

Article

Perceptions of regional origin and social attributes of phonetic variants used in Iberian Spanish

Nicholas Henriksen, Lorenzo García-Amaya, Micha Fischer, Jessica Czapla, Natalie Dakki, Amber Galvano, Sarah Khansa, Ellie G. Maly, Zoe Phillips, Vidhya Premkumar, Stepan Topouzian and Thomas Wiaduck

University of Michigan, Ann Arbor, MI, USA

Abstract

Sociodemographic information, such as a speaker's regional origin, is intimately related to the judgments and social evaluations that listeners assign to that speaker. This association between linguistic form and social information can also lead to linguistic profiling, a harmful form of discrimination. The present study examines the geographic classifications and social attitudes attributed to ten phonetic variants used within regional varieties of Iberian (i.e., European) Spanish. We are specifically interested in understanding listeners' geographical classifications and language attitudes held toward Andalusian Spanish, which is a less privileged regional variety spoken in Spain's southern region, as compared to north-central Peninsular Spanish (NCPS). The results of an online survey show that 165 listeners were fairly consistent when geographically classifying Andalusian-sounding stimuli as originating from the south of Spain. Importantly, the respondents also attributed less favorable social meaning to the Andalusian-sounding stimuli in comparison to the NCPS-sounding stimuli. We link the findings to broader themes in sociolinguistics, such as language-based discrimination, linguistic insecurity, and the social motivations of language change.

Keywords: Andalusian Spanish; linguistic discrimination; North-Central Peninsular Spanish; perceptual dialectology; regional variation

1. Introduction

Research in the phonetic sciences has long documented the connection between the acoustic properties of speech and the perception of talker-specific attributes such as social identity, individual idiosyncrasies, emotional states, and context-induced features (e.g., Abercrombie, 1967; Goldinger, Pisoni & Logan, 1991; Mullenix, Pisoni & Martin, 1989). From a sociolinguistic standpoint, researchers have hypothesized how this multifaceted indexing of phonetic form contributes to language users' interpretations of social meaning and their subsequent language choices (Campbell-Kibler, 2010; Eckert, 2008, 2019).

One harmful byproduct of the association between phonetic features and sociodemographic information is its bearing on linguistic profiling, or language-based discrimination. Linguistic profiling manifests when the assumptions made about how a person speaks bring about inequitable outcomes for the speaker themselves (Baugh, 2003, 2017; Lippi-Green, 1997; Purnell, Idsardi & Baugh, 1999; Rickford & King, 2016). In the context of US English, for instance, Baugh (2003) synthesized the deeply interwoven relationship between patterns of racial discrimination and listeners' judgments of African American Language (AAL). Such discriminatory judgments have been shown to give rise to societal barriers

including the denial of housing, insurance, mortgages, or appropriate legal treatment for Black Americans (see Wright, 2019).

Language-based discrimination can arise from ethnoracial bias, cultural prejudice, regional stereotypes, or the compounded impacts of these factors (Baugh, 2017:2–3). With respect to regional stereotypes, instances of negative bias based on a talker's perceived regional origin have been confirmed for numerous languages, including varieties of French (Giles & Billings, 2004), Dutch (Grondelaers, Van Hout & Steegs, 2010), Turkish, and German (Du Bois, 2019). Additionally, the mediating effect of a listener's regional origin, in addition to that of the speaker, can play a crucial role in the social evaluation of linguistic forms (Preston, 1989; Walker et al., 2014).

The present study extends this line of work to the language ideologies associated with the regional variation present in Iberian (i.e., European) Spanish. Gal (2016:116) defines language ideologies as language users' metalinguistic "talk about talk, or their reflections, signals, and presuppositions about linguistic forms and their use." Here, we are primarily interested in Iberian-Spanish listeners' ideologies of the linguistic variation present in southern Spain, known collectively as "Andalusian Spanish." Linguistically, Andalusian Spanish is born from a complex set of features that differentiate the southern dialect from more predominant regional varieties spoken in the northern and central regions of the country (Alvar, 1996:233–58; Moreno-Fernández, 2020:58–69). Although Andalusia is the most populous of Spain's seventeen autonomous communities, a constellation of social, historical, cultural, and political factors contributes to the stereotyping of Andalusian people as unproductive, uncultivated, and folkloric (Jaspal & Sitaridou, 2013; Santana Marrero, 2018).

Corresponding author: Nicholas Henriksen; Email: nhenriks@umich.edu.

Cite this article: Henriksen N, García-Amaya L, Fischer M, Czapla J, Dakki N, Galvano A, Khansa S, Maly EG, Phillips Z, Premkumar V, Topouzian S, Wiaduck T. (2023). Perceptions of regional origin and social attributes of phonetic variants used in Iberian Spanish. *Journal of Linguistic Geography* <https://doi.org/10.1017/jlg.2023.6>



We developed an online survey to investigate listeners' geographic classifications and social perceptions based on three linguistic contexts: syllable-initial coronal fricatives (i.e., use of the /s/ and/or /θ/ phonemes), word-medial coda-/s/, and word-final coda-/s/. We recruited participants who are speakers of two different subvarieties of Andalusian Spanish in addition to a third participant group from Castile and León, located in central Spain. This geographically expansive participant pool allows us to evaluate intraregional perceptual variation (within Andalusia) as well as interregional variation throughout Spain.

2. Background

2.1. The social meaning of linguistic form

Sociolinguistic research theorizes the dynamic process that brings about associations between linguistic forms and often interchangeable social meanings (Eckert, 2008, 2019). The “third-wave” perspective of sociolinguistics, in particular, proposes that a form's socially attributed meaning results from communicative language practices in which context prompts a language user's continual updating of fluid indexical fields (Eckert, 2008, 2019). It is proposed that the indexical assignment of a linguistic form undergoes iterative reconstruction as listeners and speakers engage in new discursive contexts (Silverstein, 2003). Once a form and an indexed social meaning become strongly associated with one another, the form itself then becomes a resource that individual speakers integrate into their linguistic repertoires, thereby prompting language change (Eckert, 2008:457). Empirically, one method for determining the social status of a linguistic form is to measure the explicit social judgments made by listeners (Cargile et al. 1994). In the present article, we will bring attention to such evaluations of forms used in Andalusian Spanish.

Until the mid-twentieth century, Andalusia, located in southern Spain (see provincial map in Map 1), had enjoyed a vibrant agricultural history tied to commerce and a strong labor movement. Any prosperity, however, was greatly altered by the outcomes of the Spanish Civil War and Francisco Franco's fascist regime, lasting from 1939 to 1975. The repression of Franco's dictatorship targeted Andalusian nationalists through financial deprivation, the forced exile of state enemies, and mass executions of suspected dissidents including leftist political leaders, intellectuals, and artists, among others (Cobo Romero, 2012). While urban areas to the north were endowed with greater educational and financial investment, Andalusia remained



Map 1. Map of Andalusia outlining its eight provinces. The three provinces in dark-gray shade are commonly categorized as “Western Andalusia,” the two provinces without shading are commonly categorized as “Central Andalusia,” and the three provinces in light-gray shade are commonly categorized as “Eastern Andalusia” (see Jaime-Jiménez, 2019).

predominantly agrarian. This disparity was further slanted as resulting from a folkloric set of stereotypes assigned to southerners, reflected in common artistic depictions of the region as more antiquated than areas to the north (Caballero-Gálvez, 2017; Gordillo, 2012). Moreover, “Castilian” Spanish, spoken in northern and central Spain, gained newfound privilege due to efforts to strengthen national cohesion and suppress divergent regional identities linked to minority languages and local dialects (Gallego & Rodríguez, 2013; Hernández-Campoy & Villena-Ponsoda, 2009:183–84). Andalusia thus continues to remain at a social and economic disadvantage compared to the rest of Spain. Sociolinguistically, the indexing of prototypical Andalusian features includes contemporarily less favorable¹ attributes such as being less educated, working-class, rural, or informal (Regan, 2022).

2.2. Sociolinguistic overview of Andalusian Spanish

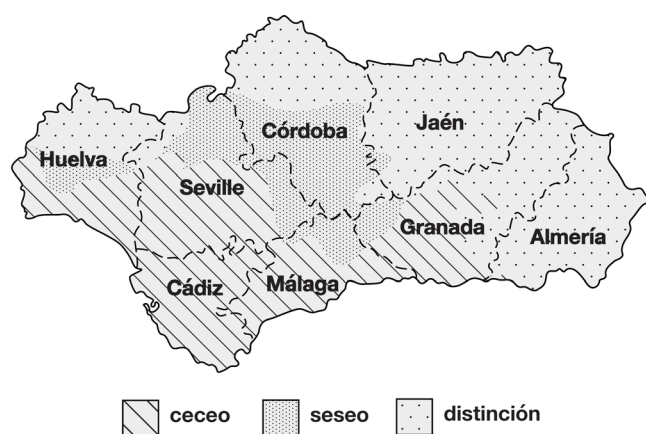
Concerning the phonetic and phonological differences between Andalusian Spanish and NCPS, the Andalusian variety is characterized by features including variable production of the coronal-fricative phonemes /s/ and /θ/ in syllable-initial position, lenition of syllable-final /s/ (in both word-medial and word-final positions), weakening of /tʃ/ to /ʃ/, dentalized pronunciation of /s/, velarization of word-final /n/, and neutralization of the syllable-final liquids /r/ and /l/, among others (Alvar, 1996:233–258; Moreno-Fernández, 2020:58–69; Villena-Ponsoda, 2008). The present study turns attention to listeners' geographic classifications and social evaluations of the variations present in two of these linguistic environments, namely syllable-initial coronal fricatives and coda-/s/ lenition (the latter will be further subclassified), which are reviewed in the following sections. Our motivation for choosing these contexts was twofold. First, within Andalusia there is considerable intraregional variation reflected in each category, with the most consistent differences falling primarily (although not exclusively) along a west-to-east continuum (Alvar, 1996). Second, the variation present within the three linguistic environments is susceptible to interrelated sociolinguistic forces (see Hernández-Campoy & Villena-Ponsoda, 2009), which allows for a more comprehensive understanding of how language users perceive and evaluate variation.

2.2.1 Syllable-initial coronal fricatives

In the initial position of a syllable, speakers of Iberian Spanish differ in their production of the graphemes <ci>, <ce>, <z>, and <s> (see Table 1). In the NCPS variety, for instance, speakers use the voiceless interdental fricative /θ/ in the <ci>, <ce>, and <z> contexts and the voiceless alveolar fricative /s/ in the <s> context. This usage, known as *distinción*, brings about phonemic contrast through minimal pairs such as *casa-caza* ‘house’-‘hunt’ (e.g., /kasa/ - /kaθa/). In Andalusian Spanish, by contrast, four systems are possible (see Map 2, based on accounts from early work in Spanish dialectology). First, *distinción* is present along Andalusia's northernmost border and in the easternmost provinces. Second, *ceceo*, which consists of producing /θ/ in all four graphemic contexts (e.g., /kaθa/ - /kaθa/), is more common in the westernmost provinces, the central province of Málaga, and in many rural areas.² A third possibility is *seseo*—namely, use of /s/ in all four contexts (e.g., /kasa/ - /kasa/)—which is prevalent in the urban centers of Córdoba and Seville. Finally, some speakers, primarily those with greater social and physical mobility, variably produce /θ/ or /s/ in the <ci>, <ce>, and <z> contexts, while maintaining /s/ in the <s> context. This fourth usage among Andalusian speakers is known as *seceo* (Dalbor, 1980; Sawoff, 1980).³

Table 1. List of the ten phonetic variants, with examples, included in the perception study

Linguistic context	Phonetic variant	Example
Syllable-initial coronal fricatives	<i>distinción</i>	[po.li.'θia.am.bi.'θjo.so] <i>policía ambicioso</i> 'ambitious police officer'
	<i>ceceo</i>	[po.li.'θi.a.am.bi.'θjo.θo] <i>policía ambicioso</i> 'ambitious police officer'
	<i>seceo</i>	[po.li.'si.a.am.bi.'θjo.so] <i>policía ambicioso</i> 'ambitious police officer'
	<i>seseo</i>	[po.li.'si.a.am.bi.'sjo.so] <i>policía ambicioso</i> 'ambitious police officer'
Word-medial /s/	/s/-retention	[pes.'ta.ɲa] <i>pestaña</i> 'eyelash'
	post-aspiration of /s/	[pe.'tʰa.ɲa] <i>pestaña</i> 'eyelash'
	/s/-deletion	[pe.'ta.ɲa] <i>pestaña</i> 'eyelash'
Word-final /s/	/s/-retention	['ne.nes] <i>nenes</i> 'children'
	/s/-deletion + tense vowel	['ne.ne] <i>nenes</i> 'children'
	/s/-deletion + lax vowel	['ne.ne] <i>nenes</i> 'children'

**Map 2.** Map of Andalusia outlining the regional use of syllable-initial coronal fricatives (adapted from Alvar, 1996:250). *Seceo* is common among *seseo* speakers from urban Córdoba and Seville.

Apparent-time studies show that *ceceo*, *seceo*, and *seseo* are currently giving way to *distinción* in many urban areas of Andalusia (Hualde, 2021; Melguizo Moreno, 2007). The ongoing sound change in favor of the national norm is led primarily by university-educated young speakers, socially and physically mobile individuals of various ages, and female speakers (Regan, 2020b). In light of this intersecting regional and sociolinguistic variation, we anticipate an especially fluid social indexing with respect to the four systems, making them apt for perceptual testing.

2.2.2 Word-medial coda-/s/

In addition to variation of syllable-initial coronal fricatives, coda-/s/ variation in the word-medial and word-final positions has received considerable attention in the literature (e.g., Carbonero Cano, 1982; Villena-Ponsoda, 2008; for overview see Gradoville, Brown & File-Muriel, 2022). In both word positions, syllable-final-/s/ may undergo lenition (i.e., weakening) through debuccalization, resulting in glottal aspiration prior to the stop closure (i.e., pre-aspiration) or coda deletion. This stands in contrast to the presence of coda-/s/ by northern and central speakers who primarily converge toward retention of the alveolar fricative (Henriksen & Harper, 2016; Momcilovic, 2009).

Concerning word-medial coda-/s/, there is now an ongoing sound change whereby speakers from the western provinces of Cádiz, Huelva, and Seville are transitioning toward productions of phonological /s/ + /p t k/ sequences in which a glottal fricative follows the stop closure, resulting in the post-aspirated stops [p^h t^h k^h] (e.g., Henriksen et al., 2023; Torreira, 2012). Thus, although a word like /pasta/ can be variably pronounced as [pasta] (with retention), [pahta] (with pre-aspiration), or [pata] (with /s/-deletion), [pat^ha] with post-aspiration is becoming the dominant pronunciation. Ruch and Harrington (2014) and Ruch and Peters (2016) show that young speakers from the western provinces, and from Andalusia's capital city Seville in particular, are leading this change. In light of this ongoing shift, there is a need for new research assessing listeners' social evaluations of post-aspirated stops in comparison to other phonetic variants. For example, it is possible that post-aspirated variants provoke more favorable social evaluations among Andalusian listeners than straightforward /s/-deletion, which would help to explain why the new sound is undergoing intraregional transmission within Andalusia (Ruch, 2018).

2.2.3 Word-final coda-/s/

In contrast to post-aspiration, which is more prevalent in Andalusia's western provinces, deletion of coda-/s/ in the word-final position is accompanied by a vowel laxing process in the eastern provinces of Almería, Granada, and Jaén, and in portions of the central provinces of Córdoba and Málaga. For this reason, historical word-final /as es os/ now have laxed or open realizations [æ e o] (Alvar, 1996; Hernández-Campoy & Trudgill, 2002; see Map 3). This word-final laxing is accompanied by a leftward vowel-harmony process that transmits the quality of the final vowel to preceding nonhigh vowels up to the stressed syllable. Thus, plural words like *camas* 'beds,' *nenes* 'boys,' or *conos* 'cones' are pronounced as ['kæ.mæ], ['ne.ne], and ['kɔ.nɔ], while their singular forms retain the respective tense vowels as in ['ka.ma], ['ne.ne], and ['ko.no]. For speakers from the central and eastern provinces, this results in an eight-vowel system /a e i o u æ e o/ distinct from the traditional five-vowel system /a e i o u/ (Jiménez & Lloret, 2019).

The current body of work focused on Andalusian vowel laxing predominantly considers the phonetic and/or phonological bases for this sound change (e.g., Henriksen, 2017; Herrero de Haro, 2016). While there is some indication that the occurrence of laxing



Map 3. Map of Andalusia outlining the regional use of tense versus lax productions of /a es os/ sequences (adapted from Alvar, 1996:245).

may be more common in smaller towns than in their respective capital cities (e.g., Martínez Melgar, 1994), further research is required to determine the complete sociolinguistic basis underlying the perceptual indexing of word-final lax vowels. Due to the unjust influence of the bias generally associated with Andalusian features (Hernández-Campoy & Villena-Ponsoda, 2009:190–191), it becomes crucial to examine listeners' social evaluations of this understudied sociophonetic feature.

2.3. The current study

We developed an online survey aiming to disambiguate the relationship between listeners' geolinguistic classifications of phonetic variants based on three linguistic contexts (see Table 1) and their related social evaluations. As outlined in Section 2.2, some of the variants within each linguistic context are distinctive features of Andalusian Spanish and thus likely play a role in shaping listeners' ideologies regarding this regional variety (for further elaboration, see Carbonero Cano, 2003). Our design builds upon previous research on listeners' perceptions of sociophonetic variation in Iberian Spanish (e.g., Harjus, 2018; Molina-García, 2020; Regan, 2019, 2020a, 2022; Ruch, 2018; Santana Marrero, 2018) by expanding the set of features included in a single survey and by including listeners of distinct regional provenances.

