Special Issue Reprint

Social Meanings of Language Variation in Spanish

Edited by
Whitney Chappell and Sonia Barnes

MDPI

Social Meanings of Language Variation in Spanish

Social Meanings of Language Variation in Spanish

Editors

Whitney Chappell
Sonia Barnes

Editors

Whitney Chappell
Modern Languages and
Literatures
The University of Texas
at San Antonio
San Antonio
United States

Sonia Barnes
Languages, Literatures
and Cultures
Marquette University
Milwaukee
United States

This is a reprint of articles from the Special Issue published online in the open access journal *Languages* (ISSN 2226-471X) (available at: https://www.mdpi.com/journal/languages/special_issues/42DJAH7QI8).

For citation purposes, cite each article independently as indicated on the article page online and as indicated below:

Lastname, Firstname, Firstname Lastname, and Firstname Lastname. Article Title. *Journal Name* **Year**, *Volume Number*, Page Range.

Cover image courtesy of Sonia Barnes

Contents

About the Editors

Whitney Chappell

Whitney Chappell is an Associate Professor of Hispanic linguistics at the University of Texas at San Antonio (UTSA), the UTSA College of Liberal and Fine Arts' 2021 Researcher of the Year, a 2020-2021 Lutcher Brown Distinguished Professor, and a 2019 Fulbright Scholar. Her research focuses on sociophonetic variation in the Spanish-speaking world, and her work has appeared in prestigious venues like *Language Variation and Change* and *Studies in Second Language Acquisition*. Her edited volumes include *Recent Advances in the Study of Spanish Sociophonetic Perception* and *Spanish Socio-Historical Linguistics: Isolation and Contact* (co-edited with Bridget Drinka).

Sonia Barnes

Sonia Barnes is an Associate Professor of Spanish at Marquette University. Her research explores the linguistic and social factors that govern language variation in language contact situations, the social meaning indexed by different language variants, and the role of linguistic variation in constructing place-based identity. Her publications include articles in venues like the *Journal of Sociolinguistics* and *Phonetica* and chapters in edited volumes like *Recent Advances in the Study of Spanish Sociophonetic Perception* (edited by Whitney Chappell) and *Spanish Phonetics and Phonology in Contact: Studies from Africa, the Americas and Spain* (edited by Rajiv G. Rao).

MDPI

Introduction to the Special Issue *Social Meanings of Language Variation in Spanish*

Sonia Barnes [1,*] and Whitney Chappell [2,*]

1   Department of Languages, Literatures and Cultures, Marquette University, Milwaukee, WI 53233, USA
2   Department of Modern Languages and Literatures, The University of Texas at San Antonio,
     San Antonio, TX 78249, USA
*   Correspondence: sonia.barnes@marquette.edu (Sonia Barnes); whitney.chappell@utsa.edu (Whitney Chappell)

Centering stances, positionalities, and style, the third wave of sociolinguistic study positions individuals at the heart of its analysis. Understood as a continuation of the first and the second waves (see Eckert 2005), which sought to elucidate correlations between linguistic features and broad socio-demographic groups (first wave) or locally relevant categories and networks (second wave), the third wave focuses on the social meaning of variable linguistic features and recognizes that speakers agentively employ the linguistic resources at their disposal to signal group memberships, construct their personae, and position themselves in interaction. To understand how linguistic features (or clusters of features) index social meanings, scholars in the third wave have adopted a variety of methodologies, including ethnographic methods as well as experimental work. For instance, in her study on language use among rival Latina groups in a California school, Mendoza-Denton (2008) conducted ethnographic research to examine how the members of each group employed various practices (including, but not limited to, linguistic behaviors) to express social alliances and identities. Similarly, Snell (2010) used an ethnographic approach to explore how possessive "me" is used in stylistic performances (and, as a result, in the construction of local identities) by students in two different schools in North East England.

Experimental research within the third wave has allowed sociolinguists to tap into social meanings and perceptions that might not be apparent by only examining speech production. Campbell-Kibler (2009), for example, used a combination of experimental and interview data to investigate the social evaluations of the English (ING) variable. The study highlighted, on the one hand, the diversity of social meanings indexed by this variable and, on the other, the flexibility of such associations, given that, for instance, perceptions of education and intelligence shifted based on other factors (e.g., perceptions of the speaker's regional origin). Subsequent perception research has clearly demonstrated that social meanings are highly dependent on factors related to the speaker, such as their perceived age, race, or nationality (Niedzielski 1999; Drager 2005; Staum Casasanto 2010), and the listener, like their gender or their own linguistic variety (Hay et al. 2006; Drager 2011). The dynamic and complex links between social meanings and linguistic practices that studies in the third wave have uncovered are created, reaffirmed, and extended in an iterative process, allowing speakers to position themselves in social space by adopting particular styles that are associated with specific linguistic features.

Third-wave approaches thus provide critical insights into the driving forces behind sociolinguistic variation at the micro-level and elucidate the connections that exist between what happens locally and broader social categories. However, research on the third wave has been, to the present, largely focused on the English language. Because the resulting sociocultural and linguistic biases are baked into our current understanding of the reasons that lead speakers to select particular linguistic and stylistic features, a broader exploration of linguistic variation and social meaning in other languages is sorely needed. To tackle the

questions raised by the third wave in other contexts (e.g., how are social meanings linked to particular linguistic forms? How is social meaning structured? How does linguistic use affect social meaning? How do social meanings relate to one another? What is the role of social meaning in language variation and change?), we have mobilized scholars at the vanguard of Spanish sociolinguistics to share their work in this Special Issue. In doing so, we seek to both further extend the third wave to the Spanish-speaking world and to deepen, develop, and enrich the scope of third-wave research.

The contexts under exploration in the studies featured in this Special Issue include monolingual (i.e., Chile, Spain) and multilingual spaces (i.e., Spanish in the U.S., Mayan Spanish, Argentine Guarani), where the interaction between multiple languages provides speakers with a greater linguistic repertoire from which to draw in socioindexical processes. The authors also adopt a variety of methodological approaches, including perception tasks, survey experiments, discourse completion tasks, as well as more qualitative methods such as discourse analysis and grounded-theory text analysis. Finally, the inherent interdisciplinarity and variability of the third wave is on full display in this Special Issue. In the articles featured here, sociolinguistic analysis interfaces with pragmatics (Czerwionka et al.), psycholinguistics (Licata), onomastics (Enríquez Duque), and computer-mediated discourse (Callesano, Pinta and Dickinson). The studies also draw heavily from fields outside of linguistics, such as anthropology (Baird), social psychology (Regan), and psychology (Licata), and they target different linguistic levels, namely phonetics, morphosyntax, pragmatics, and semantics.

Just as the third wave is complemented by other fields, it is also complemented by earlier waves. As Schilling (2013) astutely notes, while we have "sailed over the first and second waves of variation study to reach the third, we would do well to remember that the three waves are part of the same ocean, that elements of all three "waves" of study were present from the outset of variation study [...], and that the best current studies will approach the social meaning of linguistic variation from a range of perspectives" (p. 343). The third wave's connection to earlier waves is highlighted in "Spanish in the Southeast: What a Swarm of Variables Can Tell Us about a Newly Forming Bilingual Community," by Jim Michnowicz, Rebecca Ronquest, Sarah Chetty, Georgia Green, and Stephanie Oliver. In their exploration of how two generations of Spanish speakers in the Southeastern United States employ multiple linguistic variables, or a "variable swarm," Michnowicz et al. find that second-generation speakers demonstrate more correlations across contact variables than first-generation speakers, with some innovative features incorporated in tandem, and that certain linguistic features follow a pattern of adoption across generations (e.g., phonological filled pauses). This approach utilizes first-wave methodologies to explore how broad social groups (first generation vs. second generation) employ variable linguistic features, and the analysis of a swarm of factors demonstrates the order of adoption and interdependence of these features within the community. However, Michnowicz et al. build on first-wave approaches by exploring individual behavior, concluding that, unlike other variables, the most salient contact features (e.g., English discourse markers) show highly variable rates of uptake. In other words, their swarm analysis sheds light on (i) which variables demonstrate predictable behavior and uptake across the community and (ii) which are potentially available as vectors for actively constructing bilingual identities at the individual level.

As we mentioned above, sociolinguistic perception studies play a key role in extrapolating the kinds of identities speakers aim to construct through a particular variant, as they establish the social meanings its use evokes. For example, in "Clothing, Gender, and Sociophonetic Perceptions of Mayan-Accented Spanish in Guatemala," utilizing a comparative word cloud analysis of evaluations of different guises, Baird finds that Guatemalan listeners tend to associate Mayan phonetic features (/f/ fortition and the apocope of word-final unstressed vowels) with the *traje típico*, or traditional Mayan clothing. However, *traje típico* responses, serving as a proxy for a Mayan cultural identity, are significantly more likely for female voices, especially those that include Mayan-accented speech, which aligns with

gendered clothing (and identity) practices in Guatemala. Baird concludes that "the visual body-language link is significantly more essentialized for the identity of a woman than for the identity of a man in Guatemala, suggesting that gendered stereotypes, language ideologies, and embodied practices mutually reinforce one another in the collective consciousness of the region" (p. 1). In other words, bilingual linguistic features live unique social lives, depending on the social context in which they are embedded.

In addition to social context, linguistic factors can influence a variant's social meanings, and this notion serves as the focus of "Exploring the Role of Phonological Environment in Evaluating Social Meaning: The Case of /s/ Aspiration in Puerto Rican Spanish" by Christina García, Abby Walker, and Mary Beaton. Building on previous research that found that Caribbean coda /s/ reduction is associated with lower status and masculinity, the authors include the additional factors of phonological context (preconsonantal /s/ vs. prevocalic /s/) and the proportion of a particular variant ([s] vs. [h]) in their methodology to determine whether and to what extent these factors influence listeners' social evaluations. The incremental addition of [s] or [h] variants did not alter listeners' perceptions, but the phonological context did, such that [s] was only rated as less masculine than [h] in preconsonantal environments, where it is least expected, particularly in men's speech. As a result, García et al. contend that marked linguistic contexts, such as preconsonantal [s] in the speech of Puerto Rican men, invite more robust socioindexical meaning.

When it comes to analyzing speech production, studies in the third wave have often turned to stance to examine the role of differing positionalities in intra-speaker or stylistic variation. In "Creaky Voice in Chilean Spanish: A Tool for Organizing Discourse and Invoking Alignment," Mariška Bolyanatz explores the meanings and stances conveyed by creaky voice quality in Chilean Spanish. Using a conversation analytic approach to examine sociolinguistic interviews from Santiago, Bolyanatz demonstrates that creaky utterances were frequently used to organize the speaker's discourse (e.g., to signal a hedge or the end of a turn). The phonation type was also commonly employed to position speakers in the interaction, namely as a token to invoke alignment, ensure comprehension, or seek support for a position, particularly when potentially controversial information was shared. Aligning with data from other languages, Bolyanatz observes that creaky voice as a hedge may connect the bodily experience with the thought process, with the bodily expression of creak connected to one's inability to complete a thought. By focusing on intra-speaker variation, this study contributes to our understanding of how linguistic resources like phonation type might be employed for various stylistic purposes, providing a basis for future work that links these positionalities to the construction of more enduring personae and identities in Spanish-speaking communities.

With the rapid rise of computer-mediated discourse, social media has increasingly become the locus of research on identity work. In "Correntino Spanish Memes and the Enregisterment of Argentine Guarani Loanwords," Justin Pinta examines the social values of memes in Corrientes, where prolonged contact between Spanish and Guarani has led to the incorporation of numerous Guarani loan words in the local variety of Spanish. Through a corpus analysis of Instagram memes, Pinta demonstrates that, in addition to conveying nostalgia, humor, and irony, Guarani loans have undergone enregisterment, indexing local ideological stances toward social phenomena and character types, such as the hard-working, family-oriented Correntino gaucho and the sweet-talking, heavy-drinking Correntino womanizer. Pinta concludes that language contact serves as an engine of variation and social meaning-making, which can be reified and intensified through memes.

Kendra V. Dickinson's "What Does it Meme? English-Spanish Codeswitching and Enregisterment in Virtual Social Space" also delves into a corpus of memes to explore the enregisterment of a particular brand of U.S. Latinx millennial identity through linguistic practices. Dickinson finds that Spanish lexical insertions, often related to cultural practices, food, and kinship terminology, help to construct a bicultural identity revolving around a shared set of references. Similarly, Latinx characterological figures are evoked through quotatives, whereby mothers, fathers, and older family members are repeatedly

constructed as Spanish speakers, unlike quotatives attributed to others in the community. Dickinson argues that, together, these linguistic features enregister an identity that involves speaking both Spanish and English as well as having close ties to Latinx cultural norms and Spanish-speaking parents. Memes can both reflect and propagate the properties and attributes associated with a Latinx millennial identity, playing an important role in identity construction in virtual social space.

Language co-occurs with other semiotic resources in the process of enregisterment, and Salvatore Callesano sheds light on the interplay between linguistic and thematic features in "Mediated Bricolage and the Sociolinguistic Co-Construction of No Sabo Kids." More specifically, Callesano uses "hashtag communities" on TikTok to analyze portrayals of "no sabo kids," a derogatory term used to refer to a purported lack of Spanish fluency among English-dominant heritage Spanish speakers. Highlighting the notion of bricolage, the discursive themes of "performative lexical gaps," "ethnicity," and "proficiency" commonly appeared alongside phonological and lexical variation. The author contends that multimodality and semiotic bricolage can enregister the linguistic features associated with the no sabo kid persona, conveying ideologies about the purported "inauthentic ethnicity" of Latinx bilinguals. Through his analysis, Callesano illustrates how language ideologies shape the creation of collective sociolinguistic styles and ethnolinguistic identities while also highlighting the urgency of considering new media in examinations of social meaning, indexicality, and enregisterment.

The third wave's applicability to interdisciplinary research is apparent in the final four articles of this Special Issue. In "(Mis)pronunciations of Hispanic Given Names in the U.S.: Positionalities and Discursive Strategies at Play," Paola Enríquez Duque unites the third wave with onomastics and explores how the variable pronunciations of given Hispanic names are used by speakers to position themselves in a range of contexts. A qualitative analysis of the metalinguistic comments of six Hispanic participants who recognize two variants of their name demonstrates, on the one hand, their phonological awareness and, on the other, their individual criteria in determining which specific features distinguish the variants. The participants' narratives shed light on how the variants of their given names are purposefully mobilized as they negotiate social positionings, a phenomenon underscored by a case study of one participant's conflicting attitudes toward the pronunciations of her name as she constructs stances of both annoyance and resignation toward the Anglicized variant. In addition to a dynamic socioindexical tool, Enríquez Duque argues that Hispanic naming practices are a sociocultural strategy that is used to bolster intergenerational relationships.

Next, Lori Czerwionka, Bruno Staszkiewicz, and Farzin Shamloo utilized a mixed-methods approach to link the third wave with pragmatics in "Contextual Variables as Predictors of Verb Form: An Analysis of Gender and Stance in Peninsular Spanish Requests." An analysis of verb forms elicited in a Discourse Completion Task (DCT) found that male and female speakers of Madrid Spanish respond to the contextual variables of power, distance, and imposition differently when selecting verbs: men made use of imposition as a higher-level predictor, while women relied more on distance. However, the participants exhibited a great deal of individual variation within these broad social groups, with certain participants employing more categorically direct or indirect verb form choices, regardless of the contextual variables provided in the DCT. Highlighting the social moves made by speakers who deviate from the gendered norm in addition to broad gendered tendencies, the findings of this study elucidate the social meanings and distinct frames associated with different verb forms at both the micro- and macro-levels of social structure.

Gabriella Licata connects the third wave with psychology in "Indexing Deficiency: Connecting Language Learning and Teaching to Evaluations of US Spanish." Using two different research paradigms—the matched guise technique (MGT) and implicit association test (IAT)—Licata tests the attitudes of three groups (Spanish teachers, heritage language students, and second language students) toward standardized Spanish and US Spanish. These tests shed light on both deeply rooted implicit bias, as elicited by the IAT, and more

pliant attitudes, as elicited by the MGT, and the use of both helps to determine if there is a correlation between participation in different language programs and shifting explicit attitudes toward Spanish in the US. Licata discovered that, while lexical features of US Spanish were salient to all three groups, there was no significant difference in groups' attitudes in either testing condition, which suggests that biases against US Spanish are shared regardless of language program, although explicit and implicit biases seem to be different cognitive processes. These findings demonstrate the prevalence of standard language and monoglossic ideologies in the US, which are often formalized in academic settings, resulting in the racialization and discrimination of US Spanish speakers.

Finally, in "Examining Monoglossic Language Ideologies in the Social Perception of Spanish in Texas," Brendan Regan and Jazmyn Martinez continue this line of research, adopting a social psychology perspective to explore the role of ideologies in socioindexical processes. The authors conduct an experiment in which they provide the same linguistic stimuli to participants across two guises but alter the social information presented with them. More specifically, they inform both second language and heritage Spanish learners that the speakers they hear have different nationalities (from Mexico or Texas), serving as a proxy for implied monolingual or bilingual status, respectively. The results of the experiment show that the implication of whether the listener is hearing a bilingual or monolingual variety of Spanish has a significant effect on the social evaluation of the speaker, with "bilingual" speakers (i.e., the guises that were presented with a 'Texas' label) being evaluated more negatively than "monolingual" voices (i.e., those accompanied by a 'Mexico' label). Given that the indexical differences observed are solely based on the implied monolingual or bilingual status of the speakers and not directly associated with specific linguistic features, these findings highlight the importance of considering exposure to language attitudes when examining social meaning. This analysis also uncovered an effect of the listeners' linguistic background: only heritage listeners evaluated Mexico-labeled voices differently from Texas-labeled guises, giving the former more positive ratings in the categories of socioeconomic status and social affect. Such an effect illustrates the indeterminacy of social meaning, which, as this study shows, applies not only to linguistic variants but also to entire language varieties.

Having advanced sociolinguistic work from a third-wave perspective in the specific context of the Spanish-speaking world, this Special Issue also opens the door for future work. While the possible research directions are innumerable, in this paragraph we focus on three key areas that strike us as particularly vital. First, given the large amount of existing research on macro-social patterns in Spanish and Hispanic contexts, a greater focus on intra-speaker variation, style-shifting, and stance-taking would not only round out the sociolinguistic panorama but also help connect the broader sociolinguistic trends observed in existing research with individual decision-making at the interactional level and locally relevant social categories. Next, we encourage more diverse methodologies that build on the more traditional and more widely employed paradigms. While the ubiquitousness of the sociolinguistic interview, for example, is clear in Hispanic sociolinguistics, relatively few researchers have adopted ethnographic methodologies that require them to immerse themselves in a community and observe it for an extended period of time. Similarly, few scholars have looked to other social science fields to incorporate experimental tests that, as we show in this Special Issue, could shed light on social meaning at different levels. Finally, if Spanish was underrepresented in previous work on the third wave, the indigenous languages of the Americas, which are often in contact with Spanish, are virtually absent.[1] While documenting the linguistic properties of indigenous languages is a worthwhile endeavor, our research needs to explore the rich social meaning conveyed by and within them, which will challenge unfounded stereotypes about their simplicity and enrich our general understanding of how social meaning is structured.

Conflicts of Interest: The authors declare no conflict of interest.

Notes

[1]   A notable exception includes the work of Anna Babel on Quechua and Spanish.

References

Campbell-Kibler, Kathryn. 2009. The nature of sociolinguistic perception. *Language Variation and Change* 21: 135–56. [CrossRef]

Drager, Katie. 2005. From bad to bed: The relationship between perceived age and vowel perception in New Zealand English. *Te Reo* 48: 55–68.

Drager, Katie. 2011. Speaker age and vowel perception. *Language and Speech* 54: 99–121. [CrossRef] [PubMed]

Eckert, Penelope. 2005. Variation, convention, and social meaning. Paper Presented at the Annual Meeting of the Linguistic Society of America, Oakland, CA, USA, January 7.

Hay, Jennifer, Paul Warren, and Katie Drager. 2006. Factors influencing speech perception in the context of a merger-in-progress. *Journal of Phonetics* 34: 458–84. [CrossRef]

Mendoza-Denton, Norma. 2008. *Homegirls: Language and Cultural Practice among Latina Youth Gangs*. Oxford: Blackwell.

Niedzielski, Nancy. 1999. The effect of social information on the perception of sociolinguistic variables. *Journal of Language and Social Psychology* 18: 62–85. [CrossRef]

Schilling, Natalie. 2013. Investigating stylistic variation. In *The Handbook of Language Variation and Change*. Edited by J. K. Chambers, Peter Trudgill and Natalie Schilling. Malden: Blackwell, pp. 325–49.

Snell, Julia. 2010. From sociolinguistic variation to socially strategic stylisation. *Journal of Sociolinguistics* 14: 630–56. [CrossRef]

Staum Casasanto, Laura. 2010. What do listeners know about sociolinguistic variation? *University of Pennsylvania Working Papers in Linguistics* 15: 40–49.

Spanish in the Southeast: What a Swarm of Variables Can Tell Us about a Newly Forming Bilingual Community

Jim Michnowicz *, Rebecca Ronquest, Sarah Chetty, Georgia Green and Stephanie Oliver

Department of World Languages and Cultures, North Carolina State University, Raleigh, NC 27695-8106, USA
* Correspondence: michnowicz@ncsu.edu

Abstract: The southeastern United States has experienced rapid growth in the Hispanic population in recent decades, giving rise to a newly forming bilingual community. The present study builds on previous work by the authors via expansion of a "variable swarm": the analysis of multiple linguistic variables simultaneously for the same set of speakers, with the goal of understanding patterns of accommodation and change within the community. The initial study included four linguistic variables (prosodic rhythm, bilingual discourse markers, the realization of /bdg/ and vowel space), and the present study adds an additional four variables (bilingual filled pauses, subject pronoun realization, code switching, and the labiodental realization of orthographic <v>) for 23 speakers of Mexican and Central American origin across two sociolinguistic generations (G1 vs. G2). Results for individual speakers show a pattern of adoption of some features by speakers of both generations (such as English-influenced prosodic rhythm and phonological filled pauses), while other, possibly more salient forms directly integrated from English (English discourse markers and code switching) exhibit later, highly variable rates of adoption, suggesting that speakers may consciously manipulate these variables as part of a process of active identity construction. Likewise, G1 speakers show fewer correlations among linguistic variables than G2 speakers, and patterns reveal that some bilingual forms are incorporated in tandem due to shared phonological traits or discourse functions. The innovative swarm analysis further contributes to the advancement of techniques employed in sociolinguistic research by serving as a bridge between traditional first- and second-wave studies that focus on a single variable, and third-wave studies that focus more on variation at the individual level.

Keywords: Spanish in the U.S.; sociolinguistics; language contact; "New Destination" communities; "variable swarm"; interfacing variables

1. Introduction

The southeastern United States, and North Carolina in particular, is home to one of the most rapidly growing Hispanic populations in the country, having witnessed 900% growth between 1990 and 2010 (Carolina Demography 2021). Due to a booming economy among other factors, as of 2020, North Carolina has an estimated Hispanic population of over one million (Pew Research Center 2014). Unlike other regions within the U.S., such as the Southwest and Northeast, which have long-standing, well-established communities, the Southeast represents a newer community or "New Destination" community (Zúñiga and Hernández-León 2005), in which the effects of language and dialect contact are still taking shape. Such a community provides the opportunity to examine the initial stages of language and dialect contact as they emerge in real time. The present study aims to augment our understanding of the diverse and rapidly growing Hispanic communities in the southeastern United States, with particular emphasis on North Carolina, by extending our initial study (Ronquest et al. 2020) via the analysis of four additional linguistic variables in a "variable swarm" (Thomas 2015): subject pronoun expression (SPE), code switching, phonological filled pauses, and an acoustic analysis of the pronunciation of <b> and <v>.

As discussed in the original "variable swarm" analysis, many sociolinguistic studies focus on a single linguistic variable and assess the relationships between the variable and macro-sociolinguistic factors. A swarm, in contrast, analyzes multiple variables for the same set of speakers and is able to provide a more detailed and nuanced view of the speech of a community, as well as offer insight into particular patterns of speaker variation and how the variables intersect and interact. Thomas (2015, p. 3) states succinctly that "[e]ven if one variable shows noteworthy patterns that provide clues about social identities of speakers, it cannot provide a complete picture of the intersecting identities that individuals exhibit. Each linguistic variable may reveal new social meanings and patterning. What is needed is inquiry that compares a large number of diverse variables".

This study's methodology and general line of inquiry, which focuses on how multiple variables pattern across and within individual speakers and larger social groups (in particular, sociolinguistic generation), situates it at the border of second- and third-wave sociolinguistics. Eckert, in her groundbreaking 2012 article, identifies three "waves" of variationist sociolinguistics, with the second and third waves overlapping in time and currently constituting the bulk of variationist work. According to Eckert's (2012) definition, first-wave studies of the 1960s and 1970s focused on how linguistic variables index pre-determined socioeconomic groups, with variation "resulting from the effects of these categories on speakers' orientation to their assigned place in the hierarchy" (p. 90). Stylistic variation was primarily seen as the avoidance of stigmatized forms in particular contexts. Second-wave studies, as defined by Eckert (2012), returned to the ethnographic roots of the very first variationist study (Labov 1963), with a focus on how speakers within communities can actively utilize "the vernacular as an expression of local or class identity" (p. 91), and language is but one (albeit key) component in the construction of group identity (along with choices in clothing and music, for example). Finally, third-wave studies shift the focus from group behavior to individual speaker behavior, exploring how "speakers place themselves in the social landscape through stylistic practice" (Eckert 2012, p. 94), a view which emphasizes the agency that speakers have in determining not only their own systems of linguistic variation, but ultimately those of the speech communities to which they belong. The present study serves as a bridge between the (sometimes) artificial distinctions between "waves." By means of an examination of eight sociolinguistic variables through a largely second-wave methodological and theoretical lens, we are able to begin to construct a larger picture of how discrete linguistic variables from across multiple domains of language (segmental and suprasegmental phonology/phonetics, morpho-syntax, and pragmatics) pattern across and within individual speakers. The swarm approach allows us to observe which variables are adopted, either in conjunction or in isolation, by particular speakers, illuminating both individual variation as well as each speaker's linguistic behavior as part of a larger social group (in this case, primarily sociolinguistic generation). In this way, a swarm approach to language variation provides a panoramic view of a speech community that is not possible through the analysis of only one or two variables, as well as sets the stage for future third-wave studies that can further explore why speakers adopt or reject possible contact forms in a newly developing bilingual region.

We begin with a brief description of Spanish in the southeastern U.S. and summarize several of the key studies that have been conducted in the area to date. Next, the overall methodology (e.g., speakers and corpus) is described, followed by separate subsections for each of the four variables under investigation. The results of the present four variables are then combined with the initial four variables from Ronquest et al. (2020), and an exploration of the interaction of all eight variables is presented in Section 4. The paper ends with a general discussion, directions for future research, and conclusions.

2. Background Studies

2.1. The Southeast as a New Dialect Region/"New Destination" Community

Since the late 1990s, the Hispanic population in the southeastern United States has grown significantly. According to the latest data, North Carolina is one of four U.S. states

that has experienced an increase to over 1 million Hispanics/Latinos since 2010, with an estimated population of 1,118,596 in 2020 (Carolina Demography 2021). Just over half (56%) of Hispanics residing in NC were born in the U.S., and 44% are foreign born. While 55% are of Mexican origin, the demographic profile is diversifying. Central American speakers make up the second-largest group, especially those of El Salvadoran and Guatemalan descent (16%) (U.S. Census Bureau 2020).

Scholars investigating Spanish in the United States have tended to focus their attention on regions within the country that have well-established communities such as the upper-Midwest, Southwest, and Northeast (c.f. Otheguy and Zentella 2012; Poplack 1978; Silva-Corvalán 1994; among others). Within the past decade in particular, however, the rapidly growing Hispanic/Latino population in the Southeast has motivated research in this previously understudied region, encompassing a wide range of topics. Studies of rhythmic differences (Carter 2005; Ronquest et al. 2020), including Hispanic English (Wolfram et al. 2004, 2011), have revealed distinctions in rhythmic profiles among bilingual and monolingual speakers of distinct backgrounds. Limerick's (2019, 2021) investigations of subject pronoun expression (SPE), which will also be examined in the present study and are described in more detail below, have revealed that Hispanics residing in Georgia exhibit distinct patterns of usage from those residing in other U.S. regions, indicating that their system is at a different/intermediate stage of development than those in more established communities.

In the lexical domain, Michnowicz et al. (2018) reported differences in acceptance and usage of English-origin loan words among first- and second-generation speakers of varying backgrounds residing in North Carolina, as well as differences with other varieties of Spanish in the U.S. Finally, work on language attitudes (Montes-Alcalá and Sweetnich 2014; Howe and Limerick 2020; Knouse et al. 2022), identity (Carter 2007, 2013), attitudes towards inclusive language (Michnowicz et al. 2023a), and language maintenance and shift (Michnowicz et al. 2023b) suggests that a complex interplay of factors is shaping the way in which residents in the Southeast perceive and utilize their language(s), as well as strategies they employ to construct their identity in a nascent bilingual environment. In conjunction, the investigations conducted thus far confirm that the Southeast is home to diverse linguistic communities at various stages of development, and that multiple factors—linguistic, social, and attitudinal—are involved in the creation of novel contact varieties.

2.2. New Dialect Formation in Language Contact Settings

Spanish in the United States is subject to two related but distinct forces that can shape the direction of future development of the language varieties present within a community (Otheguy et al. 2007). The first of these is contact with English, as bilingual speakers adopt English-influenced forms from the majority/dominant language, not just as a way to index a bilingual identity (Zentella 1997), but also as a strategy of "lightening the cognitive load of having to remember and use two different linguistic systems" (Silva-Corvalán 1994, p. 6). Language contact often results in convergence, where bilingual speakers produce forms that are in some way intermediate between the two languages in contact, particularly at the levels of pragmatics and lexicon (Silva-Corvalán 2008) or phonology (Ronquest 2012; Carter and Wolford 2016).

The second force that plays a role in the development of Spanish in the U.S. is dialect contact with other varieties of Spanish. Dialect contact between mutually intelligible linguistic varieties can result in linguistic accommodation, as speakers negotiate differences in form across regional dialects (Britain and Trudgill 1999). Accommodation is often, but not exclusively, in the direction of the majority or prestige dialect in the community, as has been found for pronouns of address (Hernández 2002) and subject pronouns (Otheguy et al. 2007). Dialect contact often results in processes of leveling and koineization or new dialect creation, whereby after a period of heightened variation, differences between dialects are diminished across time, as the younger generations in a community converge on a new set of dialect norms (Kerswill 2013). Koineization involves the processes of mixing, leveling,

and simplification, which can result in "the reduction or attrition of marked variants" (Trudgill 1986, p. 98).

Spanish in the U.S. presents a mixture of the two processes of language and dialect contact, and the line between them can often blur. For example, in her study of Spanish in New York, Zentella (1990) found that speakers of different Spanish varieties often opted for an English loanword in order to facilitate communication across communities. In other words, speakers looked to a borrowed form in order to resolve a difficulty arising from dialect contact. Particularly with regard to Spanish in the U.S., where speakers are exposed to both English and different varieties of Spanish, language contact and dialect contact are two sides of the same coin, since the consensus or intermediate form may well involve an English loan, rather than a borrowing from one of the Spanish dialects in contact. The entire sociolinguistic context must be taken into account, as speakers negotiate and accommodate both towards English and towards other varieties of Spanish simultaneously.

Features such as markedness or salience of particular forms play an important role in which features are adopted in a new (bilingual) community. Trudgill (1986) states that "[i]n contact with speakers of other language varieties, speakers modify those features of their own varieties of which they are most aware" (p. 11). Two of Trudgill's (1986, p. 11) criteria for salience of linguistic forms which are particularly relevant to the current study are if a form is "overtly stigmatized" within a community and if two forms are "phonetically radically different." We will argue later that these criteria apply to language contact-induced forms as well. Erker (2017) explicitly tests this idea, finding dialect convergence for two highly salient variables, /s/ weakening and *voseo*, while differences based on region of origin persisted for a low-salience variable, subject pronoun expression (SPE). Erker (2017) expands on Trudgill's (1986) salience criteria, finding that the linguistic domain of a feature (e.g., phonology vs. morpho-syntax) is not the deciding factor, but rather "what accounts for the different fate of features in settings characterized by both dialectal and language contact is their varying social salience" (p. 15). As we will see in the results, social salience also appears to play an important role in determining the order in which linguistic features are adopted by the newly forming Spanish-speaking community in NC.

2.3. The Initial Variable Swarm

The initial variable swarm (Ronquest et al. 2020), which serves as the basis for the present study, analyzed the lenition of /bdg/, vowel production (including the size of a speaker's vowel space as analyzed by the Convex Hull Area—CHA), prosodic rhythm (nPVI), and bilingual discourse markers (DMs) in the speech of the same 23 participants examined herein. Each of these initial variables is detailed briefly in the sections that follow, and we refer the reader to the original study for more detail.

2.3.1. Realization of Intervocalic /bdg/

In many monolingual dialects of Spanish, /bdg/ show two realizations that exist in complementary distribution: stop [bdg] is found after pauses and homorganic consonants, whereas approximant [βðɣ] arise in all other contexts, including in intervocalic position (Hualde 2005). The lenition of /bdg/ > [βðɣ] is a gradient, acoustic phenomenon that responds to both linguistic and social factors, including dialect and whether or not the variety of Spanish is in contact with another language (Colantoni and Marinescu 2010; Hualde et al. 2011; Lipski 1994, 2020).

Mexican and Central American Spanish, the two varieties studied here, both demonstrate stronger, more stop-like realizations of /bdg/ in intervocalic position when compared to some other varieties, such as Caribbean Spanish (Lipski 1994). Regarding Spanish varieties in the United States, the question of English influence arises, since English only has the stop variants of /bdg/. Some research has shown that heritage speakers (second generation [G2] or later) largely match monolingual patterns of /bdg/ lenition (Knightly et al. 2003), while other studies show a great deal of individual variation, based at least in part on how often a G2 speaker uses and interacts with Spanish. Specifically, Rao (2015)

found that regular users of Spanish showed patterns similar to those of monolinguals, while those with less regular exposure showed larger differences, including more stop-like [bdg]. Rao (2015, p. 66) notes that this can contribute to "a potential heritage accent" for some G2 speakers.

The initial swarm analysis (Ronquest et al. 2020) analyzed the intensity difference (IntDiff, Hualde et al. 2011) between the consonant and the following vowel, where a larger intensity difference suggests a stronger, more occlusive-like variant, and a smaller intensity difference indicates a more lenited variant. A total of 15,828 tokens of /bdg/ were analyzed, with results of the mixed-effects linear regression (random intercept of speaker) showing that while there was no significant main effect for any of the social factors (region of origin, generation, sex), there was a significant interaction between consonant (/bdg/) and region/generation. Overall, Mexican speakers showed higher rates of lenition than Central American speakers, but the patterns among G1 and G2 speakers were reversed across dialect groups. Central American G2 speakers produced significantly larger intensity differences than their G1 counterparts, matching the expected result for English-language influence, while for Mexicans the opposite pattern was observed. These results show the complex interplay of factors in determining the outcome of bilingual lects.

2.3.2. Vowel Space and Convex Hull Area (CHA)

The Spanish vowel system, which consists of five phonemes /ieaou/, has traditionally been described as fairly stable across dialect regions (Hualde 2005; Navarro Tomás 1918). The presence of minor differences in vowel quality and quantity across varieties has been established (e.g., Chládková et al. 2011; Quilis and Esgueva 1983), however, not to the extent observed for English, which is characterized by a larger vowel system that is highly variable across geographic regions. Studies of bilingual vowel systems, however, have revealed that both L2 learners of Spanish and G2 speakers of Spanish differ with regard to their pronunciation of the Spanish vowels. The high back vowel /u/ in particular is subject to a more fronted articulation (i.e., higher F2) in both learner and heritage systems in comparison to monolingual norms (Alvord and Rogers 2014; Cobb and Simonet 2015; Menke and Face 2010; Ronquest 2012; Willis 2005; among others). The low mid /a/ has also been described as fronted and approximating /æ/ (Willis 2005), and /e/ is often produced farther back in the vowel space among heritage speakers (Ronquest 2012).

