more divisive evaluation.

Variable	**factor**	**coefficient**	**tokens**	**mean**
Listener Gender (p=0.005)				
	Men	5.387	400	34.2
	Women	-5.387	816	23
Phenomenon (p=0.019)				
	Sibilant	3.216	608	28.4
	Liquid	0.31	456	25.3
	Stop	-3.526	152	23.6
Region ID (p<0.023)				
	Canary Islands	10.952	8	41.1
	Murcia	5.136	12	39.4
	Extremadura	-1.06	29	33
	Other	-1.197	249	31.8
	Andalusia	-3.415	532	27
	Catalonia	-4.424	25	24.9
	Northern Central Spain	-5.993	361	21.5
Politician Gender (p=0.605)				
	Women	*0.718*	*304*	*26.9*

Men	*-0.718*	*912*	*26.6*
Token Regionalism (p=0.212)			
Regional	*1.983*	*608*	*28.4*
Normative	*-1.983*	*608*	*24.9*
Political Leaning (p=0.125)			
continuous	*coef*		
+1	*2.163*		
TokenRegionalism:Political Leaning (p=0.002)			
factor:continuous	coef		
Regional:+1	1.159		
Normative:+1	-1.159		
TokenRegionalism:PoliticianGender (p=0.001)			
Regional:Men	2.435	456	29.3
Normative:Women	2.435	152	28.2
Regional:Women	-2.435	152	25.7
Normative:Men	-2.435	456	23.8
n=1216 df=18 Log-Likelihood=-5236 AIC=10508 R^2 Fixed=0.097 R^2 Total=0.508			

Table 44: Mixed effects linear regression for classification of tokens as Educated in part 2 of the questionnaire, with Listener and Politician as random effects

Likeable-unlikeable represents the next adjective pair under consideration, with the model represented in Table 45 describing four variables that condition variation. The factor ordering for regional identification shows that northern varieties tend to be seen as more unlikeable, with Catalonian Spanish topping the list, while Andalusian Spanish is rated as the most likeable overall. Next, men are overall more likely to rate stimuli as unlikeable, while women were more likely to rate them as likeable. Following that, for party identification, politicians described by listeners as coming from the right-wing parties *Vox* and PP were described as producing the most unlikeable stimuli, while those identified as being from the PSOE (closely followed by the other two left-leaning parties) produced the most likeable tokens. Finally, listeners with a background in linguistics, meaning they had a career, a major, or a minor in the field, were more likely to rate tokens closer to 50, while those from other fields gave more likeable ratings.

Variable	factor	coefficient	tokens	mean
Region ID (p<0.001)				
	Catalonia	13.342	25	60.4
	Northern Central Spain	4.453	361	48.5
	Extremadura	3.014	29	46.3
	Canary Islands	0.011	8	43.0
	Other	-2.591	249	42.9
	Murcia	-7.129	12	41.6
	Andalusia	-11.1	532	37.6
Listener Gender (p<0.001)				
	Men	7.655	400	50.8
	Women	-7.655	816	38.7
Party ID (p=0.007)				
	Vox (R)	4.927	66	52.0
	PP (R)	0.888	428	46.0
	Ciudadanos (R)	-0.811	177	42.5
	Podemos (L)	-1	113	41.9
	PSOE (L)	-4.004	432	38.2
Informant Linguistics Background (p=0.016)				
	Linguistics	4.234	896	43.4
	Other	-4.234	320	40.7
n=1216 df=16 Log-Likelihood=-5517 AIC=11066 R^2 Fixed=0.117 R^2 Total=0.334				

Table 45: Mixed effects linear regression for classification of tokens as Likeable in part 2 of the questionnaire, with Listener and Politician as random effects

Finally, for the last adjective pair, liberal-conservative, the model had five fixed-effects and two interactions that described variable use. First, the factors for politician gender were ordered such that male voices were rated as more conservative overall while female voices were rated as more liberal. For political party identification, those stimuli that listeners described as coming from right-leaning politicians were also evaluated as much more conservative than those from ostensibly left-leaning voices. Next up, for regional identification, northern varieties were seen as being more conservative overall, as were central-southern varieties from Murcia and Extremadura, while only Andalusia and Catalonia fell below 50 on average and were described as liberal. The following variable, listener gender, had a similar division to that of the politicians, with male listeners coding tokens overall as more conservative and female listeners making more

liberal evaluations. For the last fixed effect, informant education, the more educated the listener, the higher the likelihood that they would rate tokens overall as conservative.

Two interactions were also included in the model, using fixed effects that did not reach the level of statistical significance (Table 46). The first, between token regionalism and informant age, shows that the older the participant, the more likely they are to rate normative tokens as conservative and regional tokens as liberal. Finally, the last interaction, between token regionalism and informant political activism, shows that the more active a listener identifies themselves in politics, the more likely they are to identify normative tokens as conservative and regional tokens as liberal. Both of these results suggest that greater experience with political speech leads to identification of regional variants with political affiliation. However, the fact that younger and less-active participants did not recognize this correlation could result either from a lack of familiarity, or – given that considerable work in peninsular political speech has focused on norms that are now one to two decades out of date (e.g., Hernández-Campoy & Cutillas-Espinosa, 2010; Cruz-Ortiz, 2019), unable to account for a rise of far-right populism in recent years – even changing speech norms in recent years, as the results in Chapter 4 suggest.

Variable	**factor**	**coefficient**	**tokens**	**mean**
Politician Gender (p<0.001)				
	Men	3.84	912	51.6
	Women	-3.84	304	37.9
Party ID (p<0.001)				
	Vox (R)	13.888	66	66.4
	PP (R)	11.486	428	61.5
	Ciudadanos (R)	-1.871	177	47.6
	PSOE (L)	-7.468	432	37.7
	Podemos (L)	-16.035	113	28.3
Region ID (p=0.009)				
	Canary Islands	8.253	8	57.9
	Northern Central Spain	5.575	361	54.9
	Murcia	1.531	12	54.5
	Extremadura	-2.437	29	54.1
	Andalusia	-3.533	532	45.4

	Catalonia	-3.885	25	44.1
	Other	-5.505	249	43.5
Listener Gender (p=0.041)				
	Men	3.254	400	52.4
	Women	-3.254	816	46.2
Informant Education (p=0.049)				
	continuous	coef		
	+1	2.95		
TokenRegionalism:InformantAge (p<0.001)				
	factor:continuous	coef		
	Normative:+1	0.373		
	Regional:+1	-0.373		
TokenRegionalism:InformantPoliticalActivism (p=0.015)				
	factor:continuous	coef		
	Normative:+1	1.149		
	Regional:+1	-1.149		
Informant Age (p=0.064)				
	continuous	*coef*		
	+1	*-0.487*		
Informant Political Activism (p=0.628)				
	continuous	*coef*		
	+1	*0.637*		
Token Regionalism (p=0.720)				
	Normative	*10.199*	*608*	*51.6*
	Regional	*-10.199*	*608*	*44.8*
n=1216 df=22 Log-Likelihood=-5563 AIC=11169 R^2 Fixed=0.071 R^2 Total=0.265				

Table 46: Mixed effects linear regression for classification of tokens as Liberal in part 2 of the questionnaire, with Listener and Politician as random effects

6.3.3 Comparisons of political speech

This last collection of statistical comparative models for part three of the instrument examined participants' binary responses for each of the adjectives. They identified whether they considered the regional or normative production to be more "adjective" (e.g., more Andalusian). This response was analyzed in mixed-effects logistic regression models, with selection of the regional production

as the application value, and both listener and politician still included as random effects to account for assumptions made by the model.

Figure 101 shows the overall selection for regional productions in each of the phenomena. For *ceceo, seseo,* syllable-final /s/, and lateral /l/, regional productions were selected in above 50% of contexts as being more urban and Andalusian. Meanwhile, intervocalic /d/, the affricate, the trill, and the tap saw normative productions selected as more urban and Andalusian. Intervocalic /d/ elision and affricate fronting received the highest identification as educated, followed by rhotic reductions and *seseo.* Regional productions of *ceceo, seseo,* lateral /l/, intervocalic /d/ and the affricate were seen as more likeable and liberal. Finally, in terms of overall preference (i.e., who would you vote for), regional productions of *ceceo, seseo,* syllable-final /s/, intervocalic /d/ and the affricate were all favored, although several of these are close to chance, while regional lateral and tap reduction were favored very infrequently.

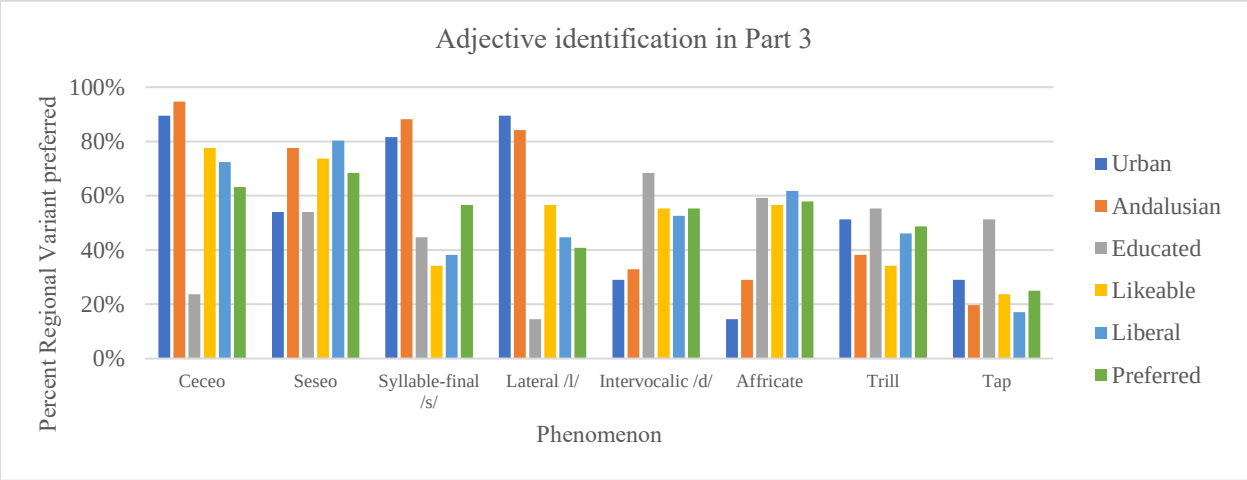

Figure 101: Binary identification by adjective and regional variant for part 3

First off, as shown in Table 47, the first regression examined identification of stimuli as the most urban. Six variables were able to describe variation for this response. The factor with the highest range in factor weight was the gender of the speaker who produced the normative token. When the normative production was uttered by a female speaker, listeners were much more likely to call the regional production urban. When the normative production was produced by men, the normative production was more likely to be described as urban. Second, for listener region, listeners who spoke WAS (i.e., either being from Seville or other surrounding provinces) called more regional tokens urban, while those from EAS were more likely to describe normative tokens as urban. Third, the regional identification of the normative token conditioned variation. Tokens identified as coming from the north of Spain or outside of Spain were seen as more urban overall, while those identified as coming from southern Spain or Andalusia were less frequently identified as urban.

Following that, for the phone, tokens with sibilants and liquids were overall more likely to have their regional production rated as urban, while normative stops were more frequently classified as urban. Finally, the last two variables were for the regional production. When the regional token was produced by a woman, it was more likely to be selected as urban, and for region identification, regional productions identified as coming from Andalusia and the south were more likely to be selected as urban. These results seem to suggest that "urban" may be associated with the urban variety of Seville, following geolinguistic descriptions of prestige that are more likely to emanate from nearer sources of linguistic standard than more distant national ones (e.g., Villena-Ponsoda, 2013).

Variable	**factor**	**log-odds**	**tokens**	**% regional selection**	**factor weight**
Normative Production Politician Gender ($p<0.001$)					

Women	1.245	456	61.6%	0.776
Men	-1.245	152	34.2%	0.224
Range				*55.2*
Listener Region (p=0.002)				
Seville	0.651	512	56.4%	0.657
West Andalusia	0.582	64	56.2%	0.641
East Andalusia	-1.233	32	25.0%	0.226
Range				*43.1*
Normative Production Region ID (p<0.001)				
North	0.922	224	78.6%	0.716
Elsewhere	0.417	166	56.0%	0.603
Andalusia	-0.577	184	29.3%	0.36
South	-0.762	34	29.4%	0.318
Range				*39.8*
Phone (p=0.012)				
Sibilant phenomena	0.812	304	59.9%	0.693
Liquid phenomena	0.029	228	56.6%	0.507
Stop phenomena	-0.841	76	28.9%	0.301
Range				*39.2*
Regional Production Politician Gender (p<0.001)				
Women	0.711	228	77.6%	0.671
Men	-0.711	380	41.1%	0.329
Range				*34.2*
Regional Production Region ID (p<0.001)				
Andalusia	0.594	256	77.3%	0.644
South	0.376	35	68.6%	0.593
Elsewhere	-0.419	138	45.7%	0.397
North	-0.552	179	26.8%	0.365
Range				*27.9*
n=608 df=15 Log-Likelihood=-280 AIC=590 R^2 Fixed=0.516 R^2 Total=0.526				

Table 47: Mixed effects logistic regression for comparisons of candidates as being more urban, with selection of the regional token as the application value, and Listener and Politician as random effects

The next model, for the most Andalusian-sounding token, included six variables that described variable responses from informants. This is laid out in Table 48. The first two factors deal with region identifications for both the regional and normative productions. When regional productions were identified as coming from politicians originating in Andalusia and the south of Spain, tokens were also identified as sounding more Andalusian. In contrast, when normative productions were perceived as southern or Andalusian, identification of regional tokens as

Andalusian fell. Two of the next variables reflect the gender of the politicians who produced each token, with normative productions by men causing the regional variant to be selected less often as Andalusian, and regional productions by women causing the regional variant to be selected more often as Andalusian. For phone, regional sibilant productions favored identification as Andalusian over normative ones, while normative liquid and stop phenomena were more often selected as Andalusian. Finally, the last selected variable was the answer to the question "who would you vote for." Regional variants identified as more Andalusian were more likely to be selected as listeners' preferred token overall, suggesting that "Andalusian-ness" plays a role in how voters connect to politicians.

Variable	**factor**	**log-odds**	**tokens**	**% regional selection**	**factor weight**
Regional Production Region ID (p<0.001)					
	Andalusia	1.72	256	89.1%	0.848
	South	0.007	35	57.1%	0.502
	Elsewhere	-0.717	138	49.3%	0.328
	North	-1.01	179	20.7%	0.267
	Range				*58.1*
Normative Production Region ID (p<0.001)					
	North	1.084	224	81.2%	0.747
	Elsewhere	0.532	166	66.3%	0.63
	South	-0.349	34	44.1%	0.414
	Andalusia	-1.268	184	25.0%	0.22
	Range				*52.7*
Normative Production Politician Gender (p<0.001)					
	Women	1.05	456	59.6%	0.741
	Men	-1.05	152	53.3%	0.259
	Range				*48.2*
Phone (p<0.001)					
	Sibilant phenomena	1.322	304	72.4%	0.789
	Liquid phenomena	-0.559	228	47.4%	0.364
	Stop phenomena	-0.762	76	32.9%	0.318
	Range				*47.1*
Regional Production Politician Gender (p<0.001)					
	Women	0.817	228	85.5%	0.694
	Men	-0.817	380	41.6%	0.306
	Range				*38.8*
Overall Preferred Token (p=0.008)					
	Regional	0.337	316	69.3%	0.583

	Normative	-0.337	292	45.9%	0.417
	Range				*16.6*
n=608 df=14 Log-Likelihood=-225 AIC=478 R^2 Fixed=0.666 R^2 Total=0.683					

Table 48: Mixed effects logistic regression for comparisons of candidates as being more Andalusian, with selection of the regional token as the application value, and Listener and Politician as random effects

Following that, the logistic regression for most educated token has four variables that were found to influence variation (Table 49). The first two reflect the regional identification of the two stimuli: regional productions identified as being from the north were more likely to be selected as educated, while normative productions identified as being from the south were less likely to be selected as educated. Following that, for the preferred token, listeners who would vote for a politician producing the regional stimuli were more likely to rate it as educated, while those who preferred the normative stimuli rated the regional token as less educated. Finally, for the gender of the speaker producing the regional variant, men were considerably more likely to see their tokens selected as educated than women.

Variable	**factor**	**log-odds**	**tokens**	**% regional selection**	**factor weight**
Normative Production Region ID (p<0.001)					
	South	0.701	34	67.6%	0.668
	Andalusia	0.411	184	59.2%	0.601
	Elsewhere	-0.546	166	44.0%	0.367
	North	-0.566	224	34.4%	0.362
	Range				*30.6*
Regional Production Region ID (p<0.001)					
	North	0.511	179	59.8%	0.625
	Elsewhere	0.284	138	51.4%	0.57
	Andalusia	-0.141	256	37.1%	0.465
	South	-0.654	35	25.7%	0.342
	Range				*28.3*
Overall Preferred Token (p=0.008)					
	Regional	0.511	316	54.1%	0.625
	Normative	-0.511	292	38.0%	0.375

	Range				*25.0*
Regional Production	Politician Gender (p<0.001)				
	Men	0.444	380	55.8%	0.609
	Women	-0.444	228	30.7%	0.391
	Range				*21.8*
n=608 df=11 Log-Likelihood=-365 AIC=753 R^2 Fixed=0.196 R^2 Total=0.239					

Table 49: Mixed effects logistic regression for comparisons of candidates as being more educated, with selection of the regional token as the application value, and Listener and Politician as random effects

The second to last model, which examined identification of tokens as likeable, was described using six variables and is displayed in Table 50. These included, first, the overall preferred token; listeners who would vote for the politician producing the regional token were more likely to rate them as likeable, while those who preferred the normative token were much less likely to rate the regional token as likeable. Next, for the regional token, both speaker gender and region identification were significant descriptors of variation; women producing the regional production were more likely to be selected as likeable, and regional variants produced by speakers perceived as coming specifically from Andalusia were most likely to be selected as likeable. The other three variables relate to information about the listener; first, men were more likely to select regional variants as likeable than women, who were around chance. Next, in two continuous variables, older participants were more likely to select regional variants as likeable, reflecting a generational difference. Meanwhile, those who identified as being more politically active were less likely to select regional variant as likeable, showing the effects of a regional persona on listeners more familiar with political speech.

Variable	**factor**	**log-odds**	**tokens**	**% regional selection**	**factor weight**
Overall Preferred Token (p<0.001)					

	Regional	1.167	316	75.6%	0.763
	Normative	-1.167	292	25.3%	0.237
	Range				*52.6*
Regional Production Politician Gender (p=0.032)					
	Women	0.538	228	69.3%	0.631
	Men	-0.538	380	40.8%	0.369
	Range				*26.2*
Regional Production Region ID (p=0.001)					
	Andalusia	0.49	256	66.4%	0.62
	South	0.426	35	45.7%	0.605
	Elsewhere	-0.399	138	46.4%	0.401
	North	-0.517	179	35.2%	0.374
	Range				*24.6*
Listener Gender (p=0.025)					
	Men	0.251	200	58.0%	0.563
	Women	-0.251	408	48.3%	0.437
	Range				*12.6*
Listener Age (p=0.010)					
	continuous	log-odds			
	+1	0.048			
Listener Political Activism (p=0.017)					
	continuous	log-odds			
	+1	-0.22			
n=608 df=11 Log-Likelihood=-303 AIC=628 R^2 Fixed=0.407 R^2 Total=0.448					

Table 50: Mixed effects logistic regression for comparisons of candidates as being more likeable, with selection of the regional token as the application value, and Listener and Politician as random effects

Finally, the last model represents tokens that were selected as sounding more liberal, with five variables representing variation in responses (Table 51). The first variable to condition variation was the overall preferred token; listeners who "voted" for the speaker producing the regional variant were more likely to rate the speaker as liberal, while those voting for the normative variant were less likely to have rated it as sounding liberal. The next two variables concern the normative variant. First, when these tokens were identified as coming from speakers in right-leaning political parties (i.e., *Vox*, PP), speakers were much more likely to rate the regional variant as liberal-sounding. Additionally, when the normative variant was produced by a male voice,

listeners were more likely to select the regional variant as sounding liberal, hinting at an association between regionalisms, socialist politics, and female politicians. The last two variables reference the regional production. When regional tokens were identified as coming from politicians in left-leaning parties, or when female voices produced them, they were more likely to be selected as liberal.

Variable	**factor**	**log-odds**	**tokens**	**% regional selection**	**factor weight**
Overall Preferred Token (p<0.001)					
	Regional	0.96	316	74.4%	0.723
	Normative	-0.96	292	27.1%	0.277
	Range				*44.6*
Normative Production Political Party ID (p<0.001)					
	Vox (R)	1.125	45	84.4%	0.755
	PP (R)	0.485	180	73.9%	0.619
	Ciudadanos (R)	0.022	86	53.5%	0.505
	PSOE (L)	-0.645	201	34.8%	0.344
	Podemos (L)	-0.987	96	28.1%	0.271
	Range				*48.4*
Normative Production Politician Gender (p=0.018)					
	Men	0.462	152	71.1%	0.614
	Women	-0.462	456	45.2%	0.386
	Range				*22.8*
Regional Production Political Party ID (p=0.035)					
	Podemos (L)	0.532	67	71.6%	0.63
	PSOE (L)	0.172	241	64.7%	0.543
	Ciudadanos (R)	0.139	76	52.6%	0.535
	PP (R)	-0.33	181	32.0%	0.418
	Vox (R)	-0.514	43	27.9%	0.374
	Range				*25.6*
Regional Production Politician Gender (p=0.018)					
	Women	0.326	228	65.8%	0.581
	Men	-0.326	380	43.2%	0.419
	Range				*16.2*
n=608 df=14 Log-Likelihood=-295 AIC=618 R^2 Fixed=0.444 R^2 Total=0.466					

Table 51: Mixed effects logistic regression for comparisons of candidates as being more liberal, with selection of the regional token as the application value, and Listener and Politician as random effects

All in all, these results suggest correlations between political affiliation, gender norms, regional belonging, and token evaluation. Before comparing these findings with previous theoretical literature discussing perceptual data and the link between production and perception, the following subsection provides a summation of the three parts of the instrument, as well as a final model examining the overall results to determine significant overall trends.

6.3.4 Overall regional perception

While each individual model in the previous three subsections treated ratings as a continuous dependent variable, a secondary approach was devised to determine how regional variants in each part of the instrument are influenced by adjective selection. For the four models in this section, the type of production was set as a binary dependent variable (i.e., regional vs. normative), and listener responses were treated as independent variables. This included region and party IDs, continuous ratings for each of the five adjectives, listener background information, and interactions between the continuous ratings and the region and party IDs.

The first model, provided in Table 52, was for the first part of the instrument. Only two variables were found to significantly explain differences between listeners' treatment of regional and normative variants. First, regional identification was found to distinguish the two token types most strongly – regional tokens were much more likely to be rated as coming from WAS, EAS, or southern and island varieties of Spanish, while normative tokens were more often rated as coming from the north. Even without the influence of political information, listeners have a strong ability to distinguish local and national speech norms. Second, the adjective pair Andalusian-Non-

Andalusian was the only continuous adjective measure to significantly track the difference between regional and normative tokens; the more a token was rated as non-Andalusian, the less likely it was to be a regional variant production.

Variable	factor	log-odds	tokens	% regional	factor weight
Region ID (p<0.001)					
	WAS	0.795	366	71%	0.689
	EAS	0.636	63	68%	0.654
	Canary Islands	0.523	25	64%	0.628
	Murcia	0.497	23	61%	0.622
	Extremadura	-0.081	45	49%	0.48
	Don't know	-0.21	140	44%	0.448
	Northern Central Spain	-0.337	402	36%	0.417
	Castilla-La Mancha	-0.486	53	36%	0.381
	Other country	-0.505	48	33%	0.376
	Catalonia	-0.831	51	26%	0.303
	Range				*38.6*
Andalusian (continuous; p<0.001)					
	continuous	log-odds			
	+1	-0.008			
n=1216 df=11 Log-Likelihood=-768 AIC=1557 R^2 =0.142					

Table 52: Mixed effects logistic regression for regional tokens in part 1, with listeners and politicians as random effects

For part 2, the model was found to have five fixed effects and one interaction between effects (Table 53). The first variable, regional identification, shows a difference between southern and northern Spain, with regional tokens being most likely identified as coming from the Canary Islands, followed by Extremadura and Andalusia; normative tokens were more often identified as coming from the north. Next, for party identification, regional variants were most often described as coming from the PSOE, although interestingly the hard-right party *Vox* was the second most likely identification for regional tokens (although it was selected much less often), suggesting a correlation between newer conservative rhetoric and regional speech. The following two variables are both continuous measures for the adjectives Andalusian-Non-Andalusian and Urban-Rural.

Tokens coded as less Andalusian were still less likely to be regional variants, while those coded as more rural were more likely to be regional variants, pointing to the correlation between AS, rurality, and covert prestige. Finally, the last variable is an interaction between the party identification and the adjective rating as Andalusian; tokens identified as being produced by left-leaning *Podemos* and the right-leaning PP and coded as less Andalusian were more likely to be regional tokens. Meanwhile, for other parties, coding as less Andalusian was more common for normative variants.