We recruited listeners from the cities of Granada, Jerez de la Frontera, and Salamanca (see Map 4) due to their comparable sizes—all are mid-size cities with populations ranging between 150,000 and 230,000. Salamanca was selected to reflect the language ideologies held by listeners from north/central Spain, while Granada and Jerez de la Frontera were selected to reflect those of eastern and western Andalusia, respectively. In doing so, the study was driven by two research questions.

Research question 1: How do listeners of Iberian Spanish perceptually categorize the geographic origin of ten phonetic variants?

1a: How do listeners categorize the variants based on a predetermined north/central/south geographic classification?

1b: How do listeners categorize the variants in an open-ended response format?

Research question 2: What social attributes do listeners of Iberian Spanish assign to each of the ten phonetic variants?



Map 4. Map of Spain indicating the three cities of origin for the respondent groups recruited for this study.

2a: What is the effect of linguistic context (i.e., syllable-initial coronal fricatives, word-medial coda-/s/, word-final coda-/s/) on the assignment of social attributes?

2b: What is the effect of a listener's geographic origin (i.e., Granada, Jerez de la Frontera, Salamanca) on the assignment of social attributes?

3. Methods

3.1. Participants

Participants ($N=199$) were recruited through advertisements published in social media groups targeting university students. The inclusion criteria were as follows: participants must (1) be 18–30 years of age at the time of taking the survey; (2) be native speakers of Spanish who were born and raised in Spain and have lived in Granada, Jerez de la Frontera, or Salamanca for the majority of their lives; and (3) have access to a quiet space, a computer with internet, and headphones to listen to audio. The data from 30 participants were excluded for various reasons, including failure to fulfill the inclusion criteria or failure to enter a validation code at the end of the survey. An additional four participants opted out of the study. Here, the data from 165 respondents were analyzed: 59 from Granada, 52 from Jerez de la Frontera, and 54 from Salamanca.

The average respondent age was 21.3 years (range = 18–30), and the majority self-identified as women ($N=104$, 63.0%). Most participants had completed university studies or were enrolled at a local university at the time of the survey ($N=135$, 81.9%), while 30 participants (18.1%) reported secondary studies as their highest level of education.

3.2. Audio stimuli

The audio stimuli, containing isolated words or noun phrases, were recorded by a native male speaker of Western Andalusian Spanish (the second author of this article). This speaker is a trained linguist who was raised in an Andalusian-speaking household and attended private elementary and secondary schools that required the use of Castilian Spanish (the term generally used to describe the socially privileged status of the NCPS variety). By all accounts, the speaker is a well-balanced bidialectal individual.

The speaker recorded 90 target stimuli in addition to 15 distractors (see Appendix A for the full stimuli list). The target items were developed per the three linguistic contexts: (1) syllable-initial coronal fricatives; (2) word-medial coda-/s/; and (3) word-final coda-/s/. For the syllable-initial coronal fricatives, four variants were featured in the survey: *distinción*, *ceceo*, *seceo*, and *seseo*. For the word-medial /s/ stimuli, three variants were featured: /s/-retention, post-aspiration of /s/, and /s/-deletion. For the word-final /s/ stimuli, three variants were featured: /s/-retention, /s/-deletion + tense vowel, and /s/-deletion + lax vowel. There were nine baseline words for each of the three contexts, and the speaker produced each baseline word using the targeted variants, yielding 90 stimuli: (9×4 variants for syllable-initial coronal fricatives) + (9×3 variants for word-medial coda-/s/) + (9×3 variants for word-final coda-/s/).

Following the recording session, a research assistant performed acoustic analysis to ensure consistency among the Voice Onset Time (VOT) values from the post-aspiration stimuli, as well as among the F1 and F2 values from the tense- and lax-vowel stimuli (see values in Appendix B). The stimuli involving more categorical differences (e.g., *distinción*, *seseo*, /s/-deletion) were reviewed by two trained linguists, both of whom agreed that they accurately reflected the targeted features. All stimulus files were normalized for duration and amplitude with PSOLA resynthesis using a series of Praat scripts. We also added 0.5 seconds of silence prior to the start of each stimulus to mitigate the effects of unexpected audio glitches.

3.3. Survey design and implementation

The survey was created using the Qualtrics platform and consisted of four parts: (1) an introduction involving participant consent and an audio-functionality verification; (2) a short demographic questionnaire; (3) the dialect survey including all 105 audio stimuli; and (4) debriefing questions asking participants to describe their experience taking the survey and to explain what they believed the study's goal to be.

In the survey, each audio stimulus was presented at the top of a new page in Qualtrics and was accompanied by an orthographic transcription (see Figures 1 and 2). The respondents were permitted to listen to each stimulus twice at most in an effort to reduce any risk of semantic satiation. Next, the listeners responded to seven prompts on the screen. The first prompt was a forced multiple-choice question requiring participants to associate the pronunciation with a geographical area of Spain (north, central, or south). The second prompt used a free-response format in which participants provided more concrete information regarding the perceived geographic origin (e.g., specific province, city, or town) of the stimulus. The five remaining questions asked participants to characterize the audio stimulus according to five social attributes: (1) work ethic; (2) education level; (3) urbanicity; (4) habitus; and (5) humor. The participants responded to the social-attribute prompts by moving a slider to the intended value between 0 and 100 (see Table 2 for the predefined values). All slider scales were initially set to a value of 50 to avoid priming effects.

Por favor haz clic en el botón de inicio para escuchar la frase "**policía ambicioso**". Puedes escuchar este archivo de sonido dos veces como máximo.



¿Con qué zona de España asocias esta pronunciación?

Norte de España

Centro de España

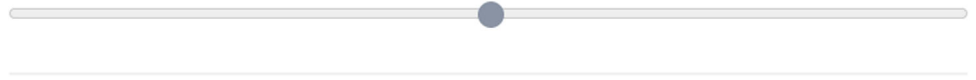
Sur de España

Específicamente, ¿cuál crees que es el origen geográfico de esta pronunciación? Puedes indicar comunidad autónoma, provincia, e incluso localidad si la conoces.

Figure 1. Image of the first two prompts that respondents saw upon playing an audio stimulus. English translation: Please click the button to hear the phrase "policía ambicioso." You may listen to the recording a maximum of two times. What region of Spain do you associate this pronunciation with? North of Spain/Central Spain/South of Spain. More specifically, what do you think the geographic origin of this pronunciation is? You can indicate the autonomous community, province, or locality if you know it.

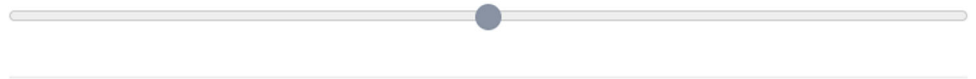
1. En función de lo que acabas de oír, ¿cómo caracterizarías a este hablante? Como alguien...

soso 0 10 20 30 40 **agradable** 50 60 70 80 **gracioso** 90 100



2. En función de lo que acabas de oír, ¿cómo caracterizarías a este hablante? Como alguien...

con pocos estudios 0 10 20 30 **con estudios secundarios** 40 50 60 70 80 **con estudios universitarios** 90 100



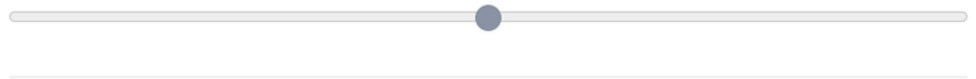
3. En función de lo que acabas de oír, ¿cómo caracterizarías a este hablante? Como alguien...

flojo 0 10 20 30 40 **regular** 50 60 70 80 **trabajador** 90 100



4. En función de lo que acabas de oír, ¿cómo caracterizarías a este hablante? Como alguien...

de pueblo/rural 0 10 20 30 40 **de ciudad pequeña** 50 60 70 80 **de ciudad grande** 90 100



5. En función de lo que acabas de oír, ¿cómo caracterizarías a este hablante? Como alguien...

cani 0 10 20 30 40 **normal** 50 60 70 80 **pijo** 90 100



Figure 2. Image of the five slider prompts. Translation: Based on what you just heard, how would you describe this speaker? Someone who is... 1. dull/boring, pleasant/nice, funny; 2. uneducated, high-school educated, or university-educated; 3. slacking, ordinary, or hard-working; 4. from a town/rural area, from a small city, or from a large city; 5. shabby, average, or flashy/posh.

To be noted, the perceived humor scale is more semantically ambiguous than the four other scales. The indexing of Andalusian Spanish in popular culture as sounding “funny” or of Andalusians themselves being joke-tellers could parallel the American “class clown” stereotype, describing those who are seen as funny yet in an inappropriate or attention-seeking manner. In the context of Andalusian stereotypes, this characteristic often goes hand in hand with unfavorable social attributes such as being uncultivated, lazy,

or originating from rural areas of Spain with lower socioeconomic status (Caballero-Gálvez, 2017; Gordillo, 2012).

The stimuli were randomized within a single block of 105 questions in Qualtrics. All participants received ten euros for their participation, and the average completion time was 41 minutes. The participants were informed that payment would not be released until three days after completing the survey, which allowed a research assistant to validate the data by confirming that

Table 2. Description of slider values at 0%, 50%, and 100% for the five social attributes

Social-attribute percept	“0” label	“50” label	“100” label
WORK ETHIC	<i>flojo</i> “slacking”	<i>regular</i> “ordinary”	<i>trabajador</i> “hard-working”
EDUCATION LEVEL	<i>con pocos estudios</i> “uneducated”	<i>con estudios secundarios</i> “high-school educated”	<i>con estudios universitarios</i> “university-educated”
URBANICITY	<i>de pueblo/rural</i> “from a town/rural area”	<i>de ciudad pequeña</i> “from a small city”	<i>de ciudad grande</i> “from a large city”
HABITUS	<i>cani</i> “shabby”	<i>normal</i> “average”	<i>pijo</i> “flashy/posh”
HUMOR	<i>soso</i> “dull/boring”	<i>agradable</i> “pleasant/nice”	<i>gracioso</i> “funny”

each participant had responded appropriately to the audio test and had entered the correct validation code.

3.4. Statistical modeling

There were six dependent variables in the original modeling: the participants’ geographical classifications based on the forced-choice question of north, central, or south (research question 1a); and the participants’ responses to the five social-attribute slider scales (research question 2). For research question 1a, we fitted a logistic mixed-effects model for the geographical classifications based on a “north/central” (due to low counts of “north” and “central” in some instances, see plots in Appendix C) versus “south” recoding. For research question 2, we fitted linear mixed-effects models (LMEMs) to the five continuous outcomes (work ethic, education level, habitus, urbanicity, humor) using the “lme4” package (Bates et al., 2015) within the statistical software package “R,” version 3.6.3 (R Core Team, 2020). The continuous outcomes were analyzed separately by each linguistic context (syllable-initial coronal fricatives, word-medial coda-/s/, word-final coda-/s/), leading to fifteen linear models (3 linguistic contexts \times 5 social-attribute percepts). To ensure that the linear-model assumptions were met, these continuous outcomes were log-transformed. In response to research question 1b, we created a series of mosaic plots and interpreted the data patterns through visual inspection of the plots.

As fixed factors, we included RESPONDENT CITY (Granada, Jerez de la Frontera, Salamanca), PHONETIC VARIANT (ten levels, see Table 1), and their two-way interaction in all models. We additionally considered each respondents’ SEX and AGE, to control for potential effects of the listeners’ sociodemographic characteristics on their individual response patterns (e.g., Cole, 2021; Walker et al., 2014). For fixed-effects selection, we followed the top-down strategy, starting with all potential effects in the model and sequentially removing all nonsignificant effects ($\alpha = 0.005$, see below), starting with the highest-order interaction (West, Welch & Galecki, 2014:39). For the 16 models presented in Sections 4.1 and 4.2, the two-way interaction of interest was significant in eight instances. In addition, the control variables SEX (binary variable) and AGE (continuous variable) were always removed from the final models due to lack of significance.

Regarding random-effects selection, for all models we included random-intercept effects for PARTICIPANT and STIMULUS to account for dependencies within data points from the same respondent and audio file. Further, for the models in Sections 4.1 and 4.2, we included by-PARTICIPANT random slopes for PHONETIC VARIANT, following Barr et al. (2013). All models with this random effect structure converged.

For each model, we computed type III F-statistics (χ^2 -statistics for the logistic model) for the two predictor variables of interest and their interaction. Additionally, as a measure of effect size, we computed marginal R^2 values (i.e., the percentage of variance explained by the reported variable adjusted for the other predictors in the model) using the R package “r2glmm” (Jaeger, 2017; Jaeger et al., 2017). In order to facilitate model-based inference regarding the fixed effects, we computed marginal predicted means based on the fitted models using the R package “ggeffects” (Lüdtke, 2018). This procedure calculates predicted values for the outcome variable based on all used observations for each category combination of the interaction of interest. To account for the multiple testing throughout the top-down modeling strategy with a beforehand-unknown number of tests, we set an adjusted significance level of $\alpha = 0.005$, following Benjamin et al. (2018).

As a follow-up analysis (see Section 4.4), we fitted models using the five social-attribute percepts as outcomes, and GEOGRAPHIC CLASSIFICATION (“north/central” versus “south”), RESPONDENT CITY, and their interaction as fixed effects, as well as random intercepts for PARTICIPANT and STIMULUS. Additionally, we included by-PARTICIPANT random slopes for GEOGRAPHIC CLASSIFICATION. The analysis followed the same modeling strategy as previously described. All models with this random effect structure converged.

4. Results

4.1. Geographical classification of variants (“north/central” versus “south”)

The first prompt for each stimulus asked participants to select from one of three options regarding the perceived regional origin of the audio file: north, central, or south. From the corresponding regression model, we uncovered a significant effect for the two-way interaction PHONETIC VARIANT * RESPONDENT CITY ($\chi^2(18) = 44.376$, $p < 0.001$, marginal $R^2 = 0.07$) on the binary outcome coded for “north/central” versus “south.” Figure 3 shows the model-derived marginal means and 95% confidence intervals for the outcome organized according to PHONETIC VARIANT (x-axis) and RESPONDENT CITY (legend). For plots of the raw data, see Appendix C.

Regarding the effect of PHONETIC VARIANT for syllable-initial coronal fricatives (leftmost panel of Fig. 3), the mean probability of being classified as “south” was overwhelmingly higher in the *ceceo*, *seceo*, and *seseo* conditions (mean values between 0.93 and 0.99) than in the *distinción* condition (mean values between 0.02 and 0.14). For word-medial /s/, the center

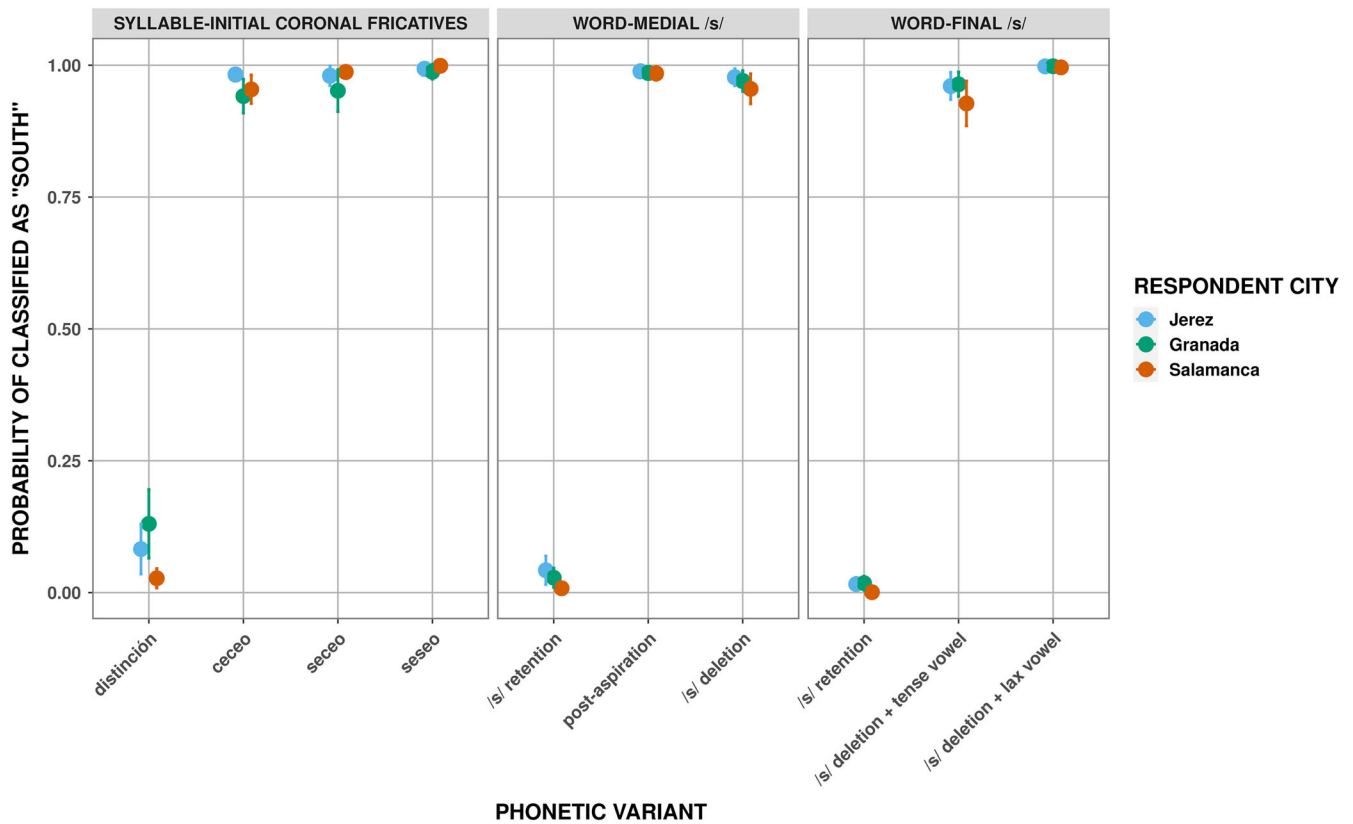


Figure 3. Model-predicted marginal means plot for the outcome GEOGRAPHIC CLASSIFICATION per the predictors PHONETIC VARIANT and RESPONDENT CITY. Error bars mark the 95% confidence intervals around each predicted mean value.

panel shows that this probability was likewise higher in the post-aspiration and /s/-deletion conditions (means between 0.94 and 0.99) than in the /s/-retention condition (means between 0.01 and 0.06). For word-final coda-/s/ (rightmost panel), the probability of being classified as “south” was again higher in the two /s/-lenition conditions (means between 0.90 and 0.99) than in the /s/-retention condition (means between 0.00 and 0.03).