Such differences in vowel quality, acoustic distribution, and organization often result in a more condensed vowel space for L2 learners of Spanish and some bilinguals (Menke and Face 2010), motivating the analysis of the overall area of the vowel space in the first version of the swarm. The Convex Hull Area (CHA), or overall geometric area of the vowel space, was calculated for each speaker. We hypothesized that contact-induced modifications in vowel production—such as centralization or less peripheral point vowels—would result in a smaller CHA, and would most likely be observed among the G2 speakers in the swarm given their greater degree of contact with English. While the mixed-effects linear regression (random intercept of speaker) did not reveal significant differences among the G1 and G2 participants, additional examination of the amount of variation in CHA did indicate more variability among the G2 in comparison to the G1, who were much more consistent in their productions. Analysis of the swarm indicated that 58% (7/12) of G2 and 17% (2/12) of G1 speakers favored a smaller vowel space.

2.3.3. Prosodic Rhythm (nPVI)

One of the well-known differences between Romance languages, such as Spanish, and Germanic languages, such as English, is in prosodic rhythm. In Spanish, a prototypical syllable-timed language, both tonic and atonic syllables have approximately the same duration, whereas in English, a prototypical stress-timed language, tonic syllables are lengthened and atonic syllables suffer reduction, including (in the case of English) both shorter durations and centralization (Low and Grabe 1995). Rather than being a binary distinction of stress- vs. syllable-timed languages, rhythm is a gradient feature that can

be measured using a series of established rhythm metrics, such as the Pairwise Variability Index (PVI) (Low and Grabe 1995; Grabe and Low 2002). The normalized PVI measurement (nPVI) takes the difference in duration between adjacent segments (in this case, vocalic segments), and divides that value by the mean duration of both vowels, to control for speech rate. Higher nPVI values indicate a more "stress-timed" pattern, whereas lower nPVI values suggest a more "syllable-timed" rhythm. In studies employing nPVI, rhythm has been shown to be susceptible to cross-linguistic influence in bilingual communities, with bilinguals showing intermediate or converged rhythm values (Carter 2005; Shousterman 2014; Carter and Wolford 2016), making prosodic rhythm an important point of inquiry in the development of bilingual lects.

The initial swarm analysis (Ronquest et al. 2020) used the Correlatore 2.3.4 (Mariano 2014) software package to make 59,311 vocalic comparisons across the data set. The results of the linear regression found no significant main effect for any of the independent variables (sex, region of origin, and generation), but some trends in the data point towards possible future change, as G2 speakers showed increased variability as well as a non-significant tendency to produce higher nPVI values (i.e., more "English-like"). Of the seven speakers who produced the highest nPVI values, five were G2 speakers. In this way, Spanish speakers in NC may be showing initial signs of following more established communities in the development of bilingual rhythm.

2.3.4. Bilingual Discourse Markers (DMs)

Discourse markers are "[p]articles that frequently occur in conversation . . . [that] contribute to the overall coherence of discourse by signaling relationships between portions of the speaker's utterances" (Torres 2011, p. 493). Examples of discourse markers from Spanish and English include *tú sabes, o sea, entonces, como pues, you know, I mean, so, like*, and *well*. In situations of language contact, DMs from the contact language are among the first and most common types of lexical borrowings (Torres 2011). For example, Spanish–English bilinguals may use both *so* and *entonces* in the same discourse, either with the same or with differing pragmatic functions. Over time, bilingual speakers may settle on one system of DMs in both languages, as speakers utilize DMs from the dominant language regardless of the language they are speaking. This has led some scholars to take the frequency of bilingual DMs as indicative of the level of integration into the dominant culture (Torres 2002; Lipski 2005). Due to their high frequency and largely subconscious use, bilingual DMs have been described as a "gateway" to other types of borrowings and code switching (Lipski 2005). In established Hispanic communities in the US, the use of English DMs in Spanish is common (50% among New York City Puerto Ricans, Torres 2002; 65% among Chicago Mexi-Ricans, Torres and Potowski 2008; and 68% in New Mexican Spanish, Aaron 2004). One of the striking findings of studies in established Latino communities is the high rates of English DMs even among Spanish-dominant G1 speakers, with English *so* being particularly common across social groups (Lipski 2005; Torres and Potowski 2008).

Following Torres (2002), the initial swarm analysis examined pairs of bilingual DMs that are relatively equivalent across English and Spanish (*you know~tú sabes, I mean~o sea, so~entonces, like~como* and *well~pues*). Every DM in each interview was coded for the independent social variables age, sex, and generation for a total of 1660 tokens.

The mixed-effects logistic regression (random intercept of speaker) found significant effects of discourse pairs ($p < 0.001$) and generation ($p < 0.001$), with G2 speakers producing significantly more English DMs. Sex approached significance ($p = 0.08$), with women producing more English DMs than men. Overall, the rate of English DMs in NC was much lower than in other areas (11.9% vs. 50% or greater in more established communities), showing an important distinction between the newly forming bilingual community in NC and more traditional Hispanic populations. The low rate of English DMs extended even to *so*, which was found to be ubiquitous among speakers of all generations in other regions of the U.S., with G1 speakers in NC producing less than 1% *so* (vs. *entonces*). Given that *so* has been described as a "core borrowing" (Torres 2002; Torres and Potowski 2008) in

U.S. Spanish, this finding suggests that "NC Spanish may represent an earlier stage of U.S. Spanish development than [more established regions]" (Ronquest et al. 2020, p. 317).

2.3.5. Summary of the Initial Swarm

When each variable was analyzed separately, as in a traditional second-wave study, only DMs differed significantly across generations, with G2 heritage speakers producing significantly more English DMs than G1 immigrants. Nevertheless, the swarm analysis, which was achieved by assessing the individual speaker coefficients associated with random and fixed-effects intercepts (Drager and Hay 2012), revealed important trends that were not apparent in the individual variable analyses. Coefficients (positive or negative) indicated if an individual speaker favored the contact-induced realization for a particular variable (i.e., more occlusive /bdg/, smaller CHA (vowel space), more stress-timed rhythm, and more English DMs). Examination of individual patterns revealed that those who favored the contact-induced realization for three or four of the variables in the swarm tended to be G2 speakers; three of the five speakers who did not favor any contact-induced realizations were G1 speakers. Ronquest et al. (2020) therefore concluded that "in this community, heritage speakers [G2] tend to precede IMs [G1 immigrants] in producing contact-induced realizations" (p. 319).

The patterns evident in the initial swarm also permitted a preliminary assessment of the relationship between variables. The only variable to show a significant main effect of generation (G1 vs. G2) was English DMs, as G1 speakers rarely integrated English DMs in NC, a difference with more established bilingual communities (e.g., Chicago, see Torres and Potowski 2008). Unlike English DMs, both G1 and G2 speakers showed evidence of English-influenced prosodic rhythm, suggesting early adoption of this feature in the development of bilingual lects. Speech rhythm and vowel production are also likely inherently linked, as greater degrees of vowel reduction are apt to result in rhythmic patterns that trend more towards stress timing as well as a smaller vowel space. Results pertaining to lenition of /bdg/ were more complex: Mexican and Central American heritage speakers exhibited opposite patterns, therefore suggesting the potential enhancement of a dialectal feature already present in Central American Spanish (e.g., more occlusive-like productions) and not solely the result of contact with English. In conjunction, Ronquest et al.'s (2020) findings suggest a complex interplay of factors that influence the formation of the linguistic systems of diverse speakers residing within the community.

The present study adds four additional linguistic variables that have been shown to vary in U.S. Spanish to the swarm analysis: subject pronoun expression (SPE), phonological filled pauses (FPs), code switching, and the realization of orthographic <b> and <v>. Growing the swarm with additional variables will offer further insight into how language and dialect contact manifest on different levels within the system, and if, how, when, and why members of the community might integrate these features. The background and motivation for the inclusion of each new variable is detailed below.

2.4. New Swarm Variables

2.4.1. Subject Pronoun Expression (SPE)

Spanish is a pro-drop language, and as such, overt subject pronouns are optional and variably appear based on a variety of morpho-syntactic and discourse-pragmatic factors (Otheguy and Zentella 2012). SPE refers to whether a finite verb appears with (*yo hablo*) or without (*Ø hablo*) an expressed subject pronoun, with both realizations meaning "I speak". SPE in Spanish varieties, particularly those in contact with other languages, has been described as a "showcase variable" in Hispanic sociolinguistics due to the large number of studies carried out on this phenomenon (Bayley et al. 2012). The abundance of research stems from SPE's existence at the interface between morpho-syntax and pragmatics, an area particularly susceptible to cross-linguistic influence and/or bilingual effects. Studies have shown that SPE responds to factors such as the person, number, and definiteness of the subject; the tense–mood–aspect (TAM) of the verb, often coded as distinctive (first

person and third person singular verbs have different forms, as in the preterit *hablé* vs. *habló*) vs. non-distinctive (first person and third person singular verbs have the same form, as in the imperfect *hablaba~hablaba*); the lexical content of the verb (estimative—opinion verbs such as *creer*, stative—verbs not involving any activity such as *ser* and *estar*, external activity—verbs that involve a physical action such as *ir* or *hacer*, or mental activity—verbs of thinking and volition such as *elegir* and *querer*); reflexivity; and switch reference, which refers to whether the subject of the target verb is the same or different than the subject or object of the preceding verb. Switches in reference may be complete (the subjects of the two verbs differ, as in *Mi amiga **fue** a clase y **(yo) fui** a la biblioteca* 'My friend **went** to class and (I) **went** to the library') or partial (the subject of the target verb is the object of the preceding verb, as in *Mi amiga **me** dijo que **(yo) tenía** que estudiar* 'My friend told **me** that **(I)** had to study'). Overall, overt subjects are realized more often when the subject or verb is singular, non-distinctive, estimative or referring to a mental activity, non-reflexive, or has a switch in reference from the preceding verb (cf. Sorace 2004; Otheguy and Zentella 2012; Otheguy et al. 2007; among many others for additional information on each of these factors).

Within the United States, many studies have found an increased rate of overt subject pronouns among bilingual populations (Otheguy et al. 2007; Otheguy and Zentella 2012; Abreu 2012; Shin 2013; among others), which may be due to indirect transfer from English (Silva-Corvalán 1994; Otheguy and Zentella 2012; Shin and Otheguy 2009, 2013; Shin 2013; among others). Other studies, however, suggest that observable differences in the surface pronoun rate and in the underlying grammar are due to general processes of bilingual simplification, whereby bilingual speakers lessen their cognitive load by simplifying or weakening underlying grammatical constraints (Sorace 2004, 2005).

In addition to surface effects on pronoun rates, studies have consistently found a weakening of underlying constraints among bilinguals, particularly of G2 and beyond. This weakening of sensitivity to underlying grammatical variables is especially pronounced with regard to switch reference (Silva-Corvalán 1994; Bayley and Pease-Alvarez 1997; Otheguy and Zentella 2012; Shin and Otheguy 2009; Shin 2013). Studies have suggested that a change in sensitivity to switch reference may precede a surface change in overt pronoun rates, as demonstrated by studies in emerging bilingual communities, such as Spanish in metropolitan Atlanta (Limerick 2019), even when there are no significant differences between generations in overall pronoun rate.

2.4.2. Code Switching

Code switching is defined as the "alternating use of two languages within a segment of discourse" (Toribio 2011, p. 532). Despite popular views that code switching is random, chaotic, and reflects a lack of proficiency in one or both of the languages being switched (see examples in Toribio 2004), studies have consistently shown that code switching is structured and rule based (Poplack 1980; Toribio 2004; Anderson and Toribio 2007) and that speakers need a high level of proficiency in both languages in order to produce the most complex, sentence-level code switches (Poplack 1980, 1988). Code switching is common in many, but not all, bilingual communities, and when code switching is accepted within a community, speakers most often code switch as a way of indexing in-group or bi-cultural identities (Myers-Scotton 1995; Zentella 1997). As Zentella (1997, p. 114) concluded, speakers utilize code switches as a "way of saying that they belonged to both worlds, and should not be forced to give up one for the other".

Code switches can take many forms that vary in their complexity and in the level of bilingualism required (Lipski 2005, 2008; Escobar and Potowski 2015). Individual words from English may be integrated into Spanish discourse by speakers with even the most basic English proficiency. Single-word code switches are distinguished from loanwords based on being spontaneously produced, rather than accepted forms within the community and on the lack of phonological integration for code switches (Escobar and Potowski 2015). Intersentential code switches, which occur between two independent phrases, are the next least complex category of switches, as they do not require speakers to respect the

grammar of both languages (i.e., the equivalence constraint, see Poplack 1980). In other words, less-proficient bilinguals are able to code switch between sentences "without fear of violating a grammatical rule of either of the languages involved" (Poplack 1980, p. 581). Finally, intrasentential code switches occur at predictable points within the same sentence where the grammar of the two languages is equivalent, generally respecting the integrity of a syntactic constituent (Poplack 1980). Intrasentential code switches occur only among the most proficient bilinguals (Poplack 1980; Lipski 2014), and therefore, the presence of more complex code switches can be interpreted as an indicator of the level of bilingualism or linguistic integration of a particular speaker (Poplack 1980). Examples from our data of each of the three types of code switches considered here are found below:

- Single word switch: *Ella estaba bien **surprised*** 'She was very **surprised**'.
- Intersentential switch: *¿Cómo te puedo decir?* ***The right path*** 'How can I say it (for you)? **The right path'**.
- Intrasentential switch: *O como cuando van **hunting in the woods*** 'Or like when they go **hunting in the woods'**.

2.4.3. Filled Pauses

Filled pauses (FPs) are "nonsilent hesitations" that serve a variety of functions within conversation, including giving speakers "time to plan utterances and a way to hold the conversational floor" (Erker and Bruso 2017, p. 205). Filled pauses may either be lexical (e.g., English *well*, *so*,[1] or Spanish *este*, *sea*) or phonological (e.g., English *u(m)*, or Spanish *e(m)*), where the different default vowels in each language (/ə/ for English, /e/ for Spanish) appear as phonological fillers. The present study focuses on phonological fillers, where there is no overlap with other categories, such as discourse markers (see Ronquest et al. 2020). FPs have not been widely studied in bilingual U.S. Spanish, with Erker and Bruso (2017) as an important exception. The authors examine three possible realizations of phonological FPs in the Spanish spoken by G1 and G2 in Boston, MA: Spanish [e], English [ə], and a third form, [a]. They find that, as contact with English increases, so do instances of both [ə] and [a]. While [ə] is interpreted as transfer from or convergence with English, [a] is seen as a possible intermediate form that is shared by the phonological systems of both languages, and therefore, may be used by bilingual speakers as a compromise form, lessening the cognitive load of maintaining two separate FP systems (Erker and Bruso 2017, p. 238). Alternatively, the authors note that in usage-based theories, the increase in [a] may stem from the increased presence of [ə] variants in bilingual speech, and speakers "might establish connections between exemplars of this category and those of the vowel category that is most like it acoustically and that is used to similar effect when speaking in Spanish, namely, [a] (and not [e])" (Erker and Bruso 2017, p. 239). Whatever the theoretical model, a decrease in the use of [e] FPs is taken to indicate higher levels of integration into the English-speaking environment.

2.4.4. The Realization of Orthographic b~v

It is widely reported that monolingual varieties of Spanish do not distinguish between orthographic <b> and <v>, as both are pronounced as variants of the phoneme /b/ (Hualde 2005). Studies of bilingual Spanish in the United States, however, have reported a distinction between bilabial /b/ corresponding to <b> and labiodental /v/ corresponding to <v>, particularly among G2 speakers (Trovato 2017). Given that English is characterized by a phonemic distinction between /b/ and /v/, the existence of an incipient distinction among U.S. Spanish speakers is taken to be due to the influence of orthographic <v> in English (Rao and Ronquest 2015; Boomershine and Ronquest 2019). Among speakers with less exposure to Spanish, Rao (2014) found more "tense" (i.e., less monolingual-like) pronunciations of orthographic <v> compared to orthographic <b>, although the lack of fricatives in his data made him doubt the direct influence of English fricative /v/.[2] In a more detailed analysis, Rao (2015) does allow for the influence of English on orthographic <v>, stating that "[t]he relatively high rates of [tense approximant] realizations associated

with increases in articulatory tension that trend toward a fricative production suggest that the English /v/ phoneme interfered with productions of /b/, which was exacerbated by seeing <v>" (p. 68).

While a few studies have undertaken acoustic analyses of Spanish <b> and <v>, researchers are not in clear agreement on which acoustic measures best capture the difference between bilabial and labiodental realizations. Trovato's (2017) study of the production and perception of <b> and <v> among bilinguals residing in Texas revealed a significant effect of segment duration and intensity difference, but not spectral center of gravity (COG), in spite of COG being one of the primary acoustic correlates utilized to distinguish fricatives. COG is a weighted mean of frequencies measured in Hertz (Hz) that indicates how high or low in the spectrum most of the energy is concentrated (Boersma and Weenink 2022). Different from Trovato (2017), Chetty (2018), who incorporated videos of lip movements, found that COG was the best predictor of labiodentalization: the presence of teeth in the video correlated with significantly higher COG values, suggesting higher rates of frication, as fricatives are produced with the bulk of their energy at higher frequencies (Ladefoged and Disner 2012). Likewise, initial analyses of the present data confirm that COG is a more reliable predictor of grapheme, as the COG of productions of <v> trends higher than that of <b>, while duration and intensity difference did not yield significant results. Based on these results, our analysis focuses on COG, although further exploration of the relationship between <b>/<v> and various acoustic properties is warranted in future studies.

2.4.5. Growing the Swarm

The addition of the four variables outlined above, in conjunction with the four variables included in the initial swarm analysis (Ronquest et al. 2020) contributes to our understanding of how bilingual communities and individual speakers integrate contact-induced forms in the early stages of contact, and the results presented here will focus on differences between G1 and G2 speakers.

To facilitate comprehension of the new swarm variables and how they connect with those analyzed in the initial swarm (Ronquest et al. 2020), we begin with a review of the general methodology, followed by methods specific to the analysis of each individual variable (SPE, code switching, FPs and <b/v>). Separate results sections for each variable are presented next, and the new integrated swarm analysis including all eight variables is presented in Section 5.

3. Materials and Methods

3.1. General Methods

The data for the present study were obtained from sociolinguistic interviews with 10 men and 13 women (23 informants total) of Hispanic/Latino descent ranging in age from 20 to 53 (average age 27.3 years). Informants were further subdivided into two groups: G1 speakers who were foreign-born immigrants (n = 12) and G2 speakers (n = 11) who were born in the United States or born outside of the United States and immigrated by the age of three. Sixteen participants were of Mexican Heritage and seven were from Central America (El Salvador and Guatemala). Detailed information regarding demographics can be viewed in Appendix A.

Informants participated in sociolinguistic interviews that lasted between 30 and 60 min and included questions pertaining to general experiences living in the United States and abroad, customs, and similarities and differences between the U.S. and their heritage countries. Interviews were recorded in quiet locations utilizing Zoom H2 digital recorders (44.1 kHz, 16 bit) and subsequently transcribed orthographically in Praat (Boersma and Weenink 2022) to facilitate forced alignment with the *Forced Alignment System for Español* (FASE; Wilbanks 2015). Alignment facilitated acoustic analysis of phonetic variables in the first swarm (i.e., lenition of /bdg/, vowel space, nPVI) and acoustic properties of <b> and <v> in the current study. The automated system significantly increased the number of

analyzable units by orders of magnitude (for example, more than 125,000 vowel tokens in the initial swarm and more than 3500 <b/v> tokens here; see Labov et al. 2013).[3]

Additional details regarding the specific coding scheme and statistical analysis for each of the four variables included in the present analysis are provided below. The swarm analysis, which focuses on analyzing individual speaker patterns and the interactions between all variables, is presented in Section 5.

3.2. Methodology: Subject Pronoun Realization (SPE)

For each of the sociolinguistic interviews, the first 100 finite verbs that fit into the envelope of variation were identified. The requirements for inclusion in the envelope of variation were based on the coding manual found in Otheguy and Zentella (2012) and included all finite verbs that could appear with an overt subject pronoun, whether or not an overt subject pronoun was present. Verbs appearing with lexical subjects, inanimate subjects, and non-personal pronouns (such as *eso* or *aquel*) were considered outside the envelope of variation and thus excluded from the study.

A total of 2265 tokens were included in the analysis. A series of mixed-effects logistic regression models was run in Rbrul (Johnson 2009), with a random intercept of speaker. The binary-dependent variable was whether the pronominal subject was null or overt, and independent variables included a variety of morpho-syntactic, pragmatic, and social factors found to be important predictors of SPE in previous studies (Otheguy and Zentella 2012; Otheguy et al. 2007; among many others), as well as speaker generation, sex, and region of origin, as seen in Table 1.

Table 1. Variables included in the analysis of SPE.

Variable	Levels
Dependent	
SPE	Overt
	Null
Independent	
	Yo
	Tú (definite)
	Tú (indefinite)
Person/number/definiteness	Él/ella (definite)
	Nosotros/nosotras
	Ellos/ellas (definite)
	Ellos/ellas (indefinite)
Distinctiveness of TAM	Distinctive
	Non-distinctive
	Complete switch
Switch reference	Partial switch
	No switch
	External activity
Lexical content	Mental activity
	Estimative
	Stative
Generation	G1
	G2
Sex	Male
	Female
Region of origin	Mexico
	Central America
Speaker	Random intercept

In order to establish the comparative constraint hierarchies across groups (Tagliamonte 2011), the relative importance of each linguistic variable was determined based on model comparison via AIC, with one independent variable removed for each of the subsequent

models.[4] The AIC values for each model were then compared, with the size of the difference in AIC value between the full model and the reduced model indicating the strength of each variable in the analysis (see Kapatsinski 2012).

3.3. Methodology: Code Switching (CS)

In each sociolinguistic interview, all code switches were identified and categorized as single word, intersentential or intrasentential, for a total of 779 instances of code switching. Observed patterns in the data were confirmed via a chi-square analysis run in R (R Core Team 2022), as well as mixed-effects logistic regressions in Rbrul (Johnson 2009) with type of switch as the dependent variable and a random intercept of speaker. Two separate regression models were run, one with sentence-level switches (vs. word-level switches) as the application value, and a second with the most complex type of switch (intrasentential vs. other) as the application value. The independent variables for both models were speaker generation, sex, and region of origin. Table 2 presents the variables included in the analysis.

Table 2. Variables included in the analysis of code switching.

Variable	Levels
Dependent	
Code-switch type	Sentence-level code switching (Intrasentential + intersentential) (analysis 1)
	Word
	Intrasentential (analysis 2)
	Intersentential + Word
Independent	
Generation	G1
	G2
Sex	Male
	Female
Region of origin	Mexico
	Central America
Speaker	Random intercept

3.4. Methodology: Filled Pauses (FP)

For this initial analysis, every phonological filled pause in each interview was impressionistically coded as a variant of [e(m)], [a(m)] or [ə(m)] based on the vowel heard in each token.[5] A total of 1925 filled pauses were included in the analysis, and a chi-square analysis was conducted in R (R Core Team 2022) to determine the significance of the overall distribution of forms across generations. For the mixed-effects logistic regression analysis fit to the data in Rbrul (Johnson 2009), the binary-dependent variable was the bilingual forms [ə] + [a] vs. the monolingual Spanish form [e]. The independent variables were speaker generation, sex, and region of origin. Speaker was included as a random intercept. The variables included in the analysis of filled pauses are presented in Table 3.

Table 3. Variables included in the analysis of FPs.

Variable	Levels
Dependent	
Filled pause vowel	[ə] + [a] [e]
Independent	
Generation	G1 G2
Sex	Male Female
Region of origin	Mexico Central America
Speaker	Random Intercept

3.5. Methodology: The Realization of Orthographic <b/v>

In each of the sociolinguistic interviews, all instances of intervocalic /b/ were identified, for a total of 3550 tokens. Following File-Muriel and Brown (2011), a Praat (Boersma and Weenink 2022) script was used to measure the COG of the middle 60% of the token, with the outer 40% being ignored in order to minimize any effect of surrounding phonemes on the COG measurement. Since initial analyses determined that the data were highly skewed to the right,[6] the COG measurements were log transformed in order to normalize the distribution of the data for analysis. Two mixed-effects linear regression models were fit to the data in Rbrul (Johnson 2009) with random intercepts of speaker and word. The dependent variable of the first was log (COG), and the independent variables were the following vowel and word position (initial vs. medial), grapheme (<b> vs. <v>), as well as speaker generation, sex, and region of origin (see Table 4). Finally, in order to better compare <b/v> patterns across speakers, the mean COG value for <b> was subtracted from the mean COG value for <v> for each speaker, with a positive value indicating a higher mean COG for orthographic <v>. The difference in mean COG values formed the dependent variable of the second linear regression model, with speaker as a fixed effect. The results of the COG difference analysis will be discussed as part of the swarm analysis in Section 5.1.

Table 4. Variables included in the analysis of <b/v>.

Variable	Levels
Dependent	
COG	Continuous variable (analysis 1) COG difference <v>—<b> (analysis 2)
Independent	
Following vowel	[i] [e] [a] [o] [u]
Word position	Initial Medial
Grapheme	<b> <v>
Generation	G1 G2
Sex	Male Female
Region of origin	Mexico Central America
Speaker	Random intercept
Word	Random intercept

4. Results

4.1. Subject Pronoun Expression (SPE)

Table 5 presents the results of the mixed-effects logistic regression. In the regression table, a positive log odds indicates a favoring of an overt subject pronoun. The overall analysis with all speakers found significant main effects of person, number and definiteness, switch reference, lexical content of the verb, and reflexivity, with patterns largely matching those found in previous studies (Table 5). Generation approached significance; however, a comparison of overall pronoun rates shows very little difference between G1 and G2 speakers in the present data. This corroborates the findings of Limerick (2019) for Roswell, Georgia, another community in the Southeast, further supporting that the southeastern U.S. may be at an earlier stage of bilingual SPE development than more established communities, where significant differences across generations are frequently found (e.g., New York City, see Otheguy and Zentella 2012).

Table 5. Results of the multivariate one-level mixed-effects regression model, speaker as a random factor. * Significant factors with a *p*-value < 0.05.

Factor	Log Odds	N	% Overt Pro	*p*-Value
Person/Number and Definiteness				<0.001 *
él-ella.definite	1.902	267	38.2	
1st singular—yo	0.847	1214	21.6	
3rd plural—ellos-ellas.definite	0.567	209	18.2	
2nd singular—tú.definite	−0.352	58	8.6	
2nd singular—tú.indefinite	−0.494	151	6.6	
1st plural—nosotros-nosotras	−0.589	292	8.6	
3rd plural—ellos-ellas.indefinite	−1.881	74	2.7	
Switch Reference				<0.001 *
Complete switch	0.643	945	25.5	
Partial switch	−0.284	181	18.2	
No switch	−0.359	1139	15.0	
Reflexivity				<0.001 *
Non-reflexive	0.392	1997	20.7	
Reflexive	−0.392	268	11.6	
Distinctiveness of TAM				<0.001 *
Non-distinctive	0.248	475	26.1	
Distinctive	−0.248	1790	17.9	
Lexical Content				0.002 *
Estimative	0.676	99	43.4	
Stative	−0.132	562	22.0	
External activity	−0.220	1229	17.4	
Mental activity	−0.324	375	17.1	
Generation				0.094
G1	0.238	1173	19.9	
G2	−0.238	1092	19.2	
Sex				0.24
Female	0.173	1282	21.6	
Male	−0.173	983	17	
Region				0.60
Central America	0.08	686	24.2	
Mexico	−0.08	1579	17.6	

N	df	intercept	overall prop	AIC	R2.fixed	R2.random	R2.total
2265	18	−2.348	0.196	1970.175	0.224	0.063	0.287

Table 6 shows the comparative constraint hierarchies for speaker subgroups when analyzed separately, broken down by origin and generation, revealing important differences

between subgroups (Full regression tables are found in Appendix B). G1 speakers show the same constraint hierarchies, indicating that they share an underlying grammar for SPE, regardless of their region of origin. All five linguistic variables were significant for Mexican G1 speakers, whereas only the first three variables were significant for Central American G1 speakers.[7]

Table 6. Variable hierarchies by generation and origin. * Significant factors with a *p*-value <0.05.

Mexican G1	Central American G1	Mexican G2	Central American G2
1. Person/Number *	1. Person/Number *	1. Person/Number *	1. Person/Number *
2. Switch *	2. Switch *	2. Lexical Content *	2. Switch *
3. Lexical Content *	3. Lexical Content *	3. TMA Distinctive *	3. Lexical Content
4. TMA Distinctive *	4. TMA Distinctive	4. Reflexivity *	4. TMA Distinctive
5. Reflexivity *	5. Reflexivity	5. Switch *	5. Reflexivity

Several important differences appear when comparing G1 and G2 speakers. Mexican G2 speakers show the same overall hierarchy and significance as Mexican G1 speakers, except that switch reference is much less important for the G2 group. This finding suggests that these speakers show a decreased sensitivity to switch reference, which has also been reported in contact varieties of Spanish around the world (Shin and Otheguy 2009; Michnowicz 2015). Central American G2 speakers show the same order of constraints as their G1 counterparts, but with fewer significant differences (i.e., lexical content of the verb is no longer significant for G2 speakers).

In summary, while there are very few differences in the use of overt subject pronouns across generations overall, a comparative analysis of constraint hierarchies reveals a simplification of the underlying grammar for G2 speakers, realized as a weakening of switch reference or a reduction in the number of significant predictors, which suggests possible future changes in rates of overtly expressed subject pronouns in NC Spanish.

4.2. Code Switching

Of the 779 total code switches in the corpus, the vast majority were single-word switches, as has also been observed in previous studies (Poplack 1980). Intersentential switches were a distant second (7%), with intrasentential switches the least frequent and comprising only 2.6% of the tokens (Table 7). In this way, the frequency of code-switch types in the corpus reflects the complexity of each type of switch, as simpler switch types are also more frequent in the data.

Table 7. Frequency of code-switch types.

Switch Type	N	% of Data
Single word	702	90.1%
Intersentential	57	7.3%
Intrasentential	20	2.6%

The distribution of code-switch types across generations follows the expected pattern, with 19/20 (95%) of the most complex intrasentential code switches appearing among G2 speakers. The chi-square analysis revealed a significantly different pattern across generations ($X2(2, N = 779) = 8.1604$, $p = 0.017$), as seen in Figure 1, which presents code-switch (CS) type by generation. The width of the box for each generation indicates the number of tokens, showing that G2 speakers produced many more code switches overall, and likewise produced the majority of inter- and intrasentential code switches.
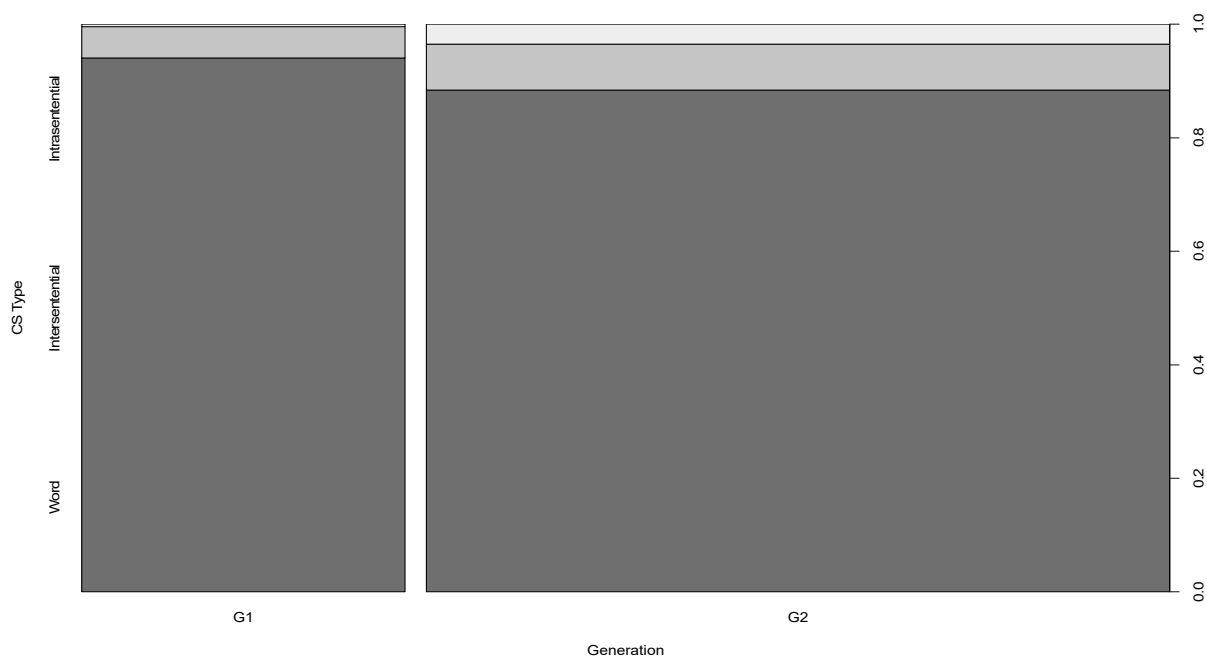
Figure 1. Code switch type by generation.

The mixed-effects logistic regression comparing sentential switches vs. single word revealed a significant effect of generation ($p = 0.046$), with G2 speakers producing significantly more sentential switches than G1. When comparing the most complex switch type (intrasentential switches) vs. other switch types (intersentential + word-level), the difference between generations was even more pronounced ($p = 0.025$), as only one of the twenty intrasentential code switches was produced by G1 speakers. Region of origin and speaker sex were not significant predictors of switch type. In summary, G2 speakers in NC produced more code switches overall than G1 speakers, a difference that was most pronounced for sentential-level code switches, suggesting the potential for increased levels of code switching among future generations of speakers.

4.3. Filled Pauses (FPs)

The analysis of FPs revealed that G2 speakers produced more than twice as many phonological filled pauses as G1 speakers (1304 vs. 621), which may suggest differences in fluency (see García-Amaya 2009 and sources therein).[8] Regarding specific forms (see Figure 2), G2 participants produced almost three times as many instances of [ə] (28% compared to 11% for G1), suggesting increased English influence in their Spanish. Conversely, G1 speakers showed higher rates of [e] (20% vs. 15% for G2). Speakers of both generations showed [a] as the majority form (69% G1, 57% G2), providing further evidence in support of Erker and Bruso's (2017) claim that [a] may serve as an intermediate form between Spanish [e] and English [ə]. A chi-square analysis found the overall distribution of filled pause variants across generations to be significant ($X^2(2, N = 1925) = 74.612, p < 0.001$).

The mixed-effects logistic regression comparing the bilingual forms [ə] and [a] to the monolingual variant [e] found a marginally significant effect of sex ($p = 0.053$), with men producing more bilingual FPs than women (93% vs. 79%). Generation approached significance ($p = 0.09$), with G2 speakers producing more bilingual FPs than G1 speakers (85% vs. 80%). Region was not a significant predictor ($p = 0.21$).

Despite the observed differences across generations, further analysis of the R^2 values for fixed and random factors indicates that individual speaker differences are the primary factor in FP variation in NC, as the random factor of speaker accounts for 43% of the observed variation compared to only 12% for the fixed effects. This is clearly seen in Figure 2, which shows the filled pause variants for each speaker divided by generation.