Variable	**factor**	**log-odds**	**tokens**	**% regional**	**factor weight**
Region ID (p<0.001)					
	Canary Islands	14.072	8	100%	0.999
	Extremadura	-0.82	29	69%	0.306
	Andalusia	-1.873	532	71%	0.133
	Other	-2.531	249	42%	0.074
	Northern Central Spain	-2.846	361	25%	0.055
	Murcia	-2.87	12	25%	0.054
	Catalonia	-3.133	25	20%	0.042
	Range				*95.7*
Party ID (p<0.001)					
	PSOE (L)	1.293	432	72.7%	0.785
	Vox (R)	0.688	66	51.5%	0.666
	Ciudadanos (R)	-0.063	177	40.1%	0.484
	PP (R)	-0.857	428	33.9%	0.298
	Podemos (L)	-1.061	113	38.9%	0.257
	Range				*52.8*
Andalusian (p<0.001)					
	continuous	log-odds			
	+1	-0.015			
Urban (p<0.001)					
	continuous	log-odds			
	+1	0.009			
PartyID:Andalusian (p=0.008)					
	factor:continuous	log-odds	vif		
	Podemos:+1	0.014	>20		
	PP:+1	0.007	>20		
	PSOE:+1	-0.003	>20		
	Ciudadanos:+1	-0.006	>20		
	Vox:+1	-0.013	>20		
n=1216 df=17 Log-Likelihood=-647 AIC=1328 R^2 Fixed=0.491 R^2 Total=0.510					

Table 53: Mixed effects logistic regression for classification of tokens as regional in part 2 of the questionnaire

Next, for the third part of the instrument, the arrangement differed slightly due to the categorical rather than continuous nature of the data. Rather than having regional variants as the dependent variable, it was set as case where regional stimuli were "preferred" (i.e., stimuli listeners would vote for). In this model, six variables were selected to describe participant responses, as shown in Table 54. Four variables reflect adjective selection. For classification as most likeable, liberal, educated, and urban, listeners who selected the regional variant as best exemplifying these adjectives (as compared to the normative variant) were also most likely to select the regional variant as being produced by a politician they would vote for. In addition, the regional identification for the normative production was significant in predicting variation; only normative variants perceived as Andalusian were preferred more often than regional variants. Finally, for political activism, listeners who rated themselves as more involved in politics were more likely to prefer regional variants overall than those who were less-involved in politics.

Variable	**factor**	**log-odds**	**tokens**	**% regional preferred**	**factor weight**
"Likeable" Regional Preference (p<0.001)					
	Regional	1.002	313	76.4%	0.732
	Normative	-1.002	295	26.1%	0.268
	Range				*46.4*
"Liberal" Regional Preference (p<0.001)					
	Regional	0.941	314	74.8%	0.719
	Normative	-0.941	294	27.6%	0.281
	Range				*43.8*
"Educated" Regional Preference (p<0.001)					
	Regional	0.88	282	60.6%	0.707
	Normative	-0.88	326	44.5%	0.293
	Range				*41.4*
Normative Region ID (p=0.002)					
	North	0.314	224	62.9%	0.578

	South	0.306	34	58.8%	0.576
	Elsewhere	0.084	166	57.2%	0.521
	Andalusia	-0.703	184	32.6%	0.331
	Range				*24.7*
"Urban" Regional Preference (p<0.017)					
	Regional	0.332	333	57.4%	0.582
	Normative	-0.332	275	45.5%	0.418
	Range				*16.4*
Listener Political Activism (p=0.005)					
	continuous	log-odds			
	+1	0.203			
n=608 df=11 Log-Likelihood=-267 AIC=557 R^2 Fixed=0.541 R^2 Total=0.547					

Table 54: Mixed effects logistic regression for classification of "voted for" candidates with those producing regionalisms as the application value for part 3 of the questionnaire

Finally, one last model, treating token regionalism as the dependent variable, compares only the regional identification and adjective ratings across all three parts of the survey (Table 55). In order to compare continuous data (i.e., Part 1 and 2) with binary data (i.e., Part 3), regional and normative productions of each phone were compared, with the stimuli then coded in a binary way comparable to Part 3. This final model includes four of the five variables under consideration. First, for Andalusian-Non-Andalusian, regional variants tended to be coded as more Andalusian across all three parts of the survey. Next, for region identification, those tokens identified as coming from Andalusian and southern speakers were more likely to be regional variants. Following that, for the urban-rural divide, tokens classified as rural were more likely to be regional productions, while normative variants were more likely to be classified as urban overall. Finally, for liberal-conservative, regional productions tended to be evaluated as more liberal, even though production in Stage 1 showed that conservatives actually use more regional tokens.

Variable	**factor**	**logodds**	**tokens**	**% regional**	**Factor Weight**
Andalusian-Non-Andalusian (p<0.001)					
	Andalusian	0.627	1967	66.4%	0.652
	Non-andalusian	-0.627	1681	30.8%	0.348

	Range				*30.4*
Region ID (p<0.001)					
	Andalusia	0.350	1401	66.8%	0.587
	South	0.289	211	55.9%	0.572
	Elsewhere	-0.241	741	43.2%	0.440
	North	-0.397	1295	34.7%	0.402
	Range				*18.5*
Urban-Rural (p<0.001)					
	Rural	0.319	1671	56.3%	0.579
	Urban	-0.319	1977	44.7%	0.421
	Range				*15.8*
Liberal-Conservative (p=0.013)					
	Liberal	0.091	1893	54.8%	0.523
	Conservative	-0.091	1755	44.8%	0.477
	Range				*4.6*
n=3648 df=7 Log-Likelihood=-2214 AIC=4441 R^2 Total=0.198					

Table 55: Mixed effects logistic regression for comparisons of regional and normative tokens across the entire instrument, with selection of the regional token as the application value

6.4 Discussion

The analysis of the three parts of the perceptual instrument demonstrates not only awareness of regional differences and gender norms on the part of the Andalusian listeners, but also evaluations influenced by stereotypes about political parties in southern Spain. Expectations differ across social groups – including between politicians and community members and across male and female speakers and listeners – and by evaluations of regional belonging of the speaker, regardless of their actual origin. The following sections return to the research questions presented at the onset of this chapter, examining how each part of the perceptual instrument contributes to our understanding of indexicality and social meaning in southern Spain for the eight regional phenomena considered in the study.

6.4.1 Community evaluations

The first question presented at the start of this chapter involved community evaluations of regional variants. Part one of the instrument best addressed this question, providing informants with no political biases and requesting that they assess tokens. Podesva et al. (2015: 74) recognize that receiving "information about speakers biases listeners," although there is also merit in understanding the role of that bias on perception, as they go on to acknowledge, and as part two of the perceptual instrument attempts to ascertain. However, for this first part of the analysis, the goal was to understand a baseline of how regional variation is perceived in Seville across the five adjective spectra.

Productions perceived as southern were seen as more rural, especially when coming from actual Madrid politicians, regardless of regional or normative production. This follows findings by Hernández-Campoy and Cutillas-Espinosa (2010: 300) for Murcia – they describe this variety of Spanish as being often associated with agriculture and ruralness, causing many locals to accommodate toward NCPS, especially in public speech, based on the higher prestige it possesses. As Villena-Ponsoda (2008) describes for Andalusia, while there has been relatively little work on phonological variation at the local and rural level, there are associations between informal, vernacular styles of speech and rural norms such as *heheo* and *ceceo*. At the same time, there is a trend in WAS (i.e., where the informants were located) in favor of greater social acceptance of Andalusian features, which encourages the embrace of supposedly rural features.

For the next adjective pair, normative stimuli, those produced by Madrid politicians, those evaluated by male listeners, and those perceived as northern were less likely to be seen as

Andalusian. This suggests that Seville listeners have a high accuracy in distinguishing conservative and innovative tokens by speaker given that Andalusian voices, regardless of the type of variant they produce, are much more likely to be perceived as Andalusian. There was also a curious interaction, in which more politically conservative listeners rated regional and normative tokens as less distinct using the adjective scale. Although Chapter 4 suggests it may not be the case overall, Hernández-Campoy and Cutillas-Espinosa (2010, 2013) argue that local variants are associated with working-class identity and allow politicians to show their connection to socialist ideas and the working-class community. In opposition, if normative features are a reflection of conservative ideas, and are used more frequently by right-leaning politicians, centrist and conservative listeners may be more likely to perceive normative variation as a part of Andalusian speech.

For evaluation of education, male and more conservative listeners tended to classify tokens as less educated, and stimuli identified as coming from Andalusia and southern Spain were described as less educated. There are connections here to the urban-rural division discussed previously, as well as the education-age correlation discussed by García-Amaya (2008), in which Spaniards educated in the post-Franco system had access to greater amounts of education, as well as more contact with NCPS norms. Especially given the skew in listeners toward college-aged students, it makes sense that speech identified as coming from NCPS would receive the highest education rating, as that would be the variety with the greatest prestige in higher education.

Following that, for likeability, stimuli identified as Andalusian were identified most highly, and women were more likely to describe all speech as likeable. Regional tokens were found to be most likeable across every phenomenon, while Madrid voices were seen as least likeable. This hints at the same division between innovative and conservative productions described by

Hernández-Campoy and Cutillas-Espinosa (2013), in that the female politician they examined was evaluated negatively by the media on scales of education and rurality, while at the same time being classified positively in terms of solidarity and accessibility. In this way, likeability emphasizes the distinction between overt prestige (e.g., education and urbanity) and covert prestige.

In terms of conservative identification, stimuli identified as northern, coming from male politicians, and intervocalic /d/ elision all received the highest identification as conservative, while male and more conservative listeners favored rating regional tokens as more conservative. This shows that intervocalic /d/, either elided or retained, is rated at the closest to chance, becoming so ubiquitous that it does not possess political meaning. It is curious to note the difference in interpretation between male and female voices – while female politicians using regional variants were seen as more liberal, males doing the same were seen as much more conservative, suggesting a bivalent interpretation of regional use that follows what Pollock & Wheeler (2022) found in Andalusia for the former president Susana Díaz. While listeners expect female socialists to use regionalisms, they also see it as being a norm for male conservatives.

These results reflect norms expected for typical members of the speech community. As Hernández-Campoy and Jiménez-Cano (2003) describe in their discussion of political voices on the radio, politicians follow community norms when employing sociolinguistic variation – however, they are not necessarily subject to the same criteria of evaluation. The contextual script that they follow in these media, Hernández-Campoy and Cutillas-Espinosa (2007: 149) argue, offers them a route to develop an appropriate identity, while at the same time providing the possibility of failure if they fall outside of the expected norms. Part two of the instrument determines which, if any variants, are seen as inappropriate in listener evaluations.

6.4.2 Political versus community evaluations

Following the identification of community norms, the next research question examined how public speech produced by politicians is evaluated by the same community. Based on a name, picture, and position, listeners evaluated speech and political affiliation. In their identification of political party, listeners were more accurate for socialist voices. About 70% of stimuli coming from PSOE politicians were identified as being produced by speakers from left-leaning parties, while only about 52% of those tokens produced by PP politicians were associated with right-leaning parties. Given that listeners selected from between five parties, these results are above chance, but this nonetheless suggests stereotypes and expectations about conservative speech that were not manifested in these tokens.

In their description of rurality, listeners still described southern varieties as more rural, although now productions with sibilant phenomena were also seen as more rural, and more conservative listeners were more likely to rate tokens, particularly regional ones, as more rural. This feeds into Hernández-Campoy and Cutillas-Espinosa's (2010) correlation between conservatism and "conservative" or normative variant production – political speech, in particular, has certain norms placed upon it, including an expectation of more urban features, and politicians (unlike community members) are seen as more rural by conservatives when using features perceived as rural.

Next, when evaluating Andalusian identity, listeners were still most likely to evaluate tokens perceived as coming from the south as most Andalusian, as well as identifying Andalusian voices and regional variants as considerably more Andalusian. Women were also still more likely

to classify all tokens as sounding more Andalusian. Unlike in the previous analysis, though, there were no interactions with token regionalism here, with listeners being more united by political leaning and gender in their evaluation of what sounds Andalusian.

For identification of education, there was a marked shift in the hierarchy based on regional identification: while community members seen as coming from Andalusia and southern Spain favored rating as "uneducated," politicians seen as coming from Andalusia more highly favored a classification as "educated," third only to Catalonian and NCPS. While non-political Andalusian speakers were seen as more uneducated, it seems more expected that educated politicians would use these variants. Other factors conditioned variation as well. Female listeners and stops favored identification as education, and there were interactions suggesting that more conservative listeners evaluated regionalisms as less educated, and that male voices producing regionalisms were seen as least educated, while male voices producing normative variants were described as the most educated sounding.

With respect to likeability, Andalusian-identified tokens, as well as tokens perceived as coming from left-leaning parties, were seen as most likeable. Female raters and non-linguists were also most likely to rate tokens as likeable. It is interesting to note that the overall ratings increased on all accounts by around 10 points, with community speech being perceived as generally more likeable than that produced by politicians, regardless of the context. This suggests that the differences in evaluation Hernández-Campoy and Cano (2003) describe for political speech, in part, comes with a degree of wariness. Anecdotally, several participants and locals, upon hearing about the focus of the project, described political speech as trying to "sell" something – this suspicion may translate into the evaluative context as well. Such a degree of suspicion seems to

push likeability closer to 50%, divesting these speakers of the benefit of the doubt offered to random voices from their community.

Finally, with respect to political affiliation, there was a similar bump of around 8 points overall: rather than all ratings falling below 50, being seen as at least partially "liberal," ratings of politicians straddled the line. Most voices were seen as conservative, while Andalusian and Catalonian identification correlated with liberal identification. Political party identification directly correlated with rating, while male voices were seen as considerably more conservative than female ones. Male listeners and more educated listeners were also more likely, overall, to rate tokens as more conservative. With respect to interactions, older participants and those more experienced in political activism were more likely to classify normative tokens as conservative, suggesting a changing approach to variation. These results suggest that conservatives in more recent years have begun to employ greater rates of regional variants, reflecting changing norms in evaluation, with those who have had greater experience with politics expecting a correlation that younger listeners no longer recognize.

These results suggest some of the changing evaluative norms applied to political and community speech. Political speech is expected to use regionalisms without sounding rural or uneducated in the same way as speakers from the community. However, at the same time, political speech is seen as less likeable and more politically polarized than community speech, in that politicians have specific goals for which they employ language, and are using language as a means of conveying political messaging. While political affiliation does influence the way that even community speech is perceived, suggesting it merits further use as a social factor to examine variation, it reveals particularly marked differences when applied to differences in political speech.

6.4.3 Success of identity work

The last research question in this chapter involves assigning political and social prestige to regional variant production, taking the comparative results from part three of the instrument. This section determined how community members distinguish between political voices, showing how identity work at the phonetic level correlates with regional production across social factors.

As a first step, we take a moment to examine the overall takeaway by phenomenon. Figure 102, pulling a single data group out of Figure 101, shows that overall preference falls off across phenomena. While regional *seseo* and *ceceo* productions and normative lateral and tap productions are more clearly preferred, the affricate, syllable-final /s/, intervocalic /d/, and reduced trill all fall very close to random chance selection. Given the production tendencies from Chapter 4, this is not surprising, given the frequency for the fronted affricate and elided intervocalic /d/ to occur across the peninsula, and the widespread nature of syllable-final /s/ and trill reduction in Andalusia – there is not a clear political affiliation with these sounds, since they are so widespread regionally or nationally. In the other cases, lateral and tap elision (often associated with EAS) is stigmatized, while *seseo* and *ceceo* (associated with Seville and WAS, respectively) both receive prestigious evaluations from listeners, suggesting an impact of geographical distance on perception (e.g., as Villena-Ponsoda & Ávila-Muñoz [2014] describe for Málaga, or Preston (1993) finds in the U.S. Midwest).

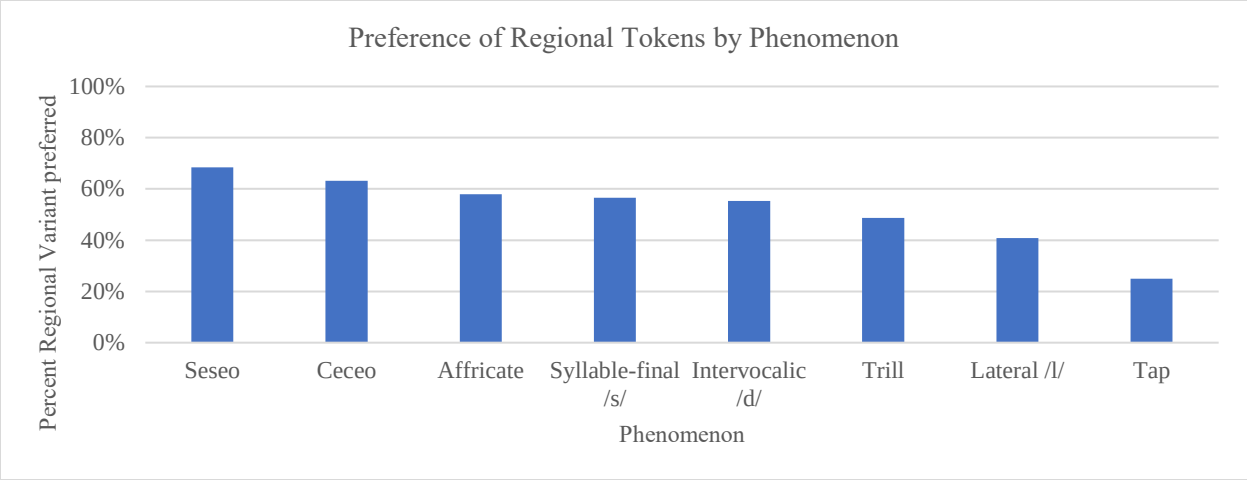

Figure 102: Listeners' preference for regional variants, by phenomenon, for part 3

The first model for adjectival preference considers urbanity, with listeners selecting the regional variant when the normative one was produced by a woman or identified as northern, when the regional one was produced by a woman or identified as southern, for sibilant phenomena, and when the listener themselves were from Seville or WAS. In addition to the influence of regional factors, and the overall preference for female voices as being more urban, it is worth emphasizing the difference between eastern and western AS. Villena-Ponsoda (2008) argues that there are separate regional processes ongoing in these two varieties that is causing WAS to embrace AS norms and dialectal differences, while EAS is moving toward a more unified koiné. This could preclude hypothetical EAS listeners from considering regional variants as urban, while Seville and WAS participants are less distinguishing in their understanding of urban versus rural varieties.

In the model for Andalusian identity, regional productions are favored when the regional variant is identified as southern or produced by a woman, when the normative variant is identified as northern or produced by a woman, and when the listener's preferred token was regional. These productions are also favored by sibilant phenomena. Thus, Andalusian identity plays a role in

Seville listeners' attitudes toward politicians, with a more favorable perspective toward those who are obviously southern peninsular (e.g., through sibilant phenomena and other acoustic measures that lead to regional identification).

Next, for education, regional variants were selected when the regional token was identified as northern or produced by a man, when the normative production was identified as southern, and when listeners' overall preference was the regional token. As with the results from part 2, male voices producing regional tokens are seen as less educated, and NCPS-identified tokens are prioritized as more educated-sounding. While listeners may consider a speaker less-educated, they are also more likely to vote for them, likely due to the regional belonging and other covert prestigious meanings associated with that manner of speaking (e.g., Andalusian identity, ruralness, etc.).

In terms of likeability, regional productions were preferred when listeners also preferred these tokens overall, as well as when the regional variant is produced by female voices or identified as Andalusian, and when the listener is male, older, and less experienced with politics. This is another case in which, although classified as less educated, rural tokens receive covert prestige as more likeable. Female voices, in particular, are more likeable when using regional variants, potentially a result of the left-leaning tendencies of the audience and the identification of male regionalisms as sounding more conservative, while female regionalisms are perceived as more liberal. As for age and activist experience, this result again suggests the possibility of changing political norms – older listeners are more likely to associate regional speech with socialist ideals, while younger ones and those more experienced with modern politics are less likely to identify a strong political association between variant selection and likeability or political party.

Lastly, for political affiliation, the regional variant was selected when listeners also preferred the variant overall, when the normative production was identified as associated with a male or right-leaning politician, and when the regional production was identified as associated with a female or left-leaning politician. Once again, female voices seem to correlate with socialist identity, whether using regional or normative variants, although there is also a marked trend for regional variation to correlate with leftist politics more broadly.

Based on the results from these comparisons, listeners may not be likely to classify regional variants as educated, but they are more likely to see them as rural, likeable, Andalusian, and liberal. Their ability to distinguish between regional and normative tokens is above chance based on these criteria, even though the actual region of the speaker is never a significant predictor of variation. Thus, instead of actually being able to identify the origin of words based on the single sound, they use they phonetic phenomena, alongside other acoustic correlates like intonation, vowel heights, durations, etc., to construct a consistent picture of who produces these tokens and what they mean, which then allows for the application of social evaluative criteria.

6.4.4 Final takeaways

The central goal of this chapter has been to address how regional and normative variants are perceived by members of the AS community, and how listeners' attitudes would influence regional variant use in identity work. Briefly, in this section, I return to the summarizing models in Section 6.3.4 to describe the main trends across the survey. First, with respect to part one of the instrument, regional tokens were more likely to receive classification as coming from southern speakers, and

to be identified as more Andalusian. Next, for part two, once listeners were aware of the political nature of stimuli, regional tokens were rated most frequently as coming from southern Spain, and as coming from the leftist establishment party PSOE or the far-right new party *Vox*. These tokens were also more likely to be identified as Andalusian and rural. For the comparisons in part three, regional tokens were preferred overall when described as likeable, liberal, educated, urban, when normative tokens were not seen as Andalusian, and when listeners were more politically active. Lastly, in the binary comparison across the survey, regional tokens were most often seen as Andalusian, rural, liberal, and were identified as coming from Andalusia or southern Spain.

These results paint a complex picture of variation, as listeners received increasing information and had to make political decisions. In the community at large, norms vary greatly across participants, although there is a strong tendency across participants to identify regional variant production as belonging to AS. Other than that, this group of participants does not share any overarching classification norm with respect to their adjective selection.

However, once the context of political speech is introduced, there is greater solidarity among responses. Listeners now see regional variants as markedly rural and Andalusian, while also correlating them with an older, established leftist party (i.e., PSOE) and a newer, far-right party that is less than a decade old (i.e., *Vox*). This suggests that *Vox*, among other political groups, may contribute to the growing correlation between conservative male speech and regional variants, in that its politicians are shaking up established linguistic expectations for political speech that has been in place for decades, as described by Hernández-Campoy and Cutillas-Espinosa (2013) and Ortiz-Cruz (2019).

When forced to select between two tokens taken from political speech, listeners are more likely to select regional variants as their preferred production type, and describe these in positive

ways. Given the major position of leftist politics in Andalusia until the 2019, it also makes sense that political experience might correlate with preference of regional variants, which have long been associated with socialist ideals. However, that reality seems to be the most in flux at the moment. There is a growing grassroots tendency embracing populism and regional identity – *Vox* became the third-largest Spanish party in 2019, only six years after it split out of the PP, and continues to see its numbers rise (Rama et al., 2021).

Turning back to the work of Podesva et al. (2015), there seems to be a much stronger and more consistent correlation between social meaning and production of several of these regional phenomena than those authors found for English politicians. Admittedly, those authors examined some of the biggest names in American politics, nearly a decade after their rise to the public eye, making individual opinions about these politicians a key part of the analysis and findings. The current chapter, on the other hand, did provide listeners with some information about politicians in part 2, but few of them were at the national level of visibility. The other two parts of the survey provided no such information, accessing political opinions without specific polarizing intuitions regarding individuals.

These results provide a methodology for future linguists, combining a composite of regional variants into a single perceptual study to determine how a host of sounds are interpreted within political speech. These results speak to the perceptual salience and social meaning of both specific variants within the community, as well as providing a holistic look at some of the key aspects of the Andalusian "accent" to determine how "sounding Andaluz" comes across to constituents in Seville. There is a clear preference for AS, particularly those norms most associated with WAS, and a general ambivalence toward features regarded as more national in their spread. In this way, this chapter helps to expand perceptual research methodologies into the southern

Iberian Peninsula, offering tools for future researchers of political speech to develop a clearer understanding of language variation and change. It also emphasizes the importance of political affiliation in understanding social meaning.

7 Final remarks

Over the course of this dissertation, political speech in Andalusian Spanish (AS) has been analyzed acoustically at the community level (Chapter 4), as well as acoustically at the individual level through Lectal Focusing in Interaction (LFI) analysis (Chapter 5), and auditively based on the evaluations of listeners from Seville (Chapter 6). The section returns to these findings, referencing some of the important takeaways that the data from each chapter presents not only on its own, but also as part of the larger picture of variation in AS political speech. The goal of this dissertation is to show the importance of combining differing methodologies. By doing this, we gain a richer understanding of variation beyond what a first-wave or third-wave sociolinguistic approach might offer on its own, including individual variation and style-shifting alongside community norms.

In the course of this chapter, that includes four discussions. First, in Section 7.1, the main contributions of the dissertation are discussed by chapter, concluding with a section that places all three chapters in dialogue with each other to complete the jigsaw puzzle of Andalusian variation. Following that, in Section 7.2, limitations are discussed for each of the individual chapters, as well as drawbacks to the composite approach. Third, Section 7.3 references future directions for each of the individual approaches in the dissertation, as well as ways to continue advancing the composite approach in political speech and beyond. Finally, Section 7.4 provides some concluding remarks to wrap up the project.