With respect to the interaction between PHONETIC VARIANT and RESPONDENT CITY, the Granada listeners displayed a higher degree of variation than the other two listener groups in the classification of the syllable-initial coronal fricatives. In particular, the Granada listeners were the most likely of the three groups to classify *distinción* as a southern feature and *ceceo* and *seceo* as northern/central features.

4.2. Open-ended perceptual classification of phonetic variants

The second prompt for each stimulus asked participants to indicate the most specific geographic origin that they associated with the audio file (see Figure 1 for the Spanish wording). We initially coded these open-ended responses based on the following three-way classification: (1) “north/central” that incorporates all responses referencing a locality in continental Spain outside of Andalusia; (2) “Andalusia-general,” which incorporates all responses referencing the autonomous community of Andalusia without listing a more specific place name (e.g., *Andalusía*, *andaluz*); or (3) “Andalusia-specific,” which incorporates all responses that reference a town or province located within Andalusia proper.⁴ Figure 4 visualizes the

response-type counts in a mosaic plot organized according to PHONETIC VARIANT and RESPONDENT CITY. A mosaic plot is a visual representation of a cross table, meaning that the interpretation relies on tile size, that is, larger areas represent higher counts of the respective response category. Importantly for this plot, the underlying data come from a near-balanced sample of participants with respect to the three respondent cities.

As Figure 4 shows, most respondents classified the *distinción* stimuli as being from outside of Andalusia, and *ceceo*, *seceo*, and *seseo* as originating from within the Andalusian region. For word-medial and word-final coda-/s/, the /s/-retention variants were overwhelmingly identified as pertaining to northern and central regions of Spain, while the /s/-lenition variants were associated with specific Andalusian localities or the general south.

To further explore the participants’ free responses, Figure 5 plots the weighted⁵ “Andalusia-specific” category based on a fine-grained classification of Andalusia’s intraregional variation (see Jaime-Jiménez, 2019):

- (1) western Andalusia (localities in the provinces of Cádiz, Huelva, and Seville);
- (2) central Andalusia (localities in the provinces of Córdoba and Málaga); and
- (3) eastern Andalusia (localities in the provinces of Almería, Granada, and Jaén).

One main finding is that the participants listed western-Andalusian localities more often than central and eastern localities combined (i.e., the overall plotted area of western

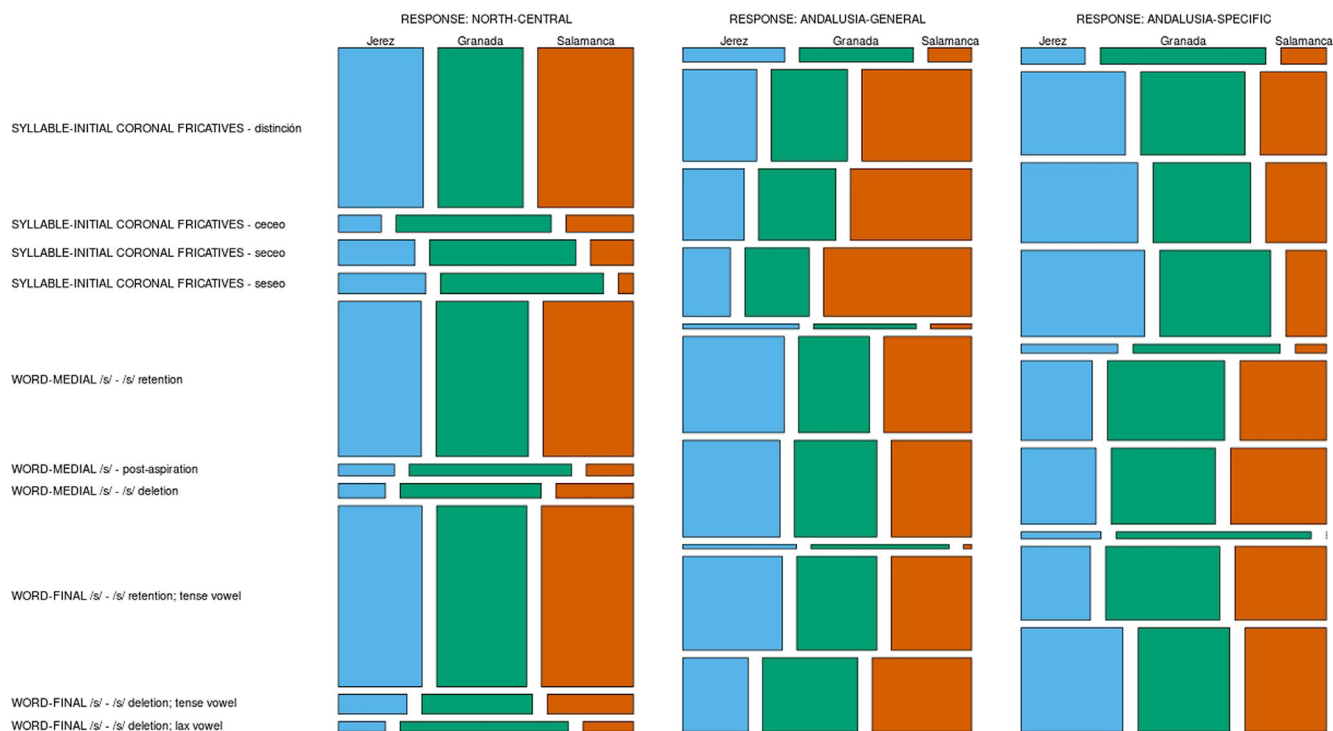


Figure 4. Mosaic plot of listeners' perceptions of speaker origin by PHONETIC VARIANT and RESPONDENT CITY.

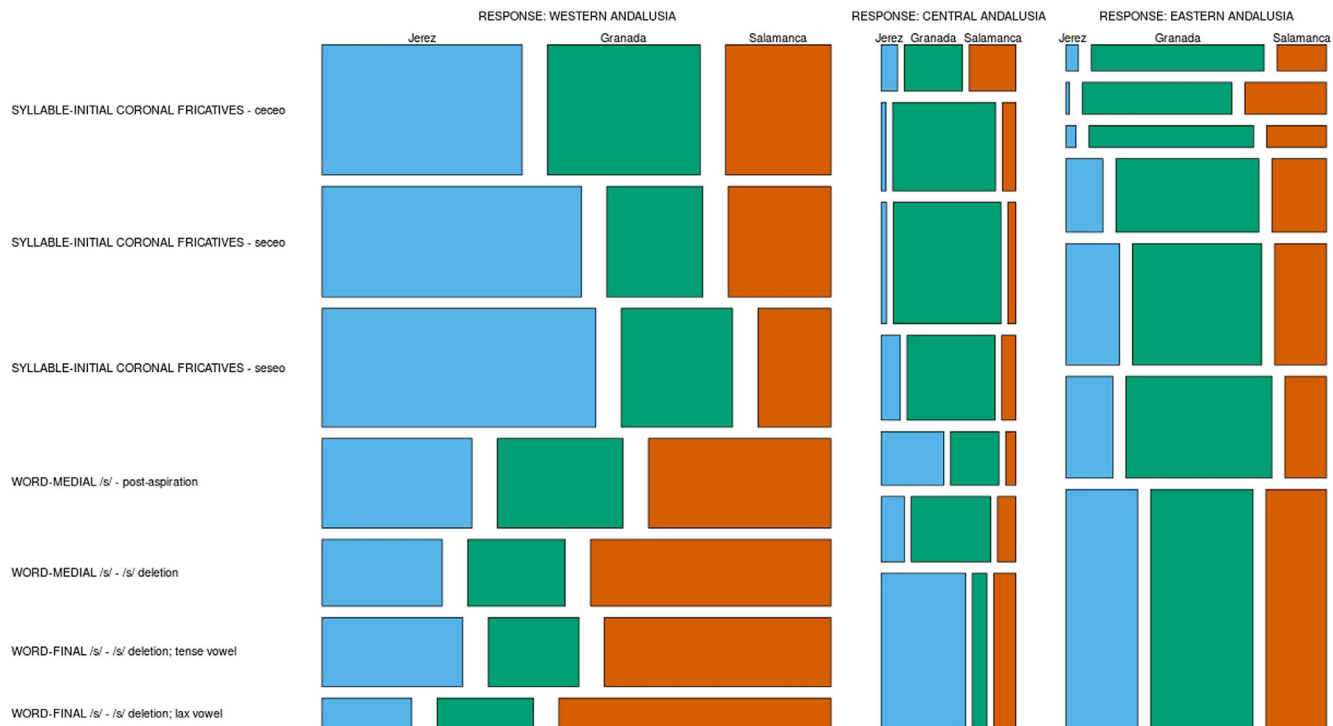


Figure 5. Mosaic plot of listeners' perceptions of speaker origin by PHONETIC VARIANT and RESPONDENT CITY. The data shown here reflect a western/central/eastern subclassification of the "Andalusia-specific" responses from Figure 4. The listeners' responses to the *distinción* and /s/-retention stimuli are excluded due to low token counts.

Andalusian responses is considerably larger than the combined area of central and eastern Andalusian responses). Figure 5 further shows that for the syllable-initial coronal fricatives, the respondents predominantly classified *ceceo*,

and *seseo* as western Andalusian features. In particular, the Jerez respondents displayed a near-categorical pattern of attributing the non-*distinción* stimuli to the western provinces.

Table 3. Results of significance tests from the fifteen models fitted for research question 2. X = significant result; n.s. = not significant result; = main effect was not tested due to the presence of the interaction in the model

Linguistic context	Outcome	Interaction RESPONDENT CITY * PHONETIC VARIANT	Main effect RESPONDENT CITY	Main effect PHONETIC VARIANT	Test result	Marginal R ²
Syllable- initial coronal fricatives	Work ethic	X	–	–	F(6, 161.009) = 3.942, p = 0.001	0.013
	Education level	n.s.	n.s.	X	F(3, 58.949) = 37.176, p < 0.001	0.123
	Habitus	X	–	–	F(6, 160.992) = 9.454, p < 0.001	0.038
	Urbanicity	X	–	–	F(6, 161.039) = 3.813, p = 0.001	0.014
	Humor	n.s.	n.s.	X	F(3, 61.250) = 36.674, p < 0.001	0.114
Word- medial coda-/s/	Work ethic	X	–	–	F(4, 161.001) = 4.081, p = 0.004	0.010
	Education level	n.s.	n.s.	X	F(2, 45.601) = 53.062, p < 0.001	0.174
	Habitus	X	–	–	F(4, 161.005) = 9.298, p < 0.001	0.042
	Urbanicity	X	–	–	F(4, 160.996) = 4.706, p = 0.001	0.018
	Humor	n.s.	n.s.	X	F(2, 67.797) = 49.352, p < 0.001	0.122
Word- final coda-/s/	Work ethic	n.s.	n.s.	X	F(2, 59.436) = 99.314, p < 0.001	0.194
	Education level	n.s.	n.s.	X	F(2, 50.419) = 73.581, p < 0.001	0.206
	Habitus	X	–	–	F(2, 65.312) = 150.191, p < 0.001	0.009
	Urbanicity	n.s.	n.s.	X	F(4, 161.940) = 3.819, p = 0.005	0.290
	Humor	n.s.	n.s.	X	F(2, 163.000) = 36.942, p < 0.001	0.071

For word-medial coda-/s/, the patterns were overall more varied. Specifically, post-aspiration evoked more western Andalusian responses than central and eastern responses, while /s/-deletion resulted in high response counts for both western Andalusia and eastern Andalusia. The Salamanca listeners were also the most likely of the three groups to perceive post-aspiration as being from western Andalusia.

For word-final /s/, all groups demonstrated a greater preference for classifying the /s/-deletion + tense-vowel stimuli as pertaining to western Andalusia than to central or eastern Andalusia. The participants were more likely to identify the lax-vowel stimuli, by contrast, as being from eastern Andalusia.

To summarize Sections 4.1 and 4.2, the respondents overwhelmingly perceived *distinción* and /s/-retention as being native to northern and central Spain, while they classified the three non-*distinción* variants and the /s/-lenition variants as being southern features. Regarding the free-response prompt, *ceceo*, *seceo*, *seseo*, and post-aspiration provoked the most western Andalusian responses, while word-final lax vowels provoked the most eastern Andalusian responses.

4.3. Social-attribute percepts

For seven of the fifteen fitted models for the derived outcomes from research question 2, we uncovered a statistically significant effect for the two-way interaction PHONETIC VARIANT * RESPONDENT CITY (see Table 3). For the remaining eight models, the main effect of PHONETIC VARIANT was significant. In the following subsections of Section 4.3, we therefore focus on describing the model-predicted marginal-means plots based on the observed effects of the significant two-way interaction or the main effect of PHONETIC VARIANT.

4.3.1. Syllable-initial coronal fricatives

This section presents the results from the listeners' social-attribute ratings for the four variants included in the context of syllable-initial coronal fricatives; the results for each social-attribute percept are graphically presented in Figures 6–10.

Collectively, Figures 6–9 show that the listeners assigned the highest (or most favorable) mean ratings to *distinción* and the lowest means to *ceceo*. *Seceo* and *seseo* fell between the more diametrically opposed *distinción* and *ceceo* (when comparing within the responses for each of the three respondent cities). Regarding the interaction between PHONETIC VARIANT and RESPONDENT CITY for perceived work ethic, habitus, and urbanicity, the Jerez listeners' ratings were the most extreme for *distinción* and *ceceo* when compared to the Granada and Salamanca listeners (i.e., the Jerez listeners assigned the highest means to *distinción* and the lowest means to *ceceo*). Notably, for perceived urbanicity (Fig. 9), all of the *ceceo* estimated means were below 50, suggesting that this variant signals more rural-sounding speech.

Regarding perceived humor, Figure 10 shows that *distinción* received lower ratings than *ceceo*, *seceo*, and *seseo*, indicating that the latter three variants are overall more “funny-sounding” than *distinción*. It is noteworthy that the numerical results for perceived humor reveal a trend opposite to that of Figures 6–9. This suggests that the respondents remained engaged while completing the survey rather than indiscriminately moving all five sliders in the same direction.

4.3.2. Word-medial coda-/s/

For word-medial coda-/s/, recall that there were three levels of PHONETIC VARIANT: /s/-retention, post-aspiration of /s/, and /s/-deletion. Figures 11–15 plot the marginal means and confidence intervals for the five social-attribute percepts.

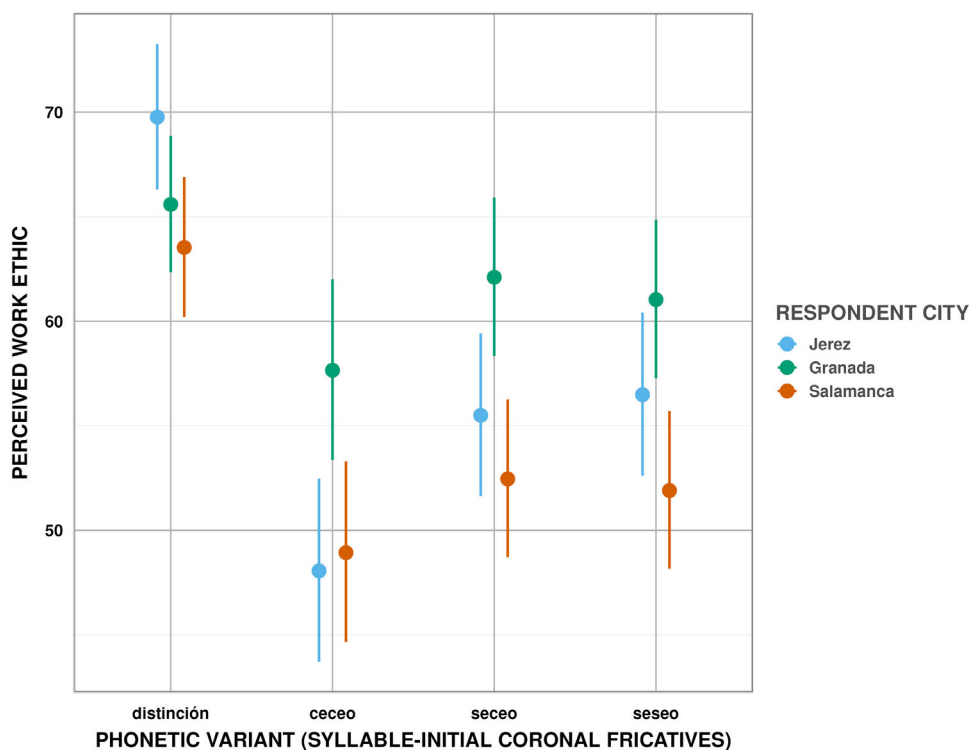


Figure 6. Model-predicted marginal means plot for the outcome PERCEIVED WORK ETHIC per the predictors RESPONDENT CITY and PHONETIC VARIANT for syllable-initial coronal fricatives. Error bars mark the 95% confidence intervals around each predicted mean value.