With the exception of one speaker (2014-13, a G1 Salvadoran male who has spent 19 years in the U.S.), G2 speakers show much greater variation than G1, both in the number of FPs and in their realization. Likewise, while most G2 speakers produced at least some instances of /ə/, the bulk of the centralized tokens was produced by two speakers (2012-15, 2013-19, a female of Salvadoran heritage and a male of Mexican heritage, respectively; both speakers report speaking English and Spanish at home). At the same time, Figure 2 shows that most speakers of both generations display intra-speaker variation, even if a majority has settled on /a/ as their default FP. These findings will be addressed further in the discussion.
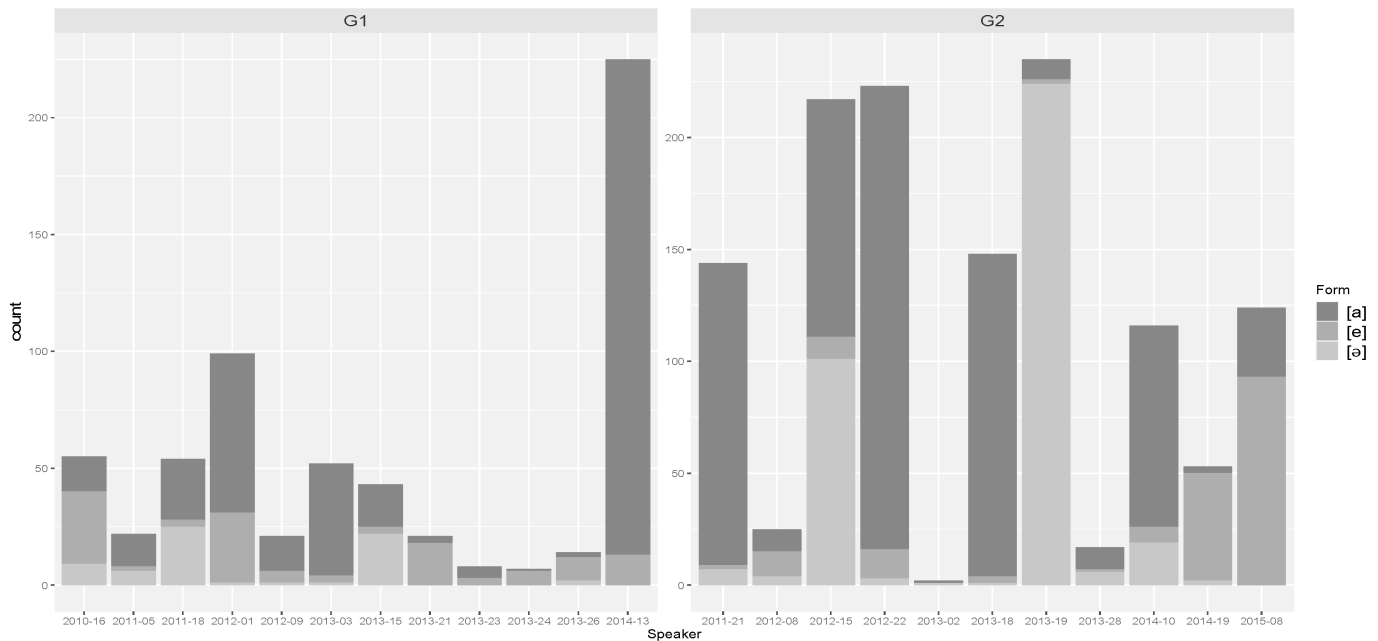

Figure 2. Filled pause variants by speaker and generation.

4.4. The Realization of Orthographic <b/v>

The mixed-effects linear regression model revealed a significant main effect of following vowel ($p < 0.001$), with grapheme being marginally significant ($p = 0.051$). COG values were higher for the grapheme <v>, suggesting a more fricative, labiodental pronunciation (Chetty 2018). Word position, speaker sex, generation, and region of origin were not significant predictors of COG in the present analysis.[9] A visual inspection of the data shows that both generations of speakers essentially show the same pattern of higher COG values for orthographic <v> than for <b>, which previous research has shown may indicate a fricative, labiodental articulation (Chetty 2018). The results by generation and grapheme are seen in Figure 3.

While the overall pattern is similar, the boxplots in Figure 3 suggest an increased separation between <b> and <v> for G2 speakers, as evidenced by less overlap in the boxes between <b> and <v>, as well as greater differences in the median values across graphemes. This observation is borne out by the results of an additional mixed-effects linear regression model that included an interaction of generation and grapheme, which indicated a slight tendency for G2 speakers to make a greater distinction between orthographic <b> and <v> in NC ($p = 0.085$).

In sum, even with a relatively small number of speakers included in the analysis, the potential for both individual variation and the impact of surrounding phonetic context, our analyses indicate a possible generational difference with regard to the articulatory and acoustic properties of <b> and <v> among NC bilinguals.
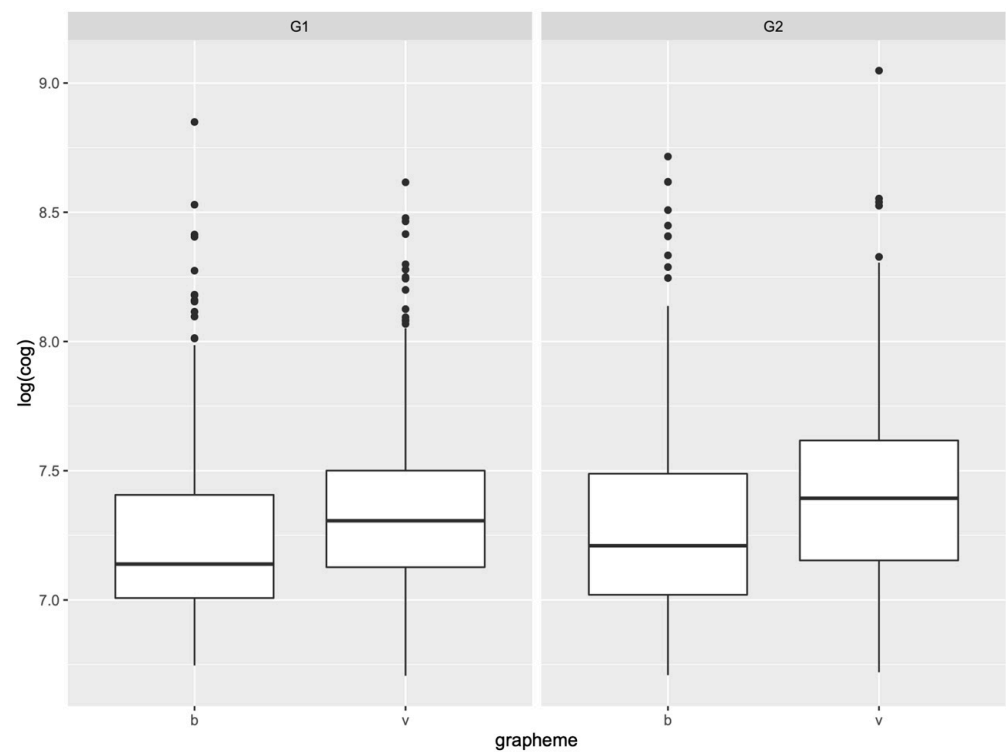
Figure 3. Log (COG) by grapheme and generation.

5. Discussion: The Swarm Analysis

5.1. Speaker Rankings

One of the primary benefits of a swarm analysis is the ability to examine how variables pattern both across and within individual speakers. Here, the results of the four linguistic variables from the present study (SPE, code switching, FPs, and <b/v>) are added to the four initial swarm variables (/bdg/ lention, vowel space via CHA, prosodic rhythm via nPVI, and bilingual DMs) analyzed in Ronquest et al. (2020) across the same group of speakers. A focus on individual speaker variation can provide additional insight into the creation of (bilingual) linguistic norms in a "New Destination" community such as NC. As in Ronquest et al. (2020), we follow Drager and Hay (2012), who show how the random effects coefficients from mixed-effects models can be used to make comparisons between individual speakers by using the individual speaker estimates from each model to rank speakers as favoring (positive coefficient) or disfavoring (negative coefficient) a particular realization for each variable. The random effects intercepts were utilized for FPs, DMs, SPE, /bdg/, and code switching. Since the analyses of nPVI, CHA, and <b/v> COG difference produced only one mean value per speaker, the fixed-effects estimates for speaker (run in separate models) were used to create the ranking for these variables. A positive coefficient indicated that a speaker favored the variant that could be considered "contact-induced" or "less monolingual": higher nPVI, FPs with a vowel other than [e], more English DMs, more overt subject pronouns, a more occlusive realization of /bdg/, higher rates of intrasentential code switching, and higher COG values for <v>, suggesting more labiodental variants. For CHA, the expected bilingual pattern is a smaller vowel space (Menke and Face 2010), which was indicated by a negative coefficient.

In Figure 4, a check mark indicates that a speaker statistically favored the contact-induced variant for that variable. Speakers were then ranked according to the number of contact variants they favored. For example, speaker 2011-21 (first row) statistically favored higher nPVI, bilingual FPs, English DMs, overt subject pronouns, greater intensity differences for /bdg/, a smaller vowel space and intrasentential code switching. This speaker did not favor higher COG values for <v> than for <b>.

Speaker	Generation	Region	Sex	nPVI	Filled Pauses not /e/	DMs	SPE	/bdg/	CHA	IntraCS	b/v.Diff
2011-21	G2	CAm	Female	✓	✓	✓	✓	✓	✓	✓	
2012-22	G2	CAm	Female	✓	✓	✓	✓		✓	✓	✓
2013-19	G2	Mex	Male	✓	✓	✓	✓	✓	✓		
2011-05	G1	Mex	Female	✓	✓	✓	✓		✓		
2012-09	G1	Mex	Female	✓	✓	✓	✓				✓
2012-15	G2	CAm	Female	✓	✓			✓	✓		✓
2013-15	G1	Mex	Male		✓	✓		✓		✓	✓
2013-18	G2	Mex	Female	✓	✓	✓			✓		✓
2012-01	G1	Mex	Female	✓	✓	✓		✓			
2013-03	G1	Mex	Male	✓	✓		✓	✓			
2013-21	G1	Mex	Female	✓				✓	✓		✓
2014-10	G2	CAm	Female	✓	✓	✓			✓		
2014-19	G2	Mex	Female	✓			✓			✓	✓
2011-18	G1	Mex	Male		✓		✓				✓
2013-02	G2	Mex	Male		✓						✓
2013-23	G1	Mex	Male				✓	✓			✓
2014-13	G1	CAm	Male	✓	✓						✓
2015-08	G2	Mex	Female	✓		✓			✓		
2013-28	G2	Mex	Male			✓				✓	
2012-08	G2	CAm	Female					✓			
2013-24	G1	Mex	Male	✓							
2010-16	G1	CAm	Female								
2013-26	G1	Mex	Male								

Note. A check mark indicate the speaker favored the contact-induced realization for that specific variable.

N of speakers favoring the contact form by generation											
			G1	7	7	4	5	5	2	1	6
			G2	8	7	7	4	4	7	4	5

Figure 4. The variable swarm variables by speaker. A check mark indicates that the speaker statistically favors the contact-induced variant.

Although the speaker rankings in Figure 4 indicate that the preference for bilingual forms in NC Spanish is highly variable, some important trends emerge when analyzing individual speakers. First, while no speaker statistically favors contact forms for all eight variables, the two speakers who favor contact forms for seven out of eight variables are both G2 (as indicated by checkmarks in Figure 4). Likewise, of the eight speakers who favor contact forms for five or more of the variables, five are G2. At the bottom of the chart, of the four speakers who favor one or fewer contact forms, three are G1. These results are similar to those for the initial swarm analysis in Ronquest et al. (2020) and speak to the robustness of these patterns even as more variables are added. The middle of the chart is characterized by substantial variation with respect to generation, as predicted by theories of new dialect formation and koineization (Kerswill 2013), as Spanish speakers in NC negotiate the newly forming norms in their community.

A comparison of patterns of use across G1 and G2 speakers reveals further detail regarding how bilingual forms are integrated into the community. The numbers at the bottom of Figure 4 indicate how many speakers of each generation favor the contact-induced form for each variable. A majority of the variables show a relative balance between G1 and G2 speakers, while three variables show a clear effect of generation: English DMs (seven G2 speakers favoring vs. four G1 speakers), smaller CHA (seven G2 vs. two G1), and intrasentential code switching (four G2 vs. one G1). These patterns can suggest an order of integration of bilingual forms, with English-like rhythm, FPs, SPE, occlusive [bdg] and labiodental productions of <v> appearing in the speech of G1 speakers, whereas English DMs, more complex code switches, and changes to the vowel space appear to require a higher level of English dominance in order to enter into the speech of NC Spanish speakers.

One interesting observation is that the bilingual forms that involve the direct integration of English words and phrases (DMs, code switching), and therefore perhaps are most salient to speakers, are also among the last to be integrated. While an interviewer effect has almost certainly played a role in suppressing code switching among some participants, as most interviews were conducted by L2 speakers from outside the Latino community, speakers' attitudes toward bilingual speech may also be an important factor. Studies have repeatedly demonstrated that many speakers hold negative attitudes toward overtly bilin-

gual forms, such as code switches (Anderson and Toribio 2007; Rangel et al. 2015; Mata 2022), and anecdotal evidence from our own outreach efforts with the Latino community in NC finds similar strong, negative reactions toward what some speakers interpret to be "Spanglish." These commonly held negative attitudes toward "mixed" speech may be reflected in speakers' reluctance to integrate overtly English forms into their speech, at least in the context of a sociolinguistic interview. Erker (2017) argues that highly salient variables are sites of convergence and overt manipulation, as speakers' conscious awareness of these features is subjected to social forces within the community. On the other hand, less-salient variables, such as many phonetic or morpho-syntactic features (e.g., SPE), respond primarily to the pressures of cognitive economy "that comes at little social cost" (Erker 2017, p. 16). The present findings reinforce this possibility, and give further weight to arguments based on salience as a deciding factor in the adoption of a feature in a bilingual community.

In light of this possibility, one interesting question that arises is why do some speakers choose to utilize overtly contact-induced forms in an interview context while others do not. For some speakers, particularly of G1, the lack of code switching and other bilingual forms in their interviews may very well reflect their normal linguistic patterns, as it has been argued that "New Destination" communities such as NC lack the critical mass of bilingual speakers required to encourage the wide-spread adoption of many contact variants (Ronquest et al. 2020). On the other hand, variationist studies in the "third wave" tradition (Eckert 2005, 2012; Mendoza-Denton 2002, 2010) emphasize the social agency of speakers to "actively [construct their identities] as they creatively and aesthetically combine linguistic elements" (Mendoza-Denton 2010, p. 189). In other words, the question becomes why have some speakers opted to utilize overtly bilingual forms (such as code switching) during a semi-formal sociolinguistic interview with an outsider? Woolard (1999) argues that bilingual speakers have the option of bivalency, "the use by a bilingual of words or segments that could 'belong' equally, descriptively and even prescriptively, to both codes" (p. 7). Michnowicz et al. (2018) have made similar arguments for English loanwords in NC Spanish, as speakers who use semantic calques such as "carpeta" (*carpet*, "rug" based on English "carpet" rather than Spanish "alfombra") have made a (semi-) conscious choice to remain in Spanish mode, instead of switching to English "carpet". In the same way, speakers who choose to integrate code switches or English DMs into their Spanish may be actively choosing to use their (bilingual) variety of Spanish, rather than simply switching to English, a language that many G2 speakers may be more comfortable using. One hypothesis would be that speakers who favor contact forms for these salient variables are actively indexing a bilingual Latino identity, showing themselves to be part of the newly forming bilingual community in NC. Although the present data do not permit us to make definitive conclusions with regard to the role of speaker agency in how bilinguals choose to employ (or not) bilingual codes, the (arguably) more salient variables such as DMs and code switching can be most overtly and easily manipulated by speakers in an interview context and could therefore be viewed as tools for identity construction.

5.2. Correlations across Variables

In addition to speaker-specific patterns of use and integration, the swarm analysis also allows for an examination of the correlations between variables. Following research in dialectology that examines the co-occurrence of variables across regional dialects (Coloma 2012), correlations between variables were run on each generation separately, and were plotted using the corrplot package (Wei and Simko 2021) in R (R Core Team 2022). Significance of correlations was determined with rquery.cormat.[10]

Close inspection of Figure 5 reveals several notable differences between variable correlations across generations. First, assessment of the statistical relationships among variables within G1 speakers indicates that there is only one moderately strong correlation: a positive correlation between FPs and DMs. The significant correlation between FPs and DMs ($p = 0.04$) indicates that as G1 speakers use more bilingual FPs ([ə] or [a]), they also use more English DMs (*you know*, *I mean*, *so*, *like*, and *well*). Given that FPs and DMs

serve similar discourse functions, such as turn taking and holding the floor (see Erker and Bruso 2017), the correlation between the two variables is not surprising and supports the finding that if G1 speakers integrate FPs and DMs into their Spanish, they tend to integrate them together.
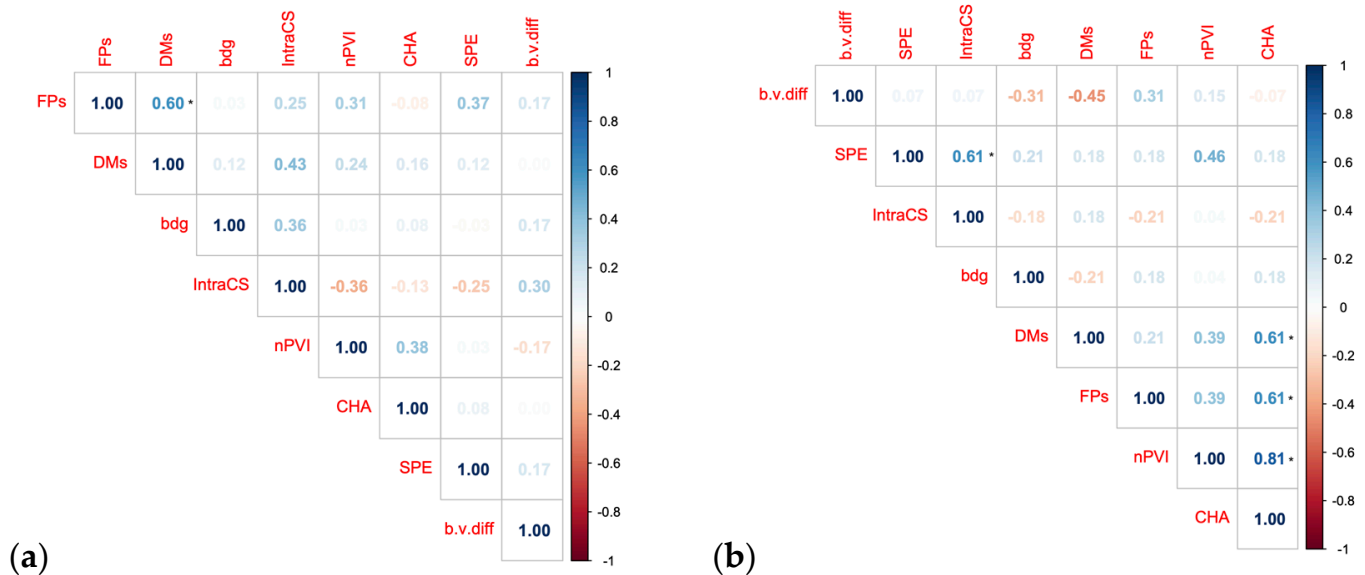

Figure 5. Correlations between variables by generation. * indicates a significant correlation at the $p \leq 0.05$ level. (**a**) G1 speakers; (**b**) G2 speakers.

G2 speakers show a higher rate of correlation, with four significant moderate or strong correlations compared to only one among G1 speakers. Several of these correlations are structurally connected; for example, as vowel space (CHA) shrinks due to increased centralization, rhythm values as shown by nPVI will also show a more stress-timed pattern, as predicted in Ronquest et al. (2020). Likewise, CHA also correlates with centralized FPs, again indicating the systematicity of bilingualism, as a change in one variable has impacts throughout the linguistic system(s). Interestingly, the number of speakers favoring bilingual forms for each of these variables (15 for nPVI, 14 for FPs, 9 for CHA) suggests that changes in prosodic rhythm and FPs act as a gateway and precede overall changes in vowel space, although we would hypothesize that once these processes begin, they likely feed off of one another.

The correlation between SPE and code switching shows how these two variables are integrated in tandem, as they both lie at the morpho-syntactic/pragmatic interface that has been shown to be particularly susceptible to cross-linguistic influence (Sorace 2004). Other connections are less obvious, such as the significant correlation between English DMs and CHA, while the correlation between DMs and FPs, significant among G1 speakers, has lost strength among G2 participants. An examination of bilingual DM and FP rates across speakers and generation provides an explanation. As seen in Figure 6, among G1 speakers, English DMs only appear among the speakers with the highest rates of bilingual FPs, thereby producing a significant correlation. Among G2 speakers, however, English DMs are present for all but one speaker, regardless of their level of bilingual FPs, and the rate of English DMs does not rise in tandem with rates of bilingual FPs, thereby weakening the correlation. As observed in Section 4.3, inter-speaker variation is an important factor in FPs, and both generations exhibit vastly different rates across speakers, although to a greater extent among G1.
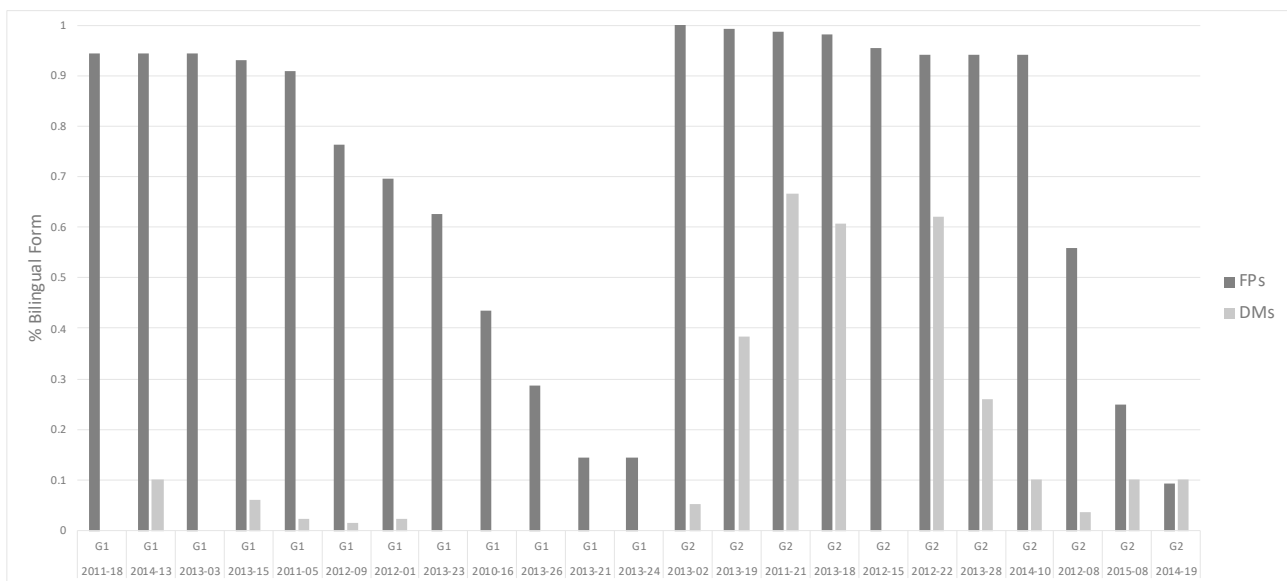
Figure 6. Rates of contact-induced FPs and DMs by generation and speaker.

In sum, connections showing how contact variables are not integrated in isolation only become apparent through a swarm approach. Furthermore, the swarm analysis and correlations between variables offer further insight into when, how, and why contact-induced forms are integrated into the bilingual system. The swarm approach combines a series of traditional, second-wave style studies on individual variables and applies a third-wave style focus on the behavior of individual speakers and variables as part of a larger, emerging bilingual system. In this way, a swarm analysis not only is able to provide a detailed panoramic picture of how variables pattern and interact within and across speakers, but is also crucial for identifying order of integration of features, which can serve as a springboard for future, more "traditional" third-wave studies.

6. Conclusions

By expanding previous research on a variable swarm in Spanish in NC (Ronquest et al. 2020), the present study allows for a more fine-grained approach to understanding language variation and the formation of bilingual norms in a newly forming community. The most important insights from the present study, including patterns of variable use across speakers and the correlations between variables that allow for a proposed order of integration of bilingual forms, would not have been possible without a swarm analysis. Additionally, the observation that English DMs and code switching—the most overt and salient strategies of English integration—are adopted later than phonetic traits, such as prosodic rhythm and vowel space modifications, speaks to the agency of individual speakers to use or not use particular variables not only as a means to index their identities, but also as a part of the process of active identity construction (Mendoza-Denton 2010). It is precisely the salience of these variables that makes them available for conscious manipulation (Erker 2017), and speakers who actively choose to employ stigmatized variants (such as code switches) in their Spanish may do so to set themselves apart from less integrated speakers in the community (see Zentella 1997). In this way, what is perceived as heightened "English influence" may actually be the actively created reflection of a bilingual/bicultural identity. The connection between contact forms and the indexation of identity was explicitly noted by Michnowicz et al. (2018) for English loanwords in NC Spanish, and the inter- and intra-speaker variation observed in the present data suggests that a complex interplay of social networks, personal experiences with English and Spanish, and perceived notions of prestige may be more important than sociolinguistic generation, at least for speakers not on the highest/lowest ends of adoption of or resistance to contact forms (see Figure 4). Social networks in particular have been found to be pivotal in the adoption of new linguistic forms

(O'Rourke and Potowski 2016; Dodsworth and Benton 2017; Carter and Lynch 2015; Carter 2007; Michnowicz et al. 2023a), and future research should focus on mapping speakers' social networks and personal and community motivations with a goal of understanding these patterns of change. While the present data cannot provide a concrete answer to the important question of speaker motivation, the insights afforded by the swarm analysis can provide researchers with a roadmap to identify areas for more fine-grained research in the future. Additionally, future research should include more in-depth analyses of the role of individual speaker choices in the formation of a new bilingual community by utilizing the panoramic analysis provided by the variable swarm as a foundation/indicator of where those choices are likely to be most meaningful.

Appendix A

Table A1. Participant demographics.

Participant	Sex	Birth Year	Generation	Region	Years in U.S.
2010-16	F	1957	G1	CAm	25
2014-13	M	1985	G1	CAm	19
2012-01	F	1957	G1	Mex	33
2013-21	F	1983	G1	Mex	1
2011-05	F	1988	G1	Mex	10
2013-03	M	1993	G1	Mex	10
2012-09	F	1992	G1	Mex	9
2013-15	M	1962	G1	Mex	35
2013-24	M	1984	G1	Mex	2
2013-23	M	1981	G1	Mex	2
2011-18	M	1992	G1	Mex	3
2013-26	M	1982	G1	Mex	1
2015-08	F	1992	G2	Mex	20
2011-21	F	1989	G2	CAm	Since birth
2014-10	F	1992	G2	CAm	Since birth
2012-15	F	1988	G2	CAm	Since birth
2012-22	F	1992	G2	CAm	Since birth
2012-08	F	1989	G2	CAm	Since birth
2013-19	M	1989	G2	Mex	Since birth
2013-18	F	1993	G2	Mex	Since birth
2014-19	F	1992	G2	Mex	Since birth
2013-28	M	1991	G2	Mex	Since birth
2013-02	M	1992	G2	Mex	Since birth

Appendix B

Table A2. Mexican G1. Results of the Multivariate One-Level Mixed-Effects Regression Model, speaker as a random factor. * Significant factors with a *p*-value < 0.05.

Factor	Log Odds	N	% Overt Pro	*p*-Value
Person/Number and Definiteness				<0.001 *
él-ella.definite	3.659	92	38.0	
yo	2.971	502	24.7	
ellos-ellas.definite	2.594	50	20.0	
tú.definite	1.860	30	13.3	
nosotros-nosotras	1.773	157	10.2	
tú.indefinite	0.883	104	3.8	
ellos-ellas.indefinite	−13.740	46	0.0	
Switch Reference				<0.001 *
Complete switch	0.693	421	24.5	
No switch	−0.210	480	16.0	
Partial switch	−0.484	80	16.2	
Lexical Content				0.0047 *
Estimative	0.916	46	52.2	
Stative	−0.164	221	23.1	
External activity	−0.358	555	16.2	
Mental activity	−0.393	159	17.6	
Distinctiveness of TAM				0.013 *
Non-distinctive	0.27	176	26.7	
Distinctive	−0.27	805	18.1	
Reflexivity				0.0185 *
Non-reflexive	0.36	865	20.7	
Reflexive	−0.36	116	12.1	

n	df	intercept	overall proportion	AIC	R2.fixed	R2.random	R2.total
981	15	−4.187	0.197	859.63	0.781	0.012	0.793

Table A3. Central American G1. Results of the Multivariate One-Level Mixed-Effects Regression Model, speaker as a random factor. * Significant factors with a *p*-value < 0.05.

Factor	Log Odds	N	% Overt Pro	*p*-Value
Person/Number and Definiteness				<0.001 *
él-ella.definite	8.721	25	56.0	
yo	7.637	107	20.6	
ellos-ellas.definite	6.251	21	9.5	
nosotros-nosotras	5.868	25	12	
tú.definite	−8.910	3	0.0	
ellos-ellas.indefinite	−9.191	10	0.0	
tú.indefinite	−10.376	1	0.0	
Switch Reference				<0.001 *
Partial switch	0.613	17	35.3	
Complete switch	0.567	87	28.7	
No switch	−1.180	88	11.4	
Lexical Content				0.009 *
Stative	5.187	62	37.1	
External activity	4.158	98	16.3	
Mental activity	3.310	28	7.1	
Estimative	−12.655	4	0.0	
Distinctiveness of TAM				0.849
Non-distinctive	0.049	53	26.4	
Distinctive	−0.049	139	119.4	
Reflexivity				0.144
Non-reflexive	0.463	158	23.4	
Reflexive	−0.463	34	11.8	

n	df	intercept	overall proportion	AIC	R2.fixed	R2.random	R2.total
192	15	−13.359	0.214	173.926	0.88	0.0	0.88

Table A4. Mexican G2. Results of the Multivariate One-Level Mixed-Effects Regression Model, speaker as a random factor. * Significant factors with a *p*-value < 0.05.

Factor	Log Odds	N	% Overt Pro	*p*-Value
Person/Number and Definiteness				<0.001 *
él-ella.definite	5.596	13	15.4	
yo	5.137	396	17.9	
ellos-ellas.definite	5.076	60	15.0	
ellos-ellas.indefinite	4.999	12	8.3	
tú.indefinite	3.910	41	4.9	
tú.definite	−11.768	24	0.0	
nosotros-nosotras	−12.951	52	0.0	
Distinctiveness of TAM				<0.001 *
Non-distinctive	0.564	106	20.8	
Distinctive	−0.564	492	12.8	
Lexical Content				0.00144 *
Estimative	1.269	34	50.0	
Mental activity	−0.247	125	15.2	
External activity	−0.348	310	10.6	
Stative	−0.674	129	12.4	
Reflexivity				0.00163 *
Non-reflexive	1.151	535	15.7	
Reflexive	−1.151	63	1.6	
Switch Reference				0.0431 *
Complete switch	0.646	236	18.2	
No switch	0.101	315	12.7	
Partial switch	−0.747	47	4.3	

n	df	intercept	overall proportion	AIC	R2.fixed	R2.random	R2.total
598	15	−7.662	0.142	420.84	0.905	0.013	0.918

Table A5. Central American G2. Results of the Multivariate One-Level Mixed-Effects Regression Model, speaker as a random factor. * Significant factors with a *p*-value < 0.05.

Factor	Log Odds	N	% Overt Pro	*p*-Value
Person/Number and Definiteness				<0.001 *
tú.definite	10.809	1	100.0	
tú.indefinite	2.161	5	80.0	
él-ella.definite	−0.570	137	37.2	
ellos-ellas.definite	−2.268	78	21.8	
yo	−2.424	209	21.5	
ellos-ellas.indefinite	−3.632	6	16.7	
nosotros-nosotras	−4.077	58	10.3	
Switch Reference				<0.001 *
Complete switch	0.885	201	34.8	
Partial switch	−0.127	37	32.4	
No switch	−0.757	256	16.8	
Lexical Content				0.481
External activity	0.404	266	28.2	
Stative	0.162	150	22.0	
Mental activity	0.053	63	23.8	
Estimative	−0.619	15	13.3	
Distinctiveness of TAM				0.713
Non-distinctive	0.05	140	29.3	
Distinctive	−0.05	354	23.7	
Reflexivity				0.428
Non-reflexive	0.153	439	25.7	
Reflexive	−0.153	55	21.8	

n	df	intercept	overall proportion	AIC	R2.fixed	R2.random	R2.total
494	15	0.308	0.253	485.02	0.322	0.133	0.455

Notes

1. Note that many lexical FPs are also analyzable as discourse markers.
2. It should be noted that in our own data, the majority of the auditorily labiodental tokens were also approximants rather than fricatives. Thus, the presence of the labiodental approximant [ʋ] complicates the classification of potential contact forms, as a bilingual speaker could produce a hybrid form that utilizes the labiodental point of articulation from English alongside the approximant mode of articulation from Spanish. In this way, bilingual labiodentals may not correspond exactly with English fricative /v/.
3. Technical aspects regarding *FASE* are described in greater depth in Wilbanks (2015) and Ronquest et al. (2020). Regarding the reliability of automatic alignment with *FASE*, Wilbanks (2015) compares *FASE* alignment to the alignment produced by trained human phoneticians. He finds that the refined, adapted *FASE* segmentation was similar to human segmentation: boundary differences between the two human coders had a mean of 14.47 ms, compared to a mean difference of 20.81 ms between the human coders and the trained *FASE* model.
4. For example, the full model AIC was compared to the AIC of a model with all variables except one (e.g., Switch Reference). The difference in AIC values indicates the importance of the variable in the full model, with larger differences denoting a stronger effect on the model. Kapatsinski (2012) demonstrates why this method is superior to the range of coefficients, the traditional method of determining the constraint hierarchy in sociolinguistic studies, which can be biased towards variables/factors with more levels.
5. An acoustic analysis of phonological filled pauses, as well as an examination of lexical pauses (e.g., *sea*, *este*) and silent pauses is forthcoming.
6. Data are frequently skewed to the right when there is a lower boundary to the measurement; in this case, COG cannot be a negative number. See https://www.itl.nist.gov/div898/handbook/eda/section3/histogr6.htm, accessed on 4 March 2023.
7. There are only two Central American G1 speakers in the present corpus, and some of these variables may achieve significance in a larger sample. Still, the match between Mexican and Central American G1 speakers speaks to the robustness of these constraints across Spanish varieties, even when few speakers are analyzed.
8. However, in some contexts, bilinguals have been found to produce fewer FPs than monolinguals but higher rates of silent pauses, which were not studied here (García-Amaya 2022). Additionally, although not coded in our data, impressionistically some monolinguals may have compensated for lower rates of phonological FPs by using more lexical FPs (*pues*, *este*, etc.). Further study is warranted and underway.
9. A more detailed analysis of <b> and <v> in NC Spanish is forthcoming.
10. http://www.sthda.com/english/wiki/correlation-matrix-an-r-function-to-do-all-you-need, accessed on 4 March 2023.

References

Abreu, Laurel. 2012. Subject pronoun expression and priming effects among bilingual speakers of Puerto Rican Spanish. In *Selected Proceedings of the 14th Hispanic Linguistics Symposium*. Edited by Kimberly Geeslin and Manuel Díaz-Campos. Somerville: Cascadilla Proceedings Project, pp. 1–8.

Alvord, Scott, and Brandon Rogers. 2014. Miami-Cuban Spanish vowels in contact. *Sociolinguistic Studies* 8: 139–70. [CrossRef]

Anderson, Tyker Kimball, and Almeida Jacqueline Toribio. 2007. Attitudes towards lexical borrowing and intra-sentential code-switching among Spanish-English bilinguals. *Spanish in Context* 4: 217–40. [CrossRef]

Aaron, Jessi Elana. 2004. The Gendered Use of salirse in Mexican Spanish: Si me salía yo con las amigas, se enojaba. *Language in Society* 33: 585–607. [CrossRef]

Bayley, Robert, and Lucinda Pease-Alvarez. 1997. Null pronoun variation in Mexican-descent children's narrative discourse. *Language Variation and Change* 9: 349–71. [CrossRef]

Bayley, Robert, Norma L. Cárdenas, Belinda Treviño Schouten, and Carlos Martin Vélez Salas. 2012. Spanish dialect contact in San Antonio, Texas: An exploratory study. In *Selected Proceedings of the 14th Hispanic Linguistics Symposium*. Edited by Kimberly Geeslin and Manuel Díaz-Campos. Somerville: Cascadilla Proceedings Project, pp. 48–60.

Boersma, Paul, and David Weenink. 2022. *Praat: Doing Phonetics by Computer [Computer Program]*, Version 6.3.02. Available online: https://www.praat.org/ (accessed on 29 November 2022).

Boomershine, Amanda, and Rebecca Ronquest. 2019. Teaching pronunciation to Spanish heritage speakers. In *Key Issues in the Teaching of Spanish Pronunciation*. Edited by Rajiv Rao. London: Routledge, pp. 288–303.