7.1 Contributions of the project

At this point, we return to a few overarching questions presented throughout the dissertation, both after the presentation of previous research, and in the discussions of guiding theoretical questions for each of the three investigative chapters. These questions also point to the gaps in our knowledge about political speech, style-shifting behavior, and composite methodologies that can be applied to gain a better understanding of these issues. Especially in the Spanish context, the application of third-wave sociolinguistic methodologies has only recently gained attention in the literature (e.g., Chappell & Barnes, in preparation; Delgado-Díaz, Galarza, & Díaz-Campos, 2021), offering ample opportunities for future researchers to connect pre-existing community-based research with novel approaches to style-shifting and identity constructive behavior.

The first research question, which guided Chapter 4, asked what variation existed for the community across the ten phenomena under consideration, both at the linguistic and social levels. Second, Chapter 5 examined which regional variables were employed in style-shifting behavior, and how this differs across social groups. Third, the question addressed by Chapter 6 was how regional and normative variants are perceived by members of the AS community, and how political affiliation influences linguistic evaluation. Finally, the last question, addressed across all three investigative chapters, has been how these findings can be related back to other research on political speech, identifying similarities that exist across language and geographical contexts. Through this combined analysis of sociolinguistic variation from various angles, this dissertation aims to offer a framework for future analysis of political speech that can allow for future comparisons across a diverse body of research.

These four questions are addressed, in light of the major investigative findings, in four sub-sections on community norms, LFI analysis, perception, and the composite approach.

7.1.1 Community norms

In the examination of community norms, three topics were examined in great detail: variable use, the identification of micro-social categories, and variation across lexical items. With respect to variable use, impressionistic and acoustic coding determined a wide range of factors influencing production, with regional speech being most favored with female interlocutors, on local channels, by AS speakers, and conservatives. Politician gender, age, and style-over-time were found to condition variation for certain phenomena. An important social category identified in this analysis was the political party, with female conservatives using extremely high regional production in interviews with women, and female socialists using higher rates in more "formal" scripted speeches. Conservative men were found to be relatively similar to socialist ones, although their use of the fronted affricate and elision of /d/ was more frequent, and in Northern Andalusian Spanish (NAS) and Northern Central Peninsular Spanish (NCPS) they tended to have greater regional production. Finally, with respect to lexical variation, terms like *militant* and *socialism* were closely associated with the left, while conservatives were much more likely to produce *Madrid* with regional variants, suggesting topical and frequency differences that may play a role in variation.

Based on these results, information is provided based on the phonetic phenomena under consideration in this study. Variants can be grouped into "Andalusian" and "national" categories based on their respective use. Affricate fronting, intervocalic /d/ elision, variable tap production, and variable trill production occur at similar rates across the varieties under consideration, although affricate fronting is more common in the north (i.e., NAS and NCPS), while variable tap

production is most frequent in the south (i.e., WAS and EAS). Meanwhile, syllable-final /s/, *seseo*, and word-final onset /s/ elision are most frequent across Andalusia, lateral reduction is rare but occurs most often in East Andalusian Spanish (EAS), and *ceceo* is essentially avoided in political speech even among Andalusians.

Variants can also be identified by their degree of prestige, based on traditional linguistic markers of social power described by Labov (1972). Given its association with female speakers and interlocutors and use in national contexts, the fronted affricate has the highest sociolinguistic prestige profile. Following that, intervocalic /d/, word-final onset /s/, and tap reduction are associated female interlocutors, suggesting a degree of prestige. Next, coda /s/, lateral, and trill reduction, as well as *seseo*, are associated with interviews with male interlocutors and male speakers, suggesting a reduced degree of prestige, as these contexts are often associated with the vernacular. Finally, neither vowel laxing nor *ceceo* had social factors associated with prestige due to the lack of social variables informing variation in the first case, and the rarity of the second. This suggests that laxing may be a sociolinguistic indicator (i.e., lacking social meaning and a prestige value), while *ceceo*, oft reported as a stigmatized variant, has too negative a value to appear in political speech.

The main contribution to the field of sociolinguistic knowledge provided by this investigative look at Andalusian political speech from the community level is the identification of variability and social trends among politicians. While some work has been done to assemble a sociolinguistic accounting of variation in southern and central Spain (e.g., Samper-Padilla, 2011:100), many of the projects referenced in Chapter 3 discuss variation at the level of individual cities and phenomena, often describing allophonic differences rather than the way that they are manifested in society (e.g., Dalbor, 1980; Moya-Corral, 1979). Based on production trends, we

see that in this community of practice, certain norms are shared across the peninsula, while others differ by social context, city, and speaker.

7.1.2 LFI analysis

While performing the analysis of LFI among ten speakers in specific speech contexts, variation was assessed via three measures: at the level of regional variant, individual style-shifting patterns, and social factors. When considering regional variants, it was clear that all seven phenomena considered in the chapter contributed to regional peaks in variation and that, as Sharma (2018) found for the public speaker Fareed Zakaria, 'real me' moments provide insight into speakers' vernacular, revealing the forms that they use in their first-learned dialect. The most "Andalusian" phenomena identified in Chapter 4 tended to be used by regional politicians in moments of lectal focusing, including *seseo*, and coda /s/ and lateral reduction, while nationwide phenomena were present across speakers, but varied based on frequency in conversation (i.e., affricate fronting and variable trill production) and individual factors including gender (i.e., tap and intervocalic /d/ reduction). In terms of style-shifting patterns, individual differences were visualized based on regional proportions by phenomena, and patterns were compared by variant type based on social factors like party, gender, and city to distinguish norms. Finally, with relation to broader social categories, political party was found to be particularly important, showing different approaches based on the relative sincerity and performativity of the topics where regional variants were employed.

The discussion of individual patterns in LFI analysis serves as one method through which this dissertation attempts to bridge qualitative and quantitative approaches. It also combines first- and third-wave perspectives, looking both at the ways that individual speakers differ over time, while also comparing these patterns in a way that allows for a better understanding of individual variation. Cruz-Ortiz (2019) and Pollock (2023) describe the importance of individual variation in understanding the use of regional features among Andalusian politicians, presenting an opposition to the controlled contexts and situations examined by Hall-Lew et al. (2010, 2017) in the U.S. and U.K. where individuals follow party politics.

While certain moments may present politicians with a very clear script of norms (e.g., produce <Iraq> like the other members of your party, speak like a Scot rather than a Brit), and other politicians may have clear ideas about the communities that they represent (e.g., Hernández-Campoy & Cutillas-Espinosa's [2013] María Martínez's connection to working class Murcia), many situations are murkier, and politicians are faced with a variety of goals. As the LFI analysis shows, from moment to moment, there are reasons to present regional features that involve divulging secrets, telling jokes, or indexing friendliness. All of these moments require differing tactics, but have at their heart similar core factors that politicians consider as they juggle norms of political speech and identity-based goals.

The LFI results suggest that norms may be changing. In an examination of Andalusian political speech, Cruz Ortiz (2019: 310-311) finds that the political affiliation of the national government does not affect the use of regional features in speech. Despite recent findings by Regan (2017b) and others regarding the encroachment of the NCPS variety into Andalusia, Cruz-Ortiz's historical approach presents data that indicates a more cyclical reality. She argues that there have

long been swings between regional and normative variants in politics, regardless of the affiliation of the governing party.

Based on descriptions indicating a correlation between progressive politics and rural and working-class linguistic norms in southern Spain, such as that which Hernández-Campoy and Cutillas-Espinosa (2010, 2013) find for 1990s Murcian politics, there seems to be a change underway. The LFI analysis presents a different picture: conservative politicians produce more regional variants, although the contexts where they do so often appear performative, using southern features to indicate a lack of familiarity with big-city norms, or to provide cautionary advice for attendees at a local fair. Meanwhile, when left-leaning politicians use the same features, they seem to do so in the way Sharma (2018) describes as a 'real me' moment, such as when discussing suffering single mothers, reducing social distance through humor, or describing public works underway. Coupled with the community examination in Chapter 4 – which shows that conservatives produce more regional variants overall, as compared with socialist politicians – this suggests a changing landscape of variable usage, which is further supported by the perceptual results.

7.1.3 Perception

In the perceptual study based on evaluations from Seville respondents, three main questions guided the discussion of results. These questions dealt with evaluations of both speakers in the community and politicians, and probed the relative success of identity work. In terms of the first question, geared toward community evaluations, informants were accurately able to associate regional

speech with southern identity, while also describing it as rural, uneducated, more likeable, liberal when produced by female voices, and conservative when produced by male voices. In terms of the evaluation of political voices, listeners were most likely to identify all speech as coming from PSOE politicians, although there was also reasonable accuracy for PP politicians. Politicians using regional variants were described as more rural by conservatives, more educated than community members, more likeable (especially when liberal, although less likeable than community members), and more liberal. There were also curious interactions in which increased age and experience correlated normative variants with conservatism. Finally, with respect to identity work, the southern-coded *seseo* and *ceceo* receive the highest support from listeners, while liquid phenomena (associated with EAS) were rated poorly by WAS participants. Additionally, when forced to choose between two speakers, politicians producing regionalisms were not seen as educated, but were seen as rural, likeable, Andalusian, and liberal.

Importantly, as is discussed in Chapter 6, these investigative results contribute to our understanding of linguistic perception by emphasizing how successful listeners are at using acoustic correlates to distinguish their own variety from that of others, as Moreno Fernández and Moreno Fernández (1999), Díaz-Campos and Navarro-Galisteo (2009), and Díaz-Campos and Killam (2017) have found for speakers from the peninsula and South America. While regional features clearly helped participants to identify politicians' home cities, the tendency for NCPS speakers to receive different norms of classification than AS speakers shows that additional acoustic correlates are used to evaluate speech, including intonation and other segmental phenomena not under consideration in the current study.

Personal and individual preference plays a role in perception as well as production. As Podesva et al. (2015) found for released /t/ in U.S. politicians' speech, one variable alone is not

sufficient to explain attitudinal responses. This is reinforced by the LFI findings, in that a combination of variants are necessary for "regional" speech to reach a noticeable peak at any given moment, meaning that no single speech phenomenon can explain perception. However, by taking into consideration several, including in particular the sibilant and intervocalic /d/ phenomena – which were evaluated more positively by listeners who spoke WAS – we can understand features of importance to Seville voters, particularly those associated with WAS.

The changing in production norms from Chapters 4 and 5 are echoed by the perceptual results in Chapter 6 with respect to political party. In the same way that conservatives make use of comparable or greater amounts of regional variation than socialists, participants in the survey also tended to show divisions suggesting that this was becoming more expected among younger individuals, as well as those less-experienced with political activism. Female socialists and male conservatives are both associated with regionalisms, and more conservative respondents are shown to downplay the difference between regional and normative tokens. Additionally, younger respondents do not expect regional features to be tied to the left in the same way that older ones do. All told, this suggests that a change in progress may be underway. Following the ongoing changes in norms Cruz-Ortiz (2019) identifies historically, a shift may be tied to the recent change in governing party in Andalusia in 2019, when the PP took control of the regional government for the first time in decades. These larger ramifications, particularly in the context of theoretical meaning and previous sociophonetic research into political speech, are discussed in greater depth in the next sub-section, which discusses the theoretical contributions of the dissertation.

To wrap up this section, perception is a vital part of understanding variation in a speech community. Not only does it reinforce assumptions about social meaning in production research by ascertaining the opinion of listeners of a given variety, it also provides access to information

about political biases. In recent years across the globe, in Hungary, the United States, Germany, the UK, Brazil, and numerous other countries, far-right parties espousing ideologies of populism and nationalism have sprung up, pushing back against globalization and progressive politics. This doubtless has an effect on the way that members of a community perceive politicians, even in the current case, where party-led differences are not always the most significant social division. In order to understand those norms of evaluation applied to politicians, as Cutillas-Espinosa and Hernández-Campoy (2007) discuss, community members living through the current moment are the best resource. With waves of radicalization occurring across the globe, these results also offer a time capsule into the year 2022, providing a point for later comparison with respect to political biases, as well as the acceptance of regional variation in Spain.

7.1.4 The composite approach

The last question raised in this dissertation, and addressed across all three investigative chapters, has been to determine how the findings in Andalusia, Spain, can apply to previous theoretical claims, as well as offering new generalizations for other language contexts. Across Chapters 4, 5, and 6, there is a clear trend of conservative politicians employing high rates of regional variants and that behavior being less unexpected to younger listeners in Seville. Table 56 shows the main takeaways from each stage of analysis, showing the particular prevalence of coda /s/ elision and intervocalic /d/ elision in Andalusian production, followed by variable trill production, affricate fronting, and variable tap production. It also demonstrates the differences in patterns collected from the LFI, with Andalusians performing regional identity with coda /s/ elision, intervocalic /d/

elision, and some degree of *seseo* and variable tap production, while NCPS speakers used higher rates of intervocalic /d/ elision, variable tap and trill production. Finally, the perceptual instrument showed how normative productions were often associated with urban and educated identity, while regional productions received more adjectives associated with regional and covert personal prestige, including "likeable" and "Andalusian," as well as liberalness.

	Stage 1*		**Stage 2****		**Stage 3*****	
Phenomenon	**Andalusia**	**Madrid**	**Andalusia**	**Madrid**	**Regional**	**Normative**
Affricate fronting	25%	37%			Andalusian Educated Likeable Liberal	Urban
Syllable-final /s/ Word-final onset /s/	89% 90%	10% 4%	X		Urban Andalusian Liberal	Educated Likeable
Seseo	13%	6%	--		Andalusian Educated Likeable Liberal	Urban
Ceceo	0.1%	0.5%			Andalusian Likeable Liberal	Urban Educated
Intervocalic /d/	49%	58%	X	--	Andalusian Likeable Liberal	Urban Educated
Variable /l/ production	7%	4%			Andalusian Likeable Liberal	Urban Educated
Variable tap production	17%	10%	--	--	Andalusian Educated Likeable Liberal	Urban
Variable trill production	52%	48%		X	Andalusian Educated Likeable Liberal	Urban
Total	**37%**	**18%**	**38%**	**21%**		

*Percentages indicate the percent of regional productions recorded by group
**X indicates values of 60-100% and -- indicates 20-59% averaged across speakers
***Adjectives represent associations averaged from all three sections of the instrument

Table 56: Representation of the main results across all three stages of the composite analysis

It is also noteworthy that the LFI revealed differences in patterns of use, at both the quantitative and qualitative levels: conservatives seemed to play to their base, performing southernness and alternately indexing likeability (e.g., Del Cid) and ruralness (e.g., Zoido) through their usage, while socialists were more likely to use regionalisms as a means of showing solidarity and connection to rural voters. Conservatives did not use regional peaks to make promises or reference the struggles of the working class as socialists did; instead, regional features emerged during verbal stumbles, initial greetings, and during moments of performative southern identity. The first context seems incidental, while the second and third come across as forced or perhaps perfunctory.

All of this comes about at an important time for Andalusia. The formerly left-leaning *Ciudadanos* party switched to right-of-center in 2018. The PP took control of the regional government in the 2019 elections[22]. That same year, far-right *Vox* came to the forefront in Spanish politics as the third largest party, and continues to grow (Rama et al., 2021). Similar to other western European nations, Spain is experiencing a surge in interest in right-wing parties, from the center-right PP to the populist *Vox*. The growing number of distinct voices on the right – the PP has been joined in recent years by *Vox* and *Ciudadanos* – requires a way for conservatives to stand out in an increasingly crowded field.

These changes also come in the aftermath of the Trump presidency in the U.S., when it became apparent that working-class liberal voters were attracted by populist rhetoric. The U.S. elected a speaker with a tell-it-like-it-is attitude, described as having a lower register and less educated sound than other politicians (Sclafani, 2017). As Sclafani argues, these speech patterns cause Trump to sound like his voters, creating a reduced degree of social distance as compared to

[22] All but five of the 94 speeches and interviews were produced before 2019.

other politicians, allowing him to cultivate a reality that is more relatable to voters than other, more polished politicians. This also comes in opposition to a speaker like Barack Obama, whose public speech uses Modern United States English (MUSE) features to sound relatable while at the same time being more associated with educated speech (Holliday, 2017). Based on both the production and perception analysis, a similar process is underway in Spain for conservative parties, as norms shift more towards making use of regional speech as a means of connecting with voters.

For the moment, however, these tendencies do not seem to be wholly successful. Among the mainly university speakers in the perceptual survey, conservative voices (even those producing regional variants) were evaluated as more unlikeable. In the LFI analysis, tendencies came across as at times artificial and performative. And in the production analysis, although they used more regional variants, conservative patterns differed from liberal ones in consistent ways. Thus, as Sharma (2018) discussed in terms of the 'real me' uncovered by LFI analysis, this is a type of biographical indexicality: rather than trying to mimic their opponents, conservatives are following the tendencies of their own first-learned dialect. Future research should examine the extent to which differing patterns of variation come across as authentic, and how perceptual evaluations differ based on the social class of politicians.

To return to the concept of the automaticity of less self-conscious style (Schilling-Estes, 2013: 15), or the normative "script" of social expectations for public speech, these results suggest how specifically conservative voices are breaking away from expected norms. As Pollock and Wheeler (2022) showed for Susana Díaz, progressive female politicians in positions of power can make use of regional variation as a means of showing solidarity and confirming their commitment to working-class values. However, politicians on the right (particularly those in *Vox*, based on perceptual results, which also merits further investigative attention) have begun to reassociate

regional identity not with the left and the working class, but rather with some of the same populist notions associated with the Trump presidency: reduced immigration, harsher penalties for crimes, and a stance against progressivism (Rama et al., 2021).

As the perceptual results show, this is leading younger and less politically savvy Seville participants to associate regional speech with both left- and right-leaning ideologies. The 2019 political success of the PP and of *Vox*, in Andalusia and the Spanish government respectively, after decades of PSOE control in Andalusia and while a PSOE Prime Minister still holds power in Madrid, shows that this type of identity work may well have struck a chord with voters. Moving away from the norms of previous decades, conservative male voices are beginning to appeal to rural and working-class speakers, even if (and perhaps because) their approach at the phonetic and topic level differs notably from their socialist counterparts.

One other contribution of this study is to show how female voices on both sides of the political spectrum match and surpass their male counterparts. Both Pollock and Wheeler (2022) and Hernández-Campoy and Cutillas-Espinosa (2010, 2013) discuss the disadvantaged situation of women in positions of political power, using regional variation as a way to reinforce their belonging in the community and party. However, at the same time that voters can perceive them positively, national news sources and NCPS politicians can describe their speech disparagingly. In perception, women tend to rate all tokens as more likeable and liberal than their male counterparts, while female voices were rated as less educated when producing regional variables. This reflects the female speech norms Chappell (2016) discusses, where vernacular variants are often much more marked for female speakers than for males. Women who produce innovative variants can be seen much less positively than their male counterparts, although they are also seen more positively when producing normative variants, placing a greater social pressure on women to comply with

expected norms. That makes the behavior of Martínez and Díaz all the more noteworthy, in the ways that they outperform male speakers in regional variation.

The production results show the wide array of regional behavior accessible to women using Andalusian features. In the LFI analysis in Chapter 5, liquid phenomena seem to carry a degree of stigma, being in relatively low use by female politicians, while both regional phenomena like *seseo* and coda /s/ elision, and national phenomena like intervocalic /d/ and affricate fronting are more common. The fronted affricate, in particular, received positive evaluations from listeners as more urban and educated sounding, and at the same time tends to appear in the lectal focusing of female speakers, showing how a potential prestige variant associates itself with female speech (even as it seems to have spread into the speech of men as well). Meanwhile, in the community norms consideration, gender is not often a significant difference between groups, except among affricate, laterals, and trills, with women favoring use of the first and disfavoring the second two.

From the perspective of production and perception, then, we can see how individuals in minority positions in Spanish society set about establishing their identity, while at the same time avoiding the pitfalls of vernacular productions that can adversely harm their image. This trend, across acoustic and auditive analyses, is important to recognize, as it can be examined in other groups that experience reduced power. Speakers in minority positions in very different contexts may have access to these same tools and patterns. They may be able to establish themselves through the measured use of prestigious innovative variants. This involves a careful balance between not falling too far out of line with respect to the political speech script, while at the same time taking advantage of opportunities for agentive identity construction to develop their identity. Barack Obama's speech, which Holliday (2014, 2017) describes in relation to coronal stop deletion, shows a similar balancing act between MUSE and African American Language norms.

Before closing, it is also important to consider how these results speak to the understanding of sociolinguistic identity. Looking back to the work of Le Page and Tabouret-Keller (1985:5), we see that language is indeed being used among these speakers as a means for identification of individuals, rather than just as a shorthand associating speakers with political parties. As both these authors and Cutillas-Espinosa and Hernández-Campoy (2007) discuss, style is a multidimensional construct that depends on a host of factors to understand speaker behavior – by comparing across a host of social factors, from gender and age to speech context and audience, this study has shown the inherent variability that authors like Cruz-Ortiz (2019) have pointed to, while at the same time trying to establish and compare between patterns of use. We can see here, as Le Page and Tabouret-Keller argue, that linguistic resources in Andalusian political speech are marked in a number of ways, referencing not only the working-class targets of socialists, but also the rural and populist identity coming into popularity at the moment. By associating themselves with a certain pattern of regional features, at certain moments, speakers are not only accommodating with their voting base: they are also projecting their political ideologies and personal identities onto their linguistic decisions.

In this way, the dissertation offers both insight into the impact of political change on Andalusian Spain, as well as a pattern of political speech behavior that may be applicable beyond the Iberian or Spanish context. On the one hand, populist rhetoric pairs with regionalisms as a means of allowing right-leaning politicians to appeal to voters who had grown disillusioned with the PSOE. On the other, female politicians navigate harsher linguistic evaluation criteria and more demanding expectations of speech in the public sphere. To forge their own trail, they must navigate a complicated indexical field, while performing and crafting a consistent identity by using speech phenomena based on the stances they take and their preferred group affiliation.

7.2 Limitations

Every investigative approach has its disadvantages. As Chapters 2 and 3 discuss, from first to third wave sociolinguistics, as well as from production perception, each theoretical framework leaves out certain types of information, and is limited to making certain types of claims about specific manifestations of variation. It was to help address some of these problems that the composite methodology proposed in this dissertation was introduced, approaching variable use in distinct ways to help address these concerns.

When quantifying norms within a speech community, certain problems arise with respect to the social categories that can be implemented. Unlike later second- and third-wave work, this approach does not require, nor can it truly make use of, ethnographic categories and differences at a community level. For example, Eckert's (2000) description of burned-out burnouts applies to a subsection of a subsection of students at a single high school – while other high schools may have similar social groupings, there is no reason to suspect they would be identical, given different socioeconomic and geographical realities. At the level of community study, individual details must give way to generally true, albeit broad, generalizations about speech within a community. At the same time, this dissertation aims to examine regional Andalusian speech norms, but offers a relatively small groupings across social categories (i.e., 8 speakers by city, subdivided by age, political affiliation, and gender). Unfortunately, the sheer scope of the data involved when including more speakers, on top of the multitude of regional phenomena under consideration, precludes the addition of more speakers within a single community. This means that the speech realities outside the three largest urban hubs in Andalusia, be it in cities or rural spaces, must be

left for future consideration. Similarly, despite the fascinating norms surrounding the new populist movement and *Vox* in Spain, only politicians from the two currently largest parties could be included. This leaves the portrait necessarily incomplete, although opening the door for considerable fruitful research in AS.

Also, with respect to community norms, it is worth mentioning that this dissertation focused on sociolinguistic variation and extralinguistic variable to the exclusion of all other types of variation. It was for that reason that phenomena like vowel laxing and *ceceo* were given less attention than other phenomena, both due to a lack of socially-governed variation and to general scarcity. It also means that the effects of linguistic factors were not as closely interrogated as social ones, despite the important role they also played in variable use. Similarly, acoustic measures, while a useful confirmation of the validity of impressionistic coding of tokens, deserve more attention than they received in the current project. Doing so would determine what phonetic differences exist and would offer a more nuanced categorization of allophones than the normative-regional binary distinction could permit. Interesting differences occurred with respect to tap occlusion reduction, variable lateral production, and word-final /s/ elision that there was not sufficient space to expand upon here.

Next, with relation to Lectal Focusing in Interaction, Sharma (2018: 7-8) has already raised several difficulties for this methodology. This includes the method of selecting variables with meaningful lectal association, articulatory differences within certain words affecting the ability to switch, and the affect of linguistic factors on variation (e.g., a regional peak occurs because of a set of linguistic contexts, rather than social ones, that condition production). To an extent, the results of Chapters 4 and 6 were used to resolve this first challenge, in that infrequent phenomena were removed, and the inclusion of other phenomena was justified based on variable trends and

regional associations of listeners. However, linguistic factors can only be partially addressed by additional information provided in Chapter 4, and merits further consideration.

Additionally, a large challenge with the implementation of the LFI analysis was ensuring that the methodology remained comparable to the work of Sharma, while at the same time accounting for the data and speaker group under consideration. Rather than looking at a single individual, ten politicians were selected to provide a better idea of social variation, leading to greatly increased coding time – but also forcing changes in the definition of LFI units, for example, to simplify the process. Despite the increased group of speakers under consideration, the group of individuals analyzed better represents tendencies and qualitative differences than quantitative ones, as it is very likely that other speakers would have still other patterns of variable use. With that said, however, the selection of speakers did allow for extrapolation, which proved especially important in allowing for a discussion of these data in terms of larger sociolinguistic theory.