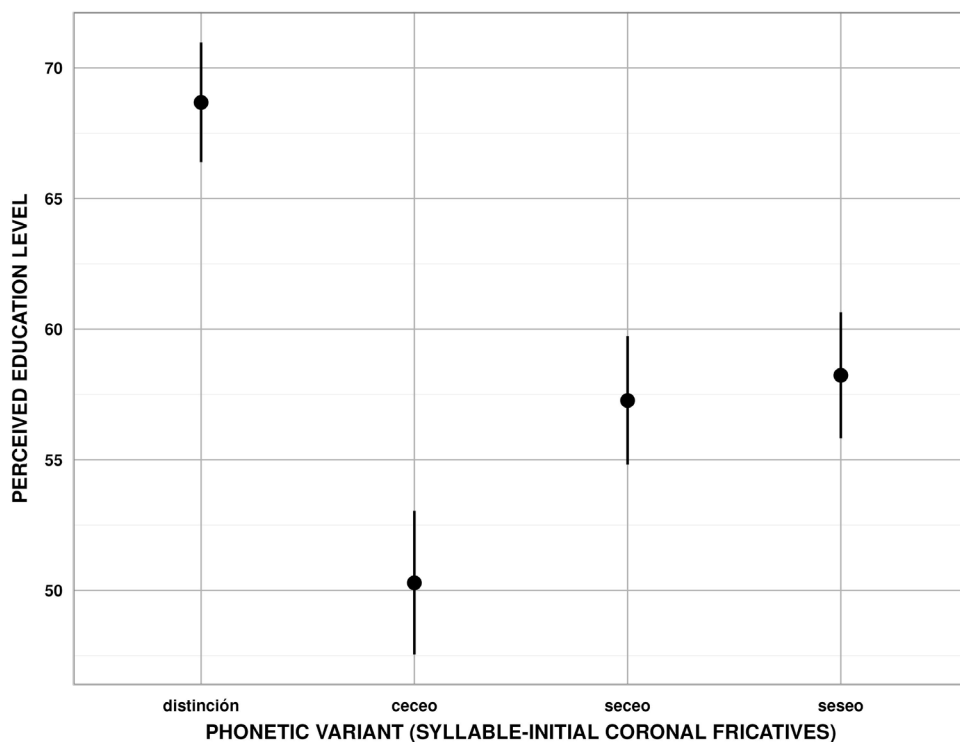


Figure 7. Model-predicted marginal means plot for the outcome PERCEIVED EDUCATION LEVEL per the predictor PHONETIC VARIANT for syllable-initial coronal fricatives. Error bars mark the 95% confidence intervals around each predicted mean value.

For the first four outcomes (i.e., work ethic, education level, habitus, and urbanicity), Figures 11-14 show that /s/-retention received the highest ratings and /s/-deletion received the lowest ratings, with post-aspiration falling intermediate. Regarding the interaction effects for perceived work ethic, habitus, and

urbanicity, the listeners from Jerez rated /s/-retention and /s/-deletion most extremely (i.e., these participants assigned the highest means to /s/-retention and the lowest means to /s/-deletion). For perceived work ethic (Fig. 11), the Granada respondents' means were the highest of the three listener groups

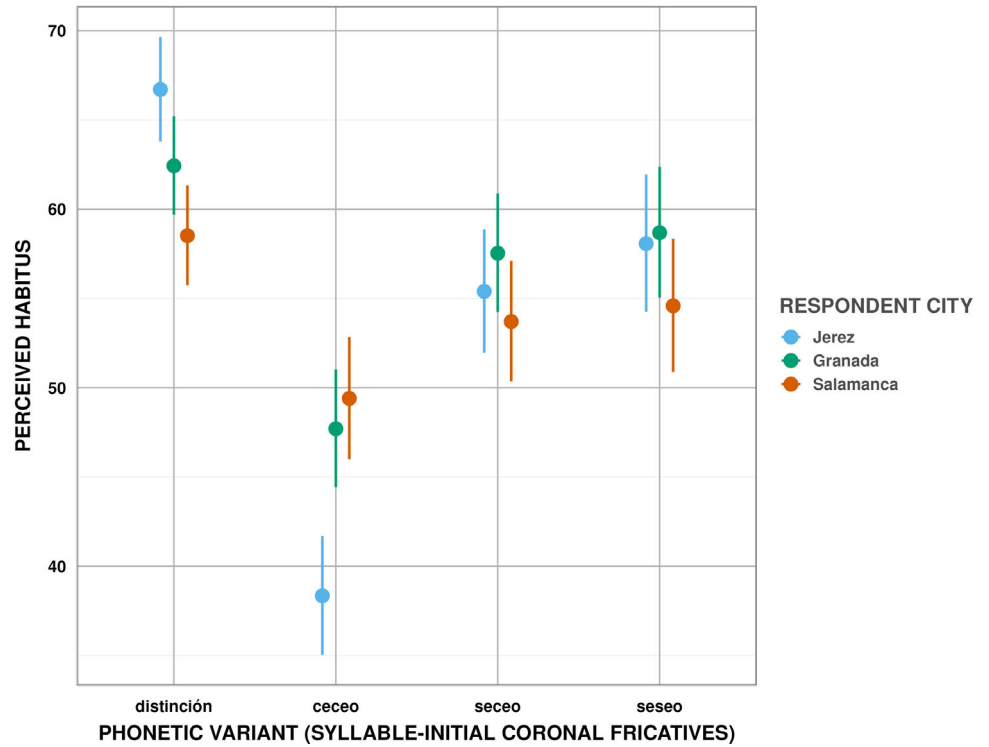


Figure 8. Model-predicted marginal means plot for the outcome PERCEIVED HABITUS per the predictors RESPONDENT CITY and PHONETIC VARIANT for syllable-initial coronal fricatives. Error bars mark the 95% confidence intervals around each predicted mean value.

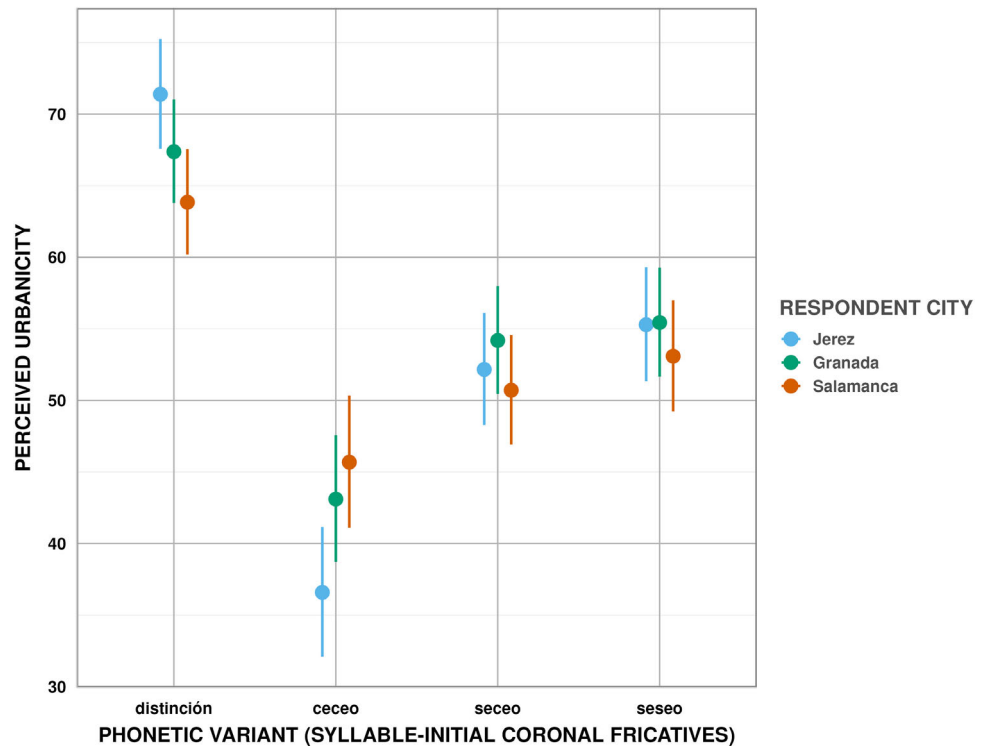


Figure 9. Model-predicted marginal means plot for the outcome PERCEIVED URBANICITY per the predictors RESPONDENT CITY and PHONETIC VARIANT for syllable-initial coronal fricatives. Error bars mark the 95% confidence intervals around each predicted mean value.

for the post-aspiration and the /s/-deletion stimuli. All groups' response ranges concerning perceived urbanicity fell below 50 (Fig. 14), emulating the effect of *ceceo* on this outcome shown in Section 4.3.1 (Fig. 9). For perceived humor (Fig. 15), post-aspiration received the highest ratings, while /s/-retention received the lowest ratings. The /s/-deletion stimuli received intermediate ratings that were closer to post-aspiration than to /s/-retention.

4.3.3. Word-final coda-/s/

This section presents the results for word-final coda-/s/ based on the three levels of PHONETIC VARIANT: /s/-retention; /s/-deletion + tense vowel; and /s/-deletion + lax vowel.

Figures 16-19. (i.e., perceived work ethic, education level, habitus, and urbanicity) show that the listeners converged in assigning the highest values to /s/-retention, followed by

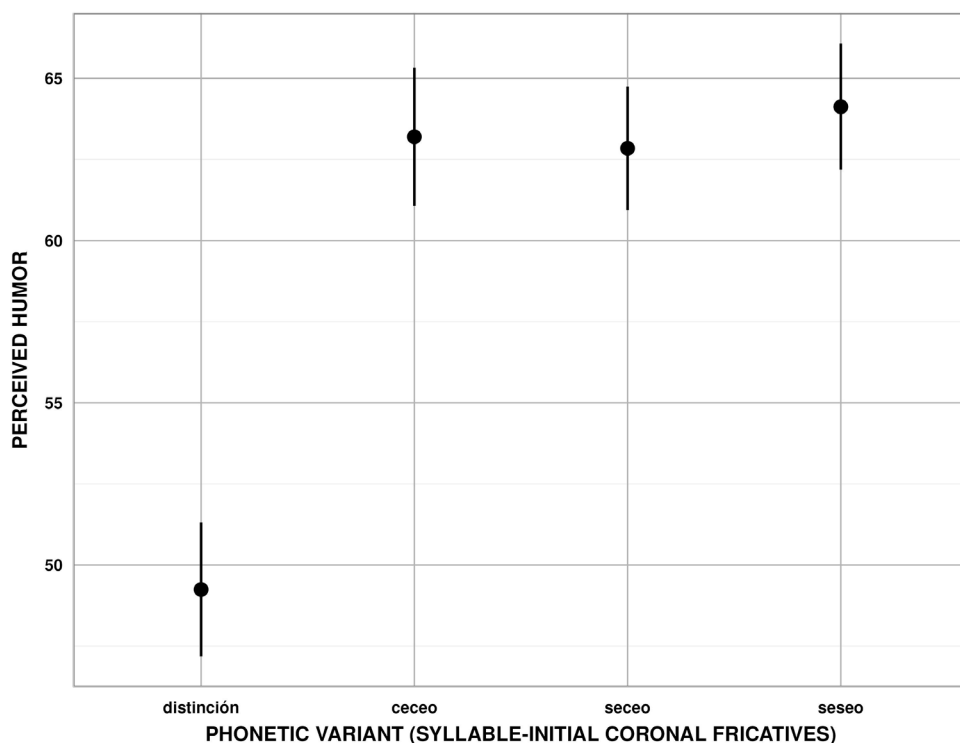


Figure 10. Model-predicted marginal means plot for the outcome PERCEIVED HUMOR per the predictor PHONETIC VARIANT for syllable-initial coronal fricatives. Error bars mark the 95% confidence intervals around each predicted mean value.

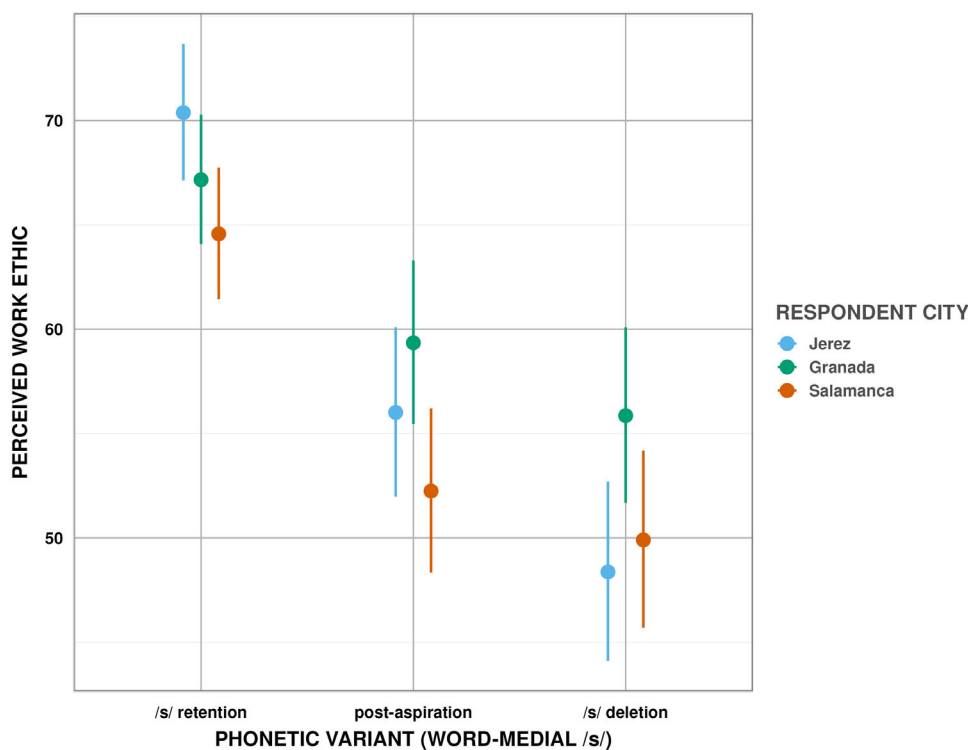


Figure 11. Model-predicted marginal means plot for the outcome PERCEIVED WORK ETHIC per the predictors RESPONDENT CITY and PHONETIC VARIANT for word-medial coda-/s/. Error bars mark the 95% confidence intervals around each predicted mean value.

/s/-deletion + tense vowel, and finally by /s/-deletion + lax vowel. When the interaction effect between PHONETIC VARIANT and RESPONDENT CITY was significant (in this instance, for perceived habitus), the Jerez listeners had assigned the highest mean values to /s/-retention. Concerning both perceived habitus

(Fig. 18) and perceived urbanicity (Fig. 19), the three groups converged in assigning means well below 50 to the /s/-deletion + lax vowel stimuli. Finally, Figure 20 shows that perceived humor was highest for /s/-deletion + lax vowel and lowest for /s/-retention.

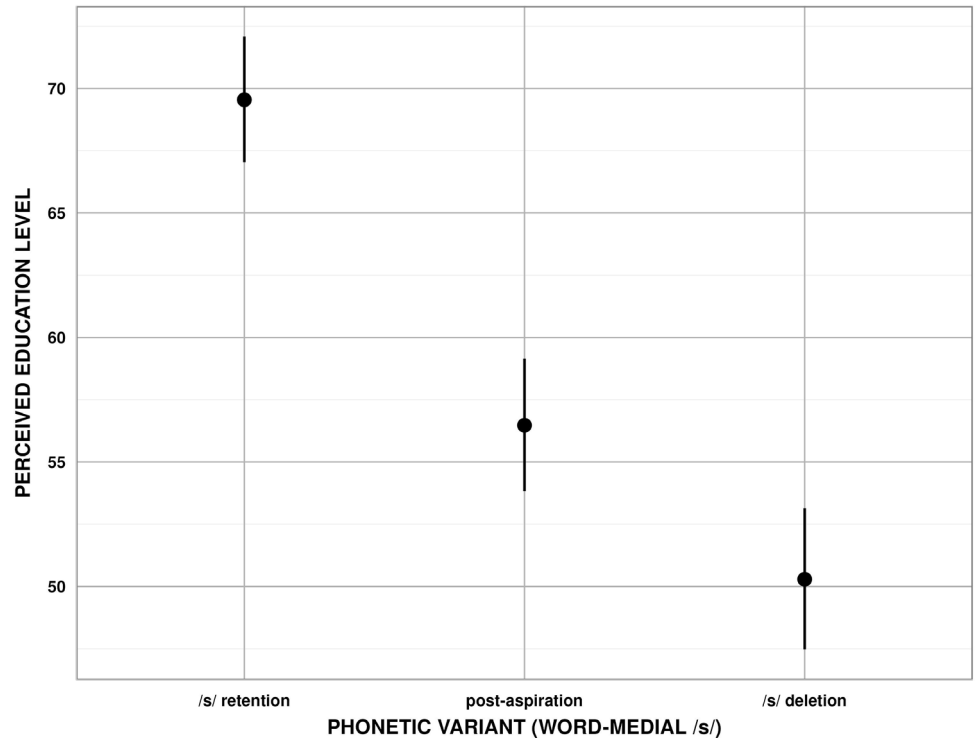


Figure 12. Model-predicted marginal means plot for the outcome PERCEIVED EDUCATION LEVEL per the predictor PHONETIC VARIANT for word-medial coda-/s/. Error bars mark the 95% confidence intervals around each predicted mean value.