Britain, David, and Peter Trudgill. 1999. Migration, new-dialect formation and sociolinguistic refunctionalisation: Reallocation as an outcome of dialect contact. *Transactions of the Philological Society* 97: 245–56. [CrossRef]

Carolina Demography. 2021. North Carolina's Hispanic Community: Snapshot (2021). Available online: https://www.ncdemography.org/ (accessed on 24 February 2023).

Carter, Phillip M. 2005. Prosodic variation in SLA: Rhythm in an urban North Carolina Hispanic community. *University of Pennsylvania Working Papers in Linguistics* 11: 59–71.

Carter, Phillip M. 2007. Phonetic Variation and Speaker Agency: Mexicana Identity in a North Carolina Middle School. *Penn Working Papers in Linguistics* 13: 1–15.

Carter, Phillip M. 2013. Shared spaces, shared structures: Latino social formation and African American English in the U.S. South. *Journal of Sociolinguistics* 17: 66–92. [CrossRef]

Carter, Phillip M., and Andrew Lynch. 2015. Multilingual Miami: Current trends in sociolinguistic research. *Language and Linguistics Compass* 9: 369–85. [CrossRef]

Carter, Phillip M., and Tonya Wolford. 2016. Cross-generational prosodic convergence in South Texas Spanish. *Spanish in Context* 13: 29–52. [CrossRef]

Chetty, Sarah. 2018. To /b/ or not to /b/: On the production of the graphemes <bv> in heritage Spanish [Poster presentation]. In *The 2018 Graduate Research Symposium*. Raleigh: NC State University. Available online: https://grad.ncsu.edu/wp-content/uploads/2018/04/sarahchettyposter.pdf (accessed on 24 February 2023).

Chládková, Kateřina, Paola Escudero, and Paul Boersma. 2011. Context-specific acoustic differences between Peruvian and Iberian Spanish vowels. *JASA* 130: 416–28. [CrossRef]

Cobb, Katherine, and Miquel Simonet. 2015. Adult Second Language Learning of Spanish Vowels. *Hispania* 98: 47–60. [CrossRef]

Colantoni , Laura, and Irina I. Marinescu. 2010. The scope of stop weakening in Argentine Spanish. In *Selected Proceedings of the 4th Conference on Laboratory Approaches to Spanish Phonology*. Edited by Marta Ortega-Llebaria. Somerville: Cascadilla Proceedings Project, pp. 100–14.

Coloma, Germán. 2012. *The Importance of Ten Phonetic Characteristics to Define Dialect Areas in Spanish*. Dialectologia: Revista Electrònica, pp. 1–26.

Dodsworth, Robin, and Richard A. Benton. 2017. Social network cohesion and the retreat from Southern vowels in Raleigh. *Language in Society* 46: 371–405. [CrossRef]

Drager, Katie, and Jennifer Hay. 2012. Exploiting random intercepts: Two case studies in sociophonetics. *Language Variation and Change* 24: 59–78. [CrossRef]

Eckert, Penelope. 2005. *Variation, Convention, and Social Meaning [Plenary Address]*. Oakland: The Annual Meeting of the Linguistic Society of America. Available online: http://www.justinecassell.com/discourse09/readings/EckertLSA2005.pdf (accessed on 19 May 2023).

Eckert, Penelope. 2012. Three waves of variation study: The emergence of meaning in the study of sociolinguistic variation. *Annual Review of Anthropology* 41: 87–100. [CrossRef]

Erker, Daniel. 2017. The limits of named language varieties and the role of social salience in dialectal contact: The case of Spanish in the United States. *Language and Linguistics Compass* 11: 1–20. [CrossRef]

Erker, Daniel, and Joana Bruso. 2017. Uh, bueno, em . . . : Filled pauses as a site of contact-induced change in Boston Spanish. *Language Variation and Change* 29: 205–44. [CrossRef]

Escobar, Anna María, and Kim Potowski. 2015. *El español de los Estados Unidos*. Cambridge: Cambridge University Press.

File-Muriel, Richard J., and Eearl K. Brown. 2011. The gradient nature of s-lenition in Caleño Spanish. *Language Variation and Change* 23: 223–43. [CrossRef]

García-Amaya, Lorenzo. 2009. New findings on fluency measures across three different learning contexts. In *Selected Proceedings of the 11th Hispanic Linguistics Symposium*. Edited by Joseph Collentine, Maryellen García, Barbara Lafford and Francisco Marcos Marín. Somerville: Cascadilla Press, pp. 68–80.

García-Amaya, Lorenzo. 2022. An investigation into utterance-fluency patterns of advanced LL bilinguals: Afrikaans and Spanish in Patagonia. *Linguistic Approaches to Bilingualism* 12: 163–90. [CrossRef]

Grabe, Esther, and Ee Ling Low. 2002. Durational variability in speech and the rhythm class hypothesis. *Papers in Laboratory Phonology* 7: 515–46.

Hernández, José Esteban. 2002. *Accommodation in a Dialect Contact Situation*. San José: Revista de filología y lingüística de la Universidad de Costa Rica, pp. 93–110.

Howe, Chad, and Philip Limerick. 2020. Perceptions of Spanish among Spanish Heritage Speakers in the Southeastern United States through Computer-Mediated Communication. In *Spanish in the United States and across Domains*. Edited by Francisco Salgado-Robles and Edwin Lamboy. Leiden: Brill, pp. 364–87.

Hualde, José Ignacio. 2005. *The Sounds of Spanish*. Cambridge: Cambridge University Press.

Hualde, José Ignacio, Miquel Simonet, and Marianna Nadeu. 2011. Consonant lenition and phonological recategorization. *Laboratory Phonology* 2: 301–29. [CrossRef]

Johnson, Daniel Ezra. 2009. Getting off the GoldVarb standard: Introducing Rbrul for mixed-effects variable rule analysis. *Language and Linguistics Compass* 3: 359–83. [CrossRef]

Kapatsinski, Vsevolod. 2012. *Towards a de-Ranged Study of Variation [Conference Presentation]*. Washington, DC: Georgetown University Round Table on Linguistics, Measured Language: Quantitative Approaches to Acquisition, Assessment, Processing, and Variation. Available online: http://pages.uoregon.edu/vkapatsi/DerangedPaper.pdf (accessed on 24 February 2023).

Kerswill, Paul. 2013. Koineization. In *The Handbook of Language Variation and Change*. Edited by J. K. Chambers and Natalie Schilling. Oxford: John Wiley & Sons.

Knightly, Leah, SunAh Jun, Janet Oh, and Terry Kit-fong Au. 2003. Production benefits of childhood overhearing. *JASA* 114: 465–74. [CrossRef] [PubMed]

Knouse, Stephanie M., Renee Neves, Erk Ortiz, and Daria Acosta-Rua. 2022. "Le está haciendo un disservice": Overt Attitudes toward Language Contact Phenomena in the Upstate of South Carolina. *Hispania* 105: 173–94. [CrossRef]

Labov, William. 1963. The social motivation of a sound change. *Word* 19: 273–309. [CrossRef]

Labov, William, Ingrid Rosenfelder, and Josef Fruehwald. 2013. One hundred years of sound change in Philadelphia: Linear incrementation, reversal, and reanalysis. *Language* 89: 30–65. [CrossRef]

Ladefoged, Peter, and Sandra Disner. 2012. *Vowels and Consonants*. Chichester: John Wiley and Sons.

Limerick, Pilip P. 2019. The discursive distribution of subject pronouns in Spanish spoken in Georgia: A weakening of pragmatic constraints? *Studies in Hispanic and Lusophone Linguistics* 12: 97–126. [CrossRef]

Limerick, Philip P. 2021. First-person plural subject pronoun expression in Mexican Spanish spoken in Georgia. *Studies in Hispanic and Lusophone Linguistics* 14: 411–32. [CrossRef]

Lipski, John M. 1994. *Latin American Spanish*. London: Longman.

Lipski, John M. 2005. Code-switching or borrowing? No sé so no puedo decir, you know. In *Selected Proceedings of the Second Workshop on Spanish Sociolinguistics*. Edited by Lotfi Sayahi and Maurice Westmoreland. Somerville: Cascadilla Proceedings Project, pp. 1–15.

Lipski, John M. 2008. *Varieties of Spanish in the United States*. Washington, DC: Georgetown University Press.

Lipski, John M. 2014. Spanish-English code-switching among low-fluency bilinguals: Towards an expanded typology. *Sociolinguistic Studies* 8: 23. [CrossRef]

Lipski, John M. 2020. Equatorial Guinea Spanish non-continuant /d/: More than a generic L2 trait. In *Spanish Phonetics and Phonology in Contact: Studies from Africa, the Americas, and Spain*. Edited by Rajiv Rao. Amsterdam: John Benjamins, pp. 13–32.

Low, Ee-Ling, and Esther Grabe. 1995. Prosodic patterns in Singapore English. In *Proceedings of the International Congress of Phonetic Sciences, Stockholm*. Edited by Kjell Elenius and Peter Branderud. Stockholm: KTH and Stockholm University Press, pp. 636–39.

Mariano, Paolo. 2014. Correlatore [Computer Program]. Available online: http://www.lfsag.unito.it/correlatore/readme_en.html (accessed on 14 September 2022).

Mata, Rodolfo. 2022. Bilingualism is good but codeswitching is bad: Attitudes about Spanish in contact with English in the Tijuana-San Diego border area. *Critical Inquiry in Language Studies* 1–22. [CrossRef]

Mendoza-Denton, Norma. 2002. Language and Identity. In *The Handbook of Language Variation and Change*. Edited by J. K. Chambers, Peter Trudgill and Natalie Schilling-Estes. Oxford: Blackwell, pp. 475–99.

Mendoza-Denton, Norma. 2010. Individuals and communities. In *The Sage Handbook of Sociolinguistics*. Edited by Ruth B. Wodak, Barbara Johnstone and Peter Kerswill. Los Angeles: Sage Publications, pp. 181–91.

Menke, Mandy R., and Timothy L. Face. 2010. Second language Spanish vowel production: An acoustic analysis. *Studies in Hispanic and Lusophone Linguistics* 3: 181–214. [CrossRef]

Michnowicz, Jim. 2015. Subject Pronoun Expression in Contact with Maya in Yucatan Spanish. In *Subject Pronoun Expression in Spanish: A Cross-Dialectal Perspective*. Edited by Ana Carvalho, Rafael Orozco and Naomi Lapidus Shin. Washington, DC: Georgetown University Press, pp. 103–22.

Michnowicz, Jim, Alex Hyler, James Shepherd, and Sonya Trawick. 2018. Spanish in North Carolina: English-origin loanwords in a newly forming Hispanic community. In *Language Diversity in the New South*. Edited by Jeffrey Reaser, Eric Wilbanks, Walt Wolfram and Karissa Wojcik. Chapel Hill: University of North Carolina Press, pp. 289–305.

Michnowicz, Jim, Rebecca Ronquest, Bailey Ambrister, Hannah Bain, Nick Chisholm, Rebecca Greene, Lindsey Bull, and Anne Elkins. 2023a. Perceptions of Inclusive Language in the Spanish of the Southeast: Data from a Large Classroom Project. *Spanish in Context*. Available online: https://www.jbe-platform.com/content/journals/10.1075/sic.00084.mic (accessed on 6 July 2023).

Michnowicz, Jim, Sonya Trawick, and Ronquest Ronquest. 2023b. Spanish language maintenance and shift in a newly-forming community in the southeastern United States: Insights from a large-class survey. *Hispanic Studies Review* 7. Available online: https://hispanicstudiesreview.cofc.edu/article/77601-spanish-language-maintenance-and-shift-in-a-newly-forming-community-in-the-southeastern-united-states-insights-from-a-large-class-survey (accessed on 7 July 2023).

Montes-Alcalá, Cecilia, and Lindsey Sweetnich. 2014. Español en el sureste de EE. UU.: El papel de las actitudes lingüísticas en el mantenimiento o pérdida de la lengua. *Revista Internacional de Lingüística Iberoamericana* 23: 77–92. [CrossRef]

Myers-Scotton, Carol. 1995. *Social Motivations for Codeswitching: Evidence from Africa*. Oxford: Oxford University Press.

Navarro Tomás, Tomás. 1918. *Manual de Pronunciación Española*, 12th ed. Madrid: CSIC.

O'Rourke, Erin, and Kim Potowski. 2016. Phonetic accommodation in a situation of Spanish dialect contact: Coda $/s/$ and $/\bar{r}/$ in Chicago. *Studies in Hispanic and Lusophone Linguistics* 9: 355–99. [CrossRef]

Otheguy, Ricardo, Ana Celia Zentella, and David Livert. 2007. Language and dialect contact in Spanish in New York: Toward the formation of a speech community. *Language* 83: 770–802. [CrossRef]

Otheguy, Ricardo, and Ana Celia Zentella. 2012. *Spanish in New York: Language Contact, Dialectal Leveling, and Structural Continuity*. Oxford: Oxford University Press.

Pew Research Center. 2014. Demographic Profile of Hispanics in North Carolina. Available online: http://www.pewhispanic.org/states/state/nc (accessed on 2 October 2022).

Poplack, Shana. 1978. Dialect acquisition among Puerto Rican bilinguals. *Language in Society* 7: 89–103. [CrossRef]

Poplack, Shana. 1980. Sometimes I'll start a sentence in Spanish y termino en español: Toward a typology of code-switching. *Linguistics* 18: 581–618. [CrossRef]

Poplack, Shana. 1988. Contrasting patterns of code-switching in two communities. *Codeswitching: Anthropological and sociolinguistic Perspectives* 48: 215–44.

Quilis, Antonio, and Manuel Esgueva. 1983. Realización de los fonemas vocálicos españoles en posición fonética normal. In *Estudios de fonética*. Edited by Manuel Esgueva and Margarita Cantarero. Madrid: Consejo Superior de Investigaciones Científicas, pp. 159–251.

Rangel, Natalie, Verónica Loureiro-Rodríguez, and María Irebe Moyna. 2015. "Is that what I sound like when I speak?": Attitudes towards Spanish, English, and code-switching in two Texas border towns. *Spanish in Context* 12: 177–98. [CrossRef]

Rao, Rajiv. 2014. On the status of the phoneme /b/ in heritage speakers of Spanish. *Sintagma* 26: 37–54.

Rao, Rajiv. 2015. Manifestations of /bdg/ in heritage speakers of Spanish. *Heritage Language Journal* 12: 48–74. [CrossRef]

Rao, Rajiv, and Rebecca Ronquest. 2015. The heritage Spanish phonetic/phonological system: Looking back and moving forward. *Studies in Hispanic and Lusophone Linguistics* 8: 403–14. [CrossRef]

R Core Team. 2022. *R: A Language and Environment for Statistical Computing*. Vienna: R Foundation for Statistical Computing. Available online: https://www.R-project.org/ (accessed on 3 October 2022).

Ronquest, Rebecca. 2012. An Acoustic Analysis of Heritage Spanish Vowels. Ph.D. thesis, Indiana University, Bloomington, Indiana.

Ronquest, Rebecca, Jim Michnowicz, Eric Wilbanks, and Cortés Claudia. 2020. Examining the (mini-)variable swarm in the Spanish of the Southeast. In *Hispanic Linguistics: Current Issues and New Directions*. Edited by Alfonso Morales-Front, Michael Ferreira, Ronald Leow and Cristina Sanz. Amsterdam: John Benjamins, pp. 303–25.

Shin, Naomi Lapidus. 2013. Women as leaders of language change: A qualification from the bilingual perspective. In *Selected Proceedings of the 6th Workshop on Spanish Sociolinguistics*. Edited by Ana M. Carvalho and Sara Beaudrie. Somerville: Cascadilla Proceedings Project, pp. 135–47.

Shin, Naomi Lapidus, and Ricardo Otheguy. 2009. Shifting sensitivity to continuity of reference: Subject pronoun use in Spanish. In *Español en Estados Unidos y en otros contextos: Cuestiones sociolingüísticas, políticas, y pedagógicas*. Edited by Manel Lacorte and Jennifer Leeman. Madrid: Iberoamericana/Vervuert Verlag, pp. 111–36.

Shin, Naomi Lapidus, and Ricardo Otheguy. 2013. Social class and gender impacting change in bilingual settings: Spanish subject pronoun use in New York. *Language in Society* 42: 429–52. [CrossRef]

Shousterman, Cara. 2014. Speaking English in Spanish Harlem: The role of rhythm. *University of Pennsylvania Working Papers in Linguistics* 20: 18.

Silva-Corvalán, Carmen. 1994. The gradual loss of mood distinctions in Los Angeles Spanish. *Language Variation and Change* 6: 255–72. [CrossRef]

Silva-Corvalán, Carmen. 2008. The limits of convergence in language contact. *Journal of Language Contact* 2: 213–24. [CrossRef]

Sorace, Antonella. 2004. Native language attrition and developmental instability at the syntax-discourse interface: Data, interpretations and methods. *Bilingualism Language and Cognition* 7: 143–5. [CrossRef]

Sorace, Antonella. 2005. Selective Optionality in Language Development. In *Syntax and Variation: Reconciling the Biological and the Social*. Edited by Leonie Cornips and Karen Corrigan. Amsterdam: John Benjamins, pp. 55–80.

Tagliamonte, Sali A. 2011. *Variationist Sociolinguistics: Change, Observation, Interpretation*. Hoboken: John Wiley and Sons.

Thomas, Erik. 2015. What a swarm of variables tells us about the formation of Mexican American English [Conference presentation]. Paper presented at LAVIS IV, Raleigh, NC, USA, August 28. Available online: https://www.academia.edu/15228944/What_a_Swarm_of_Variables_Tells_Us_about_the_Formation_of_Mexican_American_English (accessed on 1 November 2022).

Toribio, Almeida Jacqueline. 2004. Spanish/English speech practices: Bringing chaos to order. *International Journal of Bilingual Education and Bilingualism* 7: 133–54. [CrossRef]

Toribio, Almeida Jacqueline. 2011. Code-Switching among US Latinos. In *The Handbook of Hispanic Sociolinguistics*. Hoboken: Wiley, pp. 530–52.

Torres, Lourdes. 2002. Bilingual discourse markers in Puerto Rican Spanish. *Language in Society* 31: 65–83. [CrossRef]

Torres, Lourdes. 2011. Spanish in the United States: Bilingual discourse markers. In *The Handbook of Hispanic Sociolinguistics*. Edited by Manuel Díaz-Campos. Malden: Wiley-Blackwell, pp. 491–503.

Torres, Lourdes, and Kim Potowski. 2008. A comparative study of bilingual discourse markers in Chicago Mexican, Puerto Rican, and MexiRican Spanish. *International Journal of Bilingualism* 12: 263–79. [CrossRef]

Trovato, Adriano. 2017. A Sociophonetic Analysis of Contact Spanish in the United States: Labiodentalization and Labial Consonant Variation. Ph.D. thesis, University of Texas, Austin, TX, USA.

Trudgill, Peter. 1986. *Dialects in Contact*. Oxford: Blackwell.

U.S. Census Bureau. 2020. Hispanic or Latino Origin by Specific Origin. 2016–20 American Community Survey 5-Year Estimates. Available online: https://data.census.gov/cedsci/table?q=Latino%20Origin%20NC&tid=ACSDT5Y2020.B03001 (accessed on 31 October 2022).

Wei, Taiyun, and Viliam Simko. 2021. R Package 'Corrplot': Visualization of a Correlation Matrix (Version 0.92). Available online: https://github.com/taiyun/corrplot (accessed on 5 November 2022).

Wilbanks, Eric. 2015. The Development of FASE (Forced Alignment System for Español) and Implications for socioli guistic research. Paper presented at New Ways of Analyzing Variation 44, Toronto, ON, Canada, October 22–25. Available online: https://ericwilbanks.github.io/files/wilbanks_nwav_2015.pdf (accessed on 20 May 2023).

Willis, Erik. 2005. An Initial Examination of Southwest Spanish Vowels. *Southwest Journal of Lingusitics* 24: 185–98.

Wolfram, Walt, Mary Kohn, and Erin Callahan-Price. 2011. Southern-Bred Hispanic English: An Emerging Socioethnic Variety. In *Selected Proceedings of the 5th Workshop on Spanish Sociolinguistics*. Edited by Jim Michnowicz and Robin Dodsworth. Somerville: Cascadilla Proceedings Project, pp. 1–13.

Wolfram, Walt, Phillip Carter, and Beckie Morello. 2004. Emerging Hispanic English: New dialect formation in the American South. *Journal of Sociolinguistics* 8: 339–58. [CrossRef]

Woolard, Kathryn A. 1999. Strategies of simultaneity and bivalency in bilingual communication. *Journal of Linguistic Anthropology* 8: 3–29. [CrossRef]

Zentella, Ana Celia. 1990. Lexical leveling in four New York City Spanish dialects: Linguistic and social factors. *Hispania* 73: 1094–105. [CrossRef]

Zentella, Ana Celia. 1997. *Growing up Bilingual: Puerto Rican Children in New York*. Malden: Blackwell Publishers.

Zúñiga, Victor, and Rubén Hernández-León, eds. 2005. *New Destinations: Mexican Immigration in the United States: Community Formation, Local Responses and Inter-Group Relations*. New York: Russell Sage Foundation.

MDPI

Article

Clothing, Gender, and Sociophonetic Perceptions of Mayan-Accented Spanish in Guatemala

Brandon Baird

Department of Luso-Hispanic Studies & Linguistics Program, Middlebury College, Middlebury, VT 05753, USA;
bbaird@middlebury.edu

Abstract: Perceptual sociophonetic work on Guatemalan Spanish has demonstrated that listeners are more likely to link male voices with traditional Maya clothing, the *traje típico*, when their speech includes features of Mayan-accented Spanish. However, as Maya women are more likely than men to wear the *traje típico*, this matched-guise study investigates native Guatemalans' perceptions of Mayan-accented Spanish produced by female voices. The results demonstrate that guises with features of Mayan-accented Spanish were more likely to have *traje típico* as a response than guises without these features. When compared to the previous studies with male-voiced guises, the findings suggest an interaction between gender and Mayan-accented Spanish. *Traje típico* responses were more common for female-voiced guises than male-voiced guises and occurred at the highest rate among female-voiced guises with features of Mayan-accented Spanish. Thus, gendered and cultural practices are reflected in the indexical fields of Mayan-accented Spanish in Guatemala, regardless of the gender or ethnicity of the listener. That is, the visual body–language link is significantly more essentialized for the identity of a woman than for the identity of a man in Guatemala, suggesting that gendered stereotypes, language ideologies, and embodied practices mutually reinforce one another in the collective consciousness of the region.

Keywords: Maya; Guatemalan Spanish; traje típico; identity; indexicality; gender; sociophonetics

1. Introduction

Within Eckert's (2012) third wave of sociolinguistics, single linguistic variables are seen as capable of indexing various social meanings beyond the variable itself. In other words, any unique linguistic variable may represent the "embodiment of ideology in linguistic form" (Eckert 2008, p. 464), as linguistic variation has been shown to index multiple facets beyond language: race, ethnicity, intelligence, social class, emotional state, among many others. This "constellation of ideologically related meanings" (Eckert 2008, p. 454), or indexical field, is constantly evolving; different extra-linguistic variables, e.g., cultural practices, gender roles, etc., influence what social meanings may be indexed by a linguistic variable among a particular group. Conversely, the same linguistic variable may have an entirely different indexical field in another context. In fact, recent studies on Spanish have shown how single linguistic variables demonstrate "local constructions of meaning in variation" (Eckert 2008, p. 454) (see edited volumes such as Chappell 2019 and Ortíz-López and Suárez Büdenbender 2021 for some recent examples).

Aside from linguistic variables, "material style", i.e., clothing and other types of adornment (Eckert 2008, p. 457), is also a prominent indexer of social meaning and identity among different groups. For example, in her study of high school girls in California, Eckert (1980) found that certain fashion choices indexed belonging to specific social groups and that some students deliberately chose to emulate some of those fashion choices in order to claim aspects of the other social group's identity as their own. Specifically, some female students that identified as preppy chose to copy the pants style of the 'new wavers' social group in order to be perceived as both preppy and independent, while they avoided other

new waver fashion styles that they deemed to index social meanings that they did not desire in the creation of their own identities. Mendoza-Denton's (2008, pp. 156–57) work among Latina youth gangs in California describes many deliberate choices of material style and what they index, such as specific hair dye and clothing colors for different gangs, and the use of long eyeliner as a marker of the intention and willingness to fight.

Of particular interest to this study is the relationship between different linguistic variables and material style, what Calder (2019, p. 332) calls "visual body-language links". For example, Mendoza-Denton's (2008) research not only demonstrates the indexing of gang membership via clothing, hair, and eyeliner, but also by the use of phonetic variables like the raising of /ɪ/ to [i] as in *sick* [sik] and the fortition of /θ/ to [t̪] as in *things* [t̪iŋz]. In Beijing, China, those that identify as part of a group with high levels of consumerism, known locally as 'smooth operators', index this identity both in terms of more elegant clothing and lifestyles and significantly higher rates of rhotacization, the addition of /ɻ/ to the end of a syllable, than non-smooth operators (Zhang 2008). Furthermore, the connection between linguistic features and material style can go hand-in-hand with the loss of a culture and its language. This is exemplified in the small town of Oberwart, Austria; young women who wish to present a more urban identity are not only beginning to speak more German than Hungarian, but they are also changing their ways of dressing and grooming to be more German-like, as both Hungarian clothing and speaking Hungarian index a peasant lifestyle (Eckert and McConnell-Ginet 2003, p. 309).

The present study analyzes the indexical relationship between clothing and linguistic features among a minority population in Guatemala: the indigenous Maya. As with most indigenous cultures and their languages in the Americas, the Maya have endured centuries of oppression and attempts at extermination by different regimes, as they were deemed socially inferior to the Spanish-speaking upper classes (Baird 2019). Nonetheless, the somewhat recent revitalization efforts, known locally as the 'Maya movement', have been successful in their attempts to at least slow down this cultural and linguistic loss. Today, about 42% of the total population in Guatemala self-identify as Maya (Instituto Nacional de Estadística Guatemala 2018). At the forefront of the Maya movement is the documentation and preservation of Mayan languages, as the ability to speak one of the approximately 21 Mayan languages in the country is generally considered to be the single most important aspect of Maya identity (England 2003). This movement has included government-level recognition of Mayan languages in both the 1996 Peace Accords and the 2003 Law of National Languages, although both fell short of granting any Mayan language official status in Guatemala (Baird 2019).

Aside from speaking a Mayan language, the use of traditional Mayan clothing is the second most important aspect of Maya identity, and the most important visual one (Baird 2019; French 2010; Nelson 1999, 2009; among many others). The hand-made clothing, commonly referred to as the *traje típico,* greatly contrasts from more western styles of clothing used by Ladinos (non-Mayas in Guatemala), and the elaborate hand-stitched and colorful patterns often hold significant meanings that vary from town to town. The female *traje típico* from the municipality of Cantel is illustrated and described in the article from the Guatemalan newspaper *El quetzalteco* seen in Figure 1. The most commonly worn items of the *traje típico* are *el huipil*, the blouse, and *el corte*, the skirt, which is not like western-style skirts with sewn-in waists; rather, a hand-woven piece of fabric is wrapped around the body and held in place by another piece of belt-like fabric, *la faja*. The other articles shown in Figure 1, the *cinta* 'headdress', the *chachal* or *collar* 'necklace', and the *perraje* 'shawl', are typically reserved for more formal use. In the Cantel *traje típico* in Figure 1, the *cinta* represents the feathered serpent that is present in many Mesoamerican cultures, the *perraje* represents the authority of the Maya woman, the *huipil* denotes the blood of their ancestors, and the *corte* signifies Mayan altars.

Figure 1. Colop (2013). Newspaper article *El traje típico es representativo de la cultura maya* 'The traje típico is representative of the Maya culture' (13 May 2013) *El quetzalteco, Prensa Libre.* © Hemeroteca Prensa Libre. Used with the permission of Prensa Libre.

Given the significance of the different articles of clothing that constitute this style of dress, the *traje típico* has received ample attention in the ethnographic literature and is considered to be truly iconic throughout the country (Collier 1997; Hendrickson 1995, 1996;

Nelson 1999, 2009; among many others). As such, there are many nuanced layers to the use of the *traje típico* among the Maya in Guatemala, more than can adequately be addressed here. For example, for many Ladinos the *traje típico* represents antiquated cultural practices and is a sign of what is impeding progress on a national level. As mentioned above, the *traje típico* varies from town to town; the newspaper article in Figure 1 is only one of a series of articles published by *El quetzalteco* that highlight different towns' *trajes*. However, the different practices of the Maya and the meanings of their distinct *trajes* are often erased in the Ladinos' conflation of them into a single Maya ethnicity and a single *traje típico* that is a symbol of what the Ladino wants it to represent: a mark of the poor Indian whose only value is financial exploitation among tourists (Collier 1997; Nelson 1999).

In spite of the Ladinos' discrimination towards them, the continued day-to-day use of the *traje típico* demonstrates a sense of pride and even resistance among the Maya. For example, the newspaper article in Figure 1 quotes the woman in the photograph as saying, "wearing the *traje* represents respect. It is a privilege, since it allows us to make the history, culture, and richness of our country known" (Colop 2013, p. 10). In other words, representing one's own Maya identity is a choice. Some argue that, as there are no longer many significant differences between the Maya and Ladinos in terms of phenotype, a Maya may pass as Ladino by speaking Spanish and using Ladino or westernstyle clothing (Baird 2019; Nelson 1999; von den Berghe 1968). Therefore, a Maya must consciously decide to display their Maya identity via language and clothing in public. In an investigation of language and cultural attitudes among bilinguals of Spanish and the Mayan language K'iche' (also spelled K'ichee'/Quiché) from the municipalities of Nahualá and Cantel, several participants saw the loss of language and the *traje típico* as the biggest examples of abandonment of the K'iche' culture, as exemplified in 1–3 (Baird 2019, pp. 329–30):

(1) *Ri kek'ix la' kech'aw pa qach'ab'al, chuq kek'ix la' kakikoj le qajastaq.*

'They [the youth] are ashamed of our language, and they are also ashamed to use our *traje típico*.'

(2) *Donde vayamos, debemos de representar nuestra cultura sin importar con quienes estemos, debemos de vestirnos tal como nuestros abuelos lo hacían sin avergonzarnos de nuestro traje, sin avergonzarnos de nuestro idioma.*

'Wherever we go, we should represent our culture no matter who we are with, we should dress as our grandparents dressed without shame of our traje, without shame of our language.'

(3) *In kink'oxomaj chi kek'ix la' chech kakikoj chi le qajastaq, ma k'or k'o na', le, la' le a'mu'sab' e señorayib' keyoq'on che le qajastaq y xa je la' jun kak'ixik. Par are k'u le chwech in, xaq si na kink'ix ta wi, jacha chech taq le qatzij.*

'I think that they [the youth] are embarrassed to use our *trajes*, because at times the Ladinos and Ladinas make fun of our *traje típico* and that's why they get embarrassed. But I never feel embarrassed [about our *traje*] or about our language.'

Nevertheless, the Maya have constant interactions with Ladinos and Ladino culture, which are almost always in Spanish. As such, the Maya have a complicated history with the Spanish language; it represents both the language of their colonizers and oppressors and the key to protection from discrimination from Ladinos and advancement within this same society. In Baird (2019), of the 162 K'iche'-Spanish bilinguals interviewed, not a single participant expressed negative attitudes towards the K'iche' language or culture, but their attitudes towards Spanish demonstrated significant variation. Very few bilinguals in this study wanted to be seen as part of the Spanish-speaking world or to even be considered a Spanish speaker. Those that did predominately remarked that it was okay to be seen as Spanish speakers as long as it was not at the expense of their Maya identity, as in (4) (Baird 2019, p. 328):

(4) *Nabe' rajawaxik kaketa' maj le qas qach'ab'al ma xa je ri' kaqaya unimal uq'ij le kich'ab'al ri e qati't qamam, Par chuq sib'alaj rajawaxik le kaxlantzij ma are wa' kutob'ej*

wi rib' ri mayab' winaq chikiwach le kaxlan taq winaq, ma we jun na reta'm taj sib'alaj kab'ananex chech rumal la' kinchomaj chi rajawaxik le keb' ch'ab'al kak'ut chikiwach le ak'alab' on e nimaq winaq.

'First it is necessary to learn our language in order to give importance to the language of our ancestors, but the Spanish language is very necessary for the Maya people in order to be able to defend themselves from Ladino people, if we do not learn it, we will have many pains and difficulties. Because of this, I think that it is necessary to teach both languages to children and adults.'

However, even if the Maya speak Spanish, they may still be subject to prejudice and discrimination by Ladinos and other Maya according to *how* they speak Spanish. Given the centuries of contact between Spanish and Mayan languages in Guatemala, the transfer of linguistic features between the languages is common (Baird forthcoming). Historically, the social perceptions of using Mayan linguistic features in Spanish, or Mayan-accented Spanish, were, unsurprisingly, unfavorable and prescriptive in nature; the 19th century Guatemalan philologist Andrés Batrés Jáuregui (1892, p. 10) described it as containing "innumerable vulgarisms that offend good taste at every step". However, there remain few studies that analyze the social perceptions of features of Mayan-accented Spanish in the context of Guatemala and how they may index stereotypical characteristics of Maya identity, such as wearing a *traje típico*.

The two features of Mayan-accented Spanish analyzed in this study are phonetic in nature. The first is the apocope, or deletion, of word-final unstressed vowels: /ˈkasa/ → [ˈkas] 'house'. In many Mayan languages, particularly those spoken in Guatemala, stress is fixed in word-final position (England and Baird 2017). As such, many early Spanish loan words in Mayan languages were phonologically adapted to follow this pattern, e.g., /esˈpexo/ → /esˈpeχ/ 'mirror' in K'iche' (Baird 2021a, p. 225). This process of apocope occurs in the Spanish of some Spanish–Mayan bilinguals, e.g., /seˈmana/ → [seˈman] 'week', /ˈsinko/ → [ˈsiŋk] 'five' (Baird 2015). The second feature is the fortition of the voiceless labiodental fricative /f/ to the voiceless bilabial stop [p]. As no Mayan language has /f/ (England and Baird 2017), it was adapted to /p/ in Spanish loan words in Mayan languages, e.g., /kaˈfe/ → /kaˈpe/ 'coffee' in K'iche' (Baird 2023), and /f/ fortition has been reported in Mayan-accented Spanish in Guatemala by some scholars (Aleza Izquierdo 2010; García Tesoro 2008; Lentzner 1893). Even so, the mentions of apocope and /f/ fortition in the previous literature are very brief. The lone quantitative analysis demonstrates that among K'iche'–Spanish bilinguals, /f/ fortition is more common among bilinguals that are more dominant in K'iche' and that it predominantly occurs in phonological contexts that facilitate fortition (Baird and Regan forthcoming).

The social perceptions and indexical fields elicited by apocope and /f/ fortition in Guatemalan Spanish were explored in Baird (2021a, 2023), respectively. Both studies employed the matched-guise framework (Lambert et al. 1960) and an open-ended question approach. The results demonstrated that the guises with apocope and /f/ fortition were significantly more stigmatized than their counterpart guises without these features. Specially, guises with these features of Mayan-accented Spanish were rated as less honest, more ignorant, and poorer than guises sans the same features. Additionally, given the aforementioned value of the *traje típico* to Maya identity, participants in these studies were asked to respond to the open-ended question *¿Cómo se vestiría este hablante?* 'How would this speaker dress?' for each guise. Although it was not the most common response, the guises with apocope and /f/ fortition elicited responses of *traje típico* at a significantly higher rate than the guises without these features. Figure 2 illustrates an example of these results (the word cloud methodology is detailed in Section 2).

Figure 2. Sample word cloud of the responses to the question 'How would this speaker dress?' for guises with apocope and without apocope. The response of *traje típico* is highlighted in red in both world clouds. Modified from Baird (2021a, p. 232).