Turning to the study of perception, certain limitations often emerge with respect to the populations that linguistics can most easily access for such research. In this case, the majority of listeners were relatively young, female college students, with a tendency toward the political left. These were also mainly speakers from the city of Seville, which means that their evaluative criteria are influenced by urban life, and by the specific regional norms belonging to the capital of Andalusia (e.g., the tendency toward *seseo*, reduced familiarity with EAS speech, etc.). This is not so much a limitation as an indicator of the great tasks that remains ahead for sociolinguists in the study of auditive phonetics to explore the evaluations of communities that are less readily accessible and compare across norms of evaluation by different subsets of a population in both rural and urban spaces. While there has been research into certain rural varieties using sociolinguistic perspectives (e.g., Schilling-Estes, 1998; Melguizo-Moreno, 2010), these studies

are scarce. As Villena-Ponsoda (2013) argues, rural spaces provide specific norms for production, and therefore also perception, that influence speakers. Particularly in the face of populism, which is often aimed at working class and rural progressives who feel underserved by their elected officials, the results from this dissertation cannot tell us the responses of these individuals to politicians who sound like them – or if the gendered differences described by college students apply to rural populations as well.

Another important limitation of perception research comes from the tokens used for analysis themselves. Whether faced with verbal guise (i.e., where tokens are taken from pre-existing sources, like corpora), matched guise (i.e., words produced for the survey by actors), or some combination of the two using audio manipulation, there are always problems with replicability and statistical noise influencing classification. Mathced guise and manipulation were avoided here to provide listeners with access to political speech, and without running the risk of producing unnatural-sounding audio with regional phenomena added or removed. However, even under the verbal guise employed in Chapter 6, listeners from Seville distinguished between Andalusian and Madrid politicians to a surprising degree, even without the presence of the regional phenomena examined in this study. This suggests that other audio correlates, from vowel height and fronting to intonation and beyond, inevitably play a role in evaluation, providing a confound in the data that were addressed through the repeated repetitions of variants across the three parts of the survey.

The final major topic worth discussing here involves the composite nature of this dissertation project as a whole. As this section has discussed in detail, the drawbacks to each of these three investigative approaches are well-known and cannot be fully avoided, only mitigated. However, the merit of this composite approach comes about by allowing each of the individual

methodologies to offer up benefits, providing more and varied information about the community in question. In this way, the outdated representation of style offered by community-level analysis can combine with the more individualistic information provided at the level of individual LFI. The gaps in knowledge about exact interlocutor attitudes toward production and the uncertainty regarding use of regional phenomena from the viewpoint of perception are negated through the inclusion of acoustic and auditive approaches.

As such, I would argue that the whole, here, yields a product that is more than just the sum of its parts. The takeaways from each chapter combine with each other, allowing insight into ongoing changes in social meaning associated with political identity, patterns of language use available to speakers in reduced positions of power, and the ways that the power of the automaticity of public speech can be side-stepped to contribute to consistent identity construction. That this approach can provide such rich and varied data recommends it, to my mind, for wide-ranging contexts beyond only urban Andalusian Spanish, or even political speech. Ideas for future research are raised in the next section.

7.3 Recommendations for future research

Each of the three individual stages of this dissertation recommends topics for further consideration, while at the same time the composite approach offered here also provides a blueprint for research in other communities. For example, at the community level, phenomena like vowel-laxing, *ceceo*, and affricate fronting offer interesting contexts for further sociolinguistic research to determine what variability exists beyond political speech, as well as how rural speakers and politicians

evaluate and use them. In the analysis of LFI, there are exciting future possibilities involving new ways of representation, be it through the inclusion of continuous acoustic measures, or through the integration of logistic modelling to address linguistic variation. In the perceptual study, it would make sense to offer the same survey to rural speakers, as well as to those from the urban EAS, NAS, and NCPS cities under consideration, to distinguish how one's place and sociolinguistic background affects evaluation. At the composite level, meanwhile, there are a world of possibilities to begin expanding this approach to political speech in other peninsular varieties (e.g., Galicia: Pollock & Wheeler, 2023), as well as to other parts of the Spanish-speaking world, and beyond.

Hernández-Campoy and Jiménez Cano (2003), Hernández-Campoy and Cutillas-Espinosa (2010, 2013), Cruz-Ortiz (2019), and Pollock and Wheeler (2022) have all shown the usefulness and importance of considering political speech as a way of accessing information about speech communities and stylistic behavior. Acoustic correlates like COG (Hz) and duration (ms) could provide nuanced information about linguistic variation in progress that is both less reliant on impressionistic coding and more fine-grained than the binary normative-regional divide considered here. Díaz-Campos, Cole, and Pollock (2023) show how this can be carried out in Venezuelan Spanish with the affricate. While this dissertation was forced to set aside phonetic specificity in favor of social distinction, future research should examine allophonic variation in greater depth to determine if we can further improve our understanding of the political and gendered differences presented in this dissertation.

Another clear need for the near future is to examine the spread of populist ideologies through Spain using methods based in phonetic, sociolinguistic, and pragmatic analysis. While this research has intentionally focused on the two most-central parties in Spain, the recent rise of *Vox*

provides ample access to political speech from these politicians. The results of this dissertation are consistent with the reality that far-right political support is growing in the political community of practice in Spain. Production research into politicians from these parties from across the peninsula (and particularly in Andalusia) may provide interesting information about regional usage, in addition to helping confirm the theories urged forward by the perceptual and production results presented herein. A related question also involves the role of regionalisms in areas often classified as NCPS, to determine what resources speakers in northern Spain employ when engaged in LFI, and if they mirror southern tendencies.

A question raised by the LFI results can be answered with further perceptual analysis. While patterns of phenomena use during 'real me' moments have been observed and described, evaluations from AS speakers can also help determine the extent to which these combinations of regional and normative variants are socially salient. While it becomes more difficult to examine phonetic items when providing entire utterances, given the amount of additional linguistic noise that introduces into evaluation (e.g., intonation, topic, etc.), this type of instrumental design would be necessary to examine how regional authenticity is conveyed through regional use. Given the political success the right has had in Andalusia in recent years, and *Vox*'s national popularity, defining how patterns differ and the attitudes applied to them would provide more insight into the goals and speech tendencies associated with nationalist and populist movements.

7.4 Concluding Remarks

The current state of style-centric sociolinguistics offers much promise for future research. While many linguists have contributed to a rich body of English-based analysis examining style-shifting and identity construction, many other language contexts have yet to catch up. Even among these studies, third-wave research on its own is not without critique: the forest can at times be lost for the trees. Both new and old methodologies offer a way to combat that. Here, by combining a Labovian community approach with LFI analysis and verbal-guise perceptual research, Andalusian Spanish reveals an exciting web of indexicality and meaning in a current state of flux. The composite view of variation at the level of individual variants, style-shifting behavior, and interlocutor attitudes combines to offer a clearer picture of how populism, socialism, and gender relate in a changing peninsular political landscape.

By presenting a blueprint and point of comparison for future sociolinguists, this dissertation aims to further cross-linguistic pollination and comparison, particularly but by no means limited to Spanish. Third-wave sociolinguistics is often hyper-aware of speech situations and individual agency – however, I would argue that this composite methodology offers a way to expand the generalizability of findings, making them applicable beyond any single group of individuals, or even language family. As conservative politics continues to gain support across Europe, the U.S., and the rest of the world, there are clearly ideas being conveyed by these groups that appeal to large numbers of voters. It is no stretch to imagine that politicians supporting similar ideologies and goals over great geographical distance would share linguistic tools and patterns. Similarly, politics in many parts of the world has, until recent decades, belonged largely to men, with high level positions often experience considerable gender imbalance. Just as this reality extends beyond the Andalusian context, so too can the lectal focusing of female politicians – as they navigate stigmatized rural language on one hand and a desire to develop a coherent and appealing identity

on the other. The method by which both of these groups set aside the automaticity of speech at key moments to perform identity work is not just unique to their language and social context; it is part of a greater trend that generalized far beyond Seville.

References

Abercrombie, D. (1967). *Elements of General Phonetics.* Chicago: Aldine.

Almeida, M. (2019). Language hybridism: On the origin of interdialectal forms. In J. A. Villena-Ponsoda, F. Díaz-Montesinos, A. M. Ávila-Muñoz, & M. Vida-Castro (Eds.), *Language Variation - European Perspectives VII: Selected papers from the Ninth International Conference on Language Variation in Europe (ICLaVE 9), Málaga, June 2017 (Vol. 22).* Amsterdam: John Benjamins.

Alonso, A. (1951). Historia del *ceceo* y del *seseo* españoles. *Thesaurus 1*(1), 111-200.

Alonso, D., Vicente, A. Z., de Zamora, M. J. C. (1950). Vocales andaluzas. Contribución al estudio de la fonología peninsular. *Nueva Revista de Filología Hispánica 4*(3), 209-230.

Amano, S., & Yamakawa, K. (2011). Perception and Production Boundaries between Fricative [s] and Affricate [t͡s] in Japanese. In *Proceedings of the 17th ICPhS, Hong Kong,* pp. 228-232.

Amengual, M., & Chamorro, P. (2015). The Effects of Language Dominance in the Perception and Production of the Galician Mid Vowel Contrasts. *Phonetica 72*(4), 207-236.

Aristotle. (2006). *On Rhetoric* [Rhētorikḗ]. (G. A. Kennedy, Trans.). Oxford: Oxford UP

Auer, P., Hinskens, F., & Kerswill, P. (eds.) (2005). *Dialect change: Convergence and divergence in European languages*. Cambridge: Cambridge University Press.

Ávila-Muñoz, A. M. (1994) La variación reticular e individual de s/z en el vernáculo urbano malagueño. Datos del barrio de Capuchinos. *Analecta Malacitana XVII*(2), 343-367.

Barnes, S. (2016). Negotiating local identity: rural migration and sociolinguistic perception in urban Asturias. *Lengua y migración, 8*(2), 45-77.

Bell, A. (1984). Language style as audience design. *Language in Society 13*(2), 145–204.

Bell, A. (2016). Succeeding waves: Seeking sociolinguistic theory for the twenty-first century. In N. Coupland (ed.), *Sociolinguistics: Theoretical debates.* Cambridge: Cambridge University Press, 391–416.

Bishop, J. B. (2007). Incomplete neutralization in Eastern Andalusian Spanish: perceptual consequences of durational differences involved in s-aspiration. *Proceedings of the 16th International Congress of Phonetic Sciences*, pp. 1765-1768. Saarbrücken: Saarbrücken UP.

Blanco, Ana. (2004). *Estudio sociolingüístico de Alcalá de Henares*. Alcalá de Henares: Universidad de Alcalá.

Blecua, B. (2001). *Las vibrantes del español: Manifestaciones acústicas y procesos fonéticos*. (Doctoral dissertation) [Universitat Autónoma de Barcelona]. UAB Online. http://hdl.handle.net/10803/4859

Boersma, P. & Weenink, D. (2023). *Praat: doing phonetics by computer* [Computer program]. Version 6.1.40, http://www.praat.org/

Bradley, T. G. (2001). A Typology of Rhotic Duration Contrast and Neutralization. In Minjoo Kim and Uri Strauss (Eds.), Proceedings of the North East Linguistics Society (NELS 31), pp. 79-97. GSLA Publications.

Bradley, T. G. (2006). Phonetic realizations of /sr/ clusters in Latin American Spanish. Díaz-Campos, M. (Ed.) *Selected Proceedings of the 2nd Conference on Laboratory Approaches to Spanish Phonetics and Phonology*, pp. 1-13. Somerville, MA: Cascadilla Proceedings Project.

Brown, E. K. (2008). *A usage-based account of syllable- and word-final /s/ reduction in four dialects of Spanish* (Document No. 200914081) [Doctoral dissertation, University of New Mexico]. ProQuest Dissertations & Theses Global.

Bucholtz, M. & Hall, K. (2005). Identity and interaction: A sociocultural linguistic approach. *Discourse Studies 7*(4–5), 585–614.

Bucholtz, M. (2009). From stance to style: Gender, interaction, and indexicality in Mexican immigrant youth slang. In Alexandra Jaffe (ed.), *Stance: Sociolinguistic perspectives*, pp. 146–70. New York: Oxford University Press.

Burnley, D. (1983). *A Guide to Chaucer's Language.* Norman: University of Oklahoma Press.

Bybee, J. (2007). *Frequency of Use and the Organization of Language*. Oxford University Press.

Campbell-Kibler, K. (2009). The nature of sociolinguistics perception. *Language Variation and Change 21*, 135-156.

Campbell-Kibler, K. (2010). Sociolinguistics and Perception. *Language and Linguistics Compass 4/6*, 377-389.

Cano-Aguilar, R. (2001). La historia del andaluz. In C. L. Reina-Reina (Ed.), *Actas de las Jornadas "El habla andaluza: historia, normas, usos",* pp. 33-57. Estepa: Ayuntamiento.

Cano-Aguilar, R. (2006). Regionalismo, nacionalismo, lengua: el caso de Andalucía. In A. Narbona-Jiménez (Ed.,), *III Jornadas sobre el habla andaluza: diversidad y homogeneidad del andaluz*, pp. 153-177. Estepa: Ayuntamiento.

Carbonero, P. (1985). Aspectos sociolingüísticos sobre la nivelación en el español meridional. *Revista de Filología Románica 3*, 77–83.

Carbonero, P. (1992). El habla de Jerez. *Estudio sociolingüístico, Biblioteca de Urbanismo y Cultura, Cuadernos de Divulgación*. Jerez: Ayuntamiento de Jerez.

Carlson, K. M. (2012). An acoustic and perceptual analysis of compensatory processes in vowels preceding deleted post-nuclear /s/ in Andalusian Spanish. *Concentric: Studies in Linguistics 38*(1), 39-67.

Chappell, W. & Barnes, S. (eds.) (in preparation) Third-wave Spanish sociolinguistics (Special Issue). *Languages*. Basel, Switzerland: MDPI.

Chappell, W. (2016). On the social perception of intervocalic /s/ voicing in Costa Rican Spanish. *Language Variation and Change 28*(3), 357-378.

Chappell, W. (2019). *Recent Advances in the Study of Spanish Sociophonetic Perception*. Amsterdam: John Benjamins.

Chomsky, N., & Halle, M. (1968). *The Sound Pattern of English.* New York: Harper & Row.

Cicero, M. T. (*1853) On Invention* [De inventione]. (C. D. *Yonge, Trans.)*. London: Henry G. Bohn.

Clopper, C. G., & Pisoni, D. B. (2004). Homebodies and army brats: Some effects of early linguistic experience and residential history on dialect categorization. *Language Variation and Change, 16*(1), 31– 48.

Coupland, N. (1980). Style-shifting in a Cardiff work setting. *Language in Society 9*, 1–12.

Coupland, N. (2001). Language, situation, and the relational self: theorizing dialect-style in sociolinguistics. In: Eckert, Penelope, Rickford, John R. (Eds.), *Style and Sociolinguistic Variation*, pp. 185-210. Malden/Oxford: Blackwell.

Cramer, J. (2016). *Perceptual dialectology*. Oxford: Oxford University Press.

Cruz-Ortiz, Rocío. (2019). *El habla de los políticos andaluces en Madrid. Mantenimiento y pérdida del vernáculo andaluz (Tesis doctoral*). Universidad de Granada.

Cutillas-Espinosa, J. A., Hernández-Campoy, J. M., & Schilling-Estes, N. (2010). Hyper-vernacularisation and speaker design: A case study. *Folia Linguistica 44*(1), 31–52.

Cutillas-Espinosa, J. A., Hernández-Campoy, J. M., (2006). Nonresponsive performance in radio broadcasting: a case study. *Language Variation and Change 18*(3), 1–14.

Cutillas-Espinosa, J. A., Hernández-Campoy, J. M., (2007). Script design in the media: radio talk norms behind a professional voice. *Language & Communication 27*(2), 127–152.

Dalbor, J. B. (1980). Observations on present-day *seseo* and *ceceo* in Southern Spain. *Hispania 63*(1), 5-19.

Del Saz, M. (2019). From postaspiration to affrication: new phonetic contexts in western Andalusian Spanish. In *Proceedings of the 19th ICPhS, Melbourne*.

Delattre, P. (1969). An acoustic and articulatory study of vowel reduction in four languages. International Review of Applied Linguistics in Language Teaching, 7(4), 295-325.

Delgado-Díaz, G., & Galarza, I. (2015). ¿Que comiste [x]amón? A Closer Look at the Neutralization of /h/ and Posterior /r/ in Puerto Rican Spanish. In E. Willis, P. Martín Butragueño, & E. Herrera Zendejas (Eds.), Selected Proceedings of the 6th Conference on Laboratory Approaches to Romance Phonology (pp. 70-82). Somerville, MA: Cascadilla Proceedings Project.

Delgado-Díaz, G., Galarza, I., & Díaz Campos, M. (2021). Complex attitudes towards two sociolinguistic variables and their social meanings: Providing evidence from production and perception data in a speech community. In Díaz Campos, M., & Sessarego, S. (Ed.) *Aspects of Latin American Spanish Dialectology: In honor of Terrell A. Morgan* [Issues in Hispanic and Lusophone Linguistics 32], pp. 125–153. Amsterdam: John Benjamins.

Díaz-Campos, M., Cole, M., & Pollock, M. (2023). Re-conceptualizing affricate variation in Caracas Spanish. *Hispania 106*(1), 9-26.

Díaz-Campos, M. & Hernández-Campoy, J. M. (in preparation). *Enciclopedia Concisa de los Dialectos del Español*. Massachusetts: Wiley.

Díaz-Campos, M. & Killam, J. (2012). Assessing Language Attitudes through a Matched-guise Experiment: The Case of Consonantal Deletion in Venezuelan Spanish. *Hispania 95*(1), 83-102.

Díaz-Campos, M. & Killam, J. (2017). Contribuciones de los estudios de percepción en la identificación y la descripción de las variedades diatópicas y diastráticas del español contemporáneo. In Leonor Orozco and Alonso Guerrero Galván (Eds.). *Estudios de variación geolingüística*. México: Secretaría de Cultura / Instituto Nacional de Antropología e Historia.

Díaz-Campos, M. & Navarro-Galisteo, I. (2009). Perceptual categorization of dialect variation in Spanish. In Joseph Collentine et al. (Eds.), *Selected proceedings of the 11th Hispanic Linguistics Symposium*, pp. 179-195. Somerville, MA: Cascadilla Proceedings Project.

Díaz-Campos, M. & Pollock, M. (submitted). The Spanish Sibilant System: *Seseo*. In A. Gallego & C. García (eds.) *A Guide to Spanish Dialects: Descriptive and Theoretical Aspects of Linguistic Variation in the Hispanic World*. Oxford: Oxford University Press.

Díaz-Campos, M. (2008). Variable production of the trill in spontaneous speech: sociolinguistic implications. In L. Colantoni & J. Steele (Eds.), *Selected Proceedings of the 3rd Conference on Laboratory Approaches to Spanish Phonology* (pp. 47-58). Somerville, MA: Cascadilla Press.

Díaz-Campos, M. (2014). *Introducción a la Sociolingüística Hispánica.* Malden, MA: Wiley-Blackwell.

Du Bois, J. W. (2007). The stance triangle. In Robert Englebretson (ed.), *Stancetaking in discourse: Subjectivity, evaluation, interaction,* pp. 139–182. Amsterdam: John Benjamins.

Eckert, P. & McConnell-Ginet, S. (2003). *Language and Gender*. Cambridge: Cambridge University Press.

Eckert, P. & Rickford, J. R. (2001). *Style and sociolinguistic variation*. Malden, MA: Blackwell.

Eckert, P. (1989). The whole woman: Sex and gender differences in variation. *Language Variation and Change 1*, 245-267.

Eckert, P. (2000). *Linguistic variation as social practice*. Oxford: Blackwell.

Eckert, P. (2008). Variation and the indexical field. *Journal of Sociolinguistics 12*(4), 453-476.

Eckert, P. (2012). Three waves of variation study: The emergence of meaning in the study of sociolinguistic variation. *Annual Review of Anthropology 41*, 87–100.

Eckert, P. (2016). Variation, meaning, and social change. In N. Coupland (ed.), *Sociolinguistics: Theoretical Debates,* pp. 68-85. Cambridge, UK: Cambridge University Press.

Egido, A. (2018). *Por el gusto de leer a Cervantes.* Seville: Fundación José Manuel Lara.

Fafulas, S., Díaz-Campos, M., & Gradoville, M. (2018). Stable Variation or Change in Progress? A Sociolinguistic Analysis of PA(RA) in the Spanish of Venezuela. In J. King & S. Sessarego (Eds.), *Language Variation and Contact-Induced Change: Spanish across Space and Time,* pp. 223–245. Amsterdam: John Benjamins

Feldman, S. (2003). Values, ideology and the structure of political ideology. In D. Sears, L. Huddy, & R. Jervis (eds.), *Oxford Handbook of Political Psychology*. New York: Oxford University Press, 477–508.

Flores, T. (2014). *A socio-phonetic examination of co-articulation in Chilean public speech* (Document No. AAI3665473) [Doctoral dissertation, Indiana University]. ProQuest Dissertations & Theses Global.

Flores, T. (2017). Interlocutor accommodation on the variation of /tɾ/ in Chilean radio. *Discourse, Context & Media 16*, 31-38.

Flores, T. (2018). Lexical Frequency Effects in Chilean Spanish Affricate Fronting. *Southern Journal of Linguistics 41*(1), 17-39.

García, C. (2019). Regional identity in highland Ecuador: Social evaluation of intervocalic /s/ voicing. *Recent Advances in the Study of Spanish Sociophonetic Perception*, *21*, 126– 152.

García-Amaya, L. J. (2008). Variable norms in the production of /θ/ in Jerez de la Frontera, Spain, in J. F. Siegel, T. C. Nagle, A. Lorente-Lapole & J. Auger (eds.) IUWPL7: Gender in language: Classic questions, new contexts. Bloomington, IN: IULC Publications, 49–71.

García-Marcos, F. (1990). *Estratificación social del español de la Costa granadina*. Almería: Departamento de Lingüística General y Teoría de la Literatura.

Gaya, S. (1921). La r simple en la pronunciación española. *Revista de Filología Española 8*, 271-280.

Gerfen, C. (2002). Andalusian codas. *Probus 14*(2). 303-333.

Giles, H. (1980). Accommodation theory: Some new directions. *York Papers in Linguistics 9*, 105-36.

Giles, H., & Powesland, P. F. (1975). *Speech style and social evaluation*. Academic Press.

Giles, H., Coupland, N. & Coupland, J. (1991). *Contexts of accommodation: Developments in applied sociolinguistics*. Cambridge University Press.

Gil-Peña, N. (2004). *Estudio sociolingüístico de la fonética de la ciudad de Madrid*. Memoria de licenciatura inédita, Universidad de Alcalá.

Gómez-Serrano, A. (1994). *Aspectos del habla de Linares (Jaén).* Universidad de Málaga.

Gordon, M. (2012). Chapter Four: A variationist approach to language. *Labov: A Guide for the Perplexed*. Bloomsbury Publishing PLC.

Gordon, M., Barthmaier, P., & Sands, K. (2002). cross-linguistic acoustic study of voiceless fricatives. *Journal of the International Phonetic Association 32*(2), 141-174.

Gradoville, M. (2014). *A comparative usage-based approach to the reduction of the Spanish and Portuguese preposition* para (Publication No. 3611090) [Doctoral dissertation, Indiana University]. ProQuest Dissertations and Theses Global.

Guy, G. R. (1991). Explanation in variable phonology: An exponential model of morphological constraints. *Language Variation and Change 3*(1),1–22.

Guy, G., & Cutler, C. (2011). Speech style and authenticity: Quantitative evidence for the performance of identity. *Language Variation and Change 23*(1), 139–62.

Hall-Lew, L., Coppock, E., & Starr, R. L. (2010). Indexing political persuasion: variation in the Iraq vowels. *American Speech 85*(1), 91-102.

Hall-Lew, L., Friskney, R. & Scobbie, J. M. (2017). Accommodation or political identity: Scottish members of the UK Parliament. *Language Variation and Change 29*, 341-363.

Hall-Lew, L., Starr, R., & Coppock, E. (2012). Style-shifting in the U. S. Congress: The vowels of "Iraq(i)". In J. Hernández-Campoy & J. Cutillas-Espinosa (eds.), *Style-shifting in public: New perspectives on stylistic variation*. Amsterdam: John Benjamins, 45–63.

Harjus, J. (2017). Perceptual variety linguistics: Jerezano speakers' concepts and perceptions of phonetic variation in western Andalusian Spanish. *Loquens 4*(2), 1-15.

Harrington, J. (2007). Evidence for a relationship between synchronic variability and diachronic change in the Queen's annual Christmas broadcasts. Laboratory Phonology 9, 125-144.

Hay, J., Jannedy, S. & Mendoza-Denton, N. (1999). Oprah and /ay/: Lexical Frequency, Referee Design, and Style. In Proceedings of the 14th ICPhS, San Francisco, pp. 1389-1392.

Hay, J., Jannedy, S., & Mendoza-Denton, N. (2010). Oprah and /ay/: Lexical Frequency, Referee Design, and Style. In Meyerhoff, M. & Schleef, E. (eds): *The Routledge Sociolinguistics Reader*, pp. 53-59. NY: Routledge.

Henriksen, N. & Willis, E. W. (2010). Acoustic characterization of phonemic trill production in Jerezano Andalusian Spanish. In Marta Ortega-Llebaria (ed.), *Selected proceedings of the 4th Conference on Laboratory Approaches to Spanish Phonology*, pp. 115–127. Somerville, MA: Cascadilla Proceedings Project.