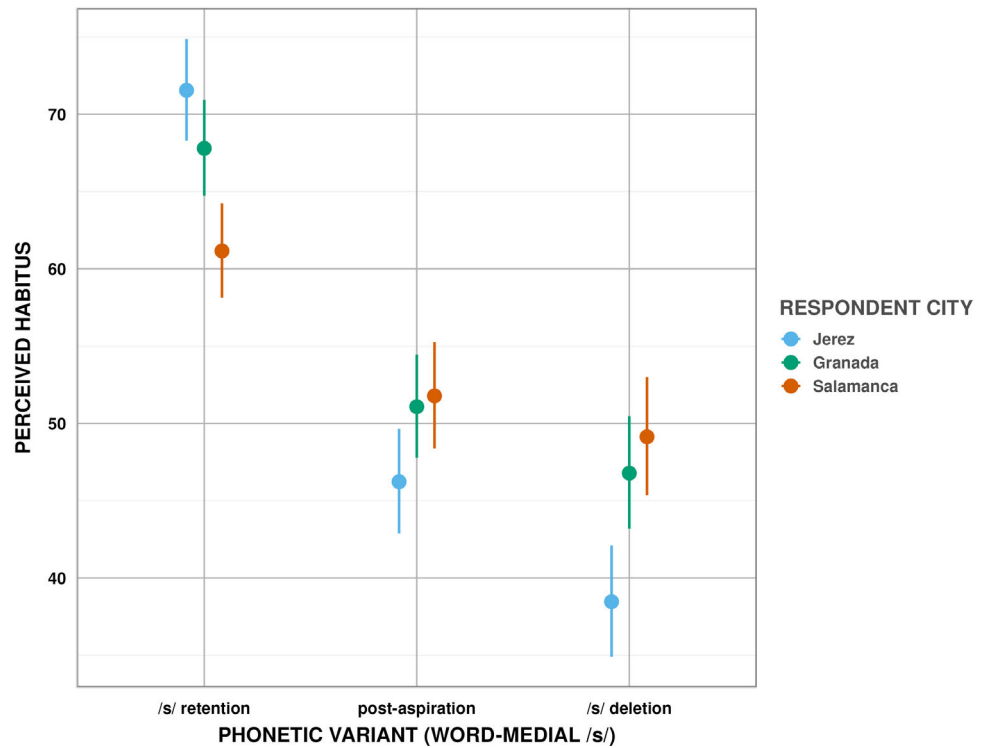


Figure 13. Model-predicted marginal means plot for the outcome PERCEIVED HABITUS per the predictors RESPONDENT CITY and PHONETIC VARIANT for word-medial coda-/s/. Error bars mark the 95% confidence intervals around each predicted mean value.

4.4. Respondents’ associations between geographic classifications and social-attribute percepts

The results from Sections 4.1 and 4.2 indicate that the non-*distinción* variants for the syllable-initial coronal fricatives and all /s/-lenition variants for both categories of syllable-final-/s/ were the most likely to be classified as originating from southern Spain.

The results from Section 4.3 indicate that this same set of variants consistently received the least favorable social-attribute ratings. To investigate the relationship between these two outcomes, we fitted LMEMs with each of the five social attributes as outcomes and RESPONDENT CITY, GEOGRAPHIC CLASSIFICATION, and their interaction as predictors. In three of the five models

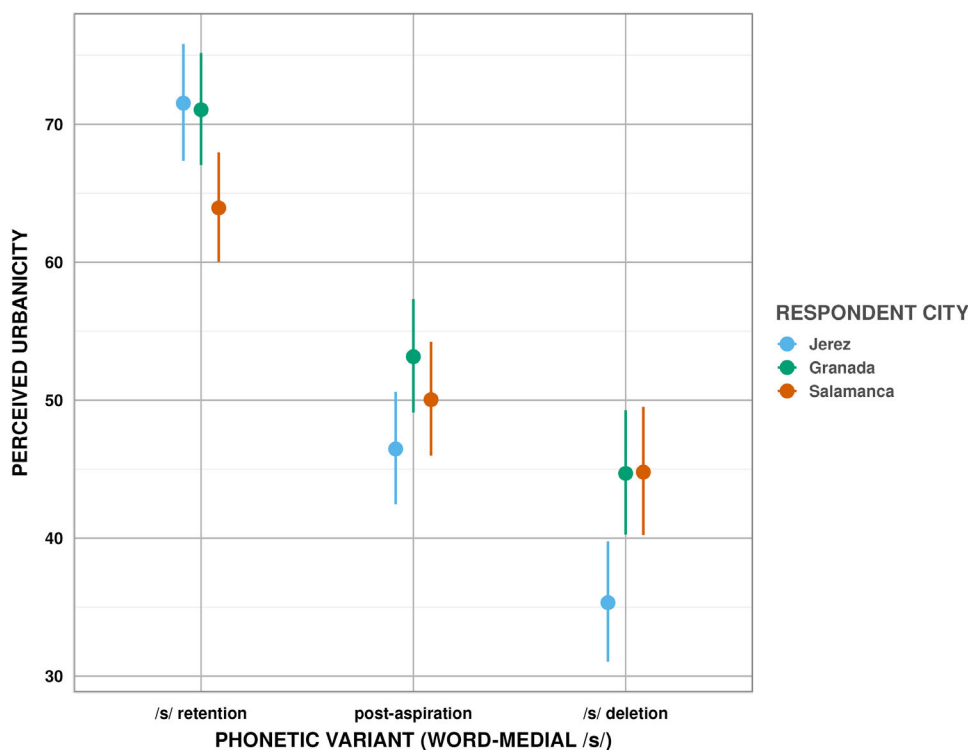


Figure 14. Model-predicted marginal means plot for the outcome PERCEIVED URBANICITY per the predictors RESPONDENT CITY and PHONETIC VARIANT for word-medial coda-/s/. Error bars mark the 95% confidence intervals around each predicted mean value.

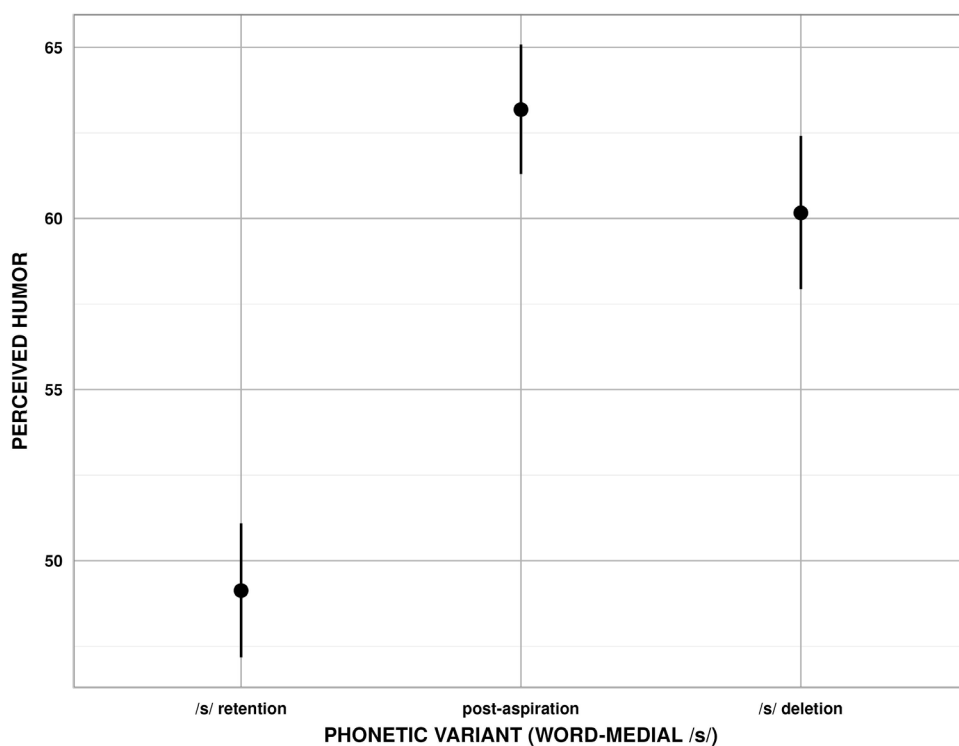


Figure 15. Model-predicted marginal means plot for the outcome PERCEIVED HUMOR per the predictor PHONETIC VARIANT for word-medial coda-/s/. Error bars mark the 95% confidence intervals around each predicted mean value.

(perceived habitus, urbanicity, and humor), we found a significant interaction between RESPONDENT CITY and GEOGRAPHIC CLASSIFICATION (see Table 4).

The results of this analysis are visually presented in Figure 21. The figure displays model-predicted marginal means of the interaction levels for RESPONDENT CITY * GEOGRAPHIC CLASSIFICATION with corresponding 95% confidence intervals

for each of the social attributes. The predicted means confirm that the stimuli classified as “south of Spain” were also more likely to be evaluated as having lower work ethic, education level, habitus, and urbanicity, as well as being more humorous than the stimuli classified as pertaining to northern or central Spain. When the interaction effect was significant (i.e., for perceived habitus, urbanicity, and humor), the Jerez respondents displayed the largest

Figure 16. Model-predicted marginal means plot for the outcome PERCEIVED WORK ETHIC per the predictor PHONETIC VARIANT for word-final coda-/s/. Error bars mark the 95% confidence intervals around each predicted mean value.

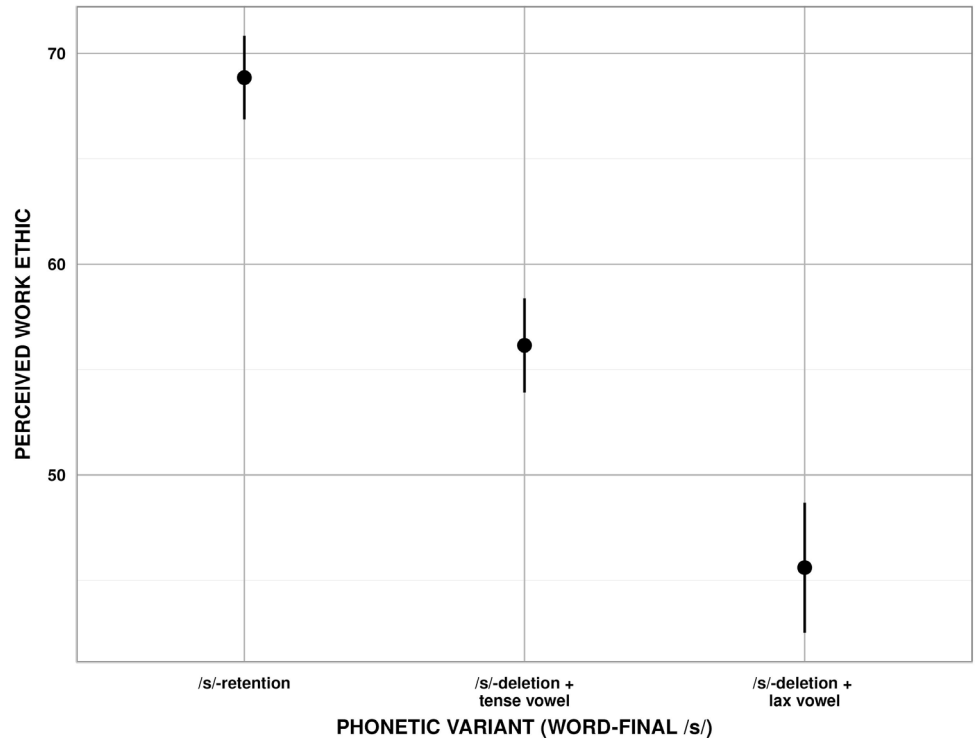
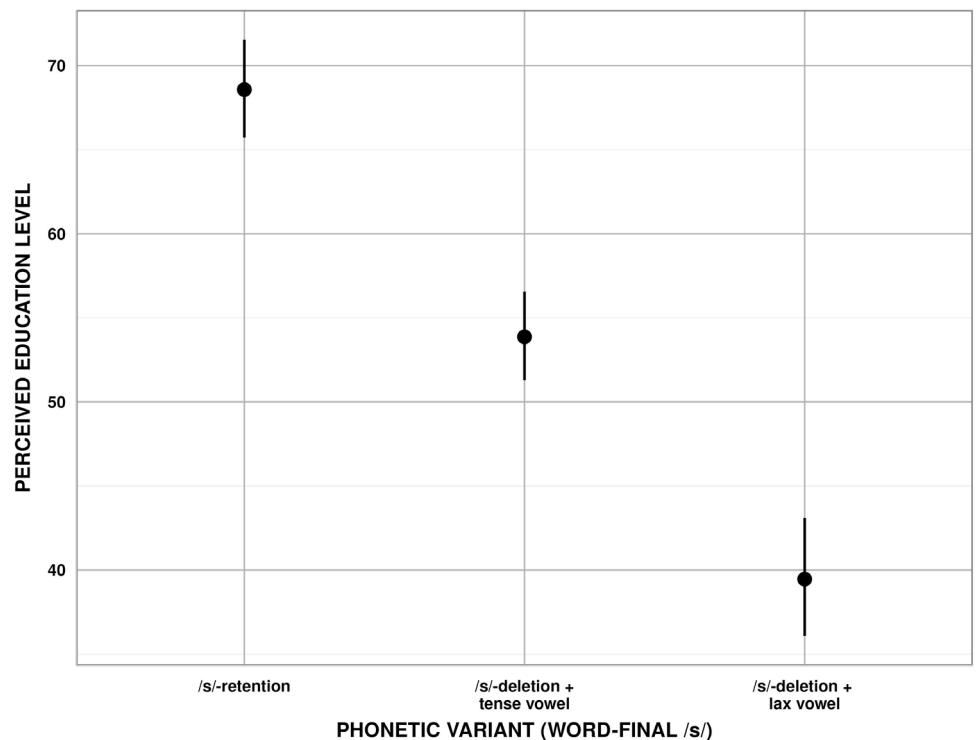


Figure 17. Model-predicted marginal means plot for the outcome PERCEIVED EDUCATION LEVEL per the predictor PHONETIC VARIANT for word-final coda-/s/. Error bars mark the 95% confidence intervals around each predicted mean value.



differences in their predicted means between “north/central” and “south,” while the Salamanca respondents showed more subtle differences on the habitus and urbanicity scales.

To summarize, from Section 4.4 we find a relationship between the participants’ geographical classifications and their social evaluations. The stimuli classified as being from the south of Spain were consistently attributed the least favorable social meanings.

5. Discussion

5.1. Responding to research question 1

Research question 1 inquired into listeners’ regional classifications of the ten phonetic variants included in the survey. Research question 1a specifically considered listeners’ responses on a forced north/central/south classification, while research

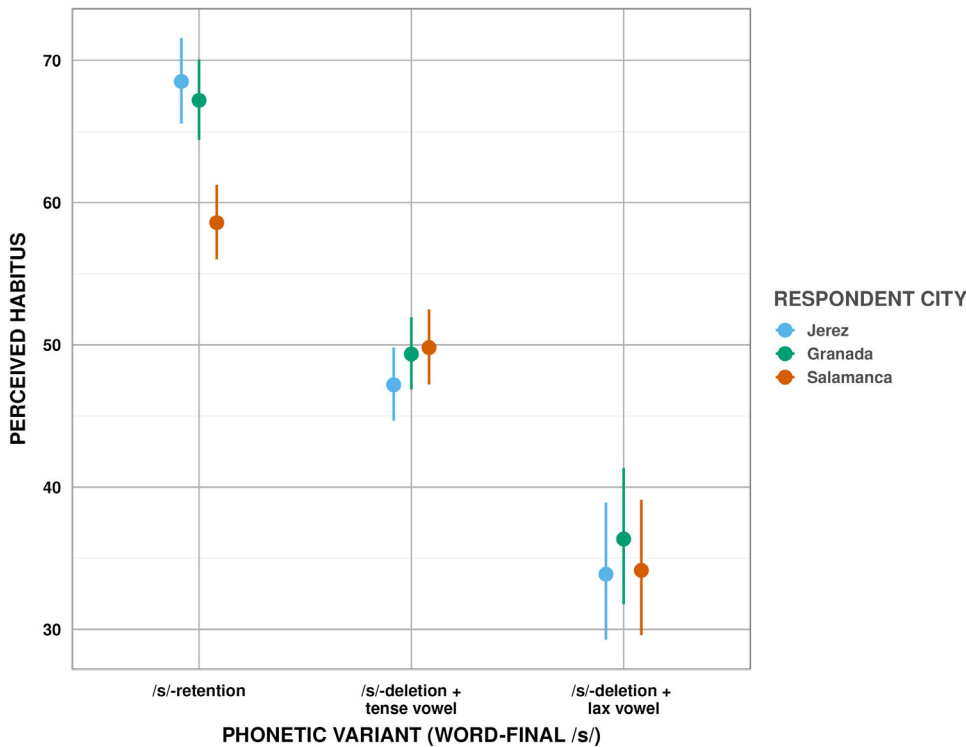


Figure 18. Model-predicted marginal means plot for the outcome PERCEIVED HABITUS per the predictors RESPONDENT CITY and PHONETIC VARIANT for word-final coda-/s/. Error bars mark the 95% confidence intervals around each predicted mean value.