Even though the results from these perception studies of apocope (Baird 2021a) and /f/ fortition (Baird 2023) demonstrate the direct indexical relationship between these linguistic variables and traditional Mayan clothing, both studies used male voices for the guises. This is noteworthy due to the traditional gender roles among the Maya in Guatemala, where the onus of cultural and language preservation tends to fall on Maya women. In other words, although not unique to the Maya, it is common for men to speak the national language (Spanish) more than their Mayan language and to be more accepting of different aspects of Ladino culture, whereas the women take on the responsibility of teaching children their Mayan language and maintaining other aspects of their culture (Collier 1997; Hawkins and Adams 2005; Hendrickson 1995, 1996; Nelson 1999, 2009; Romero 2015; von den Berghe 1968; among others). Perhaps this is most salient among the Maya in terms of the *traje típico*. Numerous anthropological ethnographies agree that it is more likely for women to wear the *traje típico* in public than men (Collier 1997; Hendrickson 1995, 1996; Nelson 1999, 2009; among many others); Nelson (1999, p. 170) states that "traditional clothing, which signifies indigenous identity in general, has become almost isomorphic with the Maya woman who weaves it and wears it far more consistently than men".

The present study is motivated by this discrepancy in the use of the *traje típico* between Maya men and women and the results of Baird (2021a, 2023) for male-voiced guises. Furthermore, although there exists an abundance of detailed work from anthropological perspectives about the uses and meanings of the *traje típico* among Maya men and women, this work is based on qualitative observations. The present study performs a sociophonetic analysis employing the matched-guise methodology to provide a different framework in which to view the relationship between the *traje típico*, gender, Maya identity, and Mayan-accented Spanish in Guatemala. The research questions guiding this study are the following: (i) Can we replicate the findings of Baird (2021a, 2023) with different guises? In other words, will guises with apocope and /f/ fortition index the speaker as wearing a *traje típico* at a higher rate than guises without these linguistic features? and (ii) Is there an effect of speaker gender on the social evaluations? How do these results with female-voiced guises compare to the results of the male-voiced guises in Baird (2021a, 2023)?

2. Materials and Methods

This study follows the matched-guise methodology used in Baird (2021a, 2023). Four native female speakers of Guatemalan Spanish produced the elicitation materials for the guises. These individuals were recorded in quiet rooms in Guatemala via a Marantz PMD661 solid-state digital voice recorder, digitized at 16 bits (44.a kHz), and a Shure SM10A dynamic head-mounted microphone. The same map task from the previous studies was used to elicit naturalistic speech from these four speakers. Specific target words in

which apocope and /f/ fortition were more likely to occur were embedded in the map as names of streets, businesses, etc. (Figure 3). During the map task, the speakers gave instructions following specific routes on the map to another speaker with a map without the route, and this was repeated four times using four different routes.
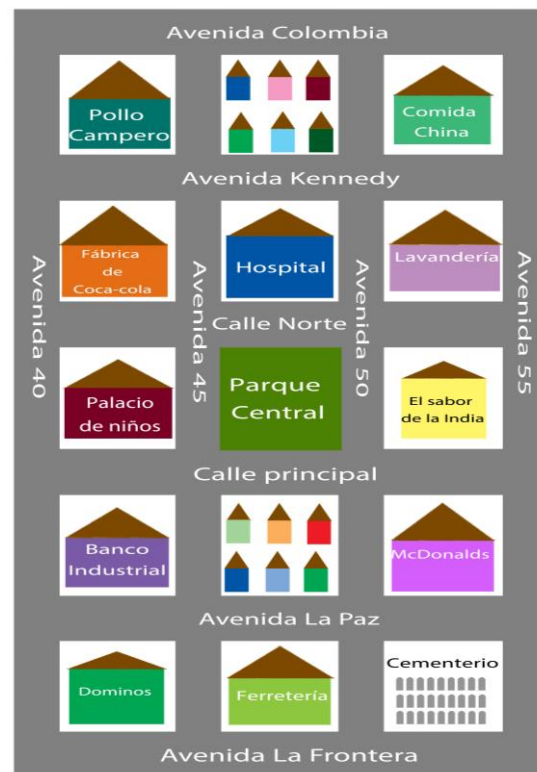
Figure 3. Map used in the elicitation of the stimuli.

Following this map task, each speaker was asked to read a list of controlled phrases aloud. These phrases included the same target words from the map task and each speaker was asked to read each phrase twice, once with the target word being produced with apocope or /f/ fortition and once with the target word being produced without these features. The [f] and [p] segments and the apocopated and non-apocopated words were then spliced into larger, more naturalistic sound clips from the map task. Hence, all the guises, regardless of whether they contained a feature of Mayan-accented Spanish or not, contained spliced segments in order to control for any effect of digital manipulation on the results. Distractor phrases for this study were produced following the same procedure. Six additional female speakers, two native speakers of Guatemalan Spanish, two non-native speakers (L1 English), and two native speakers of Peninsular Spanish, participated in the map and reading list tasks and had single segments from the reading task spliced into larger clips from the map task. Thus, all the sound clips presented to the participants of this study, including the distractor phrases, were digitally manipulated.

Following Baird et al. (2018), the spliced sound clips were judged by 28 native speakers of Guatemalan Spanish in a pilot study in order to ascertain how natural (i.e., digitally manipulated or not) each sound clip was using a 1–5 Likert scale. The participants listened to each clip in a random order, along with unspliced sound clips and sound clips with obvious digital manipulation. The sound clips chosen for this study received the highest 'natural' scores that were similar to, or higher than, the 'natural' scores given to the sound clips sans digital manipulation. The target sound clips used in this study are presented in (5–8).[1]

(5) Speaker 1: apocope

Y cruzando la Avenida Cuarenta y [ˈsiŋko/ˈsiŋk] *el parque central está al lado izquierdo...*

'And crossing 45th Avenue the central park is on the left side. . .'

(6) Speaker 2: apocope

Dos cuadras, hacia la Avenida Cincuenta y [ˈsiŋko/ˈsiŋk] *en la esquina de la lavandería. . .*

'Two blocks, towards 55th Avenue at the corner of the laundromat. . .'

(7) Speaker 3: fortition

Al llegar a la Avenida La [f/p]*rontera, tomarás la avenida cincuenta. . .*

'Upon arriving at Border Avenue, you'll take 50th avenue. . .'

(8) Speaker 4: fortition

Y después en la Avenida La Paz, para ir a dejar [f/p]*lores en el cementerio. . .*

'And after in La Paz Avenue, to go and leave flowers in the cemetery. . .'

Native speakers of Guatemalan Spanish that have only resided in Guatemala were recruited to participate in this study via social media and through announcements at universities throughout the country. Only participants that did not meet these criteria, recognized the voice of a guise speaker, or did not complete the survey were excluded. The 110 participants analyzed in this study reported being from all over the country. They included 40 that identified as men and 70 that identified as women, ages 18–70 (M: 32.6, SD: 13.1). A total of 34 participants self-identified as Maya and 76 self-identified as non-Maya.

The listening survey was presented online via the Qualtrics platform (Qualtrics, Provo, UT). Participants were given instructions and then presented with a different sound clip per webpage. The order of the clips was quasi-randomized so that two guises from the same speaker would not appear consecutively. Additionally, the quasi-random order was counterbalanced so that half of the participants would listen to the guise with the feature of Mayan-accented Spanish from one of the speakers in (5–8) first and the other half would listen to the guise sans the Mayan-accented Spanish feature first.

Although the majority of matched-guise studies employ a numeric methodology, e.g., asking participants to rank each voice on a particular social trait, this study used a more open approach to the matched-guise method. There is, of course, no methodical approach without limitations, and matched-guises such as this study do not allow for in-depth statistical analyses of multiple variables and interactions that a more quantitative analysis would provide. Nonetheless, given the specific goal of this study, to see if apocope and /f/ fortition directly index a female speaker as wearing a *traje típico*, a more qualitative method provides a more precise answer. In other words, instead of asking the participants to choose from a list of options that includes *traje típico* as a possible item of clothing or rank how likely the speaker is to wear a *traje típico*, this study simply asks the open-ended question *¿Cómo se vestiría este hablante?* 'How would this speaker dress?' for each clip.[2] Although it is perhaps impossible to completely avoid priming participants in studies such as this, and the participants are primed to think about clothing in this method, at no point during this study are Maya identity, Mayan languages, the *traje típico*, or even Guatemala mentioned. Thus, aside from being asked about clothing in general, participants are free to answer however they wish. If a participant responds that they think a speaker would wear a *traje típico*, it is even stronger evidence of the indexical relationship between apocope, /f/ fortition, and a Maya identity than if similar results occurred in one of the more numeric ranking methods mentioned previously.

An additional advantage of matched-guises with open-ended questions is that they allow us to understand more of the indexical field of a single linguistic variable. In other words, "the open-ended responses to matched-guises can expand our knowledge of these fields, as participants are not confined to the overt attitudes usually found when using a rankings method" (Baird 2021a, p. 229). The studies that have employed open-ended

questions in matched-guises primarily use word clouds to present their results, as in the previous studies on the perceptions of Mayan-accented Spanish (Baird 2021a, 2023; see also Baird et al. 2018; Drager et al. 2012; Kirtley 2011; Nance 2013). In order to better approximate the indexical field proposed by Eckert (2008), the word clouds presented in the results in Section 3 display all of the participants' answers, even those given by a single participant. As Eckert (2008) states, a single linguistic variable may index completely opposite reactions for different listeners. As such, it is common to see answers that appear to contradict each other within these word clouds. Following Baird (2021a, 2023), the answers that appear in the word clouds are left in Spanish, as translating them to English may misrepresent what the participants intended with their answers. Furthermore, the answers were only combined into broader semantic categories in specific cases, e.g., when answers were clear synonyms such as *médica* and *doctora* 'doctor'. When participants answered in complete sentences, only the descriptor words were used in the word clouds. For example, if a participant responded *Creo que es del altiplano de Guatemala y lleva traje típico* 'I believe she's from the Guatemalan Highlands and is wearing a *traje típico*', the response was analyzed as *altiplano* and *traje-típico*. Dashes are used between words in answers with more than one word so that the online word cloud generator (worditout.com, accessed on 10 November 2022) would analyze them as single answers, providing greater clarity in the word clouds.

In the word clouds presented in the results in Section 3, the size of the word represents the frequency of that response: larger words represent more frequent answers than smaller words. In order to better understand the differences in responses across guises, the same responses were analyzed via binomial (one-tailed) tests of the distribution of frequencies of these responses. In these analyses, the guise without the feature of Mayan-accented Spanish, i.e., the non-apocope and the non-fortition guises, was used as the baseline. The binomial tests' *p*-values are reported in Section 3, and the frequency tables of the proportion of the responses discussed can be found in Appendix A. If a response appeared in the word cloud of one guise but not the other, the results of the binomial tests are not reported, as they are always $p < 0.001$. Finally, the results were analyzed according to guise, participant gender, and whether the participant self-identified as Maya or not. Differences in responses across Gender and Ethnicity were examined via a series of binomial (one-tailed) tests according to the response *traje típico* by Speaker and phonetic feature, e.g., an analysis of male and female participants' responses of *traje típico* to the apocope guise for Speaker 1, an analysis of Maya and non-Maya participants' *traje típico* responses to the non-fortition guise for Speaker 3, etc. Nonetheless, the additional results for participant Gender and Ethnicity should be interpreted with caution, as running additional analyses increases the chances of false positives, especially when rates of *traje típico* are low.

3. Results

3.1. Apocope

The word clouds for the apocope and the non-apocope guises for Speakers 1 and 2 are presented in Figures 4 and 5. For both speakers, *traje típico* was the most common response to the apocope guise, whereas *jeans* was the most common response to the non-apocope guise. Although *traje típico* was also an answer for the non-apocope guises, it was significantly more frequent in the apocope guises: 45 of 110 participants responded *traje típico* to Speaker 1's apocope guise and 6 responded *traje típico* to the non-apocope guise ($p < 0.001$); for Speaker 2, there were 43 cases of *traje típico* for the apocope guise and 12 responses in the non-apocope guise ($p < 0.001$). Furthermore, additional descriptors that index Maya identity included specific pieces of the *traje típico* depicted in Figure 1.[3] Among these are *huipil* (also spelled *güipil*) 'woven blouse' and *corte* 'woven skirt'. *Huipil* and *corte* both appeared in Speaker 1's apocope guise, whereas they did not appear in the non-apocope guise. For Speaker 2, *huipil* and *corte* were significantly more frequent in the apocope guise than in the non-apocope guise ($p < 0.001$ for both comparisons). The terms *indígena* 'indigenous person' and *ladina* were found among the responses as well. *Indígena* was more common in the apocope guise for Speaker 1 ($p < 0.01$) and did not

appear in Speaker 2's non-apocope guise. *Ladina* appeared in both non-apocope guises. The difference in *ladina* replies between Speaker 2's guises was not significant ($p = 0.129$) and *ladina* did not appear as a response to Speaker 1's apocope guise. Finally, additional answers provided by single participants to the apocope guises that index Maya identity included specific Mayan-speaking areas such as *Nahualá, Quiché,* and *Totonicapán*, the *traje típico*-specific necklace, *chachal*, and shawl, *perraje*, and the derogatory remark *como indita de la montaña* 'like a little indigenous woman from the mountain'.

Figure 4. Word clouds of responses to the apocope and the non-apocope guises for Speaker 1.

Figure 5. Word clouds of responses to the apocope and the non-apocope guises for Speaker 2.

In addition to the responses to the apocope guises that directly index a Maya identity, other descriptors index a lower social class in general and may, in turn, indirectly index Maya identity. For example, responses to the apocope guise for Speaker 1 (Figure 4) included *del campo* 'from the countryside', *desgastada* 'worn out (clothing)', *hecha a mano* 'handmade', and *pobre* 'poor'. Conversely, responses that imply a higher status for the non-apocope guise included answers such as *de marca* 'name brand', *elegante* 'elegant', *tacones* 'high heels', among others. The adjectives *sencilla* 'simple (clothing)' and *sucia* 'dirty (clothing)' appeared as responses to both guises at similar rates (*sencilla*, $p = 0.426$; *sucia*, $p = 0.352$). For Speaker 2 (Figure 5), responses that index a lower social status in the apocope guise included *rural* 'rural', *pobre* 'poor', and *sucia* 'dirty (clothes)', whereas the non-apocope guise elicited more responses indexing a higher social class, such as *formal* 'formal (clothes)', *reloj* 'wristwatch', *enfermera* 'nurse', and even some that indicated the opposite of a Maya identity: *turista* 'tourist' and *tenis* 'tennis shoes'. Some responses that index a lower social class were used in reaction to both guises for Speaker 2, such as *hecha a mano* 'handmade' ($p = 1.00$).

The separate binomial tests for Gender and Ethnicity did not demonstrate any effect on the number of responses of *traje típico* in any guise. In other words, all of the participants, regardless of their gender or whether they identified as Maya or not, perceived the voices with apocope as more likely to wear the *traje típico* than non-apocope guises.[4]

3.2. [f] Fortition

The responses to guises with /f/ fortition from Speakers 3 and 4 are seen in Figures 6 and 7. As with the apocope results, these word clouds demonstrate a clear and direct visual body–language link between /f/ fortition and the speaker wearing the traditional Mayan *traje típico*. *Traje típico* was given as an answer by 51 participants for Speaker 3's fortition guise and by 47 participants for Speaker 4's fortition guise, significantly more so than in the response to the non-fortition guises: seven and nine responses, respectively ($p < 0.001$ in both comparisons). Again, the *traje típico*-specific articles of clothing *huipil* and *corte* appeared more in the responses to the fortition guise than in the non-fortition guise context. Between Speakers 3 and 4, there were 31 total responses of *huipil* and 25 of *corte*. In comparison, there was only one *huipil* response in Speaker 4's non-fortition guise and no responses of *corte* in either non-fortition guise. *Indígena* is seen in both fortition guises but is not found in either non-fortition guise. Similarly, *ladina* occurs somewhat commonly in Speaker 4's non-fortition guise but is absent from the fortition guise. The difference between Speaker 3's fortition and non-fortition guises for *ladina* responses approaches significance ($p = 0.068$). As with the apocope results, answers provided by single listeners illustrate the direct indexicality of /f/ fortition among the participants in this study. Mayan-speaking geographical areas mentioned include *Zunil, Almolonga, Nahualá,* and *Patzún.* Additional clothing responses include *traje yucateco* 'Yucatan clothes', *perraje*, and *chachal*.

Figure 6. Word clouds of responses to the fortition and the non-fortition guises for Speaker 3.

Additional descriptors that index a lower social status for Speaker 3's fortition guise include *segunda mano* 'second hand (clothes)', *hecha a mano* 'hand-made', *pobre* 'poor', and *sucia* 'dirty'. Conversely, Speaker 3's non-fortition guise elicited replies such as *maquillaje* 'makeup', *tacones* 'high heels', and *formal* 'formal (clothing)'. Interestingly, the non-fortition guise elicited answers of *del campo* 'from the countryside', whereas the fortition guise did not. Responses of *desgastada* 'worn out (clothes)' and *sencilla* 'simple (clothes)' appeared at comparable rates across both guises ($p = 1.00$ and $p = 0.691$, respectively). For Speaker 4, the fortition guise produced strong negative comments not seen in other word clouds, including *descuidada* 'sloppy, unkempt' and *descalza* 'barefoot'. Other answers that index a lower social status went beyond the question of clothing; these include *caminante de ciudad* 'street wanderer', *vendedora ambulante* 'street vendor', and *problemas para hablar español* 'trouble speaking Spanish'. Speaker 4's non-fortition guise responses were much more positive: *elegante* 'elegant', *de gala* 'gala', *tacones* 'high heels', among others. Answers found

in both guises for Speaker 4 include *desgastada* 'worn out (clothes)' ($p = 1.00$) and *formal*, which was significantly more prominent in the non-fortition guise than the fortition guise ($p < 0.05$). Again, there were no significant effects of Gender or Ethnicity in the separate binomial tests of *traje típico* responses across guises.[5]

Figure 7. Word clouds of responses to the fortition and the non-fortition guises for Speaker 4.

3.3. Comparison with Male-Voiced Guises

Aside from indexing a lower social status, via the proxy of lower quality clothing in general, the results from the female-voiced guises demonstrate that traditional Maya clothing is strongly indexed by both apocope and /f/ fortition. These findings parallel those found for male-voiced guises for both apocope (Baird 2021a) and /f/ fortition (Baird 2023). In this section, the results from the present study are compared to the results from Baird (2021a, 2023). Specifically, this section analyzes responses of *traje típico* and other answers that directly index Maya identity.

In all of these studies, the results analyzed were not conflated but kept separate for individual guise speakers. This was done to adhere to the matched-guise framework, as variables such as vocabulary, speaking rates, fundamental frequency, etc., that vary across speakers may influence the results if all of the guises from different speakers are combined and analyzed as a whole. However, it is not possible to avoid these conflations when comparing the results from the female-voiced guises to those of the male-voiced guises from the previous studies. In these comparisons, the results for Speakers 1 and 2 are combined for apocope and compared with the combined results of the two male guise speakers in Baird (2021a), and the results for Speakers 3 and 4 for /f/ fortition are combined and compared with the results of the two male-voiced guises in Baird (2023). In these comparisons, all of the responses pertaining to any specific article of the *traje típico* were combined, i.e., responses such as *huipil* and *corte* in this study and *kutin* 'male woven shirt' in the previous studies were counted as *traje típico*.[6]

The participants in this study differ from those in Baird (2021a, 2023). As stated above, the participants in this study consisted of 110 native speakers of Guatemalan Spanish that had only lived in the country: 40 men, 70 women; 34 Maya, 76 non-Maya; ages 18–70 (M: 32.6, SD: 13.1). In both Baird (2021a) and Baird (2023), the group of participants was the same: 116 native speakers of Guatemalan Spanish that have only resided in Guatemala; 57 men, 58 women, 1 non-binary; 35 Maya, 81 non-Maya; ages 18–56 (M: 27.3, SD: 6.6). The conflation of guise results across these different groups of participants was deemed necessary in order to answer the second research question about whether there is an effect of gender on the indexical fields of these two linguistic variables of Mayan-accented Spanish. These results were also analyzed via binomial (one-tailed) tests of the distribution

of frequencies of the responses with the male-voiced guises as the baseline. The frequency values (percentage rates) of these responses are reported in Appendix B.

The combined results of these studies are presented in Figure 8. These findings demonstrate that responses of *traje típico* are more likely to occur in reaction to female-voiced guises than male-voiced guises regardless of the presence of Mayan-accented Spanish: all four guise comparisons are significant at $p < 0.001$. Thus, these results suggest that it is more likely for the native Guatemalan Spanish female voices used as stimuli in this study to be perceived as Maya than the native Guatemalan Spanish male voices, with or without the features of Mayan-accented Spanish. Furthermore, the guises that directly indexed the speaker as wearing a *traje típico* the most were the female-voiced guises with Mayan-accented Spanish variables present. In other words, although these data are more qualitative in nature, there is, in a sense, an interaction between the female-voiced guises and the features of Mayan-accented Spanish analyzed in this study. Although both of these independent variables directly index a speaker as wearing traditional Maya clothing, the visual body–language link is strongest when they are combined.
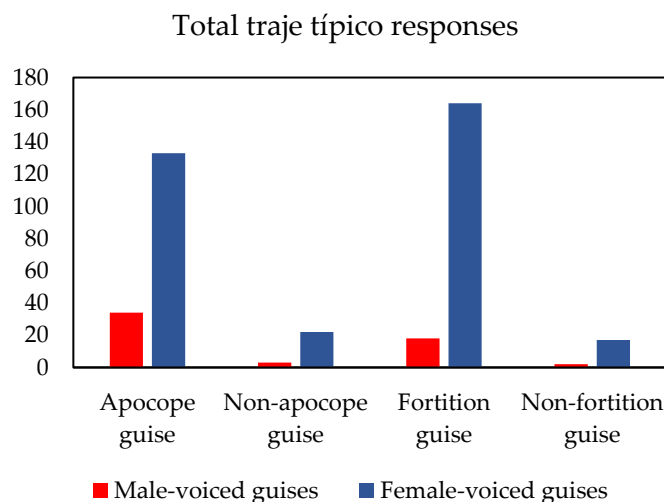
Figure 8. Total *traje típico* responses across all of the apocope/non-apocope and fortition/non-fortition guises from this study and Baird (2021a, 2023). Responses to male-voiced guises are in red and responses to female-voiced guises are in blue.

4. Discussion

The goal of this study was to replicate the findings of Baird (2021a, 2023) to see if the apocope of word-final unstressed vowels and the fortition of /f/ to [p] indexed a native speaker of Guatemalan Spanish as being Maya and, in particular, wearing traditional Maya clothing. This study goes a step further by analyzing the role of gender in these indexical fields, as it compared the results from female-voiced guises to the previous results of male-voiced guises.

The first research question was whether the results found in Baird (2021a, 2023) would hold true with different guise voices. As reported in Section 3, the results of this investigation corroborate the findings of previous studies; guises with apocope and /f/ fortition index a lower social class in general and, more specifically, the *traje típico*. Again, it is important to note that these results provide strong evidence of the indexical relationship between a single linguistic variable, apocope or /f/ fortition, and a Maya identity, as there was no mention of anything concerning the Maya in the study and the questions were open-ended. In other words, the responses concerning Maya identity were unprimed aside from asking the participants to consider clothing in general. However, as seen in a response to Speaker 4's fortition guise, one participant ignored the question about clothing and responded *problemas para hablar español* 'trouble speaking Spanish'. This type of comment shows that the indexical field of Mayan-accented Spanish in Guatemala is significantly

vaster than the specific visual body–language links explored in this paper, and the reader is referred to Baird (2021a, 2023) for a more thorough, albeit not exhaustive, review of the social meanings of these variables that extend beyond the *traje típico* in male speech.

Although we may assume that these specific phonetic features of Guatemalan Spanish in contact with Mayan languages are highly stigmatized, this study and Baird (2021a, 2023) are the first to empirically demonstrate this stigmatization. The previous literature that specifically discusses apocope and /f/ fortition is scarce and only acknowledges their existence (Aleza Izquierdo 2010; García Tesoro 2008; Lentzner 1893). Additionally, the ethnographies that reference Mayan-accented Spanish do so in passing, only briefly remarking on a few morphosyntactic features (Nelson 1999, 2009). In general, the combined results of Baird (2021a, 2023) and this study empirically demonstrate that apocope and /f/ fortition are highly marked features that elicit stereotypes associated with the Maya in Guatemala. Following Silverstein's (2003) orders of indexicality, the basic *n-order* index of both apocope and /f/ fortition appears to be a Maya identity, via the proxy of the iconic *traje típico*. As mentioned in Section 3, although the additional statistical analyses of Gender and Ethnicity are limited by the qualitative nature of the data, there were no significant differences in the responses between Maya and non-Maya participants or between female and male participants; all participants in this study, regardless of ethnicity or gender, perceived the guises with apocope and /f/ fortition as being more likely to wear a *traje típico* than the guises without these features. In other words, these results suggest that the visual body–language link between these phonetic features of Mayan-accented Spanish and the *traje típico* has moved beyond different subgroups within the country and is present among Guatemalans in general. Additionally, even though a linguistic feature may have an established *n order* index in general, different contexts and populations often result in said feature gaining new social meanings—what Silverstein (2003) calls *n + 1-order* indexes. Within the specific context of Guatemala, the social meaning of these two phonetic features is not just of a Maya identity, but, as many of the responses in the word clouds demonstrate, there is an *n + 1-order* that ties the Maya identity to a lower social status via the proxy of clothing.

In response to the second research question, which asked whether the gender of the speaker would affect the participants' responses, the results reveal that the *n-order* index of the *traje típico* in the apocope and /f/ fortition guises is strengthened when the guise voice is female. The results in Section 3.3 clearly demonstrate that female-voiced guises with Mayan-accented Spanish elicit responses of Mayan-traditional clothing more frequently than male-voiced guises with the same features. Furthermore, the results demonstrate a possible *n-order* index of female voices in general: the female-voiced guises are more likely to index a Maya identity than the male-voiced guises regardless of the presence of apocope or /f/ fortition. This statement, of course, has limitations, as it is only based on the four female speakers that created the stimuli for this study and the four male speakers from Baird (2021a, 2023). Additionally, it must be noted that there may be other linguistic variables in the voices of the female guises that index a Maya identity, e.g., the speakers' intonational patterns were not analyzed here although previous research has demonstrated ample effects of contact with Mayan languages on Guatemalan Spanish intonation (Baird 2015, 2017, 2021b).

Nonetheless, these findings provide further evidence of language and gender roles and the hierarchies of the indigenous Maya in Guatemala and amongst other populations. Although it is not universal, it is common among indigenous minority groups for the responsibility of linguistic and cultural preservation to fall on women as men shift towards the linguistic and cultural practices of the hegemonic/dominant majority. For example, in his work on the Tolowa in Northern California, Collins notes that it is "predominately the women's voices which have survived" because, throughout their history, women were the ones that were "emphatically wearing [Tolowa] clothes, sticking with [Tolowa] food, and, most importantly, speaking Tolowa with their children" (Collins 1998, p. 174). Hoffman (2008, p. 72) notes that, among the Asheli of Morocco, the cultural, clothing,

and linguistic identities of men were more "compromised" due to contact with more mainstream Moroccan society, with women seen as the "repositories" of Asheli culture and the Tashelhit language and tasked with passing it on to future generations.

As detailed in Section 1, these same gender roles concerning language and cultural maintenance are seen among the Maya in Guatemala. The Maya anthropologist Irma Otzoy (1996, p. 146) goes as far as suggesting that the Maya woman is more valiant than the Maya man because she still wears the *traje*. Hendrickson (1995, p. 89) contends that it is not just racism in Guatemala that has led men to abandon the *traje típico* more than women, but also sexism, as a man wearing a *traje* is often seen as "one who is less masculine (even less adult) in a world dominated by non-Maya values", whereas a Maya woman is often seen as more feminine in the *traje típico* (see also Nelson 1999). Various anthropologists have argued that the quintessential image of the Maya *is* a woman wearing a *traje típico* (Annis 1987; Collier 1997; French 2010; Hawkins and Adams 2005; Hendrickson 1995, 1996; Nelson 1999, 2009; Oertling 2021; Otzoy 1996; Romero 2015; von den Berghe 1968; among others). This is evidenced in the Guatemalan newspaper article about the *traje típico* featured in Figure 1, as there are no mentions of men needing to wear it. All of the onus for its continued usage is placed on women.

In order to gather some quantitative data on the daily usage of the *traje típico*, I spent two hours in the central plaza of Quetzaltenango, the second largest city in Guatemala, and two hours in the central plaza of the predominantly K'iche'-speaking and more rural municipality Nahualá observing the dress of individuals of all ages; both observations occurred on a Saturday. In Quetzaltenango, I observed the following: 114 of 168 women of all ages wore *trajes típicos* and none of the 132 men wore a *traje típico*. In Nahualá, 77 of 79 women of all ages wore a *traje típico*. The two that did not were workers at a local bank and wore their work uniform. Two of 53 men wore the *traje típico*, both of whom I know personally, and they were over 70 years old at the time. This exercise has some obvious limitations. For example, it is based on my assumptions of the gender identification of these individuals as I did not interact with them. Nevertheless, these observations corroborate the qualitative remarks from the aforementioned scholars on gender and the *traje típico*.

The results of this study demonstrate that gendered and cultural practices of the Maya are reflected in the indexical fields of Mayan-accented Spanish in Guatemala, regardless of the gender or ethnicity of the listener. In other words, this link of the visual body and language is significantly more essentialized for the identity of a woman than for the identity of a man in Guatemala. In general, listeners appear to assume that *any* woman is more likely than a man to be Maya. More specifically, while both Maya women and men are subject to negative stereotypes, the women are also more likely to be bound by Maya-specific stereotypes. Whereas a Maya man may not be seen in the same social class as a Ladino, a Maya woman is not only seen in a lower social class but also as someone who *should* be wearing the *traje típico*. These gendered stereotypes are, in turn, reinforced in the local language ideologies. Some individual Ladino participants in this study and in Baird (2021a, 2023) gave derogatory comments in response to guises with Mayan-accented Spanish. As stated in Section 3.1, in response to Speaker 2's apocope guise, one participant answered, 'like a little indigenous woman from the mountain.' In contrast, disparaging remarks made by individual participants concerning the male-voiced guises were not specifically linked with a Maya identity, e.g., in response to a guise with /f/ fortition in Baird (2023), one participant replied that they pictured the speaker as a gang member whose specific role was assaulting individuals, sans any reference to the speaker's ethnicity. Thus, these language ideologies, gendered stereotypes, and the embodied practices of the Maya appear to be mutually reinforcing one another in Guatemala.

What this matched-guise does not tell us is what the response *traje típico* signifies, besides a Maya identity, for the participants of this study. As stated earlier, the layered meanings of the female *traje típico* in Guatemala are vast and a detailed analysis is outside of the scope of this study. However, even though there were no statistical differences between the number of *traje típico* responses by Gender or Ethnicity, we cannot simply assume that

every participant in this study views the *traje típico* in the same way. As illustrated in Figure 1, there are many that celebrate the *traje típico* and the Maya culture and identity that it represents; it continues to be worn by Maya women as a source of pride in everyday life and in a growing number of other contexts, such as at the Miss Universe competition, at international film festivals, and as a uniform during athletic competitions (Baird 2019; Redacción Espectáculos 2015; Reyes 2019; Sam Chum 2019). For these Maya women, the *traje típico* has become emblematic of their resistance to the widespread attempts of their own nation to erase them. On the other hand, the *traje típico* may also be representative of the negative stereotypes that still accompany being Maya in Guatemala. For Ladinos, and even some Maya, it may represent a caricature of the static and traditional Maya that is an obstacle to national progress, something that only serves to appease the curiosity of foreign tourists. Nelson (1999) even argues that the *traje típico* is a common trope in racist and sexist jokes that Ladinos make about the Maya. Furthermore, participants' responses in this study also parallel a national process of amalgamation of the Maya into a single ethnic group. Many of the responses that indexed Maya identity in the word clouds were very general in nature, even though each community has a distinct *traje*. Hence, future studies will need to investigate listeners' attitudes towards the *traje* and the Maya in relation to their sociolinguistic reactions to features of Mayan-accented Spanish.

Finally, as in previous work, such as Mendoza-Denton's (2008) and Zhang's (2008) seminal studies, this paper reveals how linguistic variables and material style work together in the creation of an identity as visual body–language links (Calder 2019). However, a key difference here is that, whereas those investigations focus on the production of different linguistic variables and fashion choices by individuals, the present study examines this link from the point of view of perception. For example, Mendoza-Denton (2008) finds that Latina youth gang members in California use material style, such as longer eye liner, and phonetic variables, such as fortition of /θ/ to [t̪], together in the creation of their desired identities. However, what we do not known is how strong this visual body–language link is between longer eye liner and /θ/ fortition for community members. In other words, would /θ/ fortition index longer eye liner among listeners like Mayan-accented Spanish indexes the *traje típico*? As such, future studies should continue this line of research by exploring the perception of visual body–language links from a third-wave perspective.

5. Conclusions

The results of this study further demonstrate the inseparable relationship between language and society: "no language is spoken in a vacuum" (Aikhenvald 2012, p. 18). Language users, their experiences, their cultures, etc., help to mold the indexical fields of different linguistic variables. The findings of this study undoubtedly demonstrate that the gender roles and cultural practices of the Maya are so embedded in Guatemalan society that they are a critical variable in the indexical field of Mayan-accented Spanish.

According to Eckert (2008), although fashion is often a more deliberate choice than linguistic style, they are both examples of the bricolage of identity. Many Maya women consciously choose to wear the *traje típico* in both their day-to-day lives in Guatemala and even abroad because of what it represents to them. Many also choose to continue speaking their Mayan languages. As in all contexts of bilingualism and language contact, some linguistic features may transfer between languages. Future research is needed in order determine if these specific linguistic variables are consciously used by different Mayan–Spanish bilinguals along with the *traje típico* in order to index a Maya identity, as seen in other populations (Eckert and McConnell-Ginet 2003; Mendoza-Denton 2008; Zhang 2008).

Data Availability Statement: The data are not publicly available due to IRB restrictions.

Conflicts of Interest: The author declares no conflict of interest.

Appendix A

Binomial tests (one-tailed) frequency tables of compared responses of female-voiced guises.

Speaker 1 Response	apocope guise	non-apocope guise	*p*-value
traje típico	0.409	0.055	<0.001
indígena	0.036	0.009	<0.01
Speaker 2 Response	**apocope guise**	**non-apocope guise**	***p*-value**
traje típico	0.391	0.109	<0.001
huipil	0.127	0.018	<0.001
corte	0.136	0.018	<0.001
ladina	0.018	0.045	0.129
Speaker 3 Response	**[p] guise**	**[f] guise**	***p*-value**
traje típico	0.464	0.064	<0.001
ladina	0.018	0.055	0.068
Speaker 4 Response	**[p] guise**	**[f] guise**	***p*-value**
traje típico	0.427	0.082	< 0.001
huipil	0.172	0.009	< 0.001

Appendix B

Binomial tests (one-tailed) frequency tables of *traje típico* responses across male- and female-voiced guises.

Guise Comparison	Female-Voiced Guises	Male-Voiced Guises	*p*-Value
apocope	0.605	0.147	<0.001
non-apocope	0.100	0.013	<0.001
fortition	0.745	0.078	<0.001
non-fortition	0.077	0.009	<0.001

Notes

[1] As in Baird (2021a), the word *cinco* was used as it may illustrate apocope and is common in street names. As a reviewer points out, these elicitation sentences contain other words that could possibly demonstrate apocope and fortion. However, as this study specifically investigates if a single occurrence of apocope or fortition would index a Maya identity, additional examples were not included. Furthermore, it was decided that this was more natural, as production studies have shown that Spanish–Mayan bilinguals do not produce these features in every possible context (Baird and Regan forthcoming). Future research is warranted as to whether more cases of apocope and fortion in a guise elicit even stronger perceptions of Maya identity, as discussed in García et al. (2023) regarding /s/ aspiration in Puerto Rican Spanish.

[2] This question was left as the masculine, or unmarked, 'este hablante' and not the feminine gendered 'esta hablante' in order to avoid specific priming of the female gender of the guise voices to the participants. As a reviewer suggests, particpants could simply view this as a grammatical error. However, as no participant commented on this, it is not known if any did view it as such.

[3] For the articles of clothing depicted in Figure 1 and also given as responses to the matched-guises, *faja* 'sash', *cinta* 'hair ribbon', and *collar* 'necklace' are not considered to be directly indexing Maya identity here, as these words do not only signify parts of the *traje típico*, but articles of clothing in general.