Henriksen, N. (2014). Sociophonetic analysis of phonemic trill variation in two sub-varieties of Peninsular Spanish. *Journal of Linguistic Geography 2*, 4–24

Henriksen, N. (2017). Patterns of vowel laxing and harmony in Iberian Spanish: Data from production and perception. *Journal of Phonetics 63*, 106–126.

Henríquez Barahona, M. & Fuentes Grandón, D. (2018). Realizaciones de los fonemas /t͡ʃ/ /t͡ʂ/ en el chedungun hablado por niños bilingües del alto Biobío: un análisis espectrográfico. *Literatura Lingüística 37*, 253-272.

Hernández-Campoy, J. M. (2016). *Sociolinguistic styles*. Malden & Oxford: Wiley-Blackwell.

Hernández-Campoy, J. M. & Cutillas-Espinosa, J. A. (2010). Speaker design practices in political discourse: A case study. *Language & Communication 30*, 297-309.

Hernández-Campoy, J. M. & J. A. Cutillas-Espinosa. (eds.). (2012). *Style-shifting in public: New perspectives on stylistic variation*. Amsterdam/Philadelphia: John Benjamins.

Hernández-Campoy, J. M. & Cutillas-Espinosa, J. (2013). The effects of public and individual language attitudes on intra-speaker variation: case study of style-shifting. *Multilingua 31*(1), 79-101.

Hernández-Campoy, J. M., & Jiménez Cano, J. M. (2003). Broadcasting standardisation: an analysis of the linguistic normalisation process in Murcia. *Journal of Sociolinguistics 7* (3), 321–347.

Herrero de Haro, A. (2017a). Four mid back vowels in Eastern Andalusian Spanish. *Zeitschrift fur romanische Philologie 133*(1), 82-114.

Herrero de Haro, A. (2017b). The phonetics and phonology of Eastern Andalusian Spanish: review of literature from 1881 to 2016. *Íkala, Revista de Lenguaje Cultura 22*(2), 313-357.

Herrero de Haro, A. (2018). Context and vowel harmony: Are they essential to identify underlying word-final /s/ in eastern Andalusian Spanish? *Dialectologia 20*, 107-145.

Herrero de Haro, A. (2019). Consonant deletion and Eastern Andalusian Spanish Vowels: The effect of word-final /s/, /r/ and /θ/ deletion on /i/. *Australian Journal of Linguistics 39*(1), 107-

Holliday, N. (2014). "He Didn'(t) Give Up When Things Got Har(d)": Examining Barack and Michelle Obama's Rates of Coronal Stop Deletion. *University of Pennsylvania Working Papers in Linguistics, 20*(2).

Holliday, N. (2017). "My Presiden(T) And Firs(T) Lady Were Black": Style, Context, And Coronal Stop Deletion in The Speech of Barack And Michelle Obama. *American Speech 92*(4), 459-486,

Holmes-Elliott, S. & Levon, E. (2017). The substance of style: Gender, social class and interactional stance in /s/-fronting in southeast England. *Linguistics 55*(5), 1045–72.

Hualde, J. I. & Colina, S. (2014). *Los sonidos del español*. Cambridge, UK: Cambridge University Press.

Hualde, J. I. & Sanders, B. P. (1995). new hypothesis on the origin of the eastern Andalusian vowel system. *Annual Meeting of the Berkeley Linguistics Society 21*(1), 426-437.

Huddy, L. (2013). From group identity to political cohesion and commitment. In L. Huddy, D. O. Sears, & J. Levy (eds.), *Oxford handbook of political psychology*. New York: Oxford University Press, 737–773.

Hymes, D. (1974). *Foundations in Sociolinguistics. An Ethnographic Approach.* Philadelphia: University of Pennsylvania Press.

Iribar, A & Túrrez, I (2015). El *seseo* vasco: Caracterización fonética basque *seseo*: Phonetic characterization, *Estudios de Fonética Experimental 24*: 205-235.

Jiménez, J. & Lloret, M. (2007). Andalusian Vowel Harmony: Weak Triggers and Perceptibility. Presented at the *Workshop "Harmony in the Languages of the Mediterranean," OCP4*.

Johnson, D. E. (2009). Getting off the GoldVarb standard: Introducing Rbrul for mixed-effects variable rule analysis. *Language and Linguistics Compass 3*(1), 359-383.

Johnstone, B. & Kiesling, S. F. (2008). Indexicality and experience: Exploring the meanings of /aw/-monophthongization. *Journal of Sociolinguistics 12*(1), 5-33.

Johnstone, B. (2000). The individual voice in language. *Annual Review of Anthropology 29*, 405–425.

Kementchedjhieva, Y. (2016). Code-switching as strategically employed in political discourse. *Lifespans & Styles: Undergraduate Working Papers on Intraspeaker Variation 2*(1), 3–9.

Kendall, T. & Thomas, E. R. (2010). *Vowels: Vowel Manipulation, Normalization, and Plotting in R. R package, version 1.1*. [Software: http://ncslaap.lib.ncsu.edu/tools/norm/]

Kirkham, S. & Moore, E. (2016). Constructing social meaning in political discourse: Phonetic variation and verb processes in Ed Miliband's speeches. *Language in Society, 45*(1), 87-111.

Labov, W. (1966). *The Social Stratification of English in New York City.* Washington: Center for Applied Linguistics.

Labov, W. (1972). *Sociolinguistic Patterns.* Oxford: Blackwell.

Labov, W. (1981). Field Methods of the Project in Linguistic Change and Variation. *Sociolinguistic Working Paper Number 81*.

Labov, W. (1990). The intersection of sex and social class in the course of linguistic change. *Language Variation & Change 2*, 205-254.

Labov, W. (1994). *Principles of linguistic change: Internal factors*. Oxford: Blackwell.

Labov, W. (2001) *Linguistic Change Volume 2: Social Factors.* Oxford: Blackwell.

Labov, W., Cohen, P., Robins, C., & Lewis, J. (1968). A Study of the Non-standard English of Negro and Puerto Rican Speakers in New York City, I. *Final Report, Cooperative Research Project 3288*. Philadelphia: US Regional Survey.

Ladefoged, P. & Johnson, K. (2011). *Course in Phonetics* (6th ed.). Boston: Thomson Wadsworth.

Lambert, W. E., Hodgson, R. C., Gardner, R. C., & Fillenbaum, S. (1960). Evaluational reactions to spoken languages. *The Journal of Abnormal and Social Psychology* 60(1), 44–51.

Lass, N. J., Hughes, K. R., Bowyer, M. D., Waters, L. T., & Bourne, V. T. (1976). Speaker sex identification from voiced, whispered, and filtered isolated vowels. *The Journal of the Acoustical Society of America, 59*(3), 675– 678. doi:10.1121/ 1.380917

Le Page, R. B., & Tabouret-Keller, A. (1985). Acts of identity: Creole-based approaches to language and ethnicity. Cambridge: Cambridge University Press.

Li, A. & Li, Z. (2019). Spectral analysis of sibilant fricatives and the ling sound test for speakers of Chinese dialects. In *ICPHS Full Papers*. Accessed from https://icphs2019.org/icphs2019-fullpapers/

Lloret, M. R. & Jiménez, J. (2009). Un análisis "óptimo" de la armonía vocálica del andaluz. *Verba 36*, 293-325.

Lobanov, B. M. 1971. Classification of Russian vowels spoken by different listeners. *Journal of the Acoustical Society of America 49*:606-08.

Loyola, G. M. & Rafat, Y. (2022) An examination of the perceptions of Cuban and Peninsular Spanish varieties by native, second language learners of Spanish, and monolingual English speakers. *The Journal of the Acoustical Society of America 152*(4), A263.

Marín-Gálvez, R. (1995). La duración vocálica en español. E.L.U.A., 10, 213-226.

Marrero, J. S. (2016). *Seseo, ceceo* y distinción en el sociolecto alto de la ciudad de Sevilla: nuevos datos a partir de los materiales de PRESEEA. *Boletín de Filología LI*(2), 255-280.

Martín-Butragueño, P. (1995). La variable (s) en el sur de Madrid. Contribución al estudio de la frontera de las hablas meridionales del español. *Anuario de Letras 33*, 5-57.

Martínez-Celdrán, E. (1995). En torno las vocales del español: análisis reconocimiento. *Estudios de Fonética Experimental, VII, PPU*, 195-218.

Martínez-Celdrán, E. (1998). *Análisis espectrográfico de los sonidos del habla*. Barcelona: Ariel.

Melguizo-Moreno, E. (2006). La fricatización de /č/ en una comunidad de hablantes granadina. *Interlingüística 17*, 748-757.

Melguizo-Moreno, E. (2007a). La variación de /θs/: Estudio comparativo de dos muestras de población granadinas. *Estudios de Lingüística. Universidad de Alicante 21,* 1-16.

Melguizo-Moreno, E. (2007b). Convergencia y divergencia dialectal: Estudio de la variación lingüística granadina, *Linguistica Espanola Actual 2*9(2): 261-285.

Melguizo-Moreno, E. (2009). Una aproximación sociolingüística al estudio del *ceceo* en un corpus de hablantes granadinos. *Estudios de Lingüística Aplicada 49*, 57–78.

Melguizo-Moreno, E. (2010). La variación social de S y Z como consecuencia de la inmigración rural en Granada. *Sintagma: revista de lingüística 21*, 71-89.

Milla-Muñoz, A. (2020). *Beyond stereotypes and visual cues: Assessing language attitudes toward Andalusian Spanish*. (Publication No. 27960364) [Doctoral dissertation, Indiana University]. ProQuest Dissertations and Theses Global.

Milroy, J. & Milroy. L. (1985). Linguistic change, social networks, and speaker innovation. *Journal of Linguistics 21*(2), 339-84.

Molina-Martos, I. (1998). La fonética de Toledo. Contexto geográfico y social. *Alcalá de Henares: Universidad de Alcalá*.

Molina-Martos, I. & Paredes-García, F. (2014). Sociolingüística de la elisión de la dental –/d/- en Madrid (distrito de Salamanca). *Cuadernos de Lingüística del Colegio de México. Estudios de cambio y variación 2*, 55-114.

Molina-Ortés, E. (2012). Una aproximación al estudio del *seseo* en Fuente del Maestre (Badajoz). *Revista de Estudios Extremeños LXVIII*(1), 129-160.

Molina-Ortés, E. (2015). La influencia de la variable generación en la variación lingüística de Mérida (Badajoz). Análisis y resultados de nuevas actitudes. *Revista de Investigación lingüística, 18*, 65-88.

Molina-Ortés, E. (2018). El habla de Mérida y sus cercanías. Nuevos tiempos y nuevos resultados. *Anuario de letras. Lingüística y filología*, *6*(1), 119-154.

Molina-Ortés, E. (2020). ¿Cómo hablan los presentadores andaluces en televisión? Percepción de los espectadores y análisis lingüístico de la posible variación (socio)estilística en los medios. *LinRed*, 1-29.

Molina-Ortés, E. (2021). Estudio sociolingüístico de las intervenciones políticas y públicas de los presidentes del gobierno extremeño. *Cultura, Lenguaje y Representación 26*, 167-187

Moreno Fernández, J., & Moreno Fernández, F. (1999). Madrid perceptions of regional varieties in Spain. In D. Long & D. Preston (Eds.), *Handbook of Perceptual Dialectology*, (pp. 295-320). Amsterdam: John Benjamins.

Morrison, G., & Escudero, P. (2007). cross-dialect comparison of Peninsular- and Peruvian-Spanish vowels. *Proceedings of the 16th International Congress of Phonetic Sciences: Saarbrücken 2007.*

Moya-Corral, J. A. (1979). *La pronunciación del español en Jaen*. Granada: Universidad de Granada.

Moya-Corral, J. A. & García, E. (1995). *El habla de Granada y sus barrios*. Granada: Universidad.

Moya-Corral, J. A. & Sosinski, M. (2015). La inserción social del cambio: La distinction s/θ en granada. Análisis en tiempo aparente y en tiempo real. *Lingüística Española Actual 37*(1), 33-72.

Navarro-Tomás, T. (1939). Desdoblamiento de fonemas vocálicos. *Revista de Filología Hispánica 1*, 165-167.

Navarro-Tomás, T, Espinosa, A M, & Rodríguez-Castellano, L (1933). La frontera del andaluz, *Revista de Filología Española 20*(3), 225–277.

Neumann, F. & Kanwit, M. (2018). New perspectives on automatic and morphophonological alternations: harmonic processes in two peninsular varieties of Spanish. *Borealis: An International Journal of Hispanic Linguistics 7*(1), 97-112.

Niedzielski, N. A. & Preston, R. (2000). *Folk linguistics*. Berlin: De Gruyter.

Nowak, P. (2003). The role of vocalic context in the perception of Polish sibilants. In *Proceedings of the 15th ICPhS, Barcelona*, pp. 2309-2312.

O'Neill, P. (2005). Utterance final /s/ in Andalusian Spanish: The phonetic neutralization of phonological contrast. *Language Design, Journal of Theoretical and Experimental Linguistics 7*, 151-166.

O'Neill, P. (2010). Variación cambio en las consonantes oclusivas del español de Andalucía. *Estudios de fonética experimental 19*, 11-41.

Ohala, J. (1990). There is no interface between phonology and phonetics: personal view. *Journal of Phonetics 18*, 153-171.

Ohala, J. (2008). The phonetics-phonology interface...again. In E. Herrera Z. P. M. Butragueño (Eds.), *Fonología instrumental: Patrones fónicos variación*, pp. 13-23. México, DF: El Colegio de México.

Ohala, J. (2010). The Relation between Phonetics and Phonology. In Hardcastle, W. J., Laver, J., and Gibbon, F. E. (Eds.) *The Handbook of Phonetic Sciences: Second Edition*, pp. 653-677. Malden, MA: Wiley-Blackwell.

Paredes, F. (2001). *El habla de La Jara. Los sonidos (estudio sociolingüístico).* Alcalá de Henares: Universidad de Alcalá.

Parkinson, S. (1988). Portuguese. In M. Harris & N. Vincent (Eds.), *The Romance Languages*, pp. 131-169. New York: Oxford University Press.

Penny, R. (2002). History of the Spanish Language. Cambridge: Cambridge University Press.

Pérez-Martín, A. M. (2003). Estudio sociolingüístico de El Hierro. Universidad de Las Palmas

Pisoni, D.B. (1997). Some thoughts on "normalization" in speech perception. In K. Johnson and J.W. Mullennix (Eds.), *Talker variability in Speech Processing* (pp.9-32). San Diego: Academic Press.

Podesva, Robert J. (2007) Phonation type as a stylistic variable: The use of falsetto in constructing a persona. *Journal of Sociolinguistics 11*, 478–504.

Podesva, R., Reynolds, J. J., Callier, P. & Baptiste, J. (2015). Constraints on the social meaning of released /t/: A production and perception study of US politicians. *Language Variation and Change 27*, 59–87.

Pollock, M. (2020). Did you say peso or beso?: The perception of prevoicing by L2 Spanish learners. In Sessarego, S., Colomina-Almiñana, J. J. & Rodríguez-Riccelli, A. (Eds.), *Variation and Evolution: Aspects of language contact and contrast across the Spanish-speaking world*, pp. 127-161. Amsterdam: John Benjamins Publishing.

Pollock, M (2022). *Buenas no[t͡ʃ]es y mu[t͡s]ísimas gracias: Variable Affricate Production in Peninsular Spanish Political Discourse.* Presented at the 10th International Workshop on Spanish Sociolinguistics (WSS).

Pollock, M. (2023, submitted). Toeing the Party Line: Indexing Party Identity through Dialectal Phonetic Features in Spanish Political Discourse. *Languages.*

Pollock, M. & Wheeler, J. A. (2022). Syllable-final /s/ and intervocalic /d/ elision in Andalusia: The Formation of Susana Díaz's Regional Identity in Political Discourse. *Language and Communication 87*, 191-204.

Pollock, M. & Wheeler, J. (2023). Regional variation and speaker design: A sociophonetic study of Galician political norms in Castellano [Oral Conference Presentation]. *Forging Linguistic Identities, Towson University*.

Pollock, M. (in preparation). *La variación fonológica del español panameño.* In Díaz-Campos, M. & Hernández-Campoy, J. M. (Eds.) *Enciclopedia Concisa de los Dialectos del Español*. London: Wiley-Blackwell.

Pollock, M., Delgado-Díaz, G., Galarza, I., Díaz-Campos, M., & Willis, E. W. (in press). The emergence of sound change in two varieties of Spanish: Variable trill /r/ production in Puerto Rico and Venezuela. *Proceedings of the 2021 Hispanic Linguistics Symposium*.

Preston, D. R. (1989). *Sociolinguistics and second language acquisition*. Oxford: Basil Blackwell.

Preston, D. R. (1993). Folk dialectology. In D. R. Preston (Ed.), *American dialect research*, (pp. 333– 377). Philadelphia: John Benjamins.

Preston, D. R. (1999). *Handbook of perceptual dialectology*. Amsterdam: John Benjamins.

RAE/Real Academia Española (2011). *Nueva gramática de la lengua española: Fonética fonología.* Madrid: Espasa.

Rama, J., Zanotti, L., Turnbull-Dugarte, S. J., & Santana, A. (2021). *Vox: The Rise of the Spanish Populist Radical Right*. London: Routledge.

Ranson, D. L. (1991). Person marking in the wake of /s/ deletion in Andalusian Spanish. *Language Variation and Change, 3*(2), 133-152.

Reetz, H. & Jongman, A. (2009). *Phonetics: Transcription, production, acoustics and perception.* Wiley-Blackwell.

Regan, B. (2017a). study of *Ceceo* variation in Western Andalusia (Huelva). *Studies in Hispanic and Lusophone Linguistics, 10*(1), 119-160.

Regan, B. (2017b). *The Effect of Dialect Contact and Social Identity on Fricative Demerger*. [Doctoral dissertation, The University of Texas at Austin]. UTA Campus Repository. http://hdl.handle.net/2152/62339

Regan, B. (2019). Dialectology meets sociophonetics: The social evaluation of *ceceo* and *distinción* in Lepe, Spain. In Chappell, W. (Ed.), *Recent Advances in the Study of Spanish Sociophonetic Perception*, pp. 85-124. Amsterdam: John Benjamins.

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 Geographic 8*, 82-101.

Rickford, J. R. & King, S. (2016). Language and linguistics on trial: Hearing Rachel Jeantel (and other vernacular speakers) in the courtroom and beyond. *Language 92*(4), 948-988.

Rickford, J. R. & McNair-Knox, F. (1994). Addressee- and topic-influenced style shift: A quantitative sociolinguistic study. In Biber, D. & Finegan, E. (eds.), *Sociolinguistic perspectives on register*, pp. 235–76. New York: Oxford University Press.

Ruch, H. (2012). Affrication of /st/-clusters in Western Andalusian Spanish: variation and change from sociophonetic point of view. In S. Calamai, C. Celata, L. Ciucci (Eds.), *Proceedings of Sociophonetics, at the crossroads of speech variation, processing and communication* (pp. 61-64). Pisa: Edizioni della Normale.

Ruch, H. & Harrington, J. (2014). Synchronic and diachronic factors in the change from preaspiration to post-aspiration in Andalusian Spanish. *Journal of Phonetics 45*, 12-25.

Ruiz-Domínguez, M. d. M. (1997). *Estudio sociolinguistico del habla de Melilla*. [Doctoral dissertation, Universidad de Alcalá].

Ruiz-Martínez, A. M. (2003). *Estudio fonético del nordeste de la Comunidad de Madrid*. Alcalá de Henares: Universidad de Alcalá.

Ruiz-Peña, M. E. (2013). *"Alma" o "arma", evidencia de la neutralización /l/ /r/ en la variedad dialectal andaluza de Sevilla*. Unpublished MA Thesis, University of Western Ontario.

Ruiz-Sánchez, C. (2017). *Seseo, ceceo*, and distinción in Andalusian Spanish: Free variation or sociolinguistic variation? *Linguistics Vanguard 3*(1), 20160075.

Saito, Y. & Yurong-Maekawa, K. (2019). An investigation into Modern Mongolian vowel harmony using real-time magnetic resonance imaging. In *Proceedings of the 19th ICPhS, Melbourne*.

Samper-Padilla, J. A. (2011). Socio-phonological Variation and change in Spain. In Manuel Díaz Campos (ed.) *The Handbook of Hispanic Sociolinguistics*, pp. 98-120. Malden, MA: Wiley-Blackwell.

Samper-Padilla, J. A. & Hernández-Cabrera, C. (1995). La variación de -/s/ en el español culto de Las Palmas de Gran Canaria. *Philologica Canariensia 2*, 391–408.

Santana-Marrero, J. (2017). Variación de las realizaciones de /θs/ en el sociolecto bajo de la ciudad de Sevilla: datos de PRESEEA-SE. *Lingüística en la red 15*, 1-17.

Saussure, F. (1966). In Bally, C., & Sechehaye, A. (Eds.). *Course in general linguistics* (3rd ed.). New York: McGraw-Hill.

Schilling-Estes, N. (1998). Investigating self-conscious speech: The performance register in Ocracoke English. *Language in Society 27*, 53-83.

Schilling-Estes, N. (2002). Investigating Stylistic Variation. In J. K. Chambers, & N. Schilling-Estes (Eds.) *The Handbook of Language Variation and Change*, pp. 375-401. Malden, MA: Blackwell Publishing.

Schilling-Estes, N. (2013). Investigating stylistic variation. In J. K. Chambers, P. Trudgill & N. Schilling-Estes (eds.), *The handbook of language variation and change*, pp. 1–23. Oxford: Basil Blackwell.

Schoonmaker-Gates, E. (2015). Measuring Foreign Accent in Spanish: How Much Does VOT Really Matter? In Erik W. Willis, P. Martin Butragueño, E. Herrera Zendejas (eds.) *Selected Proceedings of the 6th Conference on Laboratory Approaches to Spanish Phonetics and Phonology*, pp. 95-105. Somerville, MA: Cascadilla Proceedings Project.

Sclafani, J. (2017). *Talking Donald Trump*. London: Routledge.

Sempere, J. A. (2014). Geographic and Sociolinguistic Variables in the *seseo* of Murcia, in L. Callahan (ed.) *Spanish and Portuguese Across Time, Place, and Borders: Studies in Honour of Milton M. Azevedo*, pp. 145-160. London: Palgrave Macmillan Limited.

Sharma, D. (2018). Style dominance: Attention, audience, and the 'real me.' *Language in Society 47*, 1-31.

Sharma, D. & Rampton, B. (2015). Lectal focusing in interaction: A new methodology for the study of style variation. *Journal of English Linguistics 43*(1), 3–35.

Shertz, J., Kang, Y., Kochetov, A., Kong, E. J. & Han, S. (2015). Dialectal variability in place and manner of Korean affricates. In *Proceedings of the 18th ICPhS, Glasgow*.

Silverstein, M. (2003). Indexical Order and The Dialectics of Sociolinguistic life. *Language & Communication 23*, 193-229.

Snell, J. (2018). Solidarity, stance, and class identities. *Language in Society 47*, 665-691.

Solon, M. (2015). L2 Spanish /l/: The Roles of F2 and Segmental Duration in Foreign Accent Perception. In E. W. Willis, P. Martín-Butragueño, E. Herrera (Eds.), *Selected Proceedings of the 6th Conference on Laboratory Approaches to Romance Phonology*, pp. 83-94. Somerville, MA: Cascadilla Proceedings Project.

Soriano, B. (2012). Andalusian vowel harmony and morphology-phonology interface. In Campos-Astorkiza, R. Franco, J. (Eds.) *Papers in linguistics by the BIDE generation, Anuario del Seminario de Filología Vasca <<Julio de Urquijo>> XLVI-1*, pp. 295-307.

Stewart, M., Ryan, E. B., & Giles, H. (1985). Accent and social class effects on status and solidarity evaluations. *Personality and Social Psychology Bulletin 11*(1), 98–105.

Tagliamonte, S. & Molfenter, S. (2007). How'd you get that accent? Acquiring a second dialect of the same language. *Language in Society 36*(5), 649–75.

Tagliamonte, S. A. (2006). *Analysing Sociolinguistic Variation*. Cambridge UP.

Tagliamonte, S. A. (2011). *Variationist Sociolinguistics: Change, Observation, Interpretation*. Wiley-Blackwell.

Thomas, E. (2011). *Sociophonetics: An introduction*. New York: Palgrave Macmillan.

Torrano-Moreno, L. (2017). Variación cambio sociolingüístico en una comunidad de habla local: el caso de la variable (ch) los factores ontogenético generolectal en ricote. *Tonos Digital 33*(2). Accessed from https://www.um.es/tonosdigital/znum33/indice33.htm

Torreira, F. (2006). Coarticulation between aspirated-s and voiceless stops in Spanish: An interdialectal comparison. In N. Sagarra A. J. Toribio (Eds.), *Selected proceedings of the 9th Hispanic Linguistics Symposium*, pp. 113-120. Somerville, MA: Cascadilla Press.

Torreira, F. (2007). Pre- and postaspirated stops in Andalusian Spanish. In J. Mascaró, M. J. Solé P. Prieto (Eds.) *Segmental and prosodic issues in Romance phonology*, pp. 67-82. Philadelphia: John Benjamins.

Torreira, F. (2012). Investigating the nature of aspirated stops in Western Andalusian Spanish. *Journal of the International Phonetic Association 42*(1), 49-63.