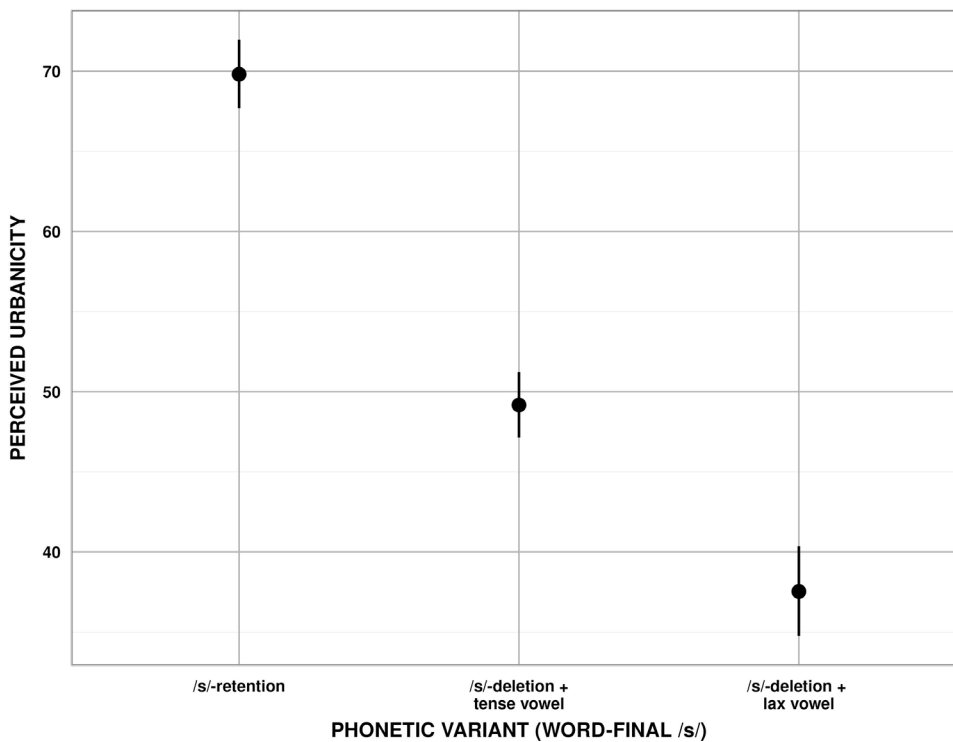


Figure 19. Model-predicted marginal means plot for the outcome PERCEIVED URBANICITY per the predictor PHONETIC VARIANT for word-final coda-/s/. Error bars mark the 95% confidence intervals around each predicted mean value.

question 1b considered the responses from an open-ended prompt.

Research question 1a

Syllable-initial coronal fricatives. *Distinción* was more likely to be classified as a northern or central feature, whereas the *ceceo*, *seceo*, and *seseo* stimuli were far more likely to be classified as southern

features. Although the listener groups showed fair agreement overall, the response patterns for the syllable-initial coronal fricatives displayed a greater level of variation than for word-medial and word-final coda-/s/. In particular, the Granada listeners were the most likely of the three groups to classify *distinción* as a southern feature and *ceceo* and *seceo* as northern features. The Granada listeners' heightened recognition of geolinguistic

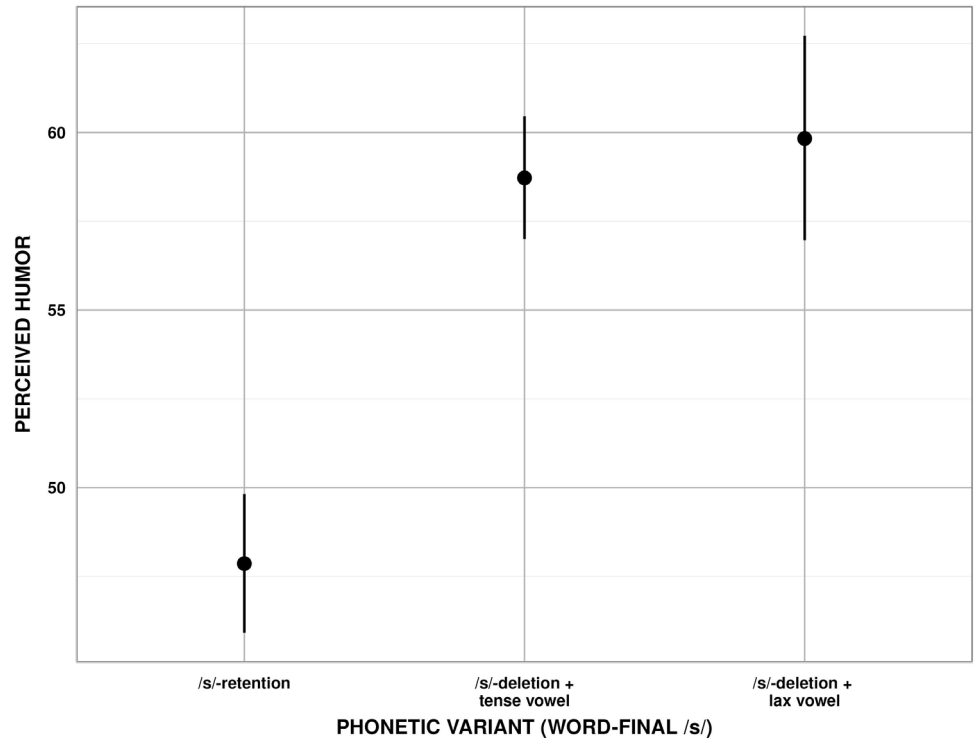


Figure 20. Model-predicted marginal means plot for the outcome PERCEIVED HUMOR per the predictor PHONETIC VARIANT for word-final coda-/s/. Error bars mark the 95% confidence intervals around each predicted mean value.

Table 4. Results of the significance tests from the five models for WORK ETHIC, EDUCATION LEVEL, HABITUS, URBANICITY, and HUMOR (Section 4.4)

Outcome	Test result for interaction of RESPONDENT CITY * GEOGRAPHIC CLASSIFICATION	Marginal R ²
Work ethic	F(2,161.0) = 4.065, p = 0.019	0.005
Education level	F(2,160.3) = 5.059, p = 0.007	0.009
Habitus	F(2,160.7) = 16.857, p < 0.001	0.023
Urbanicity	F(2,160.0) = 9.620, p < 0.001	0.012
Humor	F(2,159.8) = 7.403, p < 0.001	0.016

variation with respect to *distinción* resonates with previous studies showing that speakers from the eastern Andalusian region are more likely to use the *distinción* national standard compared to western or central Andalusians. It is hypothesized that this difference arises from sociolinguistic factors such as dialect mixing and regional dialect leveling, which are more common in the eastern provinces (Melguizo Moreno, 2007; Villena-Ponsoda, 2008).

Word-medial and word-final coda-/s/. All participant groups predominantly classified the /s/-retention stimuli as being from northern or central Spain and the /s/-lenition stimuli as pertaining to the south. With respect to group differences, the Salamanca respondents were near categorical in classifying /s/-retention as a strictly northern/central feature. Concerning the effect of PHONETIC VARIANT, the lax-vowel stimuli were the most likely to be classified as “south.”

Overall, these geographic classifications suggest that listeners of Iberian Spanish hold rigid perceptual categorizations of coda-/s/ variation. The seemingly invariant geographic indices tied to coda-/s/ suggest that /s/-lenition operates as a robust and widespread sociophonetic cue for signaling a speaker’s regional

origin (Gordillo, 2012). Such findings notwithstanding, it is important to emphasize that /s/-lenition also manifests in areas well to the north of the Andalusian border, such as in nearby Castile-La Mancha and the country’s capital Madrid (Henriksen & Harper, 2016; Momcilovic, 2009).

5.1.2. Research question 1b

Regarding the listeners’ responses to the free-response prompt, we first analyzed the data using a three-way classification coded for northern/central localities, “Andalusia” as a general category, or specific localities within Andalusia. The results closely resembled the findings from research question 1a.

We subsequently coded the “Andalusia-specific” responses as western, central, or eastern Andalusian. A main finding was that localities from western Andalusia were overall more frequent among the participants’ responses than localities from central or eastern Andalusia. For the syllable-initial coronal fricatives in particular, the respondents overwhelmingly converged in classifying *ceceo*, *seceo*, and *seseo* as western Andalusian. This finding confirms earlier proposals that listeners readily identify a speaker’s Andalusian origin based on non-*distinción* features (Regan, 2022:23). In other words, listeners have formed a collective indexical link between nonstandardized variants and the western provinces.

With respect to word-medial /s/ + /p t k/ sequences, Ruch (2018) showed that listeners from Andalusia are more likely to rate post-aspirated stops as either younger-sounding or western Andalusian rather than as older-sounding or eastern Andalusian. While our results corroborate those of Ruch (2018), a novel finding is that the Salamanca respondents were the most likely of the three groups to classify post-aspiration as a western feature and the least likely to classify it as eastern. The rigidity reflected in the Salamanca participants’ classifications suggests that while NCPS listeners may be attuned to the western Andalusian origin of the

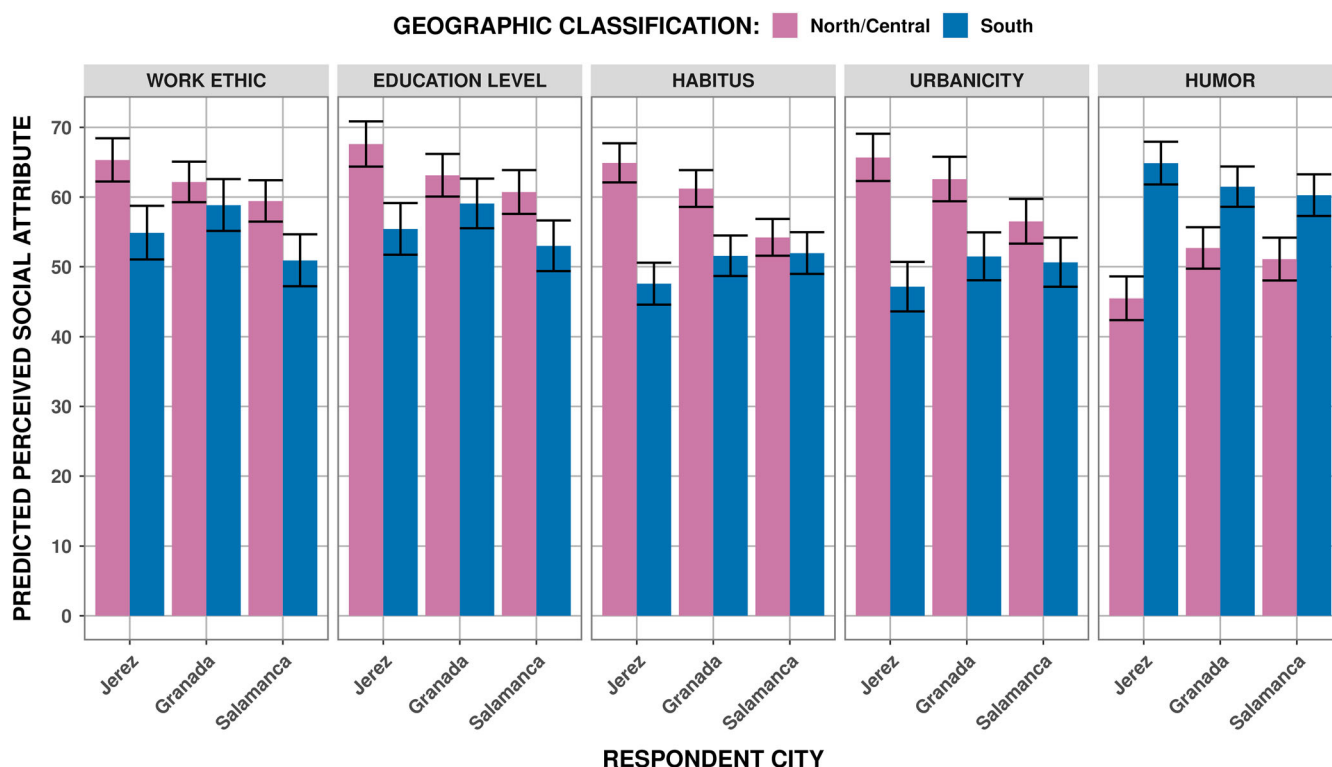


Figure 21. Model-predicted marginal-means plot for the outcomes PERCEIVED WORK ETHIC, PERCEIVED EDUCATION LEVEL, PERCEIVED HABITUS, PERCEIVED URBANICITY, and PERCEIVED HUMOR per the predictors RESPONDENT CITY and GEOGRAPHIC CLASSIFICATION. Error bars mark the 95% confidence intervals around each predicted mean value.

sound change, they remain less aware of the ongoing diffusion of post-aspiration to the eastern provinces (see Ruch & Harrington, 2014; Ruch & Peters, 2016).

Regarding word-final coda-/s/, more respondents classified the lax-vowel stimuli as having origins in eastern Andalusia than in central or western Andalusia, which supports previous work proposing that vowel laxing is a robust marker of intraregional variation in Andalusian Spanish (Hernández-Campoy & Trudgill, 2002). There was less agreement overall in classifying the lax-vowel stimuli as pertaining to the central provinces.

5.2. Responding to research question 2

Research question 2 considered Iberian Spanish listeners' social evaluations of the ten phonetic variants. The interaction of PHONETIC VARIANT * RESPONDENT CITY proved significant in seven of the fifteen LMEMs; in all other models, PHONETIC CATEGORY alone was significant.

5.2.1 Syllable-initial coronal fricatives

The *distinción* stimuli received the most favorable attribute ratings, while *ceceo* generally received the least favorable ratings (i.e., the lowest means in terms of perceived work ethic, education level, urbanicity, and habitus, and the highest mean for perceived humor); the values for *seceo* and *seseo* were generally intermediate. The Jerez listeners demonstrated the most polarized response patterns by assigning the most favorable means to *distinción* and the least favorable means to *ceceo*.

The overall response patterns for *distinción* support the findings of matched-guise perception studies (Regan, 2019; 2022) and of production studies (e.g., Melguizo Moreno, 2007;

Regan, 2020b), showing that Andalusian speakers of higher socioeconomic status draw from nationwide ideologies in their increased usage of the *distinción* system (recall that our participants were primarily university educated). In light of this, Regan (2022) has proposed that listeners' implicit language attitudes toward *distinción* and *ceceo* operate with respect to social meanings referencing socioeconomic class, urbanicity, formality, educational level, and profession. Our study emphasizes that an additional attribute, namely perceived humor, merits inclusion into the field of social meanings indexed with syllable-initial coronal fricatives.

This study additionally offers new findings concerning the respondents' evaluations of *seceo* and *seseo*, which generally fell intermediately between *distinción* and *ceceo*. One explanation for the medial ratings assigned to *seceo* and *seseo* is that listeners may be familiar with the regional prestige attached to Sevillian speech norms within the context of Andalusia proper: "the Seville urban variety has become the basis of a regional spoken standard [...] which represents an innovative alternative to the spoken national standard" (Hernández-Campoy & Villena-Ponsoda, 2009:193). A related explanation is that if the participants' initial reactions to the *seceo* and *seseo* stimuli was to associate these variants with Seville, they may have been additionally challenged by the task of conflating distinct associations found within Andalusia into a single social-attribute assignment. That is, associations with both the stigmatized rural south and the prestige of the capital city may have been simultaneously activated, yielding intermediate values. Analogously, British listeners hold conflicting views regarding the "London" label due to the sociodemographic diversity associated with different areas of the United Kingdom's capital city (Bishop, Coupland & Garrett, 2005:139).

5.2.2 Word-medial coda-/s/

All groups attributed the most favorable social meanings to /s/-retention, followed by post-aspiration, and finally /s/-deletion. The listeners' more positive ratings of post-aspiration, in comparison to /s/-deletion, provide a social basis for the ongoing change toward post-aspirated stops among speakers of Andalusian Spanish (Ruch & Harrington, 2014; Ruch & Peters, 2016). Note that Campbell-Kibler (2010) and Eckert (2008) propose that when variants become socially prized, they are more likely to become the basis of sound change. At the same time, we should underscore that both post-aspiration and /s/-deletion received markedly lower ratings than /s/-retention. This reinforces the idea that at a national level, /s/-retention continues to hold overt prestige even though post-aspiration is undergoing regional transmission across subvarieties of Andalusian Spanish (Ruch & Harrington, 2014; Ruch & Peters, 2016).

5.2.3 Word-final coda-/s/

The participant groups demonstrated considerable agreement in attributing the most favorable social meanings to /s/-retention and the least favorable meanings to the lax-vowel stimuli, with the ratings for the tense-vowel stimuli being intermediate. To our knowledge, ours is the first study to examine listeners' social evaluations of word-final vowel laxing in comparison to /s/-retention and /s/-deletion alone, as well as in comparison to phonetic variants from other phonological environments. Notably, the laxing stimuli received the most unfavorable evaluations of the ten phonetic variants targeted in this study. One possible explanation for this result is that the presence of the lax vowel in combination with deletion of syllable-final /s/ contributes to an exceptionally unfavorable socially attributed meaning.

To summarize, the results from research question 2 indicate that listeners of Iberian Spanish incorporate collective ideologies into their social assignments of regional variation. The interconnected findings suggest that the social attributes linked to the ten targeted variants derive from comparable indexical fields of meaning (Eckert, 2008, 2019). When group differences emerged, the Jerez respondents often presented a more polarized data pattern than the Granada and Salamanca respondents. The Jerez listeners' distinctly unfavorable ratings toward the Andalusian-sounding stimuli resemble what Labov (1972) calls "linguistic insecurity," which is the negative self-image held by users of their own variety vis-à-vis more overtly prestigious varieties (Moreno-Fernández, 2020:60).⁶ Labov further proposes that linguistic insecurity and regional variation are closely linked, meaning that speakers may feel less confident if they believe that a local form of speaking is less prestigious than more "correct" ways of speaking. Within the context of Andalusian Spanish, the present results suggest that such linguistically underprivileged language users remain less confident in the social favorability of their own speech forms (Preston, 1998). Yet another possibility is that the Jerez listeners responded based on how they believe the variants would be perceived on a national level, even if they hold some level of personal pride associated with them.

5.3. Social motivations of language change in Andalusian Spanish

The social-attribution results further elucidate the nature of language change in Andalusian Spanish. Eckert (2008) proposes that a linguistic form's indexical assignment undergoes cycles of reinterpretation as the form is used and heard in new social

contexts (Silverstein, 2003). Contemporarily, one way in which language users reevaluate a form's indexical assignment stems from ongoing changes affecting urban versus rural speakers (Kerswill, 2003; Trudgill, 1986). Due to societal shifts that give rise to increased social mobility and regional dialect leveling, speakers from urban areas are prone to adopt overtly prestigious forms of speech at faster rates than speakers from rural areas (Britain, 2009, 2010). A subsequent reinterpretation of overtly prestigious variants as more "urban"-sounding yields an accompanying reindexing of traditional dialect features as more local- or "rural"-sounding (Regan, 2022). One of the stronger findings of our study comes from the participants' responses to the urbanicity-rurality scale, with three variants standing out as sounding the most rural and consistently yielding mean values below 50: *ceceo* in the syllable-initial fricatives, /s/-deletion in the word-medial context, and /s/-deletion + lax vowel in the word-final context.