[4] Results by Gender: Speaker 1 apocope guise, *p* = 0.071; Speaker 1 non-apocope guise, *p* = 0.106; Speaker 2 apocope guise, *p* = 0.067; Speaker 2 non-apocope guise, *p* = 0.923. Results by Ethnicity: Speaker 1 apocope guise, *p* = 0.126; Speaker 1 non-apocope guise, *p* = 0.206; Speaker 2 apocope guise, *p* = 0.122; Speaker 2 non-apocope guise, *p* = 0.079.

[5] Results by Gender: Speaker 3 fortition guise, *p* = 0.097; Speaker 3 non-fortition guise, *p* = 0.872; Speaker 4 fortition guise, *p* = 0.841, Speaker 4 non-fortition guise, *p* = 0.849. Results by Ethnicity: Speaker 3 fortition guise, *p* = 0.578; Speaker 3 non-fortition guise, *p* = 0.071; Speaker 4 fortition guise, *p* = 0.326; Speaker 4 non-fortition guise, *p* = 0.074.

[6] Due to all of these combinations, the results presented in this section cannot be interpreted as a true matched-guise, as the comparison of different voices presents considerably more uncontrolled variables than Lambert et al. (1960) originally intended in this framework.

References

Aikhenvald, Alexandra Y. 2012. *The Languages of the Amazon*. Oxford: Oxford University Press.

Aleza Izquierdo, Milagros. 2010. Fonética y fonología. In *La Lengua Española en América: Normas y Usos Actuales*. Edited by Milagros Aleza Izquierdo and José María Enguita Utrilla. Valencia: Universitat de València, pp. 51–94.

Annis, Sheldon. 1987. *God and Production in a Guatemalan Town*. Austin: University of Texas Press.

Baird, Brandon. 2015. Pre-nuclear peak alignment in the Spanish of Spanish-K'ichee' (Mayan) bilinguals. In *Selected Proceedings of the 6th Conference on Laboratory Approaches to Romance Phonology*. Edited by Erik W. Willis, Pedro Martín Butragueño and Esther Herrera Zendejas. Cascadilla: Cascadilla Proceedings Project, pp. 163–74.

Baird, Brandon. 2017. Prosodic transfer among Spanish-K'ichee' Bilinguals. In *Mutlidisciplinary Appraoches to Bilinguaism in the Hispanic and Lusophone World*. Edited by Kate Bellamy, Michael W. Child, Paz González, Antje Muntendam and M. Carmen Parafita Couto. Amsterdam: John Benjamins, pp. 147–72. [CrossRef]

Baird, Brandon. 2019. Ciudadano maya 100%: Uso y actitudes de la lengua entre los bilingües k'iche'-español. *Hispania* 102: 319–34. [CrossRef]

Baird, Brandon. 2021a. "Para mí, es indígena con traje típico": Apocope as an indexical marker of indigeneity in Guatemalan Spanish. In *Topics in Spanish Linguistic Perceptions*. Edited by Luis A. Ortiz and Eva-María Suárez Budenbender. New York: Routledge, pp. 223–39. [CrossRef]

Baird, Brandon. 2021b. Bilingual language dominance and contrastive focus marking: Gradient effects of K'ichee' syntax on Spanish prosody. *International Journal of Bilingualism* 25: 500–15. [CrossRef]

Baird, Brandon. 2023. Social perceptions of /f/ fortition in Guatemalan Spanish. *Spanish in Context*. [CrossRef]

Baird, Brandon. forthcoming. Spanish in contact with Mayan languages in Guatemala. In *Contact Varieties of Spanish and Spanish-Iexified Contact Varieties. Volume II, Indigenized Varieties of Spanish in Multilingual Scenarios*. Edited by Leonardo Cerno, Hans-Jörg Döhla, Miguel Gutiérrez Maté, Robert Hesselbac and Joachim Steffen. Berlin: De Grutyer.

Baird, Brandon, and Brendan Regan. forthcoming. The status of /f/ in Mayan-accented Spanish. *Studies in Hispanic and Lusophone Linguistics*.

Baird, Brandon, Marcos Rohena-Madrazo, and Caroline Cating. 2018. Perceptions of Lexically Specific Phonology Switches on Spanish-origin Loan Words in American English. *American Speech* 93: 79–107. [CrossRef]

Batrés Jáuregui, Antonio. 1892. *Vicios del lenguaje y provincialismos de Guatemala*. Guatemala City: Encuadernación y tipografía nacional.

Calder, Jeremy. 2019. From Sissy to Sickening: The Indexical Landscape of /s/ in SoMa, San Francisco. *Journal of Linguistic Anthropology* 29: 332–58. [CrossRef]

Chappell, Whitney, ed. 2019. *Recent Advances in the Study of Spanish Sociophonetic Perception*. Amsterdam: John Benjamins. [CrossRef]

Collier, Jane Fishburn. 1997. *From Duty to Desire: Remaking Families in a Spanish Village*. Princeton: Princeton University Press.

Collins, James. 1998. *Understanding Tolowa Histories: Western Hegemonies & Native American Responses*. New York and London: Routledge.

Colop, Marylin. 2013. El traje típico es representativo de la cultura maya: Mujeres de Cantel mantienen uso. *El Quetzalteco, Prensa Libre*. Available online: issuu.com/elquetzalteco/docs/el_traje_tipico_representativo_de_l (accessed on 29 August 2022).

Drager, Katie, Rachel Schutz, Ivan Chik, Kate Hardeman, and Victor Jih. 2012. When Hearing Is Believing: Perceptions of Speaker Style, Gender, Ethnicity, and Pitch. Paper presented at LSA Annual Meeting, Portland, ON, USA, January 5–8; Available online: www.katiedrager.com/publications (accessed on 3 October 2017).

Eckert, Penelope. 1980. Clothing and geography in a suburban high school. In *Researching American Culture*. Edited by Conrad Phillip Kottak. Ann Arbor: University of Michigan Press, pp. 139–45.

Eckert, Penelope. 2008. Variation and the Indexical Field. *Journal of Sociolinguistics* 12: 453–76. [CrossRef]

Eckert, Penelope. 2012. Three waves of variation study: The emergence of meaning in the study of sociolinguistic variation. *Annual Review of Anthropology* 41: 87–100. [CrossRef]

Eckert, Penelope, and Sally McConnell-Ginet. 2003. *Language and Gender*. Cambridge: Cambridge University Press.

England, Nora. 2003. Mayan language revival and revitalization politics: Linguists and linguistic ideologies. *American Anthropologist* 105: 733–43. [CrossRef]

England, Nora, and Brandon Baird. 2017. Phonology and Phonetics. In *The Mayan Languages*. Edited by Judith Aissen, Nora England and Roberto Zavala. New York: Routledge, pp. 175–200.

French, Brigittine M. 2010. *Maya Ethnolinguistic Identity Violence, Cultural Rights, and Modernity in Highland Guatemala*. Tuscon: University of Arizona Press.

García, Christina, Abby Walker, and Mary Beaton. 2023. Exploring the role of phonological environment in evaluating social meaning: The case of /s/ aspiration in Puerto Rican Spanish. *Languages* 8: 186.

García Tesoro, Ana Isabel. 2008. Guatemala. In *El español en América: Contactos lingüísticos en Hispanoamérica*. Edited by Azucena Palacios. Barcelona: Ariel Letras, pp. 97–117.

Hawkins, John P., and Walter Randolph Adams. 2005. *Roads to Change in Maya Guatemala: A Field School Approach to Understanding the K'iche'*. Oklahoma City: Oklahoma University Press.

Hendrickson, Carol. 1995. *Weaving Identities: Construction of Dress and Self in a Highland Guatemala Town*. Austin: University of Texas Press.

Hendrickson, Carol. 1996. Women, Weaving, and Education in Maya Revitalization. In *Maya Cultural Activism in Guatemala*. Edited by Edward F. Fischer and R. McKenna Brown. Austin: University of Texas Press, pp. 156–64.

Hoffman, Katherine E. 2008. *We Share Walls: Language, Land, and Gender in Berber Morocco*. Malden: Blackwell Publishing.

Instituto Nacional de Estadística Guatemala. 2018. *XII Censo Nacional de Población y VII Censo Nacional de Vivienda*. Available online: www.censopoblacion.gt/ (accessed on 1 September 2020).

Kirtley, M. Joelle. 2011. Speech in the U.S. Military: A Sociophonetic Perception Approach to Identity and Meaning. Master's thesis, University of Hawaii at Manoa, Manoa, HI, USA.

Lambert, Wallace E., Richard C. Hodgson, Robert C. Gardener, and Samuel Fillenbaum. 1960. Evaluational Reactions to Spoken Languages. *Journal of Abnormal and Social Psychology* 60: 44–51. [CrossRef] [PubMed]

Lentzner, Karl. 1893. Observations on the Spanish language in Guatemala. *Modern Language Notes* 8: 41–43. [CrossRef]

Mendoza-Denton, Norma. 2008. *Homegirls: Language and Cultural Practice among Latina Youth Gangs*. Malden: Blackwell Publishing.

Nance, Claire. 2013. Phonetic Variation, Sound Change, and Identity in Scottish Gaelic. Ph.D. dissertation, University of Glasgow, Glasgow, UK.

Nelson, Diane M. 1999. *A Finger in the Wound: Body Politics in Quincentennial Guatemala*. Los Angeles: University of California Press.

Nelson, Diane M. 2009. *Reckoning: The End of War in Guatemala*. Durham: Duke University Press.

Oertling, Emily. 2021. Made in San Pedro: The Production of Dress and Meaning in a Tz'utujil-Maya Municipality. Ph.D. dissertation, Kansas State University, Manhattan, KS, USA.

Ortíz-López, Luis Alfredo, and Eva María Suárez Büdenbender, eds. 2021. *Topics in Spanish Linguistic Perceptions*. New York: Routledge.

Otzoy, Irma. 1996. Maya clothing and identity. In *Maya Cultural Activism in Guatemala*. Edited by Edward F. Fischer and R. McKenna Brown. Austin: University of Texas Press, pp. 141–55.

Redacción Espectáculos. 2015. Polémica entre Miss Costa Rica y Miss Guatemala enciende las redes. *Prensa Libre*. Available online: https://www.prensalibre.com/vida/escenario/polemica-entre-miss-costa-rica-y-miss-guatemala-enciende-las-redes/ (accessed on 17 September 2022).

Reyes, Ingrid. 2019. María Merecs Coroy viste traje de Quetzaltenango en premier de Venecia y explica por qué utiliza diferentes trajes indígenas en festivales de cine. *Prensa Libre*. Available online: www.prensalibre.com/vida/escenario/maria-mercedes-coroy-con-traje-de-quetzaltenango-en-la-premier-de-venecia-y-explica-que-la-inspira-a-presentar-diferentes-trajes-tipicos-en-los-festivales-de-cine/ (accessed on 8 December 2022).

Romero, Sergio. 2015. *Language and Ethnicity among the K'ichee' Maya*. Salt Lake City: University of Utah Press.

Sam Chum, Eduardo. 2019. María Tun Cho completa el maratón de Los Ángeles con un mensaje claro: "Las mujeres podemos". *Prensa Libre*. Available online: www.prensalibre.com/deportes/deporte-nacional/maria-tun-cho-completa-el-maraton-de-los-angeles-con-un-mensaje-claro-las-mujeres-podemos/ (accessed on 17 September 2022).

Silverstein, Michael. 2003. Indexical Order and the Dialectics of Sociolinguistic Life. *Language & Communication* 23: 193–229. [CrossRef]

von den Berghe, Pierre L. 1968. Ethnic membership and cultural change in Guatemala. *Social Forces* 46: 514–22. [CrossRef]

Zhang, Qing. 2008. Rhotacization and the "Beijing Smooth Operator": The meaning of a sociolinguistic variable. *Journal of Sociolinguistics* 12: 201–22. [CrossRef]

Article

Exploring the Role of Phonological Environment in Evaluating Social Meaning: The Case of /s/ Aspiration in Puerto Rican Spanish

Christina García [1,*], Abby Walker [2] and Mary Beaton [3]

[1] Department of Languages, Literatures, and Cultures, Saint Louis University, St. Louis, MO 63108, USA
[2] Department of English, Virginia Polytechnic Institute and State University, Blacksburg, VA 24061, USA; ajwalker@vt.edu
[3] Department of Modern Languages, Denison University, Granville, OH 43023, USA; beatonm@denison.edu
[*] Correspondence: christina.garcia@slu.edu

Abstract: Research in sociophonetic perception has suggested that linguistic factors influence the social meaning of a particular variant, such that the strength of social meaning appears to be mediated by factors like grammatical category or phonological environment. Here, we further investigate the impact of linguistic factors on the perception of sociolinguistic variables by examining evaluations of /s/ aspiration in the speech of four male Puerto Rican Spanish speakers. We look at how evaluations of this variable pattern based on the phonological context (preconsonantal vs. prevocalic), the proportion of a given variant ([s] or [h]) in the stimuli, and the listener residence (Puerto Rico vs. mainland US). Our results replicate earlier work showing that /s/ realization contributes to status and masculinity ratings. However, we do not find evidence of an effect of incremental changes in the proportions of [s]:[h] variants in an utterance or an effect of listener residence. Critically, we do find that phonological context influences the evaluations of listeners: [s] is rated as less masculine than [h] in preconsonantal environments, but in prevocalic environments, there is no effect of variant. Given that [s] is rarely found in preconsonantal contexts in Puerto Rican Spanish, and even less so in male speech, this result is consistent with studies arguing that social meaning is stronger in marked contexts. Expected patterns for gender, phonological context, and dialect interact to make an [s] realization of preconsonantal /s/ particularly rare in male speech of this variety, which opens the door for more robust socioindexical meaning.

Keywords: sociolinguistic monitor; Puerto Rican Spanish; /s/ aspiration

1. Introduction

A key contribution of work in the third wave of sociolinguistics has been the insight that the social meaning of variants is not fixed (Eckert 2012, p. 94). Different listeners respond differently to the same variant based on their linguistic background (Labov et al. 2011; Levon and Fox 2014), and even the same listener will respond differently to the same variant depending on other information they have about the speaker. This includes the regional or ethnic dialect of the speaker (Campbell-Kibler 2009; Pharao et al. 2014; Walker et al. 2014), the way the speaker pronounces other sounds in the utterance (Campbell-Kibler 2009; Levon and Buchstaller 2015), whether and/or how (in)frequently a particular talker is known to use a variant based on diverse social characteristics, such as gender (Podesva et al. 2015; Pharao and Maegaard 2017), and what the speaker is talking about (Montgomery and Moore 2018).

The cognitive mechanism responsible for speaker evaluations has been called the sociolinguistic monitor (Labov et al. 2011). Levon and Fox (2014) emphasize the role that salience has in the process of monitoring speech, both in terms of whether listeners notice a given feature and whether listeners associate that feature with a type of speaker or

trait. Crucial here is the idea of markedness; while two (or more) pronunciations may be available to a speaker, one variant (the unmarked form) is more expected, while the other (the marked form) stands out because of its unexpectedness.

Marked and unmarked variants fare quite differently in evaluative tasks. In their work to establish the profile of the sociolinguistic monitor, Labov et al. (2011) found a logarithmic response in their listeners (US undergraduate students) to the percentage of (ING) tokens realized with the alveolar variant: a single alveolar realization (10% of all (ING) tokens in the stimuli) had a considerable impact on professionalism ratings, but increasing the percentage of alveolar variants had a diminishing effect on ratings. That is, the difference between 0% and 10% alveolar realizations was larger than the difference between 10% and 20%, which was larger than the difference between 20% and 30%, etc. Critically, this pattern was in response to the marked (alveolar) variant in the context of broadcast speech, and introducing a single token of the unmarked velar variant to stimuli that were dominated by alveolar variants had no significant impact on ratings.[1]

In a follow-up study, Levon and Fox (2014) extended the idea of markedness to broader sociolinguistic norms within a community. Despite the alveolar variant also being the nonstandard of the two realizations in British English, they found no effect at all of (ING) realization on professionalism ratings by their British participants. They argued that within the British context, (ING) realization is a weaker marker of class identity than in the American context, and, therefore, the association between class and (ING) is less automatically activated for these listeners.

This work demonstrates how markedness is completely intertwined with context. While most speaker perception work has focused on markedness coming from the social context, investigations into sociolinguistic patterns usually find linguistic constraints on variation as well—what Labov (1994) calls linguistic internal factors. For example, in his famous study of rhoticity in New York City department stores, Labov (1972, p. 52) identified not only an effect of socioeconomic class and attention paid to speech but also of whether the rhotic token is coda final (floor) or not (fourth).

Several studies have suggested that linguistic factors can influence social evaluations of variants.[2] Podesva et al. (2015) found a more consistent impact of /t/ realization on speaker evaluation when the /t/ is word-medial compared to word-final. They argued that because released /t/ variants are less common word-medially, they are also "less predictable, thus more salient when heard, and therefore endowed with greater potential for carrying social meaning" (p. 81).

A similar argument was made by Vaughn (2022b), who looked at (ING) realization. It is well established in speech production studies that the likelihood of the alveolar variant of the variable (ING) depends on grammatical category; it is much more common to hear the alveolar variant on verbs (i.e., I was runnin') than on nouns or gerunds (i.e., Runnin' is fun). Vaughn and Kendall (2018) reported that listeners indicated being more surprised by the alveolar variant in less likely environments (like nouns), and, in an *-in'* monitoring task, were more likely to notice the alveolar variant in less likely environments. Vaughn (2022b) then showed how this sensitivity to variants in marked environments impacts speaker perceptions: speakers are rated as more accented when the *-in'* appears on word classes like nouns versus verbs.

The impact that linguistic factors have on the social evaluations of variables appears to depend both on the variant and the listener's experience with the dialect in question. For the (ING) variable, Vaughn (2022b) found that grammatical category interacts with perceptions of the (marked) alveolar (ING) variant, but there was no effect for the (unmarked) velar variant; *-ing* was rated as equally unaccented across all grammatical categories. Similarly, for zero copula in African American Vernacular English (AAVE), Bender (2000) reported that the listeners' race and, by extension, familiarity with copula presence/absence impacted their evaluations. This variable is grammatically conditioned, such that zero copula is more likely preceding a progressive verb (e.g., "she ['s/0] teachin' me piano") than it is preceding a noun phrase (e.g., "she ['s/0] my piano teacher"). Non-Black listeners, even

those who report being familiar with AAVE, rated the speakers differently depending on copula realization, but their ratings did not depend on the following word type, which Bender interpreted as a lack of familiarity with internal constraints of the dialect. For Black listeners who identified as *not* speaking AAVE, their evaluation of zero copula (but not copula presence) was impacted by following word type, such that zero copula received more negative ratings in its most marked environment (e.g., before a noun). However, for Black listeners who identified as speaking AAVE, it was their evaluation of copula presence (but not copula absence) that was impacted by grammatical category, such that copula presence was rated higher in its most marked environment (e.g., before a verb).[3] Bender interpreted these results as reflecting which variant is marked and/or least frequent for each group (p. 185).

Importantly, Bender's work suggests that markedness can operate independently of standardness, despite the fact that the two are often conflated. For example, Wagner and Hesson (2014) state that "listeners form an impression of the speaker as they attend to the speaker's frequency of marked (i.e., unexpected or relatively salient) features such as swear words, teenage slang, and nonstandard or nonnative-sounding grammar and pronunciation" (p. 652). For speakers of AAVE, the nonstandard variant, zero copula, also appears to be the unmarked variant based on its frequency within the dialect. It is the standardized variant, copula presence, that is marked, especially in contexts where it least frequently occurs.

In this paper, we further explore the role that linguistic factors have on the evaluation of sociolinguistic variants, focusing on another variable where the most frequent variant in the community is not the prescriptive norm: coda /s/ realization in Puerto Rican Spanish. We investigate patterns of variable evaluation depending on the percentage of a given variant in the stimuli (Labov et al. 2011; Levon and Fox 2014) and the linguistic context (Bender 2000; Podesva et al. 2015; Vaughn 2022b). Additionally, we consider how the listeners' ambient environment might influence their social sensitivity to the variable by comparing listeners who currently live in Puerto Rico (islanders) to those living in the continental US (mainlanders).

2. Spanish Coda /s/

2.1. Coda /s/ Production in Puerto Rican Spanish

Puerto Rican Spanish is among the approximately 50% of Spanish dialects described as /s/-weakening varieties (Hammond 2001), meaning that /s/ in coda position is variably pronounced as a retained sibilant [s], lenited to an aspirated variant ([h]), or deleted altogether (∅) (Cedergren 1973; Guitart 1976; Lipski 1985).[4] In Puerto Rican Spanish, the aspirated variant is overwhelmingly the most frequently used form; in its most favored linguistic environment(s), studies report rates as high as 81% (López Morales 1983) or 92% (Terrell 1978). The deleted variant is usually the next most frequent form, followed by sibilant [s] production, which is very infrequent, with studies reporting rates as low as 2% in disfavored environments (Terrell 1978).

Despite its low frequency in the dialect under study, there is evidence that sibilant [s] is the prescriptive norm (Lipski 1983) and is "generally recognized as the prestige variant" (Mack 2009, p. 27) across the Spanish-speaking world. This is reflected in speaker perception studies, where speakers using [s] are rated as being of higher status (Walker et al. 2014), and in speech production studies, which find that [s] usage increases in more formal styles (Lafford 1986), for speakers of higher social classes (Calles and Bentivoglio 1986; Lipski 1983; López Morales 1983), and for women (Cameron 2005; Terrell 1981), who commonly use more standard forms than men (Labov 1972, p. 55). However, [s] being prestigious does not mean, conversely, that [h] is low prestige; rather, aspiration ([h]) appears to be socially neutral in many weakening dialects (Lynch 2009, p. 772) and is still used more frequently than [s] in the Caribbean even by educated speakers in formal situations (Terrell 1982).[5]

Of great relevance to the present study, work on a variety of Caribbean dialects has shown that the segment following /s/ conditions variation, with the sibilant [s] being maintained more frequently when followed by a vowel (e.g., *las amigas* 'the friends') than when followed by a consonant (e.g., *las casas* 'the houses') (Alba 2000; Cedergren 1973; Lipski 1985; Lynch 2009; Poplack 1979; Terrell 1978). This appears to relate to the historical trajectory of /s/ aspiration, which began in preconsonantal contexts and from there expanded to prevocalic contexts (Lipski 1995, p. 291). Lynch (2009) found that the following segment was the strongest predictor of /s/ in his study on Cuban Spanish in Miami, with a 29% [s] retention rate before vowels and 14% before consonants. Poplack (1979) divided her study of /s/ in Philadelphia Puerto Rican Spanish into word-internal contexts, which are always preconsonantal, and word-final contexts, which can be followed by a consonant or vowel. Her statistical analysis of word-final contexts, like Lynch's overall findings, selected the following sound as the leading predictor of [s], with weakening occurring more frequently before consonants (factor weight 0.93) than before vowels (0.22). In San Juan, Puerto Rico, Terrell (1978) also found much lower rates of [s] for word-final preconsonantal (2%) than prevocalic contexts (18%) and found similarly low preconsonantal sibilant rates word-internally (3%). In an investigation of news broadcasts in the Dominican Republic (Alba 2000), rates of [s] were 50% before consonants and 70% before vowels, showing that the pattern of more frequent [s] before vowels holds even in highly formal contexts in which the prescriptive sibilant is more common than in everyday speech. Given the preponderance of /s/ aspiration and deletion in Caribbean Spanish broadly, sibilant [s] may be considered the "marked" variant in Puerto Rican Spanish, particularly in a preconsonantal environment, as it is quite infrequent.

It is important to take into account that the linguistic background of the Puerto Rican population is heterogeneous. Due to their status as US citizens since 1917, travel to and from the island is common, and generations of Puerto Ricans live in the mainland United States. In fact, the population of stateside Puerto Ricans was 5.6 million in 2017 (Pew Research Center 2019), which far outnumbers the current island population of 3.2 million (U.S. Census Bureau 2020). Puerto Ricans residing on the mainland are far more likely to both be bilingual and have frequent contact with speakers of other dialects of Spanish, most notably Mexican Spanish, which overwhelmingly retains the sibilant [s] in coda (e.g., Lipski 1994). Studies comparing the production of /s/ among Puerto Ricans residing on the island to those on the mainland have shown mixed results. Poplack (1980) and Ramos Pellicia (2012) found lower rates of coda /s/ deletion among Puerto Ricans in Philadelphia and Lorain, Ohio, respectively, than on the island, while Ghosh-Johnson's (2005) Puerto Rican participants in Chicago mirrored islander /s/ production. O'Rourke and Potowski (2016) found that generation is not a significant predictor of /s/ production for Puerto Ricans in Chicago and showed that the conditioning factors were consistent across generations. Erker and Reffel (2021) also did not find generational differences among Caribbeans in NYC and Boston. Studies on other variables in Puerto Rican Spanish, such as /r/ velarization and lateralization, have also displayed varied results in regard to whether living on the mainland influences production (Arias Alvarez 2018; Ramos-Pellicia 2007; Valentín-Márquez 2020). Given the heterogeneous linguistic experiences of Spanish speakers in the United States, conflicting results are not unexpected.

2.2. Social Meaning of /s/ Aspiration

Following the emergence of "third wave" (Eckert 2012) studies that examine the indexical meaning of linguistic variables, scholars have considered the social meanings of /s/ realization among listeners of both aspirating and non-aspirating dialects. For Puerto Rican Spanish, Mack (2009) found a connection between aspiration and perceived sexuality, such that syllable-final [s] is implicitly associated with a Puerto Rican speaker being gay. In a matched guise study, Walker et al. (2014) compared the evaluations of Puerto Rican and Mexican speakers by listeners from both dialects. In line with Mack (2009), they found that overall [s] was rated as less heteronormative than [h] regardless of speaker nationality.

However, while female Mexican listeners exhibited no significant difference in their ratings of heteronormativity, a trend in the opposite direction was observed in which they rated [s] as more heteronormative. This once again highlights the influence of context on social meaning. For measures of status, they found that overall, [s] was associated with higher status than [h]; however, the difference in status ratings between the two guises was much smaller when listeners heard a Puerto Rican speaker than when they heard a Mexican speaker. Thus, there are both local and global social meanings tied to /s/ aspiration.

The nuance of the indexical meanings of /s/ aspiration is further confirmed by Chappell (2019a), who examined Mexican listeners' evaluations of a subset of the same stimuli utilized by Walker et al. (2014). In rating Mexican and Puerto Rican speakers, these listeners again generally associated [s] with higher status; however, there was a significant interaction between variant and speaker nationality in evaluations of "goodness of Spanish." This interaction showed that regional expectations condition listener evaluations given that Mexican speakers were rated as speaking better Spanish in the [s] guise while Puerto Rican speakers were rated as speaking better Spanish in the [h] guise. Finally, Chappell (2019a) showed a strong association between coda [h] and Caribbean identity, as both Mexican and Puerto Rican speakers were rated as more Caribbean in the [h] guise.

Although previous studies have not compared mainlander and islander Puerto Ricans' perceptions of /s/ aspiration, some work has been carried out examining Mexican Americans' sociophonetic perception, providing initial indications of heritage speakers' evaluations. Using the same stimuli as Walker et al. (2014), Chappell (2021) found that overall perceptions of /s/ aspiration among second-generation Mexican Americans largely matched those of Mexicans in Mexico. For both groups of listeners, coda [s] indexed higher status and confidence, while coda [h] signaled Caribbean identity. This similarity in perceptions of monolingual Spanish listeners and bilingual US listeners is further supported when examining other sociolinguistic variables. In the evaluation of labiodentalized <v>, Chappell (2019b) observed that [v] was evaluated positively in women's speech and negatively in men's speech by both monolingual Mexican Spanish listeners and bilingual second-generation Mexican American listeners.

2.3. The Present Study

In the present study, we build on Walker et al.'s (2014) methodology, in which listeners were presented with utterances where all instances of preconsonantal coda /s/ were either aspirated (e.g., "…*queda entre la e*[h]*cuela y el ho*[h]*pital* 'it's between the school and the hospital') or sibilants (e.g., "…*queda entre la e*[s]*cuela y el ho*[s]*pital*"). In the present study, we add two new types of stimuli: (1) sentences with all [s] or all [h] in *prevocalic* environments and (2) sentences with variable /s/ realizations (i.e., both [s] and [h]) in preconsonantal environments. The former tokens shed light on the role of phonological environment on speaker evaluation, and the latter allow us to explore the impact of the proportion of different /s/ realizations. To consider the impact that listener background might have on /s/ perception, we collected data from both Puerto Ricans living in Puerto Rico and from Puerto Ricans who live in the mainland US.

Our research questions are as follows:

1. Does /s/ realization impact listeners' perception of speakers of Puerto Rican Spanish? In other words, can we replicate the findings of Walker et al. (2014)?
2. Does the impact of /s/ realization on speaker ratings depend on the phonological context in which the /s/ is produced (cf. Vaughn 2022b; Bender 2000)? Is the impact stronger in prevocalic environments (the relatively less common setting for [h] in Puerto Rican Spanish) or in preconsonantal environments (the relatively less common setting for [s])?
3. What is the shape of participant response to different proportions of preconsonantal [s] vs. [h] realizations in a single utterance? Do we see evidence of a logarithmic response (Labov et al. 2011), a linear response (Levon and Fox 2014; Vaughn 2022a), or a flat response (Levon and Fox 2014)?

4. Are any of the above effects mediated by the residential status (islander/mainlander) of the listener?

For the first research question, we expect to replicate the findings of Walker et al. (2014), showing that /s/ realization impacts social perceptions. There is abundant evidence that this variable is socially marked, and thus, it is likely that will be seen in our results. Furthermore, we expect to see an effect of phonological environment on speaker ratings, such that we observe a larger difference in ratings between [s] and [h] depending on the following environment (preconsonantal or prevocalic). Which environment results in a bigger effect depends on which of the two variants is the most marked: [s] is very infrequent, especially preconsonantally, and less expected, so we predict larger social meaning here, following previous studies showing more robust social meaning in marked contexts. Alternatively, given that [s] is considered prescriptively correct, we could find that it is responses to [h] that are most influenced by the following environment, with aspiration being relatively more unexpected in prevocalic position.

Given mixed findings in the literature to date, we enter this study without strong hypotheses for the third research question, although based on Walker et al. (2014), who showed differences between all and no [s], we do not expect a flat response. Finally, it is hard to predict whether residential status (islander/mainlander) will have an effect on ratings, as there are so few previous studies in this specific realm; however, we expect that Puerto Ricans residing in the mainland US may have different social evaluations given that they are generally exposed to more dialects of Spanish in their daily lives and may or may not have had formal schooling in Spanish.

3. Materials and Methods

3.1. Stimuli

The stimuli for the present study, consisting of a slightly modified version of the original recordings made by Walker et al. (2014), include four Puerto Rican male speakers who were between the ages of 19 and 35 and had lived in the mainland United States for one to nine years at the time of recording. The speakers were recorded giving directions in a simple map task that was designed to elicit coda /s/ in two linguistic environments: preconsonantal word-internal (i.e., *esquina* 'corner') and word-final prevocalic (i.e., *las Américas* 'the Americas'). Upon completion of the map task, each speaker was asked to repeat the target words with an alveolar fricative [s] and with aspirated [h]. Given that Puerto Ricans use both, they had no issues producing the desired variants.

Excerpts were taken from these recordings to create two sets of stimuli: "context" and "additive." "Context" stimuli were designed to test whether the following phonological environment had an effect on listener perceptions of [s]:[h]. For these stimuli, two sentences were extracted per speaker, one with two tokens of preconsonantal word-internal /s/ (such as *esquina* 'corner') and one with two tokens of word-final prevocalic /s/ (such as *vas a* 'you are going to'), as seen in examples (1) and (2).[6] There were no other tokens of /s/ in the sentences besides the target preconsonantal or prevocalic /s/. Two versions of each sentence were made: one with spliced [h] for all tokens of /s/ and one with spliced [s] for all tokens of /s/. Only three of the four speakers were included in this study because the fourth speaker did not produce a usable string of two prevocalic tokens; thus, splicing resulted in 12 "context" stimuli.

(1) Preconsonantal: *Vira a la derecha, en la esquina a la izquierda.*
    'Turn right, at the corner to the left.'
(2) Prevocalic: *Y cuando llegues a la avenida de la República, vas a virar a la derecha.*
    'When you arrive at *Avenida de la República*, you will turn right.'

The "additive" stimuli served to test whether listener evaluations would change incrementally based on the proportion of [s] to [h] in an utterance. For these stimuli, one sentence containing three tokens of preconsonantal /s/ was extracted per speaker. Eight versions of each sentence were made by splicing [s] and [h] into the three target words

in every possible combination, as seen in examples 1–8 below. This produced versions with a range of three tokens of [h] (all aspirated) to three tokens of [s] (all sibilant). In the case of one or two tokens of [s], the order was varied so that [s] and [h] appeared in all possible orders in the sentence; this also means that we have more instances of sentences with variable /s/ realizations (vs. zero [s] or three [s]). Since all four speakers were used for this set of stimuli, the splicing resulted in 32 "additive" stimuli. See Appendix A for the full set of stimuli and listen to the sound files at https://osf.io/a6nw2/?view_only=8ae951 2090af46e6a8e3ad90dd1d847a.

1. zero [s]: *E[h]tá entre el ho[h]pital y una e[h]cuela elemental.*
2. one [s]: *E[s]tá entre el ho[h]pital y una e[h]cuela elemental.*
3. one [s]: *E[h]tá entre el ho[s]pital y una e[h]cuela elemental.*
4. one [s]: *E[h]tá entre el ho[h]pital y una e[s]cuela elemental.*
5. two [s]: *E[s]tá entre el ho[s]pital y una e[h]cuela elemental.*
6. two [s]: *E[s]tá entre el ho[h]pital y una e[s]cuela elemental.*
7. two [s]: *E[h]tá entre el ho[s]pital y una e[s]cuela elemental.*
8. three [s]: *E[s]tá entre el ho[s]pital y una e[s]cuela elemental.*
   'It's between the hospital and the elementary school.'

In both sets of stimuli, spliced material came from the target word elicitation that the speakers recorded after the map task. In some cases where the /s/ was in an unstressed syllable, the duration of the spliced [s] was reduced to ensure it would sound natural. The amplitude of the spliced variants did not need to be adjusted. All stimuli were checked by four native speakers, two of whom were trained linguists, who agreed that the recordings sounded natural and would be interpreted as the desired variant ([s] or [h]) by naive native listeners.

3.2. Experimental Design

The stimuli were presented in an online survey hosted on Qualtrics (Provo, UT) that targeted native speakers of Puerto Rican Spanish. After consenting to participate in the experiment, participants completed a basic demographic questionnaire in which they reported their age, gender, where they were born, where they currently resided, in which cities they have lived and for how long, whether the majority of their friendships were with Puerto Ricans, and whether they used headphones or speakers to listen to the survey stimuli.[7]

After the demographic questionnaire, participants were presented with the stimuli and asked to evaluate each speaker, as seen in Figure 1. Following Walker et al. (2014) and work on perceptual dialectology (e.g., Preston 1999), participants rated each speaker on a six-point scale for the following social characteristics: *menos/muy masculino* ('less/very masculine'), *antipático/simpático* (roughly 'unpleasant/pleasant'), *inseguro/seguro de sí mismo* ('unconfident/confident of himself'), *menos/muy educado* ('less/very educated'), and *de clase baja/de clase alta* ('low class/high class').[8] They were also asked to estimate the age of the speaker and were given the option to leave a text comment about the speaker.