Torreira, F. & Ernestus, M. (2012). Weakening of intervocalic /s/ in the Nijmegen Corpus of Casual Spanish. *Phonetica 69*(3), 124-148.

Trudgill, P. J. (1972). Sex, covert prestige and linguistic change in the urban British English of Norwich. *Language in Society 1*, 179–195.

Valeš, M. (2013). Patrones de prestigio y cambio lingüístico en Granada. *Études romanes de Brno 30*(2), 231-240.

Valeš, M. (2014). Panorama de la sociolingüística andaluza. *Linguistica Pragensia 24*(1), 45-54.

Vida, M. (2004). *Estudio sociofonológico del español hablado en Málaga*. Alicante: Universidad de Alicante.

Villena-Ponsoda, J. A. (1996). Convergence and divergence in a standard-dialect continuum: Networks and individuals in Málaga, *Sociolinguistica 10*, 112–137.

Villena-Ponsoda, J. A. (2007). Interacción de factores internos y externos en la explicación de la variación fonológica: Análisis multivariante del patrón de pronunciación no sibilante [θ] de la consonante fricativa coronal /θs/ en el español hablado en Málaga, in J. A. Moya & M. Sosinski (eds.) *Las hablas andaluzas y la enseñanza de la lengua. Actas de las XII Jornadas sobre la Enseñanza de la Lengua Española*, pp.69-97. Granada: Universidad de Granada.

Villena-Ponsoda, J. A. (2008). Sociolinguistic patterns of Andalusian Spanish. *International Journal of the Sociology of Language*, 139-160.

Villena-Ponsoda, J. A. (2013). Actos de identidad: ¿Por qué persiste el uso de los rasgos lingüísticos de bajo prestigio social? Divergencia geográfica y social en el español urbano de Andalucía. In R. Guillén Stuil and R. Millán Garrido (Eds.) *Estudios descriptivos y aplicados sobre el andaluz*, pp. 173-207. Seville: The University of Seville Press.

Villena-Ponsoda, J. A. & Moya-Corral, J. A. (2016) Variation of intervocalic /d/ in two speech communities (Granada and Málaga). A comparative analysis of an erosive phonological change. *Boletín de Filología 51*(2), 281-321.

Villena-Ponsoda, J. A., & Ávila-Muñoz, A. M. (2014). Dialect stability and divergence in southern Spain. Stability and Divergence in Language Contact. *Factors and Mechanisms*, 207-238.

Villena-Ponsoda, J. A., Montesinos, F. D., Muñoz, A. M., & Cervantes, M. d. l. C. (2011). Interacción de factores fonéticos y gramaticales en la variación fonológica: La elisión de /d/ intervocálica en la variedad de los hablantes universitarios en la ciudad de Málaga. In *Variación lingüística y contacto de lenguas en el mundo hispánico*: in memoriam Manuel Alvar. Iberoamericana. 311-359.

Walker, A., García, C., Cortés, Y., & Campbell-Kibler, K. (2014). Comparing social meanings across listener and speaker groups: The indexical field of Spanish /s/. *Language Variation and Change 26*(2), 169-189.

Weissglass, C. (2011). An acoustic study of rhotics in onset clusters in La Rioja. In J. Herschensohn (Ed.) *Romance Linguistics 2010: Selected papers from the 40th Linguistic Symposium on Romance Languages*, pp. 49-62. Amsterdam: John Benjamins.

Weissglass, C. (2015). The production of rhotics in onset clusters by Spanish monolinguals and Spanish-Basque bilinguals. In J. Romero & M. Riera (Eds.) *The Phonetics–Phonology Interface: Representations and methodologies*, pp. 193-207. Amsterdam: John Benjamins

Willis, E. W. & Díaz-Campos, M. (2021). The perception of coda liquids in Dominican Spanish: Geographic and social variation. In L. A. Ortiz-López & E. M. Suárez-Büdenbender (eds.) Topics in Spanish Linguistic Perceptions, pp. 35-53. New York: Taylor & Francis.

Willis, E. W. (2006). Trill Variation in Dominican Spanish: An Acoustic Examination and Comparative Analysis. In Sagarra, N. & Toribio, A. J., *Selected Proceedings of the 9th Hispanic Linguistics Symposium*, pp. 121-131. Somerville, MA: Cascadilla Proceedings Project.

Willis, E. W., Delgado-Díaz, G., & Galarza, I. (2015). Contextualized Voicing of the Voiceless Posterior Fricative /h/ in Puerto Rican Spanish. In Erik W. Willis, P. Martín Butragueño, E. Herrera Zendejas (eds.). *Selected Proceedings of the 6th Conference on Laboratory Approaches to Spanish Phonetics and Phonology*, pp. 52-69. Somerville, MA: Cascadilla Proceedings Project.

Woolard, K. (1985). Language variation and cultural hegemony. *American Ethnologist 12*, 738–48.

Zahler, S., & Daidone, D. (2014). A variationist account of trill /r/ usage in the Spanish of Málaga. IULC Working Papers, 14(2).

Zampini, M. L. (1998). The relationship between the production and perception of L2 Spanish stops. *Texas Papers in Foreign Language Education 3*(3), 85-100.

Zhang, Q. (2005). A Chinese yuppie in Beijing: Phonological variation and the construction of a new professional identity. *Language in Society 34*, 431-466.

Zimman, L. (2016). Sociolinguistic Agency and the Gendered Voice: Metalinguistic Negotiations of Vocal Masculinization among Female-to-Male Transgender Speakers. In Babel, A. (ed). *Awareness and control in sociolinguistic research*, pp. 253-277. Cambridge University Press.

Zimman, L. (2017). Gender as stylistic bricolage: Transmasculine voices and the relationship between fundamental frequency and /s/. *Language in Society 46*, 339-370.

Appendices

Appendix A: Political corpus information

Table A: *List of politicians and video sources examined in the project*

ID	Name	Gender	City	Party	Context	Age	Video Year	Min.	Channel
1	Pedro Sanchez	Male	Madrid	PSOE	Discourse	46	2018	27.5	Asamblea de Naciones Unidas
					Female	44	2016	18.5	Cadena Ser: Hoy por hoy
					Male	46	2018	21	elDiario
2	Manuel Robles Delgado	Male	Madrid	PSOE	Discourse	65	2018	20.5	Ayuntamiento de Fuenlabrada
					Female	58	2011	20	Cadena Ser: Madrid
					Male	64	2017	24	Fuenlabrada noticias
3	Cristina Narbona	Female	Madrid	PSOE	Discourse	66	2017	28.5	Forum Europa: International European Economy Forum
					Female	66	2017	24.5	La Cerca
					Male	67	2018	28.5	iAgua
4	María Elena Valenciano	Female	Madrid	PSOE	Discourse	53	2013	23	PSOE
					Female	58	2018	28.5	HuffPost
					Male	53	2013	11.5	Las Mañanas Cuatro
5	José Antonio Pérez Tapias	Male	Seville	PSOE	Discourse	59	2014	15	Rueda de Prensa: Canal13
					Female	59	2014	9.5	Trece: La MariMorena
					Male	59	2014	11	Cadiz Directo
6	Antonio Pradas Torres	Male	Seville	PSOE	Discourse	52	2015	26.5	Forum Europa: International European Economy Forum
					Female	53	2016	13.5	El CorreoTV
					Male	53	2016	31.5	CanalSur: Buenos Días
7	Susana Díaz	Female	Seville	PSOE	Discourse	41	2015	23.5	Inauguration
					Female	44	2018	27	CanalSur: Buenos Días
					Male	43	2017	34.5	CanalSur: Hora de Andalucía
8	Verónica Pérez Fernández	Female	Seville	PSOE	Discourse	37	2015	20	Convivencia Navideña PSOE 2015
					Female	35	2013	21.5	CanalSur
					Male	39	2017	34.5	Por Todas: 8 TV Andalusia
9	Francisco Javier Arenas Vacas	Male	Córdoba	PSOE	Discourse	56	2013	14	PSOE: Rueda de Prensa
					Male	57	2014	33.5	Cadena Ser: Hoy por hoy
10	Antonio Ruiz Cruz	Male	Córdoba	PSOE	Discourse	46	2016	36	Europa Press Andalucía: Diputación de Córdoba
					Female	46	2016	21	PTV Córdoba
					Male	48	2018	4	Ondamezquita TV
11	Carmen Calvo Poyato	Female	Córdoba	PSOE	Discourse	61	2018	32	Acto: Andalucia PSOE
					Female	61	2018	34.5	TV3
					Male	61	2018	11.5	Cadena COPE

12	María Isabel Ambrosio Palos	Female	Córdoba	PSOE	Discourse	46	2016	37	Ondamezquita TV
					Female	47	2017	14.5	CanalSur: Buenos Días
					Male	48	2018	26.5	CanalSur: Buenos Días
13	Pío García-Escudero	Male	Madrid	PP	Discourse	61	2013	12	Congreso internacional de arquitectura y sociedad
					Female	59	2011	10.5	Zaragoza Television
					Male	59	2011	9.5	Castilla y Leon Television
14	Ángel Garrido	Male	Madrid	PP	Discourse	54	2018	31	Forum Europa: International European Economy Forum
					Female	54	2018	21	Un Café/Antena3
					Male	54	2018	13	esRadio: Es la mañana
15	Cristina Cifuentes	Female	Madrid	PP	Discourse	53	2017	29	Forum Europa: International European Economy Forum
					Female	52	2016	35	Madridiario
					Male	53	2017	42.5	OKDiario
16	Esperanza Aguirre	Female	Madrid	PP	Discourse	63	2015	24	Agencia EFE: Forum de elecciones
					Female	63	2015	38.5	Antena 3 noticias
					Male	66	2018	29	Intereconomía: El Gato Al Agua
17	Juan Manuel Moreno Bonilla	Male	Seville	PP	Discourse	48	2018	28	Forum Europa:Tribuna Andalucia
					Female	47	2017	25	CanalSur
					Male	47	2017	16	De Cerca: Málaga 24
18	Juan Ignacio Zoido Álvarez	Male	Seville	PP	Discourse	60	2017	41.5	GranCanariaTV: Comision de Interior
					Female	60	2017	22	CanalSur
					Male	60	2017	47	esRadio: Es la mañana
19	Teresa Jiménez-Becerril	Female	Seville	PP	Discourse	55	2016	17	Comisión Parlamentaria Mixta México-Unión Europea
					Female	53	2014	4	Trece: Más claro agua
					Male	55	2016	20	esRadio: Castilla y Leon
20	María Esperanza Oña	Female	Seville	PP	Discourse	55	2012	27.5	PP Málaga
					Female	61	2018	30.5	73TV Andalucia
					Male	59	2016	31	PTV Málaga
21	José Nieto Ballesteros	Male	Córdoba	PP	Discourse	47	2017	29.5	Europa Press
					Female	45	2015	42.5	PTV Córdoba
					Male	45	2015	36	PTV Córdoba
22	Adolfo Molina Rascón	Male	Córdoba	PP	Discourse	40	2015	6	LaCarlota: Rueda de Prensa
					Female	42	2017	17.5	Mira Córdoba
					Male	40	2015	19	Ondamezquita TV
23	María Luisa Ceballos	Female	Córdoba	PP	Discourse	46	2014	2.5	PP Priego: City Council Meeting
					Female	43	2011	13	PTV Córdoba
					Male	47	2015	10	E24M
24		Female	Córdoba	PP	Discourse	34	2017	10.5	Partido Popular

	María Beatríz Jurado Fernández				Female	34	2017	13	CongresoPP
					Male	33	2016	25.5	Ondamezquita TV
25	José Luis Ruiz Espejo	Male	Málaga	PSOE	Discourse	53	2018	11	Málaga 24hr: Rueda de Prensa
					Female	53	2018	28.5	Málaga 24hr: Entevista
					Male	53	2018	41.5	Antequera TV
26	José Bernal Gutiérrez	Male	Málaga	PSOE	Discourse	43	2017	13.5	PSOE Marbella
					Female	43	2017	22	OndaCero Marbella
					Male	45	2019	29	RTV Marbella
27	Fuensanta Lima Cid	Female	Málaga	PSOE	Discourse	49	2011	14.5	PSOE Mijas
					Female	52	2014	9.5	Radio Costa del Sol
28	Patricia Alba Luque	Female	Málaga	PSOE	Discourse	37	2017	27	PSOE Alhaurin de la Torre
					Female	38	2018	5.5	PTV Málaga
					Male	33	2013	49	PTV Málaga
29	Elías Bendodo Benasayag	Male	Málaga	PP	Discourse	45	2019	42	Forum Europa: International European Economy Forum
					Female	41	2015	14	Radio Television Antequera
					Male	43	2017	13.5	De Cerca: Málaga 24
30	José Francisco Salado Escaño	Male	Málaga	PP	Discourse	54	2019	22.5	Diputación de Málaga
					Female	53	2018	9	Gran Azarquía: El rincón habla
					Male	46	2011	30.5	Programa A Fondo: 24FM/Visuales Axarquia
31	Celia Villalobos Talero	Female	Málaga	PP	Discourse	62	2011	28.5	Forum Europa: International European Economy Forum
					Female	67	2016	24	OKDiario
					Male	69	2018	26	Radio 5: RNE (Radio nacional de españa) La Entrevista
32	Margarita del Carmen del Cid Muñoz	Female	Málaga	PP	Discourse	46	2019	6.5	Torremolinos TV
					Female	43	2016	7.5	Torremolinos TV
					Male	46	2019	32.5	TV: Costa del Sol

Appendix B: Praat Script for data collection

```
## Below: user provides directory (the default below is the path for my own desktop;
## you will probably want to change that), initial substring of filename (or complete
## filename minus the extension), the extension (default is .wav), and one or more tiers

form Select directory, file type, and tiers
	sentence Directory C:\Users\mbpol\Desktop\Sounds\
 sentence Extension wav
 sentence Tier(s) t1 t2 t3 t4 t5 t6
endform

Create Strings as file list... list 'directory$'*.'extension$'
file_count = Get number of strings

## Loop through files and make grids (this section partly inspired by code by Katherine
Crosswhite)

for k from 1 to file_count
 select Strings list
 current$ = Get string... k
 Read from file... 'directory$''current$'
 short$ = selected$ ("Sound")

## Below: look for grid, if found, open it, otherwise make new one
## This section inspired by code by Jen Hay

 full$ = "'directory$''short$'.TextGrid"
 if fileReadable (full$)
	Read from file... 'full$'
	Rename... 'short$'
 else
	select Sound 'short$'
	To TextGrid... "'tier$'"
 endif

## End Jen Hay inspired block

 plus Sound 'short$'

#TIER
2##############################################################################
############################

	resultfile$ = "C:/Users/mbpol/Desktop/Sounds/2CHCOG.txt"
```

```
	header_row$ = “SoundFile” + tab$ + “Word” + tab$ + “Sound” + tab$ + “Duration” + tab$
+ “COG” + newline$
	sn$ = selected$ (“Sound”)
	select Sound 'sn$'
	select TextGrid 'sn$'

	numint = Get number of intervals... 2
	for i from 1 to numint
		select TextGrid 'sn$'
		label$ = Get label of interval... 2 'i'
		if label$ <> “““
			start = Get starting point... 2 'i'
			end = Get end point... 2 'i'
			midpoint = start + ((end - start) / 2)
			select TextGrid 'sn$'
			wordint = Get interval at time... 1 'midpoint'
			select TextGrid 'sn$'
			wordlab$ = Get label of interval... 1 'wordint'

			startWord = Get starting point... 1 'wordint'
			endWord = Get end point... 1 'wordint'
			prevword = startWord - ((end-start)/100000000)
			follword = endWord + ((end-start)/100000000)
			prevint = Get interval at time... 1 'prevword'
			follint = Get interval at time... 1 'follword'
			select TextGrid 'sn$'
			prevwordlab$ = Get label of interval... 1 'prevint'
			follwordlab$ = Get label of interval... 1 'follint'

			startone = Get starting point... 1 'wordint'
			endone = Get end point... 1 'wordint'
			duration = (end - start)
			p2 = start + (1*((end - start) / 4))
			half = ((end - start) / 2)
			duration = (end - start) * 1000
			durationms = (end - start) * 1000
			tp = p2
			tpe = half
			tpn = 23
			call peakmeasure
			endif
	endfor

	procedure peakmeasure
		storeda = 0
		storedf = 0
```

```
		select Sound 'sn$'
		Edit
			editor Sound 'sn$'
			Spectrogram settings... 0 12000 0.05 50
			Move cursor to... 'tp'
			Move end of selection by... 'tpe'
			View spectral slice
			Close
		endeditor
		slice$ = selected$ (“Spectrum”)
		select Spectrum 'slice$'
		cog = Get centre of gravity... 2
		select Spectrum 'slice$'
		To Ltas (1-to-1)
		ltas$ = selected$ (“Ltas”)
		select Ltas 'ltas$'

		result_row$ = “'sn$'“ + tab$ + “'prevwordlab$'“ + tab$ + “'wordlab$'“ + tab$ +
“'follwordlab$'“ + tab$ + “'label$'“ + tab$ + “'duration'“ + tab$ + “'cog'“ + tab$ + “'start'“ + tab$
+ newline$
		fileappend “'resultfile$'“ 'result_row$'
		select all
		minus Sound 'sn$'
		minus TextGrid 'sn$'
		minus Strings list
		Remove
		select all
	endproc

	resultfile$ = “C:/Users/mbpol/Desktop/Sounds/2CHFormants.txt”
	header_row$ = “SoundFile” + tab$ + “Word” + tab$ + “Sound” + tab$ + “F1” + tab$ +
“F2” + tab$ + “TimeStamp” + newline$
	select Sound 'sn$'
	select TextGrid 'sn$'

	numint = Get number of intervals... 2
	select Sound 'sn$'
	To Formant (burg)... 0 5 5500 0.025 50
	for i from 1 to numint
		select TextGrid 'sn$'
		label$ = Get label of interval... 2 'i'
		if label$ <> “““
			start = Get starting point... 2 'i'
			end = Get end point... 2 'i'
			fifthpoint = start + ((end - start) / 5)
			select TextGrid 'sn$'
```

```
                        wordint = Get interval at time... 1 'fifthpoint'
                        select TextGrid 'sn$'
                        wordlab$ = Get label of interval... 1 'wordint'
                        select Formant 'sn$'
                        f1 = Get value at time... 1 'fifthpoint' Hertz Linear
                        f2 = Get value at time... 2 'fifthpoint' Hertz Linear
                        result_row$ = “'sn$'“ + tab$ + “'wordlab$'“ + tab$ + “'label$'“ + tab$ +
“'f1'“ + tab$ + “'f2'“ + tab$ + “'start'“ + newline$
                        fileappend “'resultfile$'“ 'result_row$'
                endif
        endfor
        select all
        minus Sound 'sn$'
        minus TextGrid 'sn$'
        minus Strings list
        Remove
        select all

#TIER
3##############################################################################
############################

        resultfile$ = “C:/Users/mbpol/Desktop/Sounds/3DIntensity.txt”
        select Sound 'sn$'
        select TextGrid 'sn$'

        numint = Get number of intervals... 3
        select Sound 'sn$'
        To Intensity... 100 0
        for i from 1 to numint
                select TextGrid 'sn$'
                label$ = Get label of interval... 3 'i'
                if label$ <> “““
                        # calculates the onset and offset
                        start = Get starting point... 3 'i'
                        end = Get end point... 3 'i'
                        half = start + ((end - start) / 2)
                        wordint = Get interval at time... 1 'half'
                        wordlab$ = Get label of interval... 1 'wordint'

                        startWord = Get starting point... 1 'wordint'
                        endWord = Get end point... 1 'wordint'
                        prevword = startWord - ((end-start)/100000000)
                        follword = endWord + ((end-start)/100000000)
                        prevint = Get interval at time... 1 'prevword'
                        follint = Get interval at time... 1 'follword'
```

```
                        select TextGrid 'sn$'
                        prevwordlab$ = Get label of interval... 1 'prevint'
                        follwordlab$ = Get label of interval... 1 'follint'

                        duration = (end - start)
                        #calculates the intensity values
                        select Intensity 'sn$'
                        min_int = Get minimum... start end Parabolic
                        min_time = Get time of minimum... start end Parabolic
                        max_int = Get maximum... start end Parabolic
                        max_time = Get time of maximum... start end Parabolic
                        meanIntensity = Get mean... start end dB
                        range = (max_int - min_int)
                        result_row$ = “'sn$'“ + tab$ + “'prevwordlab$'“ + tab$ + “'wordlab$'“ +
tab$ + “'follwordlab$'“ + tab$ + “'label$'“ + tab$ + “'duration'“ + tab$ + “'start'“ + tab$ +
“'meanIntensity'“ + tab$ + “'min_int'“ + tab$ + “'max_int'“ + tab$ + “'range'“ + tab$ + newline$
                        fileappend “'resultfile$'“ 'result_row$'
                endif
        endfor

        select all
        minus Sound 'sn$'
        minus TextGrid 'sn$'
        minus Strings list
        Remove
        select all

#TIER
4##############################################################################
############################

        resultfile$ = “C:/Users/mbpol/Desktop/Sounds/4SibsCOG.txt”
        header_row$ = “SoundFile” + tab$ + “Word” + tab$ + “Sound” + tab$ + “Duration” + tab$
+ “COG” + newline$
        select Sound 'sn$'
        select TextGrid 'sn$'

        numint = Get number of intervals... 4
        for i from 1 to numint
                select TextGrid 'sn$'
                label$ = Get label of interval... 4 'i'
                if label$ <> “““
                        start = Get starting point... 4 'i'
                        end = Get end point... 4 'i'
                        midpoint = start + ((end - start) / 2)
                        select TextGrid 'sn$'
```

```
                    wordint = Get interval at time... 1 'midpoint'
                    select TextGrid 'sn$'
                    wordlab$ = Get label of interval... 1 'wordint'

                    startWord = Get starting point... 1 'wordint'
                    endWord = Get end point... 1 'wordint'
                    prevword = startWord - ((end-start)/100000000)
                    follword = endWord + ((end-start)/100000000)
                    prevint = Get interval at time... 1 'prevword'
                    follint = Get interval at time... 1 'follword'
                    select TextGrid 'sn$'
                    prevwordlab$ = Get label of interval... 1 'prevint'
                    follwordlab$ = Get label of interval... 1 'follint'

                    startone = Get starting point... 1 'wordint'
                    endone = Get end point... 1 'wordint'
                    duration = (end - start)
                    p2 = start + (1*((end - start) / 4))
                    half = ((end - start) / 2)
                    duration = (end - start) * 1000
                    durationms = (end - start) * 1000
                    tp = p2
                    tpe = half
                    tpn = 23
                    result_row$ = “'sn$'“ + tab$ + “'prevwordlab$'“ + tab$ + “'wordlab$'“ +
tab$ + “'follwordlab$'“ + tab$ + “'start'“ + tab$ + “'label$'“ + tab$ + “'duration'“ + tab$
                    fileappend “'resultfile$'“ 'result_row$'
                    call peakmeasure
                    endif
          endfor

          procedure peakmeasure
                    storeda = 0
                    storedf = 0
                    select Sound 'sn$'
                    Edit
                              editor Sound 'sn$'
                              Spectrogram settings... 0 12000 0.05 50
                              Move cursor to... 'tp'
                              Move end of selection by... 'tpe'
                              View spectral slice
                              Close
                    endeditor
                    slice$ = selected$ (“Spectrum”)
                    select Spectrum 'slice$'
                    cog = Get centre of gravity... 2
```

```
		select Spectrum 'slice$'
		To Ltas (1-to-1)
		ltas$ = selected$ (“Ltas”)
		select Ltas 'ltas$'

		result_row2$ = “'cog'“ + newline$
		fileappend “'resultfile$'“ 'result_row2$'
		select all
		minus Sound 'sn$'
		minus TextGrid 'sn$'
		minus Strings list
		Remove
		select all
	endproc


#TIER 5 Formants##################################################################################################################################

	resultfile$ = “C:/Users/mbpol/Desktop/Sounds/5LIQsFormants.txt”
	header_row$ = “SoundFile” + tab$ + “Word” + tab$ + “Sound” + tab$ + “F1” + tab$ + “F2” + tab$ + “F3” + tab$ + “TimeStamp” + newline$
	select Sound 'sn$'
	select TextGrid 'sn$'

	numint = Get number of intervals... 5
	select Sound 'sn$'
	To Formant (burg)... 0 5 5500 0.025 50
	for i from 1 to numint
		select TextGrid 'sn$'
		label$ = Get label of interval... 5 'i'
		if label$ <> “““
			start = Get starting point... 5 'i'
			end = Get end point... 5 'i'
			duration = (end - start)
			onetenth = start + ((end - start) / 10)
			twotenth = start + (((end - start) / 10)*2)
			threetenth = start + (((end - start) / 10)*3)
			fourtenth = start + (((end - start) / 10)*4)
			fivetenth = start + (((end - start) / 10)*5)
			sixtenth = start + (((end - start) / 10)*6)
			sevententh = start + (((end - start) / 10)*7)
			eighttenth = start + (((end - start) / 10)*8)
			ninetenth = start + (((end - start) / 10)*9)
			tententh = start + (((end - start) / 10)*10)
```

```
select TextGrid 'sn$'
wordint = Get interval at time... 1 'fivetenth'
select TextGrid 'sn$'
wordlab$ = Get label of interval... 1 'wordint'

startWord = Get starting point... 1 'wordint'
endWord = Get end point... 1 'wordint'
prevword = startWord - ((end-start)/100000000)
follword = endWord + ((end-start)/100000000)
prevint = Get interval at time... 1 'prevword'
follint = Get interval at time... 1 'follword'
select TextGrid 'sn$'
prevwordlab$ = Get label of interval... 1 'prevint'
follwordlab$ = Get label of interval... 1 'follint'

select Formant 'sn$'
f11 = Get value at time... 1 'onetenth' Hertz Linear
f21 = Get value at time... 2 'onetenth' Hertz Linear
f31 = Get value at time... 3 'onetenth' Hertz Linear
f12 = Get value at time... 1 'twotenth' Hertz Linear
f22 = Get value at time... 2 'twotenth' Hertz Linear
f32 = Get value at time... 3 'twotenth' Hertz Linear
f13 = Get value at time... 1 'threetenth' Hertz Linear
f23 = Get value at time... 2 'threetenth' Hertz Linear
f33 = Get value at time... 3 'threetenth' Hertz Linear
f14 = Get value at time... 1 'fourtenth' Hertz Linear
f24 = Get value at time... 2 'fourtenth' Hertz Linear
f34 = Get value at time... 3 'fourtenth' Hertz Linear
f15 = Get value at time... 1 'fivetenth' Hertz Linear
f25 = Get value at time... 2 'fivetenth' Hertz Linear
f35 = Get value at time... 3 'fivetenth' Hertz Linear
f16 = Get value at time... 1 'sixtenth' Hertz Linear
f26 = Get value at time... 2 'sixtenth' Hertz Linear
f36 = Get value at time... 3 'sixtenth' Hertz Linear
f17 = Get value at time... 1 'sevententh' Hertz Linear
f27 = Get value at time... 2 'sevententh' Hertz Linear
f37 = Get value at time... 3 'sevententh' Hertz Linear
f18 = Get value at time... 1 'eighttenth' Hertz Linear
f28 = Get value at time... 2 'eighttenth' Hertz Linear
f38 = Get value at time... 3 'eighttenth' Hertz Linear
f19 = Get value at time... 1 'ninetenth' Hertz Linear
f29 = Get value at time... 2 'ninetenth' Hertz Linear
f39 = Get value at time... 3 'ninetenth' Hertz Linear
f110 = Get value at time... 1 'tententh' Hertz Linear
f210 = Get value at time... 2 'tententh' Hertz Linear
f310 = Get value at time... 3 'tententh' Hertz Linear
```

```
                        result_row$ = “'sn$'“ + “'prevwordlab$'“ + tab$ + “'wordlab$'“ + tab$ +
“'follwordlab$'“ + tab$ + “'label$'“ + tab$ + “'duration'“ + tab$ + “'f11'“ + tab$ + “'f21'“ + tab$ +
“'f31'“ + tab$ + “'f12'“ + tab$ + “'f22'“ + tab$ + “'f32'“ + tab$ + “'f13'“ + tab$ + “'f23'“ + tab$ +
“'f33'“ + tab$ + “'f14'“ + tab$ + “'f24'“ + tab$ + “'f34'“ + tab$ + “'f15'“ + tab$ + “'f25'“ + tab$ +
“'f35'“ + tab$ + “'f16'“ + tab$ + “'f26'“ + tab$ + “'f36'“ + tab$ + “'f17'“ + tab$ + “'f27'“ + tab$ +
“'f37'“ + tab$ + “'f18'“ + tab$ + “'f28'“ + tab$ + “'f38'“ + tab$ + “'f19'“ + tab$ + “'f29'“ + tab$ +
“'f39'“ + tab$ + “'f110'“ + tab$ + “'f210'“ + tab$ + “'f310'“ + tab$ + “'start'“ + newline$
                        fileappend “'resultfile$'“ 'result_row$'
                endif
        endfor
        select all
        minus Sound 'sn$'
        minus TextGrid 'sn$'
        minus Strings list
        Remove
        select all

#TIER                                   5                                       Intensity
##############################################################################
###########################

        resultfile$ = “C:/Users/mbpol/Desktop/Sounds/5LIQsIntensity.txt”
        select Sound 'sn$'
        select TextGrid 'sn$'

        numint = Get number of intervals... 5
        select Sound 'sn$'
        To Intensity... 100 0
        for i from 1 to numint
                select TextGrid 'sn$'
                label$ = Get label of interval... 5 'i'
                if label$ <> “““
                        # calculates the onset and offset
                        start = Get starting point... 5 'i'
                        end = Get end point... 5 'i'
                        half = start + ((end - start) / 2)
                        wordint = Get interval at time... 1 'half'
                        wordlab$ = Get label of interval... 1 'wordint'

                        startWord = Get starting point... 1 'wordint'
                        endWord = Get end point... 1 'wordint'
                        prevword = startWord - ((end-start)/100000000)
                        follword = endWord + ((end-start)/100000000)
                        prevint = Get interval at time... 1 'prevword'
                        follint = Get interval at time... 1 'follword'
                        select TextGrid 'sn$'
```

```
                        prevwordlab$ = Get label of interval... 1 'prevint'
                        follwordlab$ = Get label of interval... 1 'follint'

                        duration = (end - start)
                        #calculates the intensity values
                        select Intensity 'sn$'
                        min_int = Get minimum... start end Parabolic
                        min_time = Get time of minimum... start end Parabolic
                        max_int = Get maximum... start end Parabolic
                        max_time = Get time of maximum... start end Parabolic
                        meanIntensity = Get mean... start end dB
                        range = (max_int - min_int)
                        result_row$ = “'sn$'“ + tab$ + “'prevwordlab$'“ + tab$ + “'wordlab$'“ +
tab$ + “'follwordlab$'“ + tab$ + “'label$'“ + tab$ + “'meanIntensity'“ + tab$ + “'min_int'“ + tab$
+ “'max_int'“ + tab$ + “'range'“ + tab$ + newline$
                        fileappend “'resultfile$'“ 'result_row$'
                endif
        endfor

        select all
        minus Sound 'sn$'
        minus TextGrid 'sn$'
        minus Strings list
        Remove
        select all

#TIER                                                                                                   6
Formants#######################################################################
#####################################

        resultfile$ = “C:/Users/mbpol/Desktop/Sounds/6VowelsFormants.txt”
        header_row$ = “SoundFile” + tab$ + “Word” + tab$ + “Sound” + tab$ + “F1” + tab$ +
“F2” + tab$ + “F3” + tab$ + “TimeStamp” + newline$
        select Sound 'sn$'
        select TextGrid 'sn$'

        numint = Get number of intervals... 6
        select Sound 'sn$'
        To Formant (burg)... 0 5 5500 0.025 50
        for i from 1 to numint
                select TextGrid 'sn$'
                label$ = Get label of interval... 6 'i'
                if label$ <> “““
                        start = Get starting point... 6 'i'
                        end = Get end point... 6 'i'
                        duration = (end - start)
```

```
onetenth = start + ((end - start) / 10)
twotenth = start + (((end - start) / 10)*2)
threetenth = start + (((end - start) / 10)*3)
fourtenth = start + (((end - start) / 10)*4)
fivetenth = start + (((end - start) / 10)*5)
sixtenth = start + (((end - start) / 10)*6)
sevententh = start + (((end - start) / 10)*7)
eighttenth = start + (((end - start) / 10)*8)
ninetenth = start + (((end - start) / 10)*9)
tententh = start + (((end - start) / 10)*10)
select TextGrid 'sn$'
wordint = Get interval at time... 1 'fivetenth'
select TextGrid 'sn$'
wordlab$ = Get label of interval... 1 'wordint'

startWord = Get starting point... 1 'wordint'
endWord = Get end point... 1 'wordint'
prevword = startWord - ((end-start)/100000000)
follword = endWord + ((end-start)/100000000)
prevint = Get interval at time... 1 'prevword'
follint = Get interval at time... 1 'follword'
select TextGrid 'sn$'
prevwordlab$ = Get label of interval... 1 'prevint'
follwordlab$ = Get label of interval... 1 'follint'

select Formant 'sn$'
f11 = Get value at time... 1 'onetenth' Hertz Linear
f21 = Get value at time... 2 'onetenth' Hertz Linear
f31 = Get value at time... 3 'onetenth' Hertz Linear
f12 = Get value at time... 1 'twotenth' Hertz Linear
f22 = Get value at time... 2 'twotenth' Hertz Linear
f32 = Get value at time... 3 'twotenth' Hertz Linear
f13 = Get value at time... 1 'threetenth' Hertz Linear
f23 = Get value at time... 2 'threetenth' Hertz Linear
f33 = Get value at time... 3 'threetenth' Hertz Linear
f14 = Get value at time... 1 'fourtenth' Hertz Linear
f24 = Get value at time... 2 'fourtenth' Hertz Linear
f34 = Get value at time... 3 'fourtenth' Hertz Linear
f15 = Get value at time... 1 'fivetenth' Hertz Linear
f25 = Get value at time... 2 'fivetenth' Hertz Linear
f35 = Get value at time... 3 'fivetenth' Hertz Linear
f16 = Get value at time... 1 'sixtenth' Hertz Linear
f26 = Get value at time... 2 'sixtenth' Hertz Linear
f36 = Get value at time... 3 'sixtenth' Hertz Linear
f17 = Get value at time... 1 'sevententh' Hertz Linear
f27 = Get value at time... 2 'sevententh' Hertz Linear
```

```
                    f37 = Get value at time... 3 'sevententh' Hertz Linear
                    f18 = Get value at time... 1 'eighttenth' Hertz Linear
                    f28 = Get value at time... 2 'eighttenth' Hertz Linear
                    f38 = Get value at time... 3 'eighttenth' Hertz Linear
                    f19 = Get value at time... 1 'ninetenth' Hertz Linear
                    f29 = Get value at time... 2 'ninetenth' Hertz Linear
                    f39 = Get value at time... 3 'ninetenth' Hertz Linear
                    f110 = Get value at time... 1 'tententh' Hertz Linear
                    f210 = Get value at time... 2 'tententh' Hertz Linear
                    f310 = Get value at time... 3 'tententh' Hertz Linear
                    result_row$ = “'sn$'“ + tab$ + “'prevwordlab$'“ + tab$ + “'wordlab$'“ +
tab$ + “'follwordlab$'“ + tab$ + “'label$'“ + tab$ + “'duration'“ + tab$ + “'f11'“ + tab$ + “'f21'“ +
tab$ + “'f31'“ + tab$ + “'f12'“ + tab$ + “'f22'“ + tab$ + “'f32'“ + tab$ + “'f13'“ + tab$ + “'f23'“ +
tab$ + “'f33'“ + tab$ + “'f14'“ + tab$ + “'f24'“ + tab$ + “'f34'“ + tab$ + “'f15'“ + tab$ + “'f25'“ +
tab$ + “'f35'“ + tab$ + “'f16'“ + tab$ + “'f26'“ + tab$ + “'f36'“ + tab$ + “'f17'“ + tab$ + “'f27'“ +
tab$ + “'f37'“ + tab$ + “'f18'“ + tab$ + “'f28'“ + tab$ + “'f38'“ + tab$ + “'f19'“ + tab$ + “'f29'“ +
tab$ + “'f39'“ + tab$ + “'f110'“ + tab$ + “'f210'“ + tab$ + “'f310'“ + tab$ + “'start'“ + newline$
                    fileappend “'resultfile$'“ 'result_row$'
            endif
        endfor
        select all
        minus Sound 'sn$'
        minus TextGrid 'sn$'
        minus Strings list
        Remove
        select all

#PRINT#################################################################
##################################

        printline ##############################
        printline File 'sn$' has been processed.
        printline Your data is here: “'resultfile$'“

#########################################################################
##########################

 select all
 minus Strings list
 Remove
endfor

select Strings list
Remove
clearinfo
echo Done. 'file_count' files annotated.
```