With respect to syllable-initial coronal fricatives, one interpretation of the urbanicity findings is that the change from *ceceo* to *distinción* (Regan, 2020b) is currently more advanced in urban areas of Andalusia than in surrounding rural areas, which leads to a re-evaluation of *ceceo* as an even stronger marker of rural identity. Regarding *seceo* and *seseo*, these variants were perceived as pertaining to the speech from a small city (i.e., means near 50). Such medial ratings likely reflect the respondents' awareness that *seceo* and *seseo* are pervasive among speakers from Seville, Andalusia's capital city, and the region's chief urban area (Narbona Aguilar & Morillo, 1998). Regarding word-medial coda-/s/, the respondents converged in assigning values near 50 to postaspiration and comparatively lower values to /s/-deletion. This data pattern suggests that the straightforward deletion of the alveolar fricative in the phonological /s/ + /p t k/ context may be undergoing an indexical re-evaluation as more "rural"-sounding due to urban Andalusian speakers' heightened use of post-aspiration (Ruch & Harrington, 2014; Ruch & Peters, 2016). This subtle social differential between /s/-deletion and post-aspiration aligns with Eckert's (2008:471) proposal that "the longer a variable is around the more nuanced meanings it can take on" (recall that post-aspiration is a more recent development than /s/-deletion). Concerning listeners' assignments of lax vowels indexing rurality, the findings do not entirely support previous dialectological research claiming that Andalusia's eastern and central provinces are notoriously laxing regions (Alvar, 1996; Hernández-Campoy & Trudgill, 2002). While our perceptual data suggest that vowel laxing is indeed a pervasive process in these areas, they also give the impression that urban central and eastern speakers may be diverging away from a previously conventionalized use of lax vowels, due to an updated indexing of laxing as especially rural-sounding. Future research will need to test this hypothesis.

5.4. Implications for language-based discrimination

This study has implications for the ways in which listeners automatically and preemptively make assumptions about a speaker through combining knowledge of regional variation and implicit bias. Craft and colleagues emphasize that a harmful byproduct of language ideologies is the impression they create that language users have the liberty to "elevate or devalue the social standing and language performance of speakers according to their language choices" (2020:393). Rubin (2012:12) further explains that when listeners make "informal, on-the-spot" judgments about a speaker's voice, they utilize assumptions from stereotypical associations between linguistic variation and social identity. In the context of

Andalusian Spanish, national media and television can sometimes portray speakers from southern Spain as uncultivated, unintelligible, or producing prescriptively incorrect speech (Gordillo, 2012; León-Castro Gómez, 2016). The participants from the present study likely fell back on commonplace Andalusian stereotypes when attributing social meaning to the southern-sounding stimuli.

Although our findings do not explicitly demonstrate inequitable outcomes for Andalusians (Purnell et al., 1999), the negative judgments of the Andalusian variants call attention to a sociolinguistic climate ripe for language-based discrimination through linguistic profiling. In future work, it will be important for linguists to develop strategies that combat potential discrimination brought about by an adverse indexing of Andalusian features. In this regard, we endorse Craft and colleagues' (2020:402) call for action in response to linguistic justice across numerous institutional arenas.

5.5. Limitations

Limitations of this study include: (1) the use of a single voice in the survey; (2) the use of short noun phrases in the stimuli; and (3) the homogeneous sociodemographic background of the respondents.

Regarding the first limitation, the audio stimuli were recorded by a single male speaker. While the results are relevant for understanding how listeners incorporate global ideologies of regional differences into their social evaluations of language, we have not established whether any additional social indexing of a speaker's voice (e.g., age, speaker sex) simultaneously influences listeners' judgments (Cole, 2021; Walker et al., 2014). Second, most audio files in the survey included single noun phrases due to our attempt to mitigate the confounds of longer stimuli containing multiple Andalusian-sounding variants. However, studies argue that spontaneous-speech stimuli might activate more authentic sociolinguistic perceptions than read-speech stimuli (Campbell-Kibler, 2010:34). Finally, we limited our participant recruitment to university-level respondents with life-long experiences in Spain's national education system. Future studies aiming to present a more expansive vision of language ideologies will benefit from including a more diverse pool of sociodemographic backgrounds (Eckert, 2008:467; Henrich, Heine & Norenzayan, 2010).

6. Conclusion

This study's aim was to investigate the perceptual assessments of ten phonetic variants used in Iberian Spanish by eliciting listeners' geographic classifications and social evaluations of the variants. The audio stimuli, featuring three linguistic contexts, were evaluated by 165 listeners from Granada, Jerez de la Frontera, and Salamanca, Spain. The main findings include: (1) the listener groups were fairly consistent in their geographic classifications based on a north/central/south categorization; (2) when classifying a stimulus as originating from Andalusia, the participants more often identified localities from western Andalusia than from central or eastern Andalusia; (3) the participants attributed more favorable social meanings to variants common to northern/central areas of Spain than to variants common to southern Spain; and (4) the western Andalusian listeners from Jerez evaluated the stimuli in the most polarized manner. The latter results echo the findings of previous studies demonstrating the ways in which speakers of underprivileged language varieties can hold feelings of linguistic insecurity toward their own ways of speaking (Labov, 1972; Preston, 1989).

A major takeaway from this study is that listeners of Iberian Spanish reference shared ideologies toward regional variation when assigning social meaning to linguistic form. In this respect, this article brings attention to the largely unfavorable attitudes that language users hold toward Andalusian Spanish, which have the potential to disadvantage Andalusians themselves through instances of language-based discrimination (Baugh, 2003, 2017; Purnell et al., 1999). On a more conceptual level, the findings emphasize the dynamic relationship between linguistic variation and social evaluation (Campbell-Kibler, 2010; Eckert, 2008, 2019).

Acknowledgements. The authors would like to express their gratitude to Kelly Wright for providing the initial inspiration for this study's design. Our appreciation also extends to Kendall Cohen, Izabel Powers, and Alayna Simonds for their invaluable support in the preliminary stages of this research.

Competing interests. The authors declare none.

Notes

1 We emphasize the lack of inherent favorability or value of such social attributes (e.g., 'rurality'). In this paper, we refer to the attributes in terms of the "favorability" they have accrued within the particular socioeconomic, political, and historical contexts in which our study is couched.

2 Note that in the *ceceo* system, there is a question regarding the phonetic realization of the interdental voiceless fricative, namely, whether it is best transcribed as [θ], [s^h], or [θ^h] (see Regan, 2020b:162).

3 Note that this is different from *ceceo*, in which speakers vary between *distinción* and *ceceo* (Regan, 2022).

4 An additional category, coded as "other," includes all place names outside of continental Spain (e.g., Canary Islands, Latin America). The "other" category was dropped from the analysis due to low token counts (2.89% of the data).

5 To account for the unbalance in the subsample "Andalusia-specific," we applied a weighting procedure. The unweighted numbers of observation were as follows: 1613 from Jerez, 1771 from Granada, and 1250 from Salamanca.

6 We understand "overtly prestigious" features as supralocal linguistic forms that are prioritized in institutional or public domains, such as in educational systems or in national media (see Regan, 2022:501).

References

- Abercrombie, David. 1967. *Elements of general phonetics*. Edinburgh: Edinburgh University Press.
- Alvar, Manuel. 1996. *Manual de dialectología hispánica: El Español de España*. Barcelona: Editorial Ariel.
- Barr, Dale J., Roger Levy, Christoph Scheepers & Harry J. Tily. 2013. Random effects structure for confirmatory hypothesis testing: Keep it maximal. *Journal of Memory and Language* 68(3). 255–278. <https://doi.org/10.1016/j.jml.2012.11.001>
- Bates, Douglas, Martin Maechler, Ben Bolker & Steve Walker. 2015. Fitting linear mixed-effects models using lme4. *Journal of Statistical Software* 67(1). 1–48. <https://doi.org/10.18637/jss.v067.i01>
- Baugh, John. 2003. Linguistic profiling. In Arnetta F. Ball, Sinfree Makoni, Geneva Smitherman & Arthur K. Spears (eds.), *Black linguistics: Language, society and politics in Africa and the Americas*, 155–68. London: Routledge.
- Baugh, John. 2017. Linguistic profiling and discrimination. In Nelson Flores, Ofelia García & Massimiliano Spotti (eds.), *Oxford handbook of language and society*. New York: Oxford University Press.
- Benjamin, Daniel J., James O. Berger, Magnus Johannesson, Brian A. Nosek, E. J. Wagenmakers, et al. 2018. Redefine statistical significance. *Nature Human Behaviour* 2(1). 6–10. <https://doi.org/10.1038/s41562-017-0189-z>
- Bishop, Hywel, Nikolas Coupland & Peter Garrett. 2005. Conceptual accent evaluation: Thirty years of accent prejudice in the UK. *Acta Linguistica Hafniensia: International Journal of Linguistics* 37(1). 131–54. <https://doi.org/10.1080/03740463.2005.10416087>
- Britain, David. 2009. "Big bright lights" versus "green and pleasant land"? The unhelpful dichotomy of 'urban' versus 'rural' in dialectology. In Enam

- Al-Wer (ed.), *Arabic dialectology*, 223–247. Leiden: Brill. <https://doi.org/10.1163/ej.9789004172128.i-298.86>
- Britain, David. 2010. Supralocal regional dialect leveling. In Carmen Llamas & Dominic Watt (eds.), *Language and identity*, 193–204. Edinburgh: Edinburgh University Press.
- Caballero-Gálvez, Antonio A. 2017. The downfall of Spanish stereotypes: Andalusian, Basque and Catalan identities in “plurinational” Spain. *Catalan Journal of Communication & Cultural Studies* 9(1). 67–85. https://doi.org/10.1386/cjcs.9.1.67_1
- Campbell-Kibler, Kathryn. 2010. The effect of speaker information on attitudes toward (ING). *Journal of Language and Social Psychology* 29(2). 214–223. <https://doi.org/10.1177%2F0261927X09359527>
- Carbonero Cano, Pedro. 1982. *Norma estándar y actitud sociolingüística*. Sevilla: Universidad de Sevilla - Secretariado de Publicaciones.
- Carbonero Cano, Pedro. 2003. *Estudios de sociolingüística andaluza*. Sevilla: Universidad de Sevilla - Secretariado de Publicaciones.
- Cargile, Aaron C., Howard Giles, Ellen B. Ryan & James J. Bradac. 1994. Language attitudes as a social process: A conceptual model and new directions. *Language and Communication* 14. 211–236.
- Cobo Romero, Francisco. 2012. *La represión política en Andalucía*. Sevilla: Fundación Pública Andaluza Centro de Estudios Andaluces.
- Cole, Amanda. 2021. Disambiguating language attitudes held towards sociodemographic groups and geographic areas in South East England. *Journal of Linguistic Geography* 9. 13–27. <https://doi.org/10.1017/jlg.2021.2>
- Craft, Justin T., Kelly E. Wright, Rachel Elizabeth Weissler & Robin M. Queen. 2020. Language and discrimination: Generating meaning, perceiving identities, and discriminating outcomes. *Annual Review of Linguistics* 6(1). 389–407. <https://doi.org/10.1146/annurev-linguistics-011718-011659>
- Dalbor, John B. 1980. Observations on present-day Seseo and Ceceo in southern Spain. *Hispania* 63. 5–19. <https://doi.org/10.2307/340806>
- Du Bois, Inke. 2019. Linguistic discrimination across neighbourhoods. *Journal of Language and Discrimination* 3(2). 92–119. <https://doi.org/10.1558/jld.39973>
- Eckert, Penelope. 2008. Variation and the indexical field. *Journal of Sociolinguistics* 12. 453–476. <https://doi.org/10.1111/j.1467-9841.2008.00374.x>
- Eckert, Penelope. 2019. The limits of meaning: Social indexicality, variation and the cline of interiority. *Language* 95. 751–76. <https://doi.org/10.1353/lan.0.0239>
- Gal, Susan. 2016. Sociolinguistic differentiation. In Nikolas Coupland (ed.), *Sociolinguistics: Theoretical debates*. 113–138. New York: Cambridge University Press.
- Gallego, Juan Carlos & Anthony Rodríguez. 2013. Castilian speakers’ attitudes towards accents and regional stereotypes in Spain. *Sociolinguistic Studies* 6(3). 543–569. <https://doi.org/10.1558/sols.v6i3.543>
- Giles, Howard & Andrew C. Billings. 2004. Assessing language attitudes: Speaker evaluation studies. In Alan Davies and Catherine Elder (eds.), *The handbook of applied linguistics*, 187–209. Cornwall: Blackwell Publishing. <https://doi.org/10.1002/9780470757000.ch7>
- Golinger, Stephen D., David B. Pisoni & John S. Logan. 1991. On the nature of talker variability effects on recall of spoken word lists. *Journal of Experimental Psychology: Learning, Memory, and Cognition* 17(1). 152–162. <https://doi.org/10.1037/0278-7393.17.1.152>
- Gordillo, Inmaculada. 2012. La autorrepresentación del andaluz en web-series. *Palabra Clave* 15(1). 54–81.
- Gradoville, Michael, Earl Kjar Brown & Richard File-Muriel. 2022. The phonetics of sociophonetics: Validating acoustic approaches to Spanish /s/. *Journal of Phonetics* 91. 13–32. <https://doi.org/10.1016/j.wocn.2021.101125>
- Grondelaers, Stefan, Roeland van Hout & Mieke Steegs. 2010. Evaluating regional accent variation in Standard Dutch. *Journal of Language and Social Psychology* 29(1). 101–116. <https://doi.org/10.1177%2F0261927X09351681>
- Harjus, Jannis. 2018. *Sociofonética andaluza y lingüística perceptiva de la variación: El español hablado en Jerez de la Frontera* (Lengua y sociedad en el mundo hispánico). Madrid: Iberoamericana Editorial.
- Henrich, Joseph, Steven Heine & Ara Norenzayan. 2010. The weirdest people in the world? *Behavioral and Brain Sciences* 33(2–3). 61–83. <https://doi.org/10.1017/S0140525X0999152X>
- Henriksen, Nicholas. 2017. Patterns of vowel laxing and harmony in Iberian Spanish: Data from production and perception. *Journal of Phonetics* 63. 106–126. <http://doi.org/10.1016/j.wocn.2017.05.001>
- Henriksen, Nicholas, Amber Galvano & Micha Fischer. 2023. Sound change in Western Andalusian Spanish: Investigation into the actuation and propagation of post-aspiration. *Journal of Phonetics* 98. 1–21. <https://doi.org/10.1016/j.wocn.2023.101238>
- Henriksen, Nicholas & Sarah Harper. 2016. Investigating lenition patterns in south-central Peninsular Spanish /sp st sk/ clusters. *Journal of the International Phonetic Association* 46(3). 287–310. <https://doi.org/10.1017/S0025100316000116>
- Hernández-Campoy, Juan Manuel & Peter Trudgill. 2002. Functional compensation and southern peninsular Spanish /s/ loss. *Folia Linguistica Historica* 36. 31–58. <https://doi.org/10.1515/flih.2002.23.1-2.31>
- Hernández-Campoy, Juan Manuel & Juan Andrés Villena-Ponsoda. 2009. Standardness and nonstandardness in Spain: Dialect attrition and revitalization of regional dialects of Spanish. *International Journal of the Sociology of Language* 196/197. 181–214. <https://doi.org/10.1515/IJSL.2009.021>
- Herrero de Haro, Alfredo. 2016. Four mid front vowels in Western Almería: The effect of /s/, /r/, and /θ/ deletion in Eastern Andalusian Spanish. *Zeitschrift für Romanische Philologie* 132(1). 118–148.
- Hualde, José Ignacio. 2021. Spanish. In Christoph Gabriel, Randall Gess & Trudel Meisenburg (eds.), *Manual of Romance phonetics and phonology*. 779–807. Berlin: De Gruyter. <https://doi.org/10.1515/9783110550283>
- Jaeger, Byron C., Lloyd J. Edwards, Kalyan Das & Pranab K. Sen. 2017. An R 2 statistic for fixed effects in the generalized linear mixed model. *Journal of Applied Statistics* 44(6). 1086–1105. <https://doi.org/10.1080/02664763.2016.1193725>
- Jaeger, Byron. 2017. *r2glmm: Computes R squared for mixed (multilevel) models*. R package version 0.1.2. <https://github.com/bcjaeger/r2glmm>
- Jaime-Jiménez, Elena. 2019. Social and pragmatic effects on the variable use of plural address forms in Andalusian Spanish. *Revista Internacional de Lingüística Iberoamericana* 34. 33–54. <https://doi.org/10.1080/15283488.2012.747439>
- Jaspal, Rusi & Ionna Sitaridou. 2013. Coping with stigmatized linguistic identities: Identity and ethnolinguistic vitality among Andalusians. *Identity* 13(2). 95–119. <https://doi.org/10.1080/15283488.2012.747439>
- Jiménez, Jesús & Maria-Rosa Lloret. 2019. Vowel harmony. In Sonia Colina & Fernando Martínez-Gil (eds.), *The Routledge handbook of Spanish phonology*, 100–128. Oxford: Routledge. <https://doi.org/10.4324/9781315228112>
- Kerswill, Paul. 2003. Dialect levelling and geographical diffusion in British English. In David Britain & Jenny Cheshire (eds.), *Social dialectology: In honour of Peter Trudgill*, 223–243. Amsterdam: John Benjamins. <https://doi.org/10.1075/impact.16.16ker>
- Labov, William. 1972. *Sociolinguistic patterns*. Philadelphia: University of Pennsylvania Press.
- León-Castro Gómez, Marta. 2016. La presencia del andaluz en los medios de comunicación. In Rosalba Mancinas-Chávez (ed.), *Actas del I congreso internacional comunicación y pensamiento*, 1583–1600. Sevilla: Egregius.
- Lippi-Green, Rosina. 1997. *English with an accent: Language, ideology, and discrimination in the United States*. London: Routledge.
- Lüdtke, Daniel. 2018. ggeffects: Tidy data frames of marginal effects from regression models. *Journal of Open Source Software* 3(26). 772. <https://doi.org/10.21105/joss.00772>
- Martínez Melgar, Antonia. 1994. El vocalismo del andaluz oriental. *Estudios de fonética experimental* 6. 11–64.
- Melguizo Moreno, Elisabeth. 2007. La variación de /θs/: estudio comparativo de dos muestras de población granadinas. *ELUA: Estudios de Lingüística Universidad de Alicante* 21. 245–259. <https://doi.org/10.14198/ELUA2007.21.12>
- Molina-García, Alvaro. 2020. Percepción comunitaria de la distinción fonemática de la /s/ y la /θ/ en hablantes andaluces: Estudio en la ciudad de Málaga. *Spanish in Context* 17(3). 415–437. <https://doi.org/10.1075/sic.19007.mol>
- Momcilovic, Natasa B. 2009. *A sociolinguistic analysis of s-aspiration in Madrid Spanish*. Munich: Lincom Europa.