While there were 44 unique stimuli, each participant only heard 20. This was carried out both to keep the online experiment short to decrease attrition and because, of those 44 stimuli, some were eight versions of the same original sentence with different combinations of [s] and [h] variants (see Section 3.1), which we believed might highlight the repetition in voices and potentially frustrate or bore the participants. The 20 tokens that participants heard consisted of all 12 context tokens (both the [s] and [h] variants of the prevocalic and preconsonantal sentences from three speakers) and 8 additive tokens (two of the eight versions of the sentences produced by the four speakers). Which two additive tokens participants heard was semi-randomized (for speakers 1 and 2, it was either the sequence of tokens (1,3,5,7) or (2,4,6,8) and, for speakers 3 and 4, either (1,2,3,4) or (5,6,7,8)). The experiment was split into two halves, and each half had a set trial order, such that if a listener heard speaker 1's additive sentence first in the list, the second time they heard

the same sentence (with a different version), it would be eleventh in the list (10 trials later). Which half the participants began with was randomized, and there were two different lists (into which participants were assigned) so that the order of trials somewhat differed across listeners. We decided not to include any filler stimuli in order to keep the experiment concise. We did not feel this would be problematic given that the vast majority of participants in Walker et al. (2014) did not exhibit knowledge of which variable was being studied. In the present study, no participants specifically pointed out /s/ in the comment field at the end of the survey, which suggests it is unlikely that they were aware that coda /s/ was the variable under investigation.
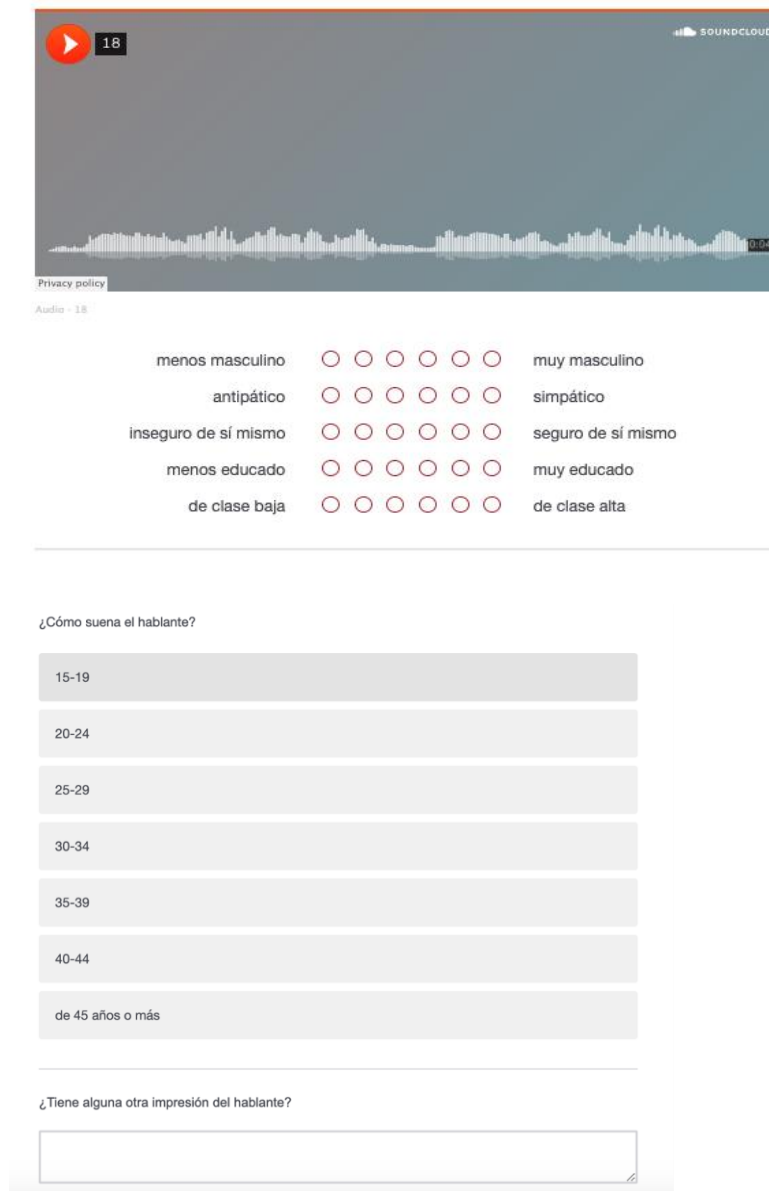
Figure 1. Screenshot of the online experiment on Qualtrics.

It is worth noting that this experimental design differs markedly from the classic paradigm that Labov et al. (2011) used to investigate the impact of different proportions of variant realization on speaker perceptions, which has also been used by researchers directly extending that work (Levon and Fox 2014; Levon and Buchstaller 2015; Wagner and Hesson 2014; Vaughn 2022a). This was a pragmatic choice (our study uses stimuli from Walker et al. 2014) but also reflects concerns expressed in the literature about the design of

Labov et al. (2011), where participants heard the same recording 10 times in a row, starting with no unmarked variants and changing only by the addition of one marked variant on each listen. As Vaughn notes, this design means that the researcher's variant of interest is "fairly transparent after hearing several versions of the passage" (Vaughn 2022a, p. 517). In our study, participants heard only two versions of each utterance, with nine other trials in between, and which version they heard was randomized. It is also worth noting that the stimuli in Labov et al.'s study had 10 tokens of the variable of interest, while our additive stimuli had only 3. Given that Labov et al.'s (2011) results suggest that the greatest changes were observed with the introduction of a single unmarked variant and were discussed in terms of proportions of variants, not raw counts, we believe that our findings can still offer insights into discussions of the sociolinguistic monitor.

3.3. Participants

Initial data collection was conducted from fall 2015 to fall 2018, recruiting participants through friends and colleagues via email. In order to complete the study, participants had to be at least 18 years old and be native speakers of Puerto Rican Spanish. We indicated that we were interested in participants from both the mainland and the island. Completion of the survey was compensated with a USD 5 gift card/voucher through Amazon, Paypal, or Venmo. While collecting additional data in October 2018, we realized the survey had been hacked by bots given that the demographic information collected was nonsensical. We increased security on the survey by adding a password and attempted to recruit more participants by posting the survey in Facebook groups targeting Puerto Ricans in May 2021; however, the survey was hacked again, this time by people who took the survey multiple times, as evidenced by repeated IP addresses and answers on the demographic questionnaire. We examined the data from October 2018 and May 2021 to differentiate real and fake responses, and suspect data, containing repeated IP addresses or nonsensical information, were not included in our analysis. We also removed participants who completed less than 75% of the study and one participant who responded with the exact same answer to every attribute for every speaker. In total, we include data from 89 participants from this initial data collection.

To protect against hacking, beginning in the summer of 2022, we used the platform Positly to continue data collection by distributing our Qualtrics survey through both Amazon's Mechanical Turk (MTurk) and via a traditional panel platform (TPP). Participants were compensated USD 3.80–6.50 through Positly; compensation varied according to platform standards. Because the TPP distributes the survey to companies who recruit workers from specific zip codes, this method was particularly effective to recruit Islanders. Positly blocks suspicious and duplicate IP addresses and allows researchers to include attention checks and language questions to filter out unsuitable participants (see Appendix B). To proceed to our study, participants on both MTurk and TPP had to pass all attention checks, claim to be a speaker of Puerto Rican Spanish, and correctly name two of four images with the Puerto Rican term for the depicted item. This last step was especially important for US participants in order to exclude speakers of other dialects of Spanish from our dataset. In our analysis, we include data from 145 participants who met these criteria, finished at least 75% of the study, and did not respond with the same answer to every attribute for all speakers (1 participant). As explained in Section 3.2, participants who finished 100% of the study only heard 20 of the total 44 stimuli; thus, participants who completed 75% of the study did not hear a significantly smaller portion of the stimuli than those who completed the survey in its entirety.

Combining the two datasets, the responses of a total of 234 listeners were included in our analysis. The demographic information of the listeners is presented in Table 1. There are over twice as many listeners residing on the island as there are on the mainland US, which is a byproduct of the fact that it was easier to recruit island residents on Positly. Unsurprisingly, the vast majority of the island residents were also born on the island; however, it is notable that most of the mainland residents were also born on the island. The

mean age of the two populations is very similar, as well as the proportion of responses from male versus female listeners, with female listeners representing nearly 60% of the data in both populations. The proportion of life in Puerto Rico was calculated by dividing the number of years the listener had spent living on the island by their age. While the average proportion of life in Puerto Rico is much higher (0.97) for island residents compared to mainland residents (0.45), the range for both populations is quite large, which reflects the migratory nature of Puerto Ricans mentioned in Section 2.1.

Table 1. Listener demographics.

	Island Residents	**Mainland Residents**
No. Total	161	73
No. Born (PR/US/Other)	151/9/1	54/19/0
Mean Age * (range)	35.5 (18–67)	35.6 (18–72)
Gender (Female/Male/Non-Binary)	94/67/0	43/28/2
Proportion of life in PR * (range)	0.97 (0.14–1)	0.45 (0–0.96)

* Age (and thus 'Proportion of life in PR') is missing for 4 participants.

3.4. Data Analysis

Walker et al. (2014) found that some of the seven different attributes on which participants were asked to rate speakers were correlated, reflecting broader latent variables. Specifically, class and education loaded highly onto a factor called "status," and masculinity and sexuality loaded onto a factor called "heteronormativity." In the present study, we similarly ran a factor analysis, this time on six attributes (because we did not include a question on sexuality), and the Kaiser rule and parallel analysis (Bandalos and Boehm-Kaufman 2008; Weatherholtz et al. 2014) indicated that our factor analysis should have 3 factors. When we ran this model,[9] the first factor was again status, with high loadings for education (0.93) and class (0.72) and a loading above 0.4 for confidence. The second factor appeared to primarily capture pleasantness ratings (0.99), and the third captured masculinity (0.62). In the following analyses, we default to using the status factor as a dependent variable in our model, but given that the last two factors consist of one factor, we revert to raw responses for pleasantness and masculinity ratings, as well as age ratings, which did not load above 0.3 on any factor.[10]

We used linear mixed-effects regression models, with random effects for the participant and speaker. We tested for main effects of the /s/ variant (Section 4.1), the number of [s] in an utterance (Section 4.2), and an interaction between the /s/ variant and phonological environment (Section 4.3). In all models, we also test for interactions between these factors and listener residency.[11] We chose the best-fit models by starting with a full model and reducing the model by log likelihood comparisons of more and less complex models.

4. Results

4.1. [s] vs. [h]: Replicating Walker et al. (2014)

We begin our analysis by examining whether we can replicate the basic finding of Walker et al. (2014) that [s] variants are rated as less masculine, higher status, and more pleasant than [h] variants in preconsonantal environments. For this analysis, we include all tokens where the /s/ realizations were either all [s] or all [h]: the extremes of the additive tokens (with zero [s] or three [s]) and the preconsonantal subset of the context tokens (total *N* = 1796). There were no significant effects of /s/ realization or listener residency on pleasantness or age ratings in this analysis. Thus, we only report masculinity and status results here.

Looking first at masculinity (Figure 2), we find a robust effect of /s/ realization on masculinity ratings in the expected direction: tokens with all [s] are rated as less masculine (mean = 4.27) than tokens with all [h] (mean = 4.45). This difference is significant (Table 2) and holds in subset islander and mainlander models as well. It also holds if we use the masculinity factor values produced from the factor analysis. Listener residency did not

improve the model as a main effect or in interaction with /s/ realization. That is, there is no statistical difference between islanders' and mainlanders' ratings of masculinity.

Figure 2. Aggregate masculinity scores by listener residence and /s/ realization (higher rating = more masculine).

Table 2. Summary of best mixed-effects model for masculinity (*N* = 1796).

	Estimate	*SE*	*t* **Value**	*p* **Value**
Intercept	4.28508	0.25747	16.643	<0.001
Variant = [s]	−0.16797	0.04727	−3.553	<0.001

Note: Random effects = (1 | participant) + (1 | speaker).

Figure 3 shows the impact that /s/ realization has on status ratings: tokens with [s] are rated as higher status than tokens with [h]. This is significant in our mixed-effects model (Table 3). Again, there is no effect of listener residency, and the main effect of /s/ realization holds in islander and mainlander subsets of the data.
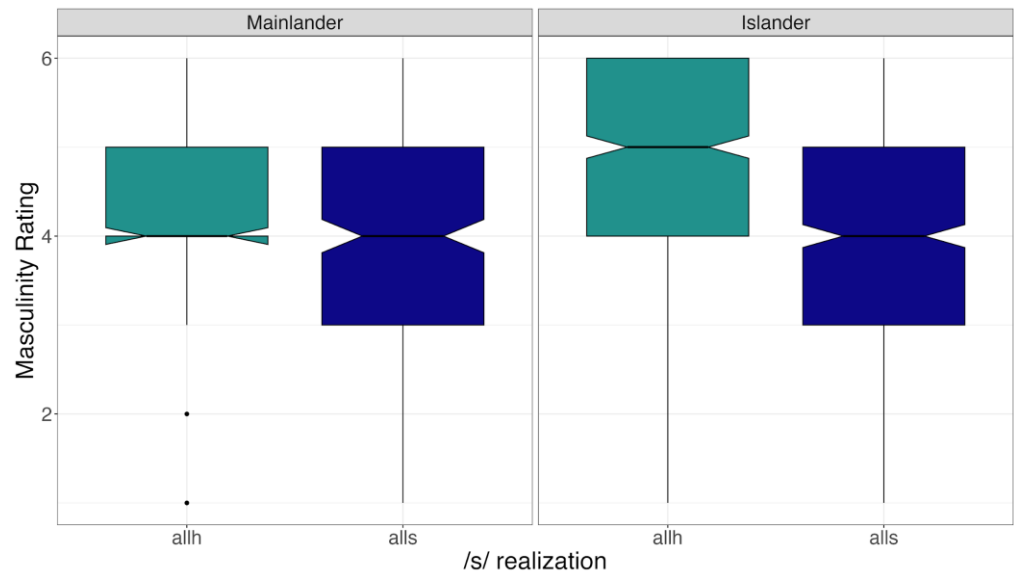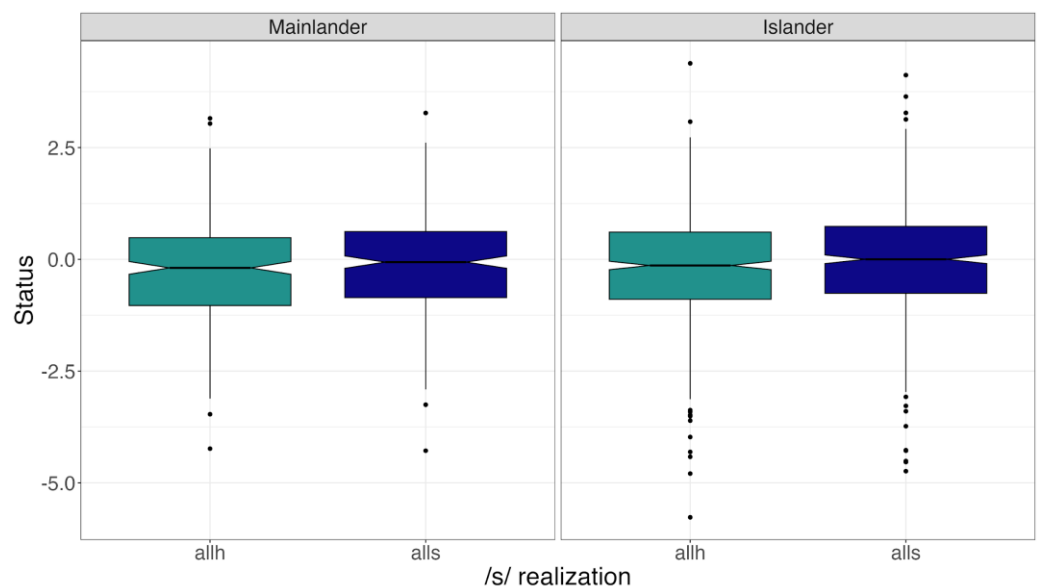

Figure 3. Aggregate status scores by listener residence and /s/ realization (higher rating = higher status).

Table 3. Summary of best mixed-effects model for status factor (*N* = 1796).

	Estimate	*SE*	*t* **Value**	*p* **Value**
Intercept	−0.17216	0.10648	−1.617	0.1950
Variant = [s]	0.13171	0.04727	2.786	0.0054

Note: Random effects = (1 | participant) + (1 | speaker).

To better understand the effect size and more easily compare status and masculinity, we can look separately at class and education ratings, the two components that contributed the most to the status factor in the factor analysis. In models with these ratings as the dependent variable, /s/ realization is not a significant factor (though it approaches significance in the education model (*p* = 0.07)). Thus, while speakers producing all [s] are rated as being of a higher class (mean = 3.84) and more highly educated (mean = 3.98) than the same speakers producing all [h] (mean class = 3.8, mean education = 3.9), these differences are very small (less than 0.08 on the six-point scale used). It is fair to say that the effect of /s/ realization on perceptions of speaker status is not as robust as the effect on perceptions of masculinity given its particularly small size and the fact that it is determined by which dependent variable is used.

4.2. The Effect of Phonological Environment on Ratings

For the next analysis, we consider only the tokens designed to test the role that phonological context might have in how much /s/ impacts speaker evaluation (N = 2724 given that all context stimuli are included). Specifically, we are comparing the impact of /s/ realization when the /s/ is prevocalic to when it is preconsonantal. In this and the following analyses, we report only on status and masculinity ratings, as pleasantness and age ratings exhibit no significant effect of /s/ realizations in this dataset.

Figure 4 shows the impact of /s/ realization on masculinity ratings (left) and status ratings (right) by the phonological environment. The results look very similar to the figures in our first analysis—this is unsurprising considering the two analyses share some data—and the effect of /s/ is again significant as a main effect for both masculinity and status ratings (Tables 4 and 5). For masculinity, there is a significant interaction between /s/ realization and phonological context (Table 4): using [s] makes speakers sound less masculine but only in preconsonantal environments.
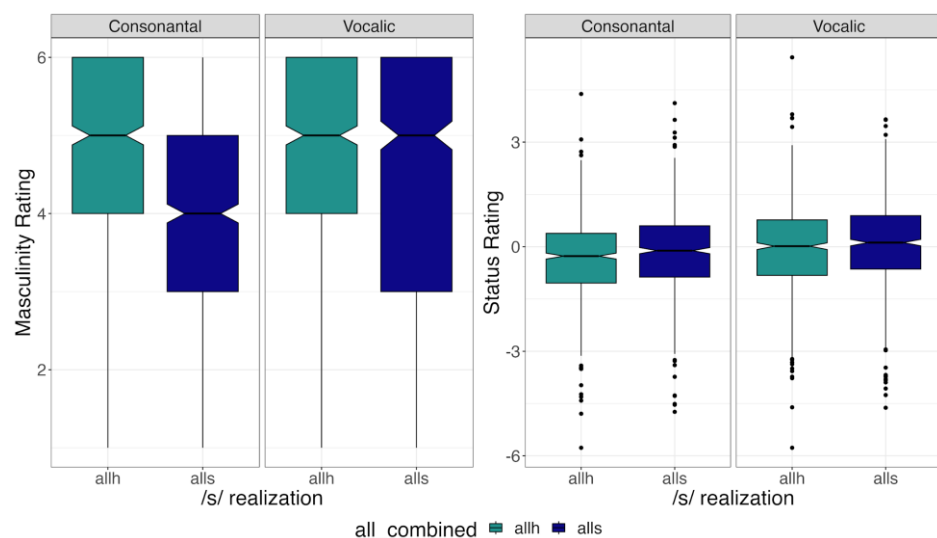
Figure 4. Aggregate masculinity (**left**) and status (**right**) scores by variant and phonological context for context stimuli (higher rating = more masculine/higher status).

Table 4. Summary of best mixed-effects model for masculinity for context stimuli (*N* = 2724).

	Estimate	SE	*t* Value	*p* Value
Intercept	4.52999	0.26459	17.121	0.002
Variant = [s]	−0.19852	0.05650	−3.514	<0.001
Type = Vocalic	−0.11540	0.05671	−2.035	0.042
Variant = [s]: Type = Vocalic	0.17323	0.08011	2.162	0.031

Note: Random effects = (1 | participant) + (1 | speaker).

In the status model (Table 5), there is a significant interaction between phonological environment and listener residence, such that Islanders rate the utterances with prevocalic /s/ as higher overall than those with preconsonantal /s/. Importantly, however, this effect is independent of /s/ realization and instead may reflect the fact that preconsonantal and prevocalic utterances were slightly different in other ways too. For instance, the prevocalic stimuli were more often made by joining together two clauses from different sentences, while the preconsonantal sentences were typically a single string of audio. Nevertheless, the content of preconsonantal and prevocalic stimuli was very similar (giving directions), and there are multi-clausal sentences in both types of stimuli.

Table 5. Summary of best mixed-effects model for status factor for context stimuli (*N* = 2724).

	Estimate	SE	*t* Value	*p* Value
Intercept	−0.35640	0.11028	−3.232	0.015
Variant = [s]	0.14377	0.03982	3.611	<0.001
Residence = Puerto Rico	0.02448	0.08958	0.273	0.785
Type = Vocalic	0.08802	0.07141	1.232	0.218
Residence = Puerto Rico: Type = Vocalic	0.28463	0.08603	3.309	<0.001

Note: Random effects = (1 | participant) + (1 | speaker).

4.3. Additive Effects of [s] and [h]

In this final analysis, we consider only the additive tokens (*N* = 1790). In these tokens, the utterance had three cases of preconsonantal /s/, and we are interested in the impact that the number of [s]/[h] realizations have on ratings. Figure 5 shows the effect of the number of /s/ realized as [s] on masculinity ratings (top) and status ratings (bottom) by listener residency. The lines are effectively horizontal, which indicates that the number of [s] realizations does not appear to impact speaker perception. This is supported by our statistical analyses, which show no effect of /s/ realization whether we treat the number of [s] as a numeric, categorical, or binary variable that splits (1) zero [s] vs. one or two [s], (2) three [s] vs. one or two [s], or (3) one or less [s] vs. two or more [s].

In fact, we do not even find an effect of all (three) [s] vs. zero [s], the comparison we made in our first analysis (if we subset our tokens to this comparison). This may be surprising, but it is important to remember that in the eight additive stimuli (see Section 3.1), there was only one version of each sentence with zero tokens of [s] and one version with all three [s], as the other six stimuli had different combinations of one or two [s] tokens. Thus, the lack of effect of an all:none comparison here probably reflects the much smaller number of observations in this subset (*N* = 427) and that the effect size is very small: the strongest effect we saw in Section 4.1 (for masculinity ratings) was an approximately .2 change on a six-point scale.
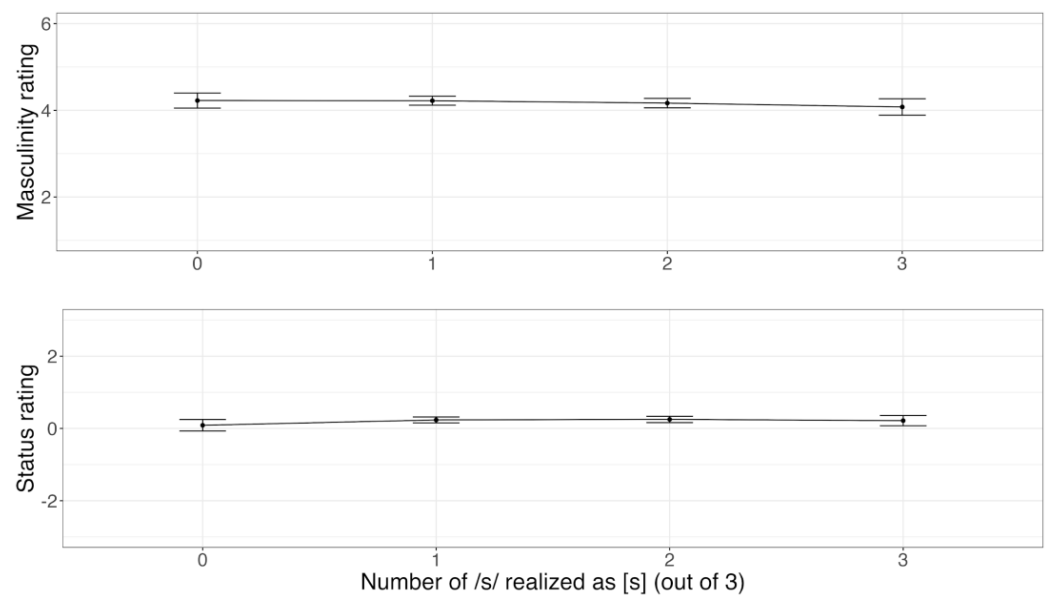
Figure 5. Aggregate masculinity (**top**) and status (**bottom**) scores by number of /s/ realized as [s] for additive stimuli (higher rating = more masculine/higher status).

5. Discussion

In the present study, we find that /s/ realization can impact masculinity and status ratings of male Puerto Rican speakers, such that they are heard as more masculine and of lower status with the [h] variant compared to the [s] variant. For masculinity ratings, this effect is only seen in preconsonantal environments; however, there is no effect of phonological context on status ratings. In examining the number of realizations of [s] or [h] that participants heard, we find that ratings of status and masculinity remain the same whether listeners heard zero, one, two, or three tokens of [s] in the stimuli.

Upon returning to our first research question, which considers the overall impact of /s/ realization on speaker evaluations, our analysis largely replicates the findings of Walker et al. (2014), who found that Puerto Rican listeners rated speakers using all [h] for preconsonantal /s/ as more masculine, lower status, and more pleasant than speakers using all [s]. We observe the same effect of /s/ realization on masculinity and status ratings but find no impact on pleasantness ratings. This may reflect the fact that the Walker et al. (2014) dataset also included ratings of Mexican speakers, and, while there was not a significant effect of speaker nationality in the pleasantness model in that study, it is possible that responses to Mexican speakers were driving the effect. The fact that the current study does not include Mexican speakers or listeners may then explain the lack of a significant result for pleasantness ratings.

In response to our second research question, concerning the influence of phonological context on social meaning, we do find an effect of phonological context on /s/ ratings such that listeners are more sensitive to /s/ realization in assessing masculinity in preconsonantal environments than they are in prevocalic environments. This effect appears to be driven by responses to the [s] variant (Figure 4, left panel): masculinity ratings are lower when there is an [s] in the preconsonantal environment. In interpreting these results, it is important to emphasize the Puerto Rican context for this study to disentangle notions of markedness and prescriptiveness. There is abundant evidence that, of the two variants we examine in this study, [s] is associated with prescriptive correctness: it is used more by speakers of higher social class (López Morales 1983) and in formal contexts (Lafford 1986) while also being rated as higher status in tasks like ours (including our findings in Sections 4.1 and 4.2 and those of Walker et al. 2014). At the same time, [h] is the overwhelmingly dominant variant in Puerto Rican Spanish (López Morales 1983; Terrell 1978), which means that most speakers that people encounter—be they particularly masculine, educated,

or not—will use [h] most of the time. This is supported by the fact that several listeners in Walker et al.'s (2014) study explicitly commented on the "foreignness" of [s] in Puerto Rican voices, and Chappell (2019a) found that Puerto Ricans (but not Mexicans) were rated as speaking "better Spanish" using [h].

While [h] is the expected realization for coda /s/ in Puerto Rican Spanish overall, this is especially true in preconsonantal environments: production studies find rates of preconsonantal [s] as low as 2–3% of the time (Terrell 1978). This renders [s] as particularly unexpected before a consonant. Bender (2000) argues that social meanings of variants are more salient in marked linguistic environments (see also Podesva et al. 2015; Vaughn 2022b), and our results are consistent with this: [s] more strongly indexes low masculinity when it appears in a particularly disfavored phonological environment. Since women tend to use [s] more than men (Terrell 1981; Cameron 2005), an [s] realization is particularly unexpected for our stimuli, as they feature only male speakers. As such, what is "expected" here depends on the relationship between dialect, gender, and phonological context. This interaction renders a preconsonantal [s] production in a male Puerto Rican voice as noteworthy, which opens the door for more robust indexical meaning to be tied to this variant in this particular context. However, in an environment where [s] is more likely, even if still a minority variant (prevocalically), masculinity ratings are not significantly affected by /s/ realization.[12]

Given that [s] is the minority variant in all contexts, it could be considered surprising that [s] is not always carrying out significant social work. That is, we might expect it to have a stronger meaning before consonants, because that is where it is particularly disfavored, but still to be carrying out some social work prevocalically. The fact that this is not the case might be related to something observed in the literature on lexical access: words with canonical but infrequent variants (i.e., released /t/ in "center") are recognized as easily as words containing the most frequent variant (i.e., /t/ deletion in "center"). Sumner et al. (2014) propose that such results are the product of a dual route of processing, where infrequent but prestigious variants are stored differently than other variants. As more studies, including our own, suggest effects of frequency on the social salience of a variant, it will also be valuable to consider other ways in which variants may become relatively unmarked to listeners.

Importantly, our work, like Bender's (2000), highlights that markedness can be distinct from prescriptive norms, primarily reflecting the frequency of use. It also highlights the contextual sensitivity of social meaning, and we predict that evaluations of female Puerto Rican Spanish speakers would be quite different given their overall higher production of coda [s] and the fact that female speakers are typically evaluated positively when using prestige variants (e.g., Chappell 2019b). Furthermore, we may expect a different pattern of markedness if we examined the impact of linguistic context on social meaning in a different variety, for example, Mexican Spanish. In such an /s/-retaining dialect, we predict that it is the social meaning of infrequent [h] that is more sensitive to phonological context, such that social meaning is intensified *prevocalically* (where [h] is less common across aspirating dialects (Hammond 2001)) or possibly that social meaning is not dependent on phonological context at all, as coda /s/ aspiration does not occur in either context in the vast majority of Mexican dialects (e.g., Lipski 1994, pp. 279–83).

Seeing as the present study included only word-internal preconsonantal /s/ and word-final prevocalic /s/ before an unstressed syllable, future studies examining phonological context should more carefully control for its interaction with stress and word position. Rates of /s/ aspiration differ before stressed and unstressed syllables and in word-internal versus word-final position (see Section 2.1), so it is possible that speakers may be attuned to this difference in their social evaluations of this variable. Vaughn (2022a) has also suggested that there might be lexically specific effects of phonological context, such that a word-final [s] could be rated differently not simply because it is currently, in this utterance, before a consonant but if it is a word that typically appears before a consonant and therefore is especially likely to be realized as [h] (regardless of the current phonological environment, cf. Guy et al. 2008).

Our third research question concerned the different proportions of [s] versus [h] realizations in a single utterance, and we do not see evidence of a logarithmic (Labov et al. 2011) or linear (Levon and Fox 2014; Vaughn 2022a) response but rather find a flat response (i.e., no effect), similar to evaluations of professionalism for (ING) in Levon and Fox (2014). It is worth emphasizing that this is a null result, which could simply reflect poor experimental design, insensitive tools, and, quite likely in this case, an underpowered statistical model. At the same time, in their cautious interpretation of their null (i.e., flat) result, Levon and Fox hypothesize that a weak correlation between (ING) and social class across British English (where it has more regional significance) leads to a weak association between this variable and professionalism, such that the null result may be truly informative: listeners simply do not find (ING) particularly meaningful when assessing professionalism. This reasoning may hold true for coda /s/ as well, given that aspiration may be considered neutral in Puerto Rican Spanish and Caribbean Spanish more broadly. For instance, Lafford (1986, p. 73) claims that /s/ aspiration "cannot be considered a significant social marker due to its weak stratification power." In the current study, the effect of /s/ realization on status ratings is small and very sensitive to the particular way we measure status, and even the larger effect size for masculinity is approximately 0.2 on a six-point scale. Therefore, even though we can see an effect of all [s] vs. all [h] realizations when we have more data (see Sections 4.1 and 4.3), these effects are so small that they suggest a weak role of /s/ realization in speaker evaluation. Moreover, as mentioned in Section 3.2, our methodology differed from the Labov et al. (2011) paradigm such that our participants' attention was likely less drawn to /s/ realization. Future work could compare how response profiles depend on the experimental design and listener awareness of the variable under study. This would force us to think more carefully about the sociolinguistic monitor in terms of what listeners are able to pay attention to and find meaning in (i.e., ability), as opposed to what they more typically pay attention to and find meaning in (i.e., proclivity).

Future work on perceptions of /s/ realization should also consider responses to *deletion*, as the deletion of coda /s/ is more socially stratified than aspiration (e.g., Lafford 1986) and is often stigmatized in Puerto Rican Spanish (Valentín-Márquez 2006). Thus, we could expect very different results in both the additive and contextual analyses if, instead of comparing sibilant and aspirated variants, we compared sibilant, aspirated, *and* deleted variants.

Our final research question addressed the role of listener residence on listener perceptions, and we do not find any differences between islanders and mainlanders in regard to their perception of /s/ realization. One might wonder if this is because the current residence of our listeners is not indicative of their full residential history, particularly given the fact that many Puerto Ricans often move back and forth between the island and mainland US (see Section 2.1). Table 1, however, shows that the islander listeners have resided a notably larger portion of their life (0.97) on the island than the mainlander listeners (0.45), but future studies could consider a more nuanced conceptualization of 'residence.' Regardless, we may expect not to find differences between the two populations, in line with some studies that show similar coda /s/ production between Puerto Ricans on the island and mainland (Ghosh-Johnson 2005; O'Rourke and Potowski 2016; Erker and Reffel 2021) and the fact that sociophonetic perception is found to be similar between second-generation Mexican Americans and Mexicans (Chappell 2019b; Chappell 2021). We *do* find an effect of listener residence independent of /s/ realization, such that Puerto Ricans rate sentences with prevocalic /s/ as higher status than sentences with preconsonantal /s/ regardless of variant. More research is needed to further elucidate this unexpected difference, but it could suggest that mainlanders and islanders do respond to Puerto Rican Spanish differently along certain dimensions.

6. Conclusions

The most important result of this study is that listeners take phonological context into consideration in their evaluations of speaker masculinity, such that social meaning is more robust in more marked environments, corroborating the findings of Bender (2000), Podesva et al. (2015), and Vaughn (2022b). Crucial to our analysis is the disentangling of prescriptiveness and markedness—in Puerto Rican Spanish, it is the norm to aspirate coda /s/, so a sibilant realization ([s]), though prescriptively correct, is more marked and unexpected, particularly in preconsonantal position in men's speech, where it is especially rare. Additionally, we find no differences in the evaluations of listeners residing on the island versus those on the mainland US in regard to their perception of /s/, which adds to the growing body of evidence that sociolinguistic norms are maintained in Spanish-speaking communities in the US (Chappell 2019b, 2021; Erker and Reffel 2021).

Appendix A

Appendix A.1. Context Stimuli

Speaker	Type	Sentence
1	Pre-C	*Y a la derecha, en la esquina a la derecha está.*
		To the right, it's on the corner to the right.
	Pre-V	*Y luego vas a ver el número diez al lado del banco.*
		And then you will see number ten next to the bank.
2	Pre-C	*Va a estar al lado del hospital.*
		It's going to be next to the hospital.
	Pre-V	*Sigue directo, llegas al final de la calle Bolívar y vas a encontrar. . .*
		Continue straight, you will arrive at the end of *Bolívar* street and you will find. . .
3	Pre-C	*Eh, vira a la derecha, en la esquina a la izquierda.*
		Turn right, at the corner to the left.
	Pre-V	*Y cuando llegues a la avenida de la República, vas a virar a la derecha.*
		And when you arrive at *Avenida de la República*, you will turn right.

Appendix A.2. Additive Stimuli

Speaker	Sentence
1	*Está entre el hospital y una escuela elemental.*
	It's between the hospital and the elementary school.
2	*Y va a estar entre, en, a la derecha, en la esquina de la Avenida de la República y Colón, al lado del hospital.*
	And it's going to be between, on, to the right, on the corner of *Avenida de la República* and *Colón*, next to the hospital.
3	*A la izquierda, queda entre la escuela y el hospital.*
	To the left, it's in between the school and the hospital.
4	*Y luego de pasar el hospital a tu izquierda, está el lugar.*
	And after passing the hospital on your left, there is the place.

Appendix B

Appendix B.1. Attention Checks

These questions are multiple-choice questions designed to eliminate users who are randomly choosing answers as well as bots. Correct answers are bolded.

Antes de empezar, conteste estas tres preguntas:

Before we begin, please answer these three questions:

En tu vida, ¿cuántas veces Ud. ha estado en la luna?

(a)   5 o más veces
(b)   2–4 veces
(c)   1 vez
(d)   nunca

In your lifetime, how many times have you been to the moon?

(a)   5+ times
(b)   2–4 times
(c)   1 time
(d)   0 times

¿Con qué frecuencia Ud. experimenta ataques cardíacos fatales?

(a)   nunca
(b)   a veces
(c)   a menudo

How often do you experience fatal heart attacks?

(a)   Never
(b)   Sometimes
(c)   Often

Verdadero o falso: Cuando Ud. se enoja, se aumenta de tamaño y se vuelve un verde brillante.

(a)   verdadero
(b)   falso

True or false: When you get angry, you increase in size and turn bright green.