Appendix C: Inter-rater reliability guide

Thank you for agreeing to take part in the coding for my dissertation. Below, I have assembled suggestions for working on the interrater reliability for this project across five files (Table C). If you have any further questions, feel free to ask!

Table C: Speakers coded in inter-rater reliability task

#	Code	Name	Origin	Start Time	Age
1	2MMS2F	Manuel Robles Delgado	Madrid	575sec	58
2	8FSS2F	Verónica Pérez Fernández	Seville	590sec	35
3	19FSC1F	Teresa Jiménez-Becerril	Seville	95sec	53
4	23FCC1M	María Luisa Ceballos	Córdoba	273sec	47
5	29MLC1M	Elías Bendodo Benasayag	Málaga	387sec	43

You will be writing down the allophones you hear for the second half of each file (beginning at the "Start Time" listed in Table C), using the pre-existing intervals I have set up on tiers 2, 3, 4, and 5 where words are present in tier 1. The phones under consideration include /t͡ʃ/ (tier 2), /d/ (tier 3), /s/ and /θ/ (tier 4), and /ɾ/, /r/, and /l/ (tier 5). The coding key I have used is as follows:

1. Tier 1: words
    a. I have removed words from all areas you do not need to code.
2. Tier 2: affricates
    a. Two allophones: fronted [t͡s] (to avoid special symbols, I used <ts> in Praat) and normative [t͡ʃ] (<ch>). Elision or deaffrication was rare, from what I saw. Instances have three intervals: the first for the occlusion, second for frication, and third for the following vowel.
    b. Acoustic measure: Center of gravity of frication (lower→fronting).
3. Tier 3: intervocalic /d/
    a. Two allophones: production (<d>) and elision (<0>).
    b. Acoustic measure: occlusion presence based on range in intensity from peaks to valleys.
4. Tier 4: fricatives
    a. Four allophones of /s/: production (<s>), aspiration (<h>), *ceceo* (<θ>), and elision (<0>)
    b. Three allophones of /θ/: production (<θ>), *seseo* (<s>), and elision (<0>)
    c. Acoustic measure: Center of gravity and visible presence of sibilant aperiodicity.
5. Tier 5: liquids
    a. Five allophones for all three phones (/r/, /ɾ/, and /l/): tap (<r>), trill (<rr>), lateral (<l>), assibilation (<ASSIB>), reduced approximant/fricative variant with 0 occlusions (<rfric>), and elision (<0>).
    b. Acoustic measure: Presence of occlusions in rhotics (e.g., intensity dip), and formant range for laterals (e.g., some sign of movement in F3).
6. Tier 5: vowels
    a. Note: As these rely on continuous acoustic measures rather than categorical impressionistic coding, these do not need to be coded.

Include the allophone that best fits in each interval. If you note other types of productions than those listed above, include the most appropriate coding from the list, then add a parenthetical describing how you would identify it.

Appendix D: IRB Approval Letter

To: Díaz-Campos, Manuel
Protocol #: 10718
Protocol Title: Crossing the Party Line: A sociophonetic comparison of the production and perception of Andalusian dialectal features in peninsular Spanish political discourse
Type of Submission: Initial
Level of Review: Exempt
Approval Date: Friday, March 5th, 2021
Expiration Date: Saturday, March 5th, 2022
The admin approved the above-referenced submission. Conduct of this study is subject to the IU HRPP Policies, as applicable.
Additional Notes:
This research is exempt under the following categories: - Category 2(i) - Category 4(i)
Documents approved with this submission:
Attachments

Data Collection Instrument	Pollock IRB Data Collection Form.docx
Other	Pollock Stage One Video Sources.docx

You should retain a copy of this letter and all associated approved study documents in your research records.
If you have any questions or require further information, please contact the HRPP via email at irb@iu.edu or via phone at (317) 274-8289.

Figure A: IRB Approval for dissertation data collection

PROTOCOLS

INDIANA UNIVERSITY

#10718 - Crossing the Party Line: A sociophonetic comparison of the production and perception of Andalusian dialectal features in peninsular Spanish political discourse

Protocol Information

Review Type	Status	Approval Date	Continuing Review Date
Exempt	**Exempt**	**Mar 05, 2021**	**Mar 05, 2022**

Expiration Date	Initial Approval Date	Initial Review Type
--	**Mar 05, 2021**	**Exempt**

Feedback

Approval Comment

Remove unnecessary Expiration Date. Add Continuing Review Date.

This research is exempt under the following categories:
- Category 2(i)
- Category 4(i)

Appendix E: Qualtrics Questionnaire

Informed consent document[23]

¡Bienvenidos!
Le invitamos participar en un estudio de investigación científica. Como científicos llevamos a cabo investigaciones que nos proporcionan con más información sobre el mundo. En este texto, le explicamos del estudio que llevamos a cabo ahora y le ayudamos a decidir si le interesa tomar parte.

Nos interesa su percepción del habla de políticos y de gente normal. Esta encuesta se divide en cuatro secciones. Primero pedimos algunos datos básicos de identificación. Después usted recibirá audio archivos de españoles que describirás con ciertos adjetivos. En la próxima sección escuchará varios políticos andaluces y madrileños que clasificará del mismo modo, además de adivinar la afiliación política del hablante además de su ciudad de origen. Por fin tendrá que comparar entre varios políticos según su manera de hablar.

Esta encuesta debe durar unos 15 minutes. Las respuestas que proporcionará serán confidenciales puesto que ni guardaremos ni compartiremos información identificativa suya. No recibe compensación por su participación en este estudio. Usted participa en este estudio voluntariamente. Tiene la derecha de retirarle del estudio en cualquier momento, por cualquiera razón, sin prejuicio. Si le interesa contactar al investigador principal de este estudio, favor de escribir a Matthew Pollock a mbpolloc@indiana.edu. Si quiere discutir problemas, quejas o preocupaciones sobre este estudio, recibir información, o ofrecer sugerencias, favor de contactar la oficina del Programa de Protección en Investigaciones Humanas de IU a irb@iu.edu.

Cuando hace clic en el siguiente botón, usted reconoce que su participación es voluntaria, que tiene más de 18 años, y que es consciente de que tiene la posibilidad de terminar su participación en este estudio en cualquier momento y por cualquiera razón.

| Continúe |

English gloss of informed consent document

Welcome!
We invite you to participate in a scientific research study. As scientists we carry out investigations that provide us with more information about our world. In this brief text, we explain the study that we are currently carrying out and help you to decide if you are interested in participating.

We are interested in your perception of the speech of politicians and normal speakers. This survey is divided into four sections. First, we ask basic background questions. Next, you will receive audio files of Spaniards that you will describe with certain adjectives. In the next section, you will listen to Andalusian and Madrilenian politicians that you will classify in the same way, in addition to guessing their political affiliation and city of origin. Finally, you will have to compare between various politicians according to their manner of speaking.

This survey should last about 15 minutes. The responses you provide will be confidential, as we will neither save nor share your identifying information. You will not receive compensation for your participation in

[23] English gloss provided on page 2. Participants (who will be monolingual Spanish L1 speakers or else Spanish dominant bi- or multi-lingual speakers) will only receive the text in Spanish.

this study. You participate in this study voluntarily. You have the right to withdraw from this study at any time, for any reason, and without prejudice. If you are interested in contacting the principal researcher of this study, please write to Matthew Pollock at mbpolloc@indiana.edu. If you wish to discuss problems, complaints, or concerns about the study, receive information, or offer suggestions, please contact the IU Human Research Protection Program office at irb@iu.edu.

When you click the following button, you acknowledge that your participation is voluntary, that you are older than 18 years of age, and that you are aware that you are able to terminate your participation at any time and for any reason.

_____
| Continue |

Section 1: Background questionnaire
¿De dónde es usted? _____
Where are you from? _____
¿Cómo le identifica? Hombre, Mujer, Otro
How do you identify? Male, Female, Other
¿Qué edad tiene? _____
How old are you? _____
¿Por cuánto tiempo ha vivido en la provincia de Sevilla? Menos de 1 año / 1-9 años / 10-14 años / 15-19 años / 20-24 años / más de 24 años _____
How long have you lived in the province of Seville? Less than a year / 1-9 years / 10-14 years / 15-19 years / 20-24 years / more than 24 years _____
¿Hasta qué nivel de educación ha cumplido? Primaria / Secundaria / Licenciatura / Maestría / Doctorado
What level of education do you have? Primary / Secondary / BA / MA / PhD
Si ha asistido la universidad, ¿qué especialización tiene/tenía?
If you attended university, what is/was your major?
¿Con cuál partido político se identifica? _____
With which political party do you identify? _____
¿Cómo se identifica políticamente? Muy liberal, liberal, ligeramente liberal, Políticamente centrista o moderado, ligeramente conservador, conservador, muy conservador
How do you identify politically? Very liberal, liberal, slightly liberal, politically centrist or moderate, slightly conservative, conservative, very conservative
¿Cuántas veces ha votado en las últimas tres elecciones españolas? Cero, uno, dos, tres
How many times did you vote in the last three Spanish elections? Zero, one, two, three
¿Cuánto contenido político consume usted en línea, en la radio o la televisión? ¿De dónde proviene esta información (p.ej., canal, sitio)? Explique. – Poco _____ / Intermedio _____ / Mucho _____
How much political content do you consume online, by radio, or on tv? Where does your information come from? Explain. – Little _____ / Intermediate _____ / A lot _____

Part 1: Normal Speech
Sección 1: Ahora, escuchar las siguientes 20 palabras, producidas por hablantes españoles, una vez y clasificarlas con adjetivos. [24]
Section 1: Now listen to the following 20 words, produced by Spanish speakers, a single time and classify them with adjectives.
Urbano - *Rural*
Andaluz - *No Andaluz*
Culto - *Inculto*

[24] Participants received 20 audio files to classify on each of the five 100-point scales.

Simpático - *Antipático*
Liberal - *Conservador*
Según su opinión, ¿de dónde es este individuo?
According to your opinion, where is this individual from? (Participants respond on a heat map)

Part 2: Politicians' Speech
Sección 2: En muchas partes del mundo hoy en día, las amenazas del extremismo político van aumentando. Esta sección intenta medir diferencias en la percepción del habla política. Escuchar las siguientes 20 palabras, producidas por políticos españoles, una vez y clasificarlas con los adjetivos.[25]
Section 2: In many part of the world today, the threats of political extremism are growing. This section measures differences in the perception of political speech. Listen to the following 20 words, produced by Spanish politicians, a single time and classify them with adjectives.
Rural - *Urbano*
Andaluz - *No Andaluz*
Culto - *Inculto*
Simpático - *Antipático*
Liberal - *Conservador*
¿De dónde es el hablante? _____
Where is the speaker from?
¿A qué partido político pertenece el hablante? Podemos, Ciudadanos, PSOE, PP, Vox
What political party does the speaker belong to? *Podemos, Ciudadanos*, PSOE, PP, *Vox*

Part 3: Comparing Politicians
Sección 3: *En esta sección, comparar el habla de parejas de políticos de partidos distintos. Escuchar las dos palabras y clasificarlas con los adjetivos:*[26]
Section 3: In this section, compare the speech of pairs of politicians from different parties. Listen to the two words and classify them with adjectives:
¿Qué politico le parece más ***rural****? A o B*
Which politician seems more rural?
¿Qué politico le parece más ***Andaluz****? A o B*
¿Qué politico le parece más ***culto****? A o B*
¿Qué politico le parece más ***simpático****? A o B*
¿Qué politico le parece más ***izquierdista****? A o B*
Si estuviera en una cabina de votación y tuviera que escoger, ¿para cuál político votaría?
If you were in a voting booth and had to choose, which politician would you vote for?
Según su opinión, ¿de dónde es el Hablante 1?
In your opinion, where is Speaker 1 from?
Según su opinión, ¿de qué partido politico es el Hablante 1?
In your opinion, what political party is Speaker 1 affiliated with?
Según su opinión, ¿de dónde es el Hablante 2?
In your opinion, where is Speaker 2 from?
Según su opinión, ¿de qué partido politico es el Hablante 2?
In your opinion, what political party is Speaker 2 affiliated with?

[25] Participants received 20 audio files to classify on each of the five 100-point scales, and were shown a picture of each speaker, their name, and the political position they held most recently.
[26] Participants received 20 audio files, presented in 10 pairs, to compare, without any speaker identifying information.

CURRICULUM VITAE
MATTHEW POLLOCK

mbpollock42@gmail.com
mbpollock.wordpress.com

EDUCATION

PhD **Indiana University, Bloomington** July 2023
Double Major: Linguistics and Spanish with a concentration in Hispanic Linguistics
Doctoral minor: Germanic Studies
Dissertation: *Performing Andalusian in Political Speech: Political Party And Sociophonetic Patterns Across Production And Perception*
Committee: Manuel Díaz-Campos (Co-Chair), Samuel Obeng (Co-Chair), Erik Willis, Stuart Davis, Tracy Hall

MA **Indiana University, Bloomington** May 2018
Spanish with a concentration in Hispanic Linguistics
Areas of study: Sociolinguistics, Pragmatics, Phonology

BA **Auburn University** May 2015
Concurrent Degrees: English Literature, Spanish
Minors: German, Linguistics
Honors Scholar

ACADEMIC APPOINTMENTS

2022-2023 **Dissertation Research Fellow,** College of Arts and Sciences, Indiana University, 1 academic year (AY)
2019-2022 **Graduate Research Assistant to Chair (Dr. Manuel Díaz-Campos)**, Indiana University, Department of Spanish and Portuguese, 3 AYs
2021-2022 **Associate Instructor,** Indiana University, Department of Spanish and Portuguese, summer semester
2019 **Apagito Rey Fellow,** Indiana University, Department of Spanish and Portuguese, spring semester
2016-2018 **Associate Instructor,** Indiana University, Department of Spanish and Portuguese, 2.5 AYs
2015-2016 **Fulbright English Teaching Assistant**, Schleswig-Holstein, Germany, 1 AY

RESEARCH INTERESTS:

Sociolinguistics: Style, agency, political discourse, sociophonetics, usage-based approaches, and bilingualism
Phonetics & phonology: Sociophonetics, affricates, Andalusian Spanish, Caribbean Spanish

Usage-based linguistics:	Token and type frequency, cognitive linguistics, corpus approaches
Bilingualism:	Acquisition of phonetic boundaries, code-switching
2nd language acquisition:	L2 acquisition of Spanish and English, socio-pragmatic competence

PUBLICATIONS

Articles in refereed journals

Díaz-Campos, M., Cole, M., & Pollock, M. (2023). Re-conceptualizing affricate variation in Caracas Spanish. *Hispania 106*(1), 9-26.