- Moreno-Fernández, Francisco. 2020. *Varietades de la lengua española*. Oxford: Routledge.
- Mullenix, John W., David B. Pisoni & Christopher S. Martin. 1989. Some effects of talker variability on spoken word recognition. *Journal of the Acoustical Society of America* 85. 365–378. <https://doi.org/10.1121/2F1.397688>
- Narbona, Antonio, Rafael Aguilar & Ramón Morillo. 1998. *El español hablado en Andalucía*. Barcelona: Editorial Ariel.
- Preston, Dennis R. 1989. *Perceptual dialectology: Nonlinguists' views of areal linguistics*. New York: De Gruyter Mouton. <https://doi.org/10.1515/9783110871913>
- Preston, Dennis R. 1998. They speak really bad English down south and in New York City. In Laurie Bauer & Peter Trudgill (eds.), *Language myths*, 139–149. London: Penguin.
- Purnell, Thomas, William Idsardi & John Baugh. 1999. Perceptual and phonetic experiments on American English dialect identification. *Journal of Language and Social Psychology* 18(1). 10–30. <https://doi.org/10.1177/0261927X99018001002>
- R Core Team. 2020. *R: A language and environment for statistical computing*. Vienna, Austria: R Foundation for Statistical Computing. Retrieved from <https://www.r-project.org/>
- Regan, Brendan. 2019. Dialectology meets sociophonetics: The social evaluation of ceceo and distinción. In Whitney Chappell (ed.), *Recent advances in the study of Spanish sociophonetic perception*, 85–121. Amsterdam: John Benjamins.
- Regan, Brendan. 2020a. Intra-regional differences in the social perception of allophonic variation: The evaluation of [tʃ] and [ʃ] in Huelva and Lepe (Western Andalucía). *Journal of Linguistic Geography* 8(2). 82–101. <https://doi.org/10.1017/jlg.2020.7>
- Regan, Brendan. 2020b. The split of a fricative merger due to dialectal contact and societal changes: A sociophonetic study on Andalusian Spanish read-speech. *Language Variation and Change* 32(2). 159–190. <https://doi.org/10.1017/S0954394520000113>
- Regan, Brendan. 2022. The social meaning of a merger: The evaluation of an Andalusian Spanish consonant merger (ceceo). *Language in Society* 51(3). 481–510. <https://doi.org/10.1017/S0047404521000543>
- Rickford, John & Sharese King. 2016. Language and linguistics on trial: Hearing Rachel Jeantel (and other vernacular speakers) in the courtroom and beyond. *Language* 92. 948–988. <https://doi.org/10.1353/lan.2016.0078>
- Rubin, Donald. 2012. The power of prejudice in accent perception: Reverse linguistic stereotyping and its impact on listener judgments and decisions. In John Levis & Kimberly LeVelle (eds.), *Proceedings of the 3rd pronunciation in second language learning and teaching conference*, 11–17. Ames: Iowa State University.
- Ruch, Hanna. 2018. Perception of speaker age and speaker origin in a sound change in progress: The case of /s/-aspiration in Andalusian Spanish. *Journal of Linguistic Geography* 6(1). 40–55. <https://doi.org/10.1017/jlg.2018.4>
- Ruch, Hanna & Jonathan Harrington. 2014. Synchronic and diachronic factors in the change from pre-aspiration to post-aspiration in Andalusian Spanish. *Journal of Phonetics* 45. 12–25. <https://doi.org/10.1016/j.wocn.2014.02.009>
- Ruch, Hanna & Sandra Peters. 2016. On the origin of post-aspirated stops: Production and perception of /s/ + voiceless stop sequences in Andalusian Spanish. *Laboratory Phonology* 7(1). 1–36. <https://doi.org/10.5334/labphon.2>
- Santana Marrero, Juana. 2018. Beliefs and attitudes of the Sevillian university students towards the Spanish normative varieties: Andalusian, Canarian and Castilian. *Pragmática sociocultural: Revista internacional sobre Lingüística del Español* 6(1). 72–97. <https://doi.org/10.1515/soprag-2018-0003>
- Sawoff, Adolf. 1980. A sociolinguistic appraisal of the sibilant pronunciation in the city of Seville. *Grazer Linguistische Studien* 11(12). 238–262.
- Silverstein, Michael. 2003. Indexical order and the dialectics of sociolinguistic life. *Language and Communication* 23. 193–229.
- Torreira, Francisco. 2012. Investigating the nature of aspirated stops in Western Andalusian Spanish. *Journal of the International Phonetic Association* 42(1). 49–63. <https://doi.org/10.1017/S0025100311000491>
- Trudgill, Peter. 1986. *Dialects in contact*. Oxford: Blackwell.
- Villena-Ponsoda, Juan Andrés. 2008. Sociolinguistic patterns of Andalusian Spanish. *International Journal of the Sociology of Language* 2008(193/194). 139–160. <https://doi.org/10.1515/ijsl.2008.052>
- Walker, Abby, Christina García, Yomi Cortés & Kathryn Campbell-Kibler. 2014. Comparing social meanings across listener and speaker groups: the indexical field of Spanish /s/. *Language Variation and Change* 26. 169–189. <https://doi.org/10.1017/S0954394514000088>
- West, Brady T., Kathleen B. Welch & Andrzej T. Galecki. 2014. *Linear mixed models: A practical guide using statistical software*. Boca Raton, FL: Chapman Hall/CRC Press.
- Wright, Kelly. 2019. *Experiments on linguistic profiling of three American dialects*. Qualifying research paper. Ann Arbor, MI: University of Michigan.

Appendix A. List of baseline token used in the perception study

Linguistic context	Baseline tokens
Syllable-initial coronal fricatives (target consonants are underlined)	<i>cereza mohosa</i> 'moldy cherry'; <i>diferencia necesaria</i> 'necessary difference'; <i>la azotea de su vecino</i> 'your neighbor's porch'; <i>moza preciosa</i> 'precious young woman'; <i>policia ambicioso</i> 'ambitious police officer'; <i>precipicio peligroso</i> 'dangerous precipice'; <i>sacerdote imbécil</i> 'idiotic priest'; <i>una oficina sencilla</i> 'a simple office'; <i>zapatazo simple</i> 'simple shoe-stomp'
Word-medial coda-/s/ (target phonemic sequences are underlined)	/sp/ targets: <i>espina</i> 'spine'; <i>espada</i> 'sword'; <i>despierto</i> 'awake'; /st/ targets: <i>pestaña</i> 'eyelash'; <i>Estela</i> 'Estela'; <i>estaba</i> 'I/he/she was' /sk/ targets: <i>ascota</i> 'pet'; <i>escuela</i> 'school'; <i>buscando</i> 'looking'
Word-final coda-/s/ (target vowels are underlined)	<i>breves</i> 'brief-plural'; <i>deberes</i> 'tasks'; <i>nenes</i> 'kids'; <i>niveles</i> 'levels'; <i>papeles</i> 'papers'; <i>paredes</i> 'walls'; <i>poderes</i> 'powers'; <i>redes</i> 'networks'; <i>teles</i> 'televisions'
Distractors	<i>Cuba y Puerto Rico</i> 'Cuba and Puerto Rico'; <i>el equipo y el grupo</i> 'the team and the group'; <i>el libro y la gorra</i> 'the book and the cap'; <i>el muñeco y la obra</i> 'the doll and the play'; <i>la amiga de Mario</i> 'Mario's friend'; <i>la cama y la lámpara</i> 'the bed and the lamp'; <i>la comida y la bebida</i> 'the food and the drink'; <i>la pelea y la bandera</i> 'the fight and the flag'; <i>Perú y Uruguay</i> 'Peru and Uruguay'; <i>un lago y un río</i> 'a lake and a river'; <i>un niño y una niña</i> 'a boy and a girl'; <i>un perro y un gato</i> 'a dog and a cat'; <i>un rico y un pobre</i> 'a rich person and a poor person'; <i>una planta y una flor</i> 'a plant and a flower'; <i>una rata y un búho</i> 'a rat and an owl'

Appendix B. Phonetic measures taken for the word-medial coda-/s/ and word-final coda-/s/ stimuli

Table B1. Mean segmental durations for word-medial coda-/s/ stimuli (all values in ms)

	/s/-retention stimuli	post-aspiration stimuli	/s/-deletion stimuli
/s/ duration	109	n/a	n/a
Closure duration	91	134	108
VOT duration	21	66	21

Table B2. Mean F1 and F2 values for word-final coda-/s/ stimuli (all values in Hz)

	/s/-retention stimuli	/s/-deletion + tense vowel stimuli	/s/-deletion + lax vowel stimuli
F1	470	487	630
F2	1738	1716	1589

Appendix C. Bar charts of raw counts from responses to the three-way 'north of Spain/center of Spain/south of Spain' classification

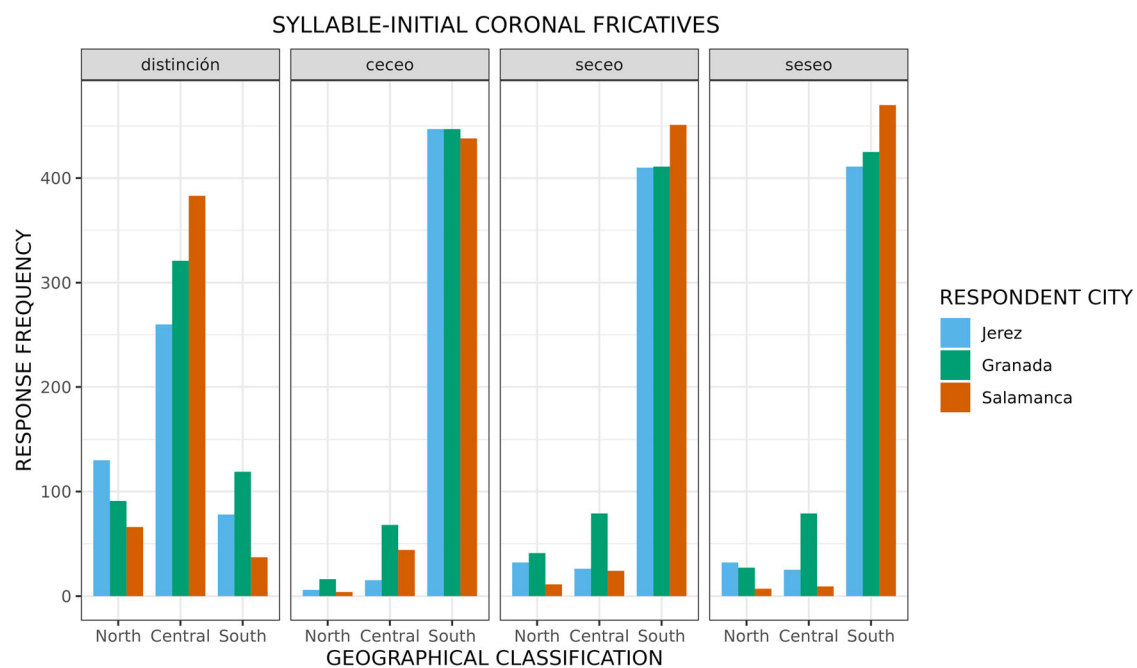


Figure C1. Frequency of geographical classification responses for each variant of syllable-initial coronal fricatives within each RESPONDENT CITY.

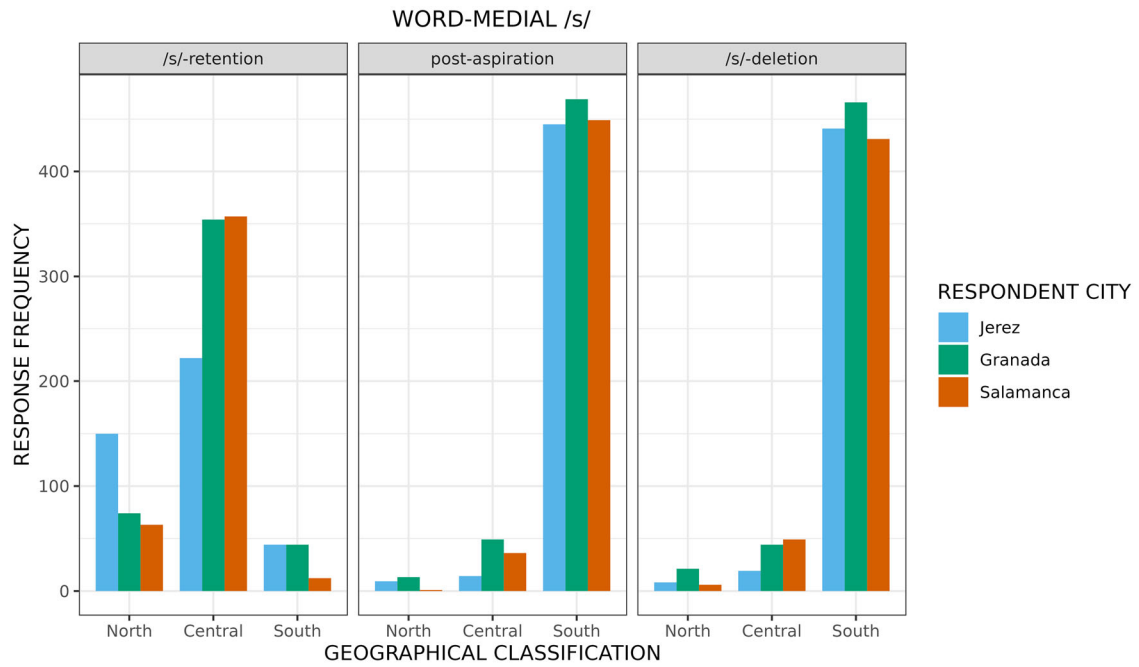


Figure C2. Frequency of geographical classification responses for each variant of word-medial coda-/s/ within each RESPONDENT CITY.

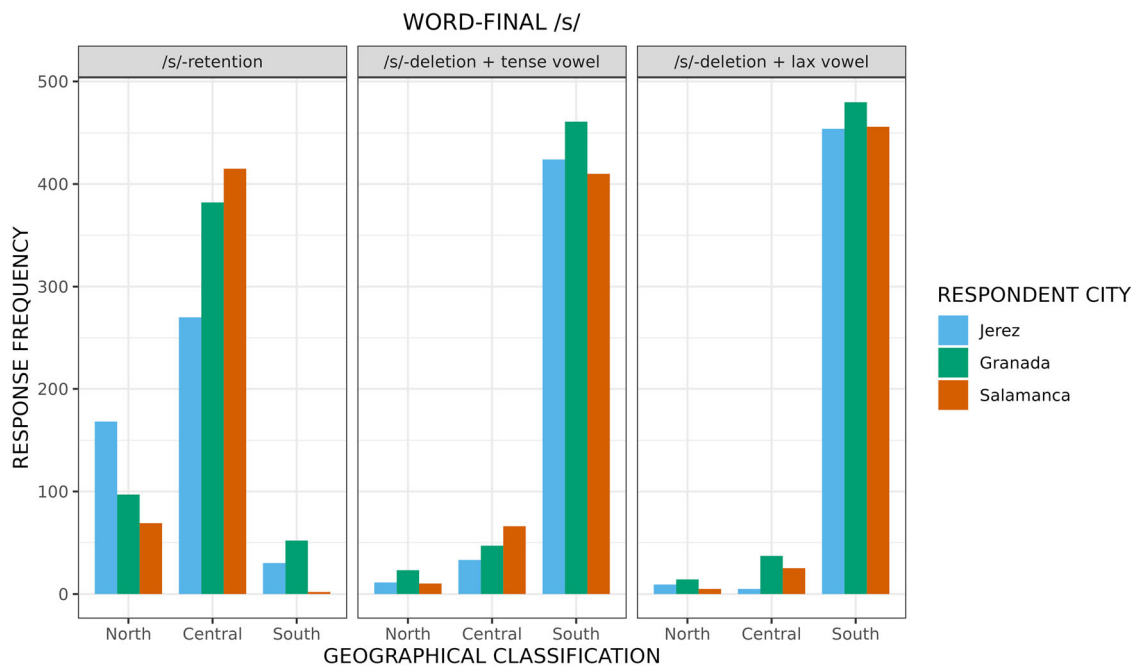


Figure C3. Frequency of geographical classification responses for each variant of word-final coda-/s/ within each RESPONDENT CITY.