(a)   True
(b)   False

Appendix B.2. Language Screening

In order to pass the language screening, participants needed to answer the first question in the affirmative and then correctly identify two of the four pictures by typing the Puerto Rican name for the object. Misspelled answers were accepted as correct.

¿Ud. habla español puertorriqueño? Para aclarar, si Ud. nació en la isla o si nació en los Estados Unidos a una familia puertorriqueña, puede decir que sí.

Sí, hablo español puertorriqueño.

No, no hablo español puertorriqueño.

Do you speak Puerto Rican Spanish? To clarify, if you were born on the island or if you were born in the United States to a Puerto Rican family, you can say yes.

Yes, I speak Puerto Rican Spanish.

No, I do not speak Puerto Rican Spanish.
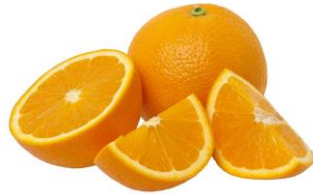
En Puerto Rico, ¿cómo se llama esta fruta?

In Puerto Rico, what is this fruit called?

(correct answer: china)

¿Cuál es la palabra puertorriqueña para este objeto?

What is the Puerto Rican word for this object?

(correct answer: zafacón)

¿Cómo se llama este plato puertorriqueño?

What is the name of this Puerto Rican dish?

(correct answer: mofongo)
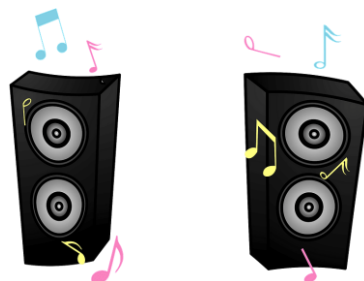
¿Cuál es la palabra puertorriqueña para estos aparatos que se usan para escuchar música?

What is the word used in Puerto Rico for these devices used to listen to music?

(correct answer: bocinas)

Notes

1.   This logarithmic response was replicated by looking at the same variable and using the same paradigm for US listeners by Wagner and Hesson (2014) but not by Vaughn (2022a), who instead found a more linear effect of the proportion of alveolar tokens on speaker ratings.

2.   It is also worth noting that work in speech production suggests that speakers account for linguistic factors (word frequency, neighborhood density, and lexical constraints) in stylistic choices (Hay et al. 1999; Munson 2007; Lin and Chan 2022).

3.   Participants in Bender's study were asked to evaluate speakers in terms of how good they thought the person's job was and how educated, likable, confident, polite, reliable, and comical they sounded. The presence or absence of copula most impacted ratings of education and job, such that copula presence led to impressions that the speaker was more educated and had a better job, and had the least impact on comical ratings. However, in the analysis examining the impact of the following grammatical category, Bender looks at any scale where a given listener was impacted by copula presence/absence.

4.   Most sociolinguistic studies (Alba 2000; Cedergren 1973; Guitart 1976; Lipski 1985; Lynch 2009; among many others) divide /s/ realizations into these three categories ([s], [h], and [∅]). Studies that take into account more phonetic detail note that other weakened variants exist. For example, aspiration is often voiced ([ɦ]) (Luna 2010). Gemination of the following consonant is common, particularly in Cuban Spanish (*estar* > [et.taɾ]) (Terrell 1979). /s/ is also sometimes realized as a glottal stop [ʔ], particularly before vowels (vamos a > [b a.moʔ.a]), but has also been documented before consonants in Puerto Rican Spanish (Mohamed and Muntendam 2020).

5.   Unlike aspiration, deletion is socially stratified in many /s/-weakening dialects (Alfaraz 2000; Lafford 1986; Lynch 2009) and is thus often stigmatized, including in Puerto Rican Spanish (Valentín-Márquez 2006).

6.   It is important to note here the interaction between phonological context, word position, and syllable stress in our stimuli. All of the preconsonantal tokens of /s/ are word-internal and are mostly followed by stressed vowels, with the exception of *hospital*. On the other hand, the prevocalic tokens of /s/ are word-final and followed by unstressed vowels. In his comparison of /s/ aspiration rates in several dialects, Lipski (1985) found minimal differences between aspiration in word-medial versus word-final preconsonantal /s/. He did find a difference between aspiration in prevocalic contexts based on stress (more aspiration before an unstressed vowel); however, this fact should not impact greatly our findings, as the prevocalic tokens included in the stimuli are homogenous in terms of stress (all before unstressed vowels).

7.   As will be explained in Section 3.3, listeners completing the survey through Positly also filled out attention checks and a language screening before completing the demographic questionnaire.

8.   One difference between the questions in Walker et al. (2014) and the present study is that, here, most of the participants were not asked to evaluate the speakers' sexuality. After the first round of data collection, we decided to take this social characteristic out of the survey given that two Puerto Rican informants mentioned that this could be a sensitive question to ask. In this paper, we do not analyze the responses of the 56 participants who did answer this question about speakers.

9.   Both Bartlett's test of sphericity and the Kaiser–Meyer–Olkin (KMO) measure of sampling adequacy suggested that our data was adequate for factor analysis. We used an oblique rotation method (oblimin), seeing as it did not assume our variables were uncorrelated, and in our data, it resulted in the simplest structure (for a discussion, see Brown 2009).

10.   We did also run models using second and third factors instead of the raw pleasantness and masculinity ratings but did not find qualitatively different results.

11.   If we substitute listener residency (a categorical factor based on where participants currently lived) with the proportion of their life they have spent in Puerto Rico (a numeric factor), it makes no qualitative difference in the models—the proportion of life lived on the island is only a significant factor in the model presented in Table 5. This is likely because listener residency and proportion of life in PR are correlated (see Table 1). At the editors' request, we conducted a post hoc test of whether listener age or listener gender has any impact on /s/ ratings. We find a main effect of listener age on masculinity ratings, such that older speakers are more likely to rate speakers as more masculine sounding. Critically, this is regardless of /s/ realization, and so it is not of particular interest in our study. The inclusion of listener age in our masculinity models does not qualitatively change our results regarding /s/ realization. We find no effect of listener gender on ratings.

12.   We confirmed this by changing the order of the factor levels of phonological context in the model presented in Table 4 (such that the /s/ realizations default to prevocalic environments): /s/ realization was no longer significant as a main effect.

References

Alba, Orlando. 2000. *Nuevos aspectos del español en Santo Domingo*. Santo Domingo: Librería la Trinitaria.

Alfaraz, Gabriela. 2000. Sound Change in a Regional Variety of Cuban Spanish. Ph.D. dissertation, Michigan State University, East Lansing, MI, USA.

Arias Alvarez, Alba. 2018. Rhotic Variation in the Spanish Spoken by Puerto Ricans in Puerto Rico and Western Massachusetts. Ph.D. dissertation, University of Massachusetts Amherst, Amherst, MA, USA.

Bandalos, Deborah L., and Meggen R. Boehm-Kaufman. 2008. Four common misconceptions in exploratory factor analysis. In *Statistical and Methodological Myths and Urban Legends*. Edited by Charles E. Lance and Robert J. Vandenberg. London: Routledge, pp. 81–108.

Bender, Emily. 2000. Syntactic Variation and Linguistic Competence: The Case of AAVE Copula Absence. Ph.D. dissertation, Stanford University, Stanford, CA, USA.

Brown, James Dean. 2009. Choosing the right number of components or factors in PCA and EFA. *Shiken: JALT Testing & Evaluation SIG Newsletter* 13: 19–23.

Calles, Ricardo, and Paola Bentivoglio. 1986. Hacia un perfil sociolingüístico del habla de Caracas. In *Actas del II Congreso Internacional sobre el Español de América*. Edited by José G. Moreno de Alba. Mexico City: Universidad Nacional Autónoma de México, pp. 111–14.

Cameron, Richard. 2005. Aging and gendering. *Language in Society* 34: 23–61. [CrossRef]

Campbell-Kibler, Kathryn. 2009. The nature of sociolinguistic perception. *Language Variation and Change* 21: 135–56.

Cedergren, Henrietta. J. 1973. On the nature of variable constraints. In *New Ways of Analyzing Variation in English*. Edited by Charles-James N. Bailey and Roger W. Shuy. Washington, DC: Georgetown University Press, pp. 13–22.

Chappell, Whitney. 2019a. Caribeño or mexicano, profesionista or albañil?: Mexican listeners' evaluations of /s/ aspiration and maintenance in Mexican and Puerto Rican voices. *Sociolinguistic Studies* 12: 367–93. [CrossRef]

Chappell, Whitney. 2019b. The sociophonetic perception of heritage Spanish speakers in the United States: Reactions to labiodentalized <v> in the speech of native and heritage voices. In *Recent Advances in the study of Spanish Sociophonetic Perception*. Edited by Whitney Chappell. Amsterdam: John Benjamins, pp. 239–64.

Chappell, Whitney. 2021. Heritage Mexican Spanish Speakers' Sociophonetic Perception of /s/ Aspiration. *Spanish as a Heritage Language*, 167–97.

Eckert, Penelope. 2012. Three waves of variation study: The emergence of meaning in the study of sociolinguistic variation. *Annual Review of Anthropology* 41: 87–100. [CrossRef]

Erker, Daniel G., and Madeline Reffel. 2021. Describing and analyzing variability in Spanish /s/: A case study of Caribbeans in Boston and New York City. In *Sociolinguistic Approaches to Sibilant Variation in Spanish*. Edited by Eva Núñez-Méndez. London: Routledge, pp. 131–63.

Ghosh-Johnson, Elka. 2005. Mexiqueño? A case study of dialect contact. *Penn Working Papers in Linguistics* 11: 91–104.

Guitart, Jorge M. 1976. *Markedness and a Cuban Dialect of Spanish*. Washington, DC: Georgetown University Press.

Guy, Gregory R., Jennifer Hay, and Abby Walker. 2008. Phonological, lexical, and frequency factors in coronal stop deletion in Early New Zealand English. Paper presented at Laboratory Phonology 11, Wellington, New Zealand, June 30.

Hammond, Robert M. 2001. *The Sounds of Spanish: Analysis and Application (with Special Reference to American English)*. Somerville: Cascadilla Press.

Hay, Jennifer, Stefanie Jannedy, and Norma Mendoza-Denton. 1999. Oprah and /ay/: Lexical Frequency, Referee Design and Style. In *Proceedings of the 14th International Congress of Phonetic Sciences*. Edited by John Ohala. Berkeley: University of California, pp. 1389–92.

Labov, William, Sharon Ash, Maya Ravindranath, Tracey Weldon, Maciej Baranowski, and Naomi Nagy. 2011. Properties of the sociolinguistic monitor. *Journal of Sociolinguistics* 15: 431–63. [CrossRef]

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

Labov, William. 1994. *Principles of Linguistic Change*. vol. 1: Internal Factors. Oxford: Basil Blackwell.

Lafford, Barbara. 1986. Valor diagnóstico-social del uso de ciertas variantes de /s/ en el español de Cartagena, Colombia. In *Estudios sobre la fonología del español del Caribe*. Edited by Rafael A. Nuñez Cedeño, Iraset Páez Urdaneta and Jorge M. Guitart. Caracas: Ediciones La Casa de Bello, pp. 53–75.

Levon, Erez, and Isabelle Buchstaller. 2015. Perception, cognition, and linguistic structure: The effect of linguistic modularity and cognitive style on sociolinguistic processing. *Language Variation and Change* 27: 319–48. [CrossRef]

Levon, Erez, and Sue Fox. 2014. Social Salience and the Sociolinguistic Monitor: A Case Study of ING and TH-fronting in Britain. *Journal of English Linguistics* 42: 185–217. [CrossRef]

Lin, Yuhan, and Marjorie K. M. Chan. 2022. Linguistic constraint, social meaning, and multi-modal stylistic construction: Case studies from Mandarin pop songs. *Language in Society* 51: 603–26. [CrossRef]

Lipski, John M. 1983. La norma culta y la norma radiofónica: /s/ y /n/ en español. *Language Problems and Language Planning* 7: 239–62. [CrossRef]

Lipski, John M. 1985. Reducción de /s/ y/ n/ en el español isleño de Luisiana: Vestigios del español canario en Norteamérica. *Revista de Filología de la Universidad de La Laguna* 4: 125–33.

Lipski, John M. 1994. *Latin American Spanish*. New York: Longman.

Lipski, John M. 1995. Blocking of Spanish /s/-aspiration: The vocalic nature of consonantal disharmony. *Hispanic Linguistics* 6: 287–327.

López Morales, Humberto. 1983. Lateralización de /r/ en el español de Puerto Rico: Sociolectos y estilos. In *Philologica hispaniensia in honorem Manuel Alvar. Tomo I*. Madrid: Gredos, pp. 387–98.

Luna, Kenneth V. 2010. The Spanish of Ponce, Puerto Rico: A Phonetic, Phonological, and Intonational Analysis. Ph.D. dissertation, University of California, Los Angeles, CA, USA.

Lynch, Andrew. 2009. A sociolinguistic analysis of final /s/ in Miami Cuban Spanish. *Language Sciences* 31: 766–90.

Mack, Sara. 2009. Socially Stratified Phonetic Variation and Perceived Identity in Puerto Rican Spanish. Ph.D. dissertation, University of Minnesota, Twin Cities, MN, USA.

Mohamed, Sherez, and Antje Muntendam. 2020. The use of the glottal stop as a variant of /s/ in Puerto Rican Spanish. *Studies in Hispanic and Lusophone Linguistics* 13: 391–417. [CrossRef]

Montgomery, Chris, and Emma Moore. 2018. Evaluating S(c)illy voices: The effects of salience, stereotypes, and co-present language variables on real-time reactions to regional speech. *Language* 94: 629–61. [CrossRef]

Munson, Benjamin. 2007. Lexical characteristics mediate the influence of talker sex and sex typicality on vowel-space size. In *Proceedings of the 16th International Congress on Phonetic Sciences*. Edited by Jürgen Trouvain and William J. Barry. Saarbrücken: University of Saarland, pp. 885–88.

O'Rourke, Erin, and Kim Potowski. 2016. Phonetic accommodation in a situation of Spanish dialect contact: Coda /s/ and /r/ in Chicago. *Studies in Hispanic & Lusophone Linguistics* 9: 355–99.

Pew Research Center. 2019. Facts on Hispanics of Puerto Rican origin in the United States, 2017. Available online: https://www.pewresearch.org/hispanic/fact-sheet/u-s-hispanics-facts-on-puerto-rican-origin-latinos/ (accessed on 15 November 2022).

Pharao, Nicolai, and Marie Maegaard. 2017. On the influence of coronal sibilants and stops on the perception of social meanings in Copenhagen Danish. *Linguistics* 55: 1141–67.

Pharao, Nicolai, Marie Maegaard, Janus S. Møller, and Tore Kristiansen. 2014. Indexical meanings of [s+] among Copenhagen youth: Social perception of a phonetic variant in different prosodic contexts. *Language in Society* 43: 1–31.

Podesva, Robert J., Jermay Reynolds, Patrick Callier, and Jessica Baptiste. 2015. Constraints on the social meaning of released /t/: A production and perception study of U.S. politicians. *Language Variation and Change* 27: 59–87. [CrossRef]

Poplack, Shana. 1979. Function and Process in Variable Phonology. Ph.D. dissertation, University of Pennsylvania, Philadelphia, PA, USA.

Poplack, Shana. 1980. Deletion and disambiguation in Puerto Rican Spanish. *Language* 56: 371–85. [CrossRef]

Preston, Dennis R. 1999. *Handbook of Perceptual Dialectology: Volume 1*. Amsterdam: John Benjamins.

Ramos Pellicia, Michelle F. 2012. Retention and deletion of /s/ in final position: The disappearance of /s/ in the Puerto Rican Spanish spoken in one community in the US Midwest. *Southwest Journal of Linguistics* 31: 161–66.

Ramos-Pellicia, Michelle F. 2007. Lorain Puerto Rican Spanish and 'r' in three generations. In *Selected Proceedings of the Third Workshop on Spanish Sociolinguistics*. Edited by Jonathan Holmquist, Gerardo A. Lorenzino and Lotfi Sayah. Somerville: Cascadilla Proceedings Project, pp. 53–60.

Sumner, Meghan, Seung K. Kim, Ed King, and Kevin B. McGowan. 2014. The socially weighted encoding of spoken words: A dual-route approach to speech perception. *Frontiers in Psychology* 4: 1–13. [CrossRef]

Terrell, Tracy D. 1978. La aspiración y elisión de /s/ en el español porteño. *Anuario de Letras* 16: 41–66.

Terrell, Tracy D. 1979. Final /s/ in Cuban Spanish. *Hispania* 62: 599–612. [CrossRef]

Terrell, Tracy D. 1981. Diachronic reconstruction by dialect comparison of variable constraints: S-aspiration and deletion in Spanish. *Variation Omnibus*, 115–24.

Terrell, Tracy D. 1982. Current trends in the investigation of Cuban and Puerto Rican phonology. In *Spanish in the United States: Sociolinguistic Aspects*. Edited by Jon Amastae and Lucía Elías-Olivare. Cambridge: University of Cambridge Press, pp. 47–70.

U.S. Census Bureau. 2020. Quick Facts: Puerto Rico. Available online: https://www.census.gov/quickfacts/PR (accessed on 15 November 2022).

Valentín-Márquez, Wilfredo. 2006. La oclusión glotal y la construcción lingüística de identidades sociales en Puerto Rico. In *Selected Proceedings of the 9th Hispanic Linguistics Symposium*. Edited by Nuria Sagarra and Almeida Jacqueline Toribio. Somerville: Cascadilla Proceedings Project, pp. 326–41.

Valentín-Márquez, Wilfredo. 2020. The sociolinguistic distribution of Puerto Rican Spanish /r/ in Grand Rapids, Michigan. In *Dialects from Tropical Islands: Caribbean Spanish in the United States*. Edited by Wilfredo Valentín-Márquez and Melvin González-Rivera. London: Routledge, pp. 88–110.

Vaughn, Charlotte, and Tyler Kendall. 2018. Listener sensitivity to probabilistic conditioning of sociolinguistic variables: The case of (ING). *Journal of Memory and Language* 103: 58–73.

Vaughn, Charlotte. 2022a. The role of internal constraints and stylistic congruence on a variant's social impact. *Language Variation and Change* 34: 331–54. [CrossRef]

Vaughn, Charlotte. 2022b. What carries greater social weight, a linguistic variant's local use or its typical use? In *Proceedings of the 44th Annual Conference of the Cognitive Science Society*. Edited by Jennifer Culbertson, Andrew Perfors, Hugh Rabagliati and Véronica Ramenzoni. Berkeley: University of California, pp. 512–18.

Wagner, Suzanne E., and Ashley Hesson. 2014. Individual sensitivity to the frequency of socially meaningful linguistic cues affects language attitudes. *Journal of Language and Social Psychology* 33: 651–66. [CrossRef]

Walker, Abby, Christina García, Yomi Cortés, and Kathryn Campbell-Kibler. 2014. Comparing social meanings across listeners and speaker groups: The indexical field of Spanish /s/. *Language Variation and Change* 26: 169–89. [CrossRef]

Weatherholtz, Kodi, Kathryn Campbell-Kibler, and T. Florian Jaeger. 2014. Socially-mediated syntactic alignment. *Language Variation and Change* 26: 387–420. [CrossRef]

Article

Creaky Voice in Chilean Spanish: A Tool for Organizing Discourse and Invoking Alignment

Mariška Bolyanatz

Department of Spanish & French Studies, Occidental College, Los Angeles, CA 90041, USA; mbolyanatz@oxy.edu

Abstract: This study relies on an interactional, conversational–analytic approach to elucidate what meanings Chilean Spanish speakers convey via creaky voice quality in informal conversations. Highly creaky utterances produced by 18 speakers were derived from a larger corpus of sociolinguistic interview speech from Santiago, Chile, and examined via an interactional approach that accounted for how creaky voice figured in the process of meaning-making and meaning negotiations throughout the conversation. Results indicate that approximately 40% of highly creaky utterances were used to organize the speaker's discourse, signaling the end of turns, hedges or uncertainty, and a change in communicative purpose, while the majority of the highly creaky utterances were used to invoke alignment with the listener via ensuring that their messages or stances were understood and potentially endorsed. This study offers evidence from a non-English language for creaky voice as a tool for both discursive organization and interactional alignment.

Keywords: creaky voice; sociolinguistic interview; alignment; interactional sociolinguistics; conversation analysis

1. Introduction

Cross-linguistically, non-modal phonation has been shown to have multiple meanings. Non-modal phonation may indicate phonological contrast in languages such as Hmong (Esposito and Khan 2012), Zapotec (Esposito 2003), and Gujarati (Esposito and Khan 2012). Creaky voice specifically has been shown to indicate phrase position as in English (Henton and Bladon 1988; Podesva 2013) and Finnish (Ogden 2001). Finally, and most importantly to the present paper, Podesva and Callier (2015) assert that non-modal phonation is a "rich phonetic resource through which speakers can display affect and take stances in interaction" (Podesva and Callier 2015, p. 173). They point to recent research in the fields of linguistic anthropology and sociolinguistics that exemplifies non-modal voice quality as a resource for the construction of social meaning (Podesva 2007; Sicoli 2010; Hildebrand-Edgar 2016).

The aim of this paper is to identify how speakers of Chilean Spanish use creaky voice to position themselves in interactions, with particular attention paid to the meaning being constructed in conversation. As Acton (2021) claims, "what speakers say and how they say it is not merely the product of exogenous forces but also depends on speakers' beliefs about how listeners would evaluate a given variant—and, in turn, how useful the variant would be in helping them achieve their desired ends" (As Acton 2021, p. 2). This attention to the specific communicative goals of the speaker, as well as how their utterances would be perceived, is centrally located within the realm of third-wave sociolinguistics (Eckert 2012) as the speaker's meaning as well as agency are foregrounded. This also aligns with Du Bois' (2007, p. 171) definition of stance as a public action, not something you can perform individually or privately. Taking a stance entails, therefore, that speakers "simultaneously evaluate objects, position subjects (themselves and others), and align with other subjects, with respect to any salient dimensions of the sociocultural field" (Du Bois 2007, p. 163). As Podesva (2007, p. 498) indicates, situating phonetic forms in their discursive contexts and attempting to infer the social meaning of linguistic features

is a new analytical tactic for determining the meaning of variation that my work aims to propel. It is also well aligned with conversation analysis, which identifies conversation as the most basic level of social organization that humans have (Sacks et al. 1978). In this view, the interlocutor is a necessary participant in the creation of meaning within the discourse, as is a thorough account of the broader sociocultural, political economic realities of the conversation (Sidnell 2013, p. 85).

In the remainder of this section, I first examine previous findings related to characteristics and uses of creaky voice in English and in other languages. I then review the literature related to my methodological approach.

1.1. Creaky Voice and Prosodic Phrasing, Cross-Linguistically

As Davidson (2021) describes, creaky phonation is not a monolithic phenomenon, and a variety of acoustic characteristics (such as low fundamental frequency, irregular spacing of glottal pulses in the spectrogram, and decreased amplitude in comparison with modal phonation) may signal what listeners perceive as creaky voice. Garellek and Keating (2011, p. 187) refer to this as the "plethora of physiological dimensions of voice quality variation". In spite of this variation, one consistent finding related to the production of creak has to do with the aerodynamic requirements for speech production that also co-occur with higher prosodic domains. That is, several studies have found that creak is more likely to occur at the ends of utterances. For instance, Henton and Bladon (1988) examined United Kingdom English speech and found that the occurrence of creak was greater on the final syllable as compared to the other syllables in a sentence. In United States English, Redi and Shattuck-Hufnagel (2001) found a high rate of glottalization of words at the end of utterances and phrases, similar to the results of Slifka (2006). Podesva (2013) finds that of the six types of non-modal phonation analyzed in sociolinguistic interviews with English speakers, creaky phonation is more likely to occur in the intonation phrase-final position. Podesva (2013) relates this to declination, the gradual diminishing of fundamental frequency over the course of an utterance. Indeed, phrase endings have been shown to co-occur with the end of a breath, phonation-ending laryngeal gestures, and a drop in alveolar pressure, which presents a challenge for modal voicing (Redi and Shattuck-Hufnagel 2001; Slifka 2006; Wolk et al. 2012).

Creak has been shown to have similar functions in other languages. For instance, Ogden (2001) reported that in Finnish, a glottal stop was used as a turn-holding strategy (indicating that the speaker intended to repair or continue their discourse), while creaky syllables typically marked the end of a turn. Ogden (2001) demonstrated that creak co-occurred with other cues for turn transition such as intonation, tempo, and duration. Garellek and Keating (2015) and González et al. (2022) have provided evidence for creak in phrase-final position in Spanish, and Garellek (2014) also documented an utterance-initial glottalization effect in read Spanish akin to findings from English (Garellek 2014; Redi and Shattuck-Hufnagel 2001; Dilley et al. 1996).

1.2. Voice Quality and Its Uses beyond Prosodic Phrasing

Podesva and Callier (2015) delineated three ways that voice quality can serve as an index of identity. The first is as a signifier of belonging to a particular group or category such as gender, race, class, and linguistic identities. Podesva (2013) identified both gender- and ethnicity-based groupings of non-modal phonation in sociolinguistic interview speech among white and African American residents of Washington, D.C. Podesva and Callier (2015) also pointed to voice quality as a way for listeners to distinguish individual speakers. The third function of voice quality delineated by Podesva and Callier (2015) is that of stancetaking, particularly of affective or emotional stances (Du Bois 2007). While the authors cautioned against directly associating voice qualities with particular affective displays, they advocated for turning the analytical focus on voice quality to the various indexical possibilities for particular, situated productions of marked voice qualities (Podesva and Callier 2015, p. 183). For instance, Lee (2015) found that creaky voice was used in so-

ciolinguistic interviews when a speaker temporarily suspended the ongoing discursive frame by inserting a sequence that was identified as belonging to a different frame, which she denoted parenthetical use. The parenthetical functions in Lee's (2015) data included jokes, off-the-record comments, inner thoughts, and additional or preemptive provision of information. According to Lee (2015, p. 295), using creak allowed a speaker to both mark a proposition as parenthetical and also to create a parenthetical effect in order to perform some kind of stancetaking. Similarly, Torres et al. (2020) examined the production of creaky voice quality situated within patient/provider interactions. The authors found that English-speaking patients in medical interactions used both low pitch and creaky voice when narrating symptoms or describing pain and when requesting opiates. The authors centered the discussion of their findings on the national conversation about opioids in the United States, positing that using creaky voice when requesting opioids is a way to counter the stigmatization of their use.

1.3. Voice Quality Variation as Stancetaking in Non-English-Speaking Contexts

Podesva and Callier (2015) noted that little research has investigated the ways in which affective indexicalities of voice quality differ cross-culturally. They pointed to the work of Sicoli (2010) as an exception to the focus on English-speaking communities and as an illustration of other kinds of stances that can be constructed from voice material. Specifically, Sicoli (2010) examined the speech of Lachixío Zapotec speakers in various situations with different interlocutors to demonstrate that variations in voice quality copattern with audiences, pragmatic functions such as requesting, and voice registers. According to Sicoli's analysis, falsetto phonation showed respect, breathy phonation indicated authority, creaky phonation sought commiseration, and whisper phonation marked urgency (Sicoli 2010, p. 545), revealing that voice qualities are enregistered through their co-occurrences with the participant roles taken up in speech situations. Similarly, Wilce (1997) examined patient and provider interactions in Bangladesh and showed that creaky voice frames pain in the patients' talk. Wilce argued that creaky voice is an "icon" of weakness and low energy and also functions as a way to "underline and lend credibility to the speaker's reference to her own pain" (Wilce 1997, p. 363).

The accounts above weave together the social context and linguistic functions of creak and other non-modal voice qualities. In the next subsections, I highlight the importance of conversational interaction for examining the interplay between creaky voice, a speaker's communicative purpose and social context, and the role of the interlocutor.

1.4. Organizing Conversational Interaction

Earlier work in discourse and conversation analysis focused on the delineation of turns or determining which party talks at which time in a conversation (Sacks et al. 1978). As Clayman (2012) summarizes, turns are incrementally built out of a succession of turn-constructional units (TCUs), such as words, phrases, clauses, and sentences, and each TCU ends in a transition-relevant place (TRP) (Sacks et al. 1978) where an interlocutor may optionally retake the floor. More recent work has focused on deconstructing turns into higher-level discourse units. For instance, Selting (2000) argued that "big projects" such as storytelling create a single TCU that is organized internally into smaller units of other kinds. In the first study of its kind, rather than focusing on a particular or specialized speech genre such as a narrative or a joke, Biber et al. (2021) conducted a bottom-up analysis of a large English-speaking corpus that revealed that most conversational talk consists of sequences of coherent discourse units that have identifiable communicative purposes (p. 34). Biber et al. (2021) defined a communicative purpose as the more general communicative actions or tasks that interlocutors hope to achieve in conversation (Biber et al. 2021, p. 24). To make this claim, the authors first divided their conversational data consisting of informal face-to-face conversations between two or three people into discourse units (hereafter DUs). The authors operationalized the DUs as recognizably self-contained (their boundaries are noted by a shift in the conversation's overarching communicative

goal), coherent for their overarching communicative goal (that is, interlocutors have a particular goal in the conversation, such as complaining about annoying co-workers or making plans for buying holiday gifts), and at least five utterances (or 100 words) (Biber et al. 2021, p. 23).

From these DUs, the authors then developed a taxonomy of nine communicative purposes: situation-dependent commentary (when speakers comment on things that are present or occurring in their shared situational context); joking around (conversation intended to be humorous); engaging in conflict (including disagreement of any type); figuring things out (discussion aimed at exploring or considering options or plans); sharing feelings and evaluations (including feelings, evaluations, opinions, and beliefs, including the airing of grievances); giving advice and instructions (offering directions, advice, or suggestions to another speaker); describing or explaining the past (including narrative stories about true events from the past or references to people or events from the past); describing or explaining the future (including descriptions or speculations about future events and intentions); and describing or explaining (time neutral) (consisting of description or information about facts, information, people, or events in which time is either irrelevant or unspecified). Each DU could have up to three primary and secondary purposes, and the authors determined that two broad categories of sharing feelings and evaluations and of conveying information accounted for over 80% of the general communicative purposes. This approach is similar in some ways to Goffman's (1974) notion of frame, which delineates how an individual categorizes or makes sense of their social interactions. However, in my view, Biber et al. (2021) offer a more clearly defined, empirically verified approach for determining the communicative purposes of discourse units within informal conversation. These categorizations are exhaustive, accounting for each of the components of the discourse, and the authors demonstrate that nearly all discourse is organized at a more micro level (albeit within a particular linguistic and social context, the English-speaking United Kingdom). The present paper demonstrates how creaky voice can be used to signal a change in communicative purpose types in Chilean Spanish conversational data.

1.5. The Relevance of the Interlocutor's Response and Alignment

In addition to the organization of discourse into units, another key tenet of conversation analysis is the importance of the interlocutor to the construction of meaning (Schegloff et al. 1996, p. 40; Duranti 1986). Uniting the concepts of discourse organization and interlocutor response, Stivers (2008) posits that the TRP is the appropriate place for an interlocutor to respond to a speaker's stance (though nodding and providing other short feedback such as "mhm" and "yeah" mid-turn is a way for the interlocutor to recognize that the speaker's turn is still ongoing), suggesting that the "preferred response to a story-telling is the provision of a stance toward the telling that mirrors the stance that the teller conveys having [. . .] whether that is as funny, sad, fabulous, or strange" (Stivers 2008, p. 33). Stivers (2008) calls this alignment, or supporting the ongoing talk of the speaker, in contrast with affiliation, denoting that a hearer displays support of and endorses the teller's conveyed stance. Clayman and Raymond (2021) use alignment in a different sense. Instead of an appropriate response to an ongoing storytelling turn, Clayman and Raymond posit that what they term alignment may initialize a conversation, in that it may invoke "a state of alignment without asserting it as such, which prompts validating responses without making them obligatory" (Clayman and Raymond 2021, p. 294). In their analysis of the *you know* particle in English, they term this usage an alignment token (Clayman and Raymond 2021), an element that invokes a convergent orientation between recipient and speaker. As Clayman and Raymond (2021, p. 293) argue, this alignment token may be intersubjective (that the listener correctly grasps the speaker's meaning), affiliative (that the recipient supports the action or stance being taken), or both.

Clayman and Raymond's work illuminates the potential for conversation analysis to move the field of interactional sociolinguistics forward and highlights the importance of acknowledging that the sociolinguistic interview is essentially a conversation, bound by

norms and rules, and subject to the disfluencies and speech production difficulties that arise in everyday conversation.[1] For instance, as Clayman and Raymond (2021) describe, the *you know* particle can be used by speakers to ensure that in spite of their hesitation, disfluency, or self-repair, the recipient will still be able to grasp their meaning. *You know* may also be used when the speaker recognizes that what they are saying is not exactly what they mean to say, but they want to ensure that the recipients understand their meaning. Clayman and Raymond (2021, p. 300) define these as sub-optimal formulations, in that they "[fall] short of the speaker's expressive aims, or [that they are] potentially problematic for recipients to grasp as intended, or both of these in combination. In these cases, understanding on the recipient's part is important to continue the forward conversational movement.

Additionally, *you know* is often recruited for actions that aim to elicit or garner some sort of response, whether acceptance or rejection, such as assessment, recruitment actions and accounts, and misdeeds (Clayman and Raymond 2021, p. 303). For instance, assessments and other evaluative comments, especially those that are "relatively 'extreme,' negatively valenced, or have a critical edge, are recurrently accompanied by alignment tokens" (p. 303). *You know* may also be used when the speakers are engaged in or reporting misdeeds or acts that may be regarded as mischievous or in breach of societal norms. In these cases, *you know* "invokes support for the questionable action and often elicits confirmation" (p. 305).

I use the data presented below to argue that creaky voice in Chilean Spanish primarily serves as a tool to "invoke a convergent orientation between speaker and listener" (cf. Clayman and Raymond 2021, p. 294) via an interactional approach. Like the particle *you know* in Clayman and Raymond's (2021) analysis, the domain of creak is basically unrestricted. That is, because it consists of a laryngeal configuration that can be deployed by all human language users, creak could appear anywhere, on any syllable. Therefore, when creak appears in certain places, we may assume that it serves a particular purpose. In the data below, I examine the purposes of creak in these Chilean Spanish sociolinguistic interviews. I show that creak has two primary functions: to organize discourse and to invoke alignment.

2. Materials and Methods

2.1. Participants and Initial Coding

The data for this analysis derive from a larger corpus of sociolinguistic interviews with Chilean Spanish speakers about a variety of topics, such as how they had celebrated Chilean Independence Day, how they met their partners/spouses, and whether they had had strict upbringings. The speakers resided in two relatively homogeneous neighborhoods of Santiago, belonged to three age groups (18–25, 26–41, 42+),[2] and comprised all six of the Chile-specific socioeconomic groups defined by Sadowsky (2021).[3] In Bolyanatz (2023), I describe how I followed Podesva (2013) in selecting two topics of conversation that had been spoken on at length by nearly all 41 speakers: talk about relationships and talk about the local community. I hypothesized that both of these topics would be "non-neutral" (cf. Yuasa 2010) or be more likely to elicit increased emotional or affective stancetaking. For instance, the talk about the local community was of particular interest to residents of these two neighborhoods, both of which are characterized as socioeconomic extremes and perceived to be facing recent increases in crime.

Following the delineation of topics, I then coded utterances. In the transcriptions below, each line is a separate utterance organized by pauses at the beginning and end (Du Bois et al. 1993; Chafe 1993). Subsequently, I coded the first approximately 100 syllable nuclei within each topic for a total of 9030 coded syllables. Each syllable was coded for several features including position within the utterance and six types of voice quality (modal, breathy, creaky, whisper, harsh, and falsetto following Podesva (2013)). I used auditory–visual methods (Dallaston and Docherty 2020) to code for creaky voice. The auditory parameters included perception of rough or gravel-like quality, and the visual parameters included increased irregularity of pulses in the waveforms and/or widely spaced vertical striations in the spectrograms. Using binary coding (creaky vs. not creaky)

accounts for the variation in types of creaky voice established cross-linguistically (Garellek and Keating 2011; Keating et al. 2015; Davidson 2021), but that are of less importance for the present project focused on the meaning of creak. The examples below show a few key elements of the coding.

First, Crowhurst (2018) showed that Mexican Spanish listeners were sensitive to