Pollock, M. & Wheeler, J. A. (2022). Syllable-final /s/ and intervocalic /d/ elision in Andalusia: The Formation of Susana Díaz's Regional Identity in Political Discourse. *Language and Communication 87*, 191-204.

Pollock, M. & Blaker, N. (2021) ¿Qué me recomiendas?: Teaching the pragmatics of recommendations in the L2 Spanish classroom. *Indiana University Linguistics Club Working Papers (IULCWP).*

Chapters in refereed volumes

Díaz-Campos, M. & Pollock, M. (in press). The Spanish Sibilant System: *Seseo*. In A. Gallego & C. García (eds.) *A Guide to Spanish Dialects: Descriptive and Theoretical Aspects of Linguistic Variation in the Hispanic World*. Oxford: Oxford University Press.

Pollock, M., Díaz-Campos, M., Jonard, K., & González, S. (in press). Vowel variation in Spanish: Acoustic phonetics and speech processing in a diachronic Caracas corpus. *Proceedings of the International Congress of Phonetic Sciences (ICPhS) 2023: Intermingling Communities and Changing Cultures.*

Díaz-Campos, M., Cole, M., & Pollock, M. (in press). Sociolinguistic approaches to bilingual phonetics and phonology. *The Cambridge Handbook of Bilingual Phonetics and Phonology.*

Díaz-Campos, M., & Pollock, M. (2023). The future of usage-based sociolinguistics. *The Wiley-Blackwell Handbook of Usage-Based Linguistics.*

Díaz-Campos, M., & Pollock, M. (2023). Hispanic South America. In M. J. Ball (Ed.), *The Routledge Handbook of Sociolinguistics Around the World*, pg. XXX-XXX. Philadelphia, PA: Routledge.

Pollock, M., Delgado-Díaz, G., Galarza, I., Díaz-Campos, M., & Willis, E. W. (2023). The emergence of sound change in two varieties of Spanish: Variable trill /r/ production in Puerto Rico and Venezuela. *Proceedings of the 2021 Hispanic Linguistics Symposium*.

Pollock, M. (2020). Did you say peso or beso?: The perception of prevoicing by L2 Spanish learners. In Sessarego, S., Colomina-Almiñana, J. J. & Rodríguez-Riccelli, A. (Eds.), *Variation and Evolution: Aspects of language contact and contrast across the Spanish-speaking world*, pp. 127-161. Amsterdam: John Benjamins Publishing.

PUBLICATIONS IN PREPARATION & SUBMITTED

Articles in refereed journals

Pollock, M., Díaz-Campos, M., & Willis, E. W. (submitted). The perception of coda /ɾ/ and /l/ in Dominican Spanish: Diversity, geography, and sociolinguistic variation. *Revista Internacional de Lingüística Iberoamericana.*

Guo, J., Jarret, D., Wheeler, J., & Pollock, M. (submitted). Differential object marking in Spanish. *Language.*

Pollock, M. (submitted). Toeing the Party Line: Indexing Party Identity through Dialectal Phonetic Features in Spanish Political Discourse. *Languages.*

Wheeler, J., Pollock, M., & Díaz-Campos, M. (submitted). ¿(Está/Es) difícil?: Variable use of ser and estar by heritage learners of Spanish. *Languages.*

Davis, S. & Pollock, M. (submitted). [menu] with [tɔɾtæ] in Southern Spain: An Optimality Theoretic Analysis of Eastern Andalusian Vowel Harmony. *Languages.*

Pollock, M. (submitted). Buenas no[tʃ]es y mu[t͡s]isimas gracias: Variable Affricate Production in Peninsular Spanish Political Discourse. *Spanish in Context.*

Pollock, M. & Wheeler, J. (in preparation). Regional variation and speaker design: A sociophonetic study of Galician political norms in Castellano. *Language and Communication.*

Pollock, M., Díaz-Campos, M., & Willis, E. W. (in preparation). Perceiving the social variation of coda liquids in Dominican Spanish through education, L2 exposure and travel. *Journal of Sociolinguistics.*

Díaz-Campos, M., Willis, E. W., & Pollock, M. (in preparation). The Usage-Based Trill in Caracas Venezuela: An analysis of bigram frequency. *Nueva Revista de Filología Hispánica.*

Chapters in refereed volumes

Díaz-Campos, M., Willis, E. W., & Pollock, M. (submitted 3/21/22). La variación sociofonética de la vibrante multiple venezolana: un análisis basado en el uso. *Proceedings of the Retorno al Español del Caribe Conference*.

Pollock, M. & Díaz-Campos, M. (submitted 2/3/23). A sociolingustic study of the palatal fricative in Venezuelan Spanish. *Conference proceedings: Hispanic Linguistics Symposium 2022.*

Pollock, M. (**submitted** 2/11/23). Second person plural forms' role in peninsular Spanish identity construction. *Conference proceedings: Hispanic Linguistics Symposium 2022*.

Pollock, M. (**submitted** 4/11/23). *La variación fonológica del español panameño.* In Díaz-Campos, M. & Hernández-Campoy, J. M. (Eds.) *Enciclopedia Concisa de los Dialectos del Español*. London: Wiley-Blackwell.

CONFERENCE PRESENTATIONS

Pollock, M., Díaz-Campos, M., Jonard, K., & González, S. (**submitted**). *Mapping the Venezuelan Vowel Space: Sociolinguistics and speech processing methods in a diachronic Caracas corpus* [Oral Conference Presentation]. **New Ways of Analyzing Variation** (NWAV 51), Queens College, City University of New York.

Pollock, M., Wheeler, J. (**submitted**). *Language, politics and power: A sociophonetic comparison of Galician political and community norms in Castellano* [Oral Conference Presentation]. **Hispanic Linguistics Symposium** (HLS), Brigham Young University.

Pollock, M., Díaz-Campos, M., Jonard, K., & González, S. (**submitted**). *Social predictors of Venezuelan vowel variation: Sociolinguistics and speech processing methods in a*

diachronic Caracas corpus [Oral Conference Presentation]. **Hispanic Linguistics Symposium** (HLS), Brigham Young University.

Pollock, M., Díaz-Campos, M., Jonard, K., & González, S. (**accepted**; 2023, August 7-11). *Vowel variation in Spanish: Acoustic phonetics and speech processing in a diachronic Caracas corpus* [Oral Conference Presentation]. **International Congress of Phonetic Sciences** (ICPhS 2023), Prague Conference Center.

Pollock, M. & Wheeler, J. (2023, March 17-18). *Regional variation and speaker design: A sociophonetic study of Galician political norms in Castellano* [Oral Conference Presentation]. **Forging Linguistic Identities**, Towson University.

Pollock, M. (2022, October 13-15). *Crossing the Party Line: A sociophonetic analysis of Andalusian dialectal features in peninsular Spanish political discourse* [Oral Conference Presentation]. **New Ways of Analyzing Variation** (NWAV 50), (Virtual) Stanford University.

Pollock, M., Díaz-Campos, M., & Willis, E. (2022, October 13-15). *Perceiving the social variation of coda liquids in Dominican Spanish: Exposure through travel and family connections* [Oral Conference Presentation]. **New Ways of Analyzing Variation** (NWAV 50), (Virtual) Stanford University.

Pollock, M. (2022, November 3-5). *Politics and Politeness under COVID: Second person plural forms' role in peninsular Spanish identity construction* [Oral Conference Presentation]. **Hispanic Linguistics Symposium** (HLS), (Virtual) Arizona State University.

Díaz-Campos, M. & Pollock, M. (2022, November 3-5). *A sociolingustic study of the palatal fricative in Venezuelan Spanish* [Oral Conference Presentation]. **Hispanic Linguistics Symposium** (HLS), (Virtual) Arizona State University.

Pollock, M. (2022, April 7-9 – Postponed from April 2020 due to Coronavirus). *Buenas no[t͡ʃ]es y mu[t͡s]ísimas gracias: Variable Affricate Production in Peninsular Spanish Political Discourse* [Oral Conference Presentation]. **The Workshop on Spanish Sociolinguistics** (WSS), (Virtual) Georgia Tech.

Pollock, M., Díaz-Campos, M., & Willis, E. (2022, April 7-9 – Postponed from April 2020 due to Coronavirus). *A usage-based account of sociophonetic variation: The case of the variable trill /r/ in Venezuelan Spanish* [Oral Conference Presentation]. **The Workshop on Spanish Sociolinguistics** (WSS), (Virtual) Georgia Tech University.

Pollock, M. (2022, March 31-April 2). *Closing the Door on Grammar as a Gatekeeping Device: SLA across modalities in the Writing Center* [Oral Conference Panel Presentation]. **East Central Writing Centers Association** (ECWCA), (Virtual) Michigan State University.

Díaz-Campos, M., Pollock, M., & Willis, E. (2021, Oct 19-24). *The perception of coda /ɾ/ and /l/ in Dominican Spanish: Geographic and sociolinguistic variation* [Oral Conference Presentation]. **New Ways of Analyzing Variation** (NWAV 49), (Virtual) University of Texas at Austin.

Wheeler, J. & Pollock, M. (2021, October 7-9). *The Formation and Maintenance of Political Identity: Regional Linguistic Patterns in Susana Díaz's Twitter* [Oral Conference Presentation]. **Hispanic Linguistics Symposium** (HLS), (Virtual) Wake Forest University.

Pollock, M., Delgado-Díaz, G., Díaz-Campos, M., Galarza, I., & Willis, E. (2021, October 7-9). *The emergence of sound change in two varieties of Spanish: Variable trill /r/ production in Puerto Rico and Venezuela* [Oral Conference Presentation]. **Hispanic Linguistics Symposium** (HLS), (Virtual) Wake Forest University.

Díaz-Campos, M., Pollock, M., & Willis, E. (2020, November 7). *La variación sociofonética de la vibrante multiple venezolana: un análisis basado en el uso* [Oral Conference Presentation]. **Congreso Retorno al Español del Caribe (RealEC)**, (Virtual) Pontificia Universidad Católica Madre y Maestra, Dominican Republic.

Pollock, M. (2020, April – Accepted, cancelled due to Coronavirus). *'Culero, I speak four other languages as well to offend Trump': Lexical Participation and Impoliteness in Facebook comments* [Oral Conference Presentation]. **20th International Conference on Pragmatics and Language Learning (PLL)**, Indiana University.

Pollock, M., & Blaker, N. (2020, April – Accepted, cancelled due to Coronavirus). *¿Qué me recomiendas? Teaching Recommendations in the L2 Spanish Classroom* [Conference Poster Presentation]. **20th International Conference on Pragmatics and Language Learning (PLL)**, Indiana University.

Guo, J., Jarrett, D., Pollock, M., & Wheeler, J. (2019, October 24-26). *Differential object marking in Spanish* [Oral Conference Presentation]. **Hispanic Linguistics Symposium** (HLS), University of Texas, El Paso.

Pollock, M. & Wheeler, J. (2019, October 24-26). *Regional Identity Formation over Time in Political Discourse* [Oral Conference Presentation]. **Hispanic Linguistics Symposium** (HLS), University of Texas, El Paso.

Pollock, M. (2019, October 10-12). *Toeing the Party Line: Indexing Party Identity through Dialectal Phonetic Features in Spanish Political Discourse* [Oral Conference Presentation]. **New Ways of Analyzing Variation (NWAV 48)**, University of Oregon.

Pollock, M. (2019, April 22-23). *Toeing the Party Line: Indexing Party Identity through Dialectal Phonetic Features in Spanish Political Discourse* [Oral Conference Presentation]. **Diálogos Conference**, Indiana University.

Pollock, M. (2018, October 27-29). *The perception of pre-voicing by Spanish learners* [Oral Conference Presentation]. **Hispanic Linguistics Symposium** (HLS), University of Texas, Austin.

Pollock, M. & Wheeler, J. (2018, March 1-3). *Regional identity formation in political discourse* [Oral Conference Presentation]. **Diálogos Conference**, Indiana University.

Pollock, M. (2017, April 8). *Code-switching as a response to impoliteness on FB* [Oral Conference Presentation]. **Indiana University Linguistics Conference**, Indiana University.

Pollock, M. (2015, April 10). *Tolkien's Interpretation of Sound Symbolism* [Oral Conference Presentation]. **Auburn University Research Week**, Auburn University.

Pollock, M. (2014, October 30-November 1). *Foreign Language Tutoring: The Multilingual WC* [Oral Conference Presentation]. **National Conference on Peer Tutors in Writing**, Orlando, Florida.

Pollock, M. (2013, November 1-2). *Creative Writing in the Writing Center* [Oral Conference Presentation]. **National Conference on Peer Tutors in Writing**, Tampa, Florida.

UNIVERSITY PRESENTATIONS

Pollock, M. (2023, March 10). *Crossing the Party Line: A sociophonetic comparison of the production and perception of Andalusian dialectal features in peninsular Spanish political discourse* [Oral Presentation]. Research Round Robin: **Linguistics Department Professionalization Event**, Indiana University.

Pollock, M. (2023, February 15). *Third wave sociolinguistics and peninsular Spanish politics* [Oral Presentation]. Job Talk: **Department of English and Modern Languages**, Louisiana State University, Shreveport.

Pollock, M. (2022, November 11). *Crossing the Party Line: A sociophonetic comparison of the production and perception of Andalusian dialectal features in peninsular Spanish political discourse* [Oral Presentation]. Job Talk: **Hispanic Linguistics Monthly Brown Bag**, Indiana University.

Pollock, M. (2022, March 28). *Buenas no[t͡ʃ]es y mu[t͡s]isimas gracias: Variable Affricate Production in Peninsular Spanish Political Discourse.* [Oral Presentation]. **Sociolinguistics Discussion Group**, Indiana University.

Pollock, M. (2021, November). *Tutoring Modalities: Online and In-person tutoring* [Oral Presentation]. **Writing Tutorial Services Working Group**, Indiana University.

Willis, E., Díaz-Campos, M., & Pollock, M. (2021, October 27). *The Perception of Coda /ɾ/ and /l/ in Dominican Spanish: Diversity, Geography, and Sociolinguistic Variation* [Oral Presentation]. **Hispanic Linguistics Monthly Brown Bag**, Indiana University.

Pollock, M. (2021, October 15). *What we do: An Introduction to Writing Tutorial Services* [Oral Presentation]. **Spanish and Portuguese Professionalization Workshop Series**, Indiana University.

Willis, E., Díaz-Campos, M., & Pollock, M. (2021, March 12). *La variación socio-fonética de la vibrante múltiple venezolana: un análisis basado en el uso* [Oral Presentation]. **Hispanic Linguistics Monthly Brown Bag**, Indiana University.

Pollock, M. (2021, March 10). *Crossing the Party Line: A sociophonetic comparison of the production and perception of Andalusian dialectal features in peninsular Spanish political discourse* [Oral Presentation]. **Phonetics and Phonology Research Group (PhlegMe)**, Indiana University.

Pollock, M. (2021, February 21). *Hispanic Linguistics: Breaking down degree requirements* [Oral Presentation]. **Spanish and Portuguese Professionalization Workshop Series**, Indiana University.

Pollock, M. (2018, December 6). *Hispanic Linguistics MA Exam Prep Session* [Oral Presentation]. **Department of Spanish and Portuguese**, Indiana University.

RESEARCH EXPERIENCE

- **Graduate Research Assistant** to Professor Manuel Díaz-Campos, Department Chair, for the academic years 2019-2020, 2020-2021, and 2021-2022.
- **Editor, organizer, and contact person** for authors on behalf of Dr. Díaz-Campos in developing layout, theme, and introduction; carrying out peer-review process, author contact, revision, and proposal writing for the following edited volumes with Dr. Díaz-Campos from 2019-2023.

  Díaz-Campos, M. & Sessarego, S. (2021). *Aspects of Latin American Spanish Dialectology: In honor of Terrell A. Morgan*. Amsterdam: John Benjamins.

  Díaz-Campos, M. & Balasch, S. (submitted, 2022). *Handbook of Usage-based linguistics*. Wiley.

  Díaz-Campos, M. & Hernández-Campoy, J. M. (in preparation). *Enciclopedia Concisa de los Dialectos del Español*. Wiley.

- **Editor** for Professor César Félix-Brasdefer on revisions of chapters for two books in 2018:
  Félix-Brasdefer, J. C. (2019). *Pragmática del español: contexto, uso y variación*. London & New York: Routledge. (Also developed companion site for this text.)
  Koike, D. A. & Félix-Brasdefer, J. C. (2021). *The Routledge Handbook of Spanish Pragmatics Foundations and Interfaces*. Philadelphia: Routledge.

HONORS & AWARDS

2023 **Travel Award,** IU College of Arts and Sciences, April: travel to the ICEP conference ($400)

2023 **Travel Award,** IU College of Arts and Sciences, April: travel to the ICEP conference ($400)

2022-2023 **Dissertation Completion Fellowship**, IU College of Arts and Sciences, March: stipend, teaching release, and tuition waiver for Academic Year 2022-2023 to finalize dissertation research ($25,000)

2022 **College Arts and Humanities Institute Graduate Research Travel Grant,** IU, November: travel to Seville, Spain for dissertation data collection ($750)

2022 **Graduate Student Advisory Council Travel Grant,** IU Department of Spanish and Portuguese, April: travel to the HLS 2021 conference ($100)

2021 **Grant-in-Aid of Doctoral Research,** IU University Graduate School, November: funding for data collection for dissertation ($865)

2019 **Graduate Student Advisory Council's Travel Grant,** IU Department of Spanish and Portuguese, October: travel to the HLS 2019 conference ($100)

2019 **Travel Award,** IU College of Arts and Sciences, October: travel to the NWAV 2019 conference ($150)

2019 **Travel Grant,** IU Department of Spanish and Portuguese, October: travel to the HLS 2019 conference ($350)

2019 **Travel Grant,** IU Department of Linguistics, October: travel to the NWAV 2019 conference ($250)

2019-2022 **Graduate Research Assistantship,** IU Department of Spanish and Portuguese, 3 academic years: tuition waiver, teaching release, and yearly stipend awarded to serve as Teaching Assistant, researcher, and assistant for the department chair, Prof. Manuel Díaz-Campos ($18,000/annum)

2019 **Summer Fellowship,** IU Department of Spanish and Portuguese, May: summer grant funding to advance graduate studies ($5000)

2019 **Householder Best Research Paper,** *honorable mention,* IU Department of Linguistics, May: "Toeing the Party Line: Indexing Party Identity through Dialectal Phonetic Features in Spanish Political Discourse"

2019 **Agapito Rey Semester Fellowship,** IU Department of Spanish and Portuguese, January: funding for tuition and waiving teaching to pursue research and coursework ($7500 & tuition waiver)

2018 **Conference Travel Grant,** IU College of Arts and Humanities Institute**,** October: travel to the HLS 2018 conference ($400)

2016-2018 **Associate Instructorship,** IU Department of Spanish and Portuguese, 2.5 AYs: tuition waiver and yearly stipend to teach 3 courses/year ($15,500 & tuition waiver)

2015-2016	**Fulbright English Teaching Assistant**, Schleswig-Holstein, Germany, 1 AY: assistant teacher of English for 5th to 12th grade with 10-month stipend (8,000€)
2011-2015	**National Merit Presidential Scholarship,** Auburn University (AU), 4 academic years: tuition, housing, and study abroad stipend ($107,000)
2014	**Undergraduate Research Fellowship,** AU, fall semester: funding to develop the linguistics project *Tolkien's Interpretation of Sound Symbolism* ($1500)
2013	**Academic Enrichment Scholarship,** AU College of Liberal Arts, December: Research award ($3000)
2013	**Castanoli Abroad Award,** AU Foreign Language Department, December: Study abroad award **(**$1000)

PROFESSIONAL APPOINTMENTS

2016-2023	**Writing Consultant**, Indiana University, WTS, 7 academic years
2021	**Assistant Graduate Director, Writing Tutorial Services & Campus Writing**, Indiana University, Writing Tutorial Services, fall semester
2018	**Spanish Linguistics Instructor/Leader,** IU Honors Program in Foreign Languages, Ciudad Real, Spain, foreign language program for high schoolers
2015-2016	**Fulbright English Teaching Assistant**, Schleswig-Holstein, Germany, 1 AY
2014-2015	**Lead Writing Consultant,** Auburn University, Miller Writing Center, 1 AY
2013	**Writing Consultant,** Auburn University, Miller Writing Center, 1 AY

TEACHING EXPERIENCE

Indiana University. Teaching Assistant to Manuel Díaz-Campos, 2019-2022.

SU22,21	**T560 Introduction to Hispanic Linguistics,** IU, Bloomington	(2 sections)
SP22,21	**S513 Intro to Hispanic Sociolinguistics**, IU	(2 sections)
F20,21	**S429 Spanish Sociolinguistics and Pragmatics**, IU	(2 sections)
SP20	**S495 Language and Cultural Diversity**, IU	(1 section)
F19	**S612 Variationist Approaches to Spanish Corpora**, IU	(1 section)

Indiana University. Associate Instructor. 2016-2018, summer 2018/2021/2022.

SU22, 23	**S492 Graduate Spanish for Reading,** IU	(2 sections)
SU21	**S326 Introduction to Hispanic Linguistics**, IU	(1 section)
F18	**S280 Spanish Grammar in Context**, IU	(2 sections)
F17, SU18	**S250 Intermediate Spanish II**, IU	(3 sections)
SP17	**S200 Intermediate Spanish I**, IU	(2 sections)
F16	**S105 First Year Spanish**, IU	(1 section)

Indiana University Honors Program in Foreign Languages, Instructor. June-July 2018.

SU18	**Intro to Spanish Linguistics,** IU Abroad, Ciudad Real, Spain	(3 sections)

Fulbright English Teaching Assistant. September 2015-July 2016.

S16	**Grade 6**, Gymnasium Trittau, Hamburg, Germany	(1 section)
F15, S16	**Grade 9**, Gymnasium Trittau, Hamburg, Germany	(2 sections)
F15, S16	**Grade 10**, Gymnasium Trittau, Hamburg, Germany	(2 sections)
F15, S16	**Grade 11**, Gymnasium Trittau, Hamburg, Germany	(2 sections)

F15, S16 **Grade 12**, Gymnasium Trittau, Hamburg, Germany (1 section)

PROFESSIONAL AFFILIATIONS

American Association of Teachers of Spanish and Portuguese (AATSP): Student Member (2022)

RESEARCH LANGUAGES & SKILLS

Languages
English: Native speaker
Spanish: Near-native fluency; reading, writing, speaking
German: Conversational speech fluency; B2 level reading and writing
French, **Arabic, Gothic, Old English**: Basic reading proficiency

Programs
Statistical Analysis: R, Rbrul, SPSS, Language Variation Suite, novice: Python
Phonetic Analysis: PRAAT, Adobe Premiere, novice: Montreal Forced Alignment Tool
Discourse Analysis: AntConc, Visual DTA
Publishing: Microsoft Publisher, Adobe InDesign

OVERSEAS EXPERIENCE

Spain
Seville, Andalusia: University of Seville, *Dissertation Data Collection*, Nov 2022
Ciudad Real, Castilla-La Mancha: IU Honors Teaching, *Instructor*, June-July 2018
Alcalá de Henares, Madrid: Instituto Franklin study abroad, *Student*, Jan-May 2014
Germany
Großhansdorf, Schleswig-Holstein: Fulbright English Teaching Assistant, TA, Sept 2015-Jul 2016

SERVICE TO THE PROFESSION

- **Conference Abstract Reviewer (x6):** New Ways of Analyzing Variation (NWAV 51). (English). On: Sociolinguistic variation. June 2023.
- **Article Reviewer:** *Indiana University Linguistics Club Working Papers.* (English). On: Iberian phonetics. April 2023.
- **Article Reviewer:** *Indiana University Linguistics Club Working Papers.* (English). On: Sociolinguistic attitudes. August 2022.
- **Conference Abstract Reviewer (x5):** New Ways of Analyzing Variation (NWAV 50). (English). On: Sociolinguistic variation. July 2022.

- **Chapter Reviewer**: *Interfaces of Phonology* (Phonology and Phonetics: De Gruyter Mouton). (English). On: Spanish stop perception. June 2022.
- **Article Reviewer:** *Indiana University Linguistics Club Working Papers.* (English). On: Spanish discourse marking. April 2022.
- **Chapter Reviewer (x2):** *Aspects of Latin American Spanish Dialectology: In Honor of Terrell Morgan* (Issues in Hispanic and Lusophone Linguistics, 32: John Benjamins). (English). On: Spanish forms of address. April 2020.
- **Article Reviewer:** *Indiana University Linguistics Club Working Papers.* (English). On: Spanish clitic acquisition. February 2020.
- **Chapter Reviewer:** *The Routledge Handbook of Variationist Approaches to Spanish.* (English). On: Spanish voicing. October 2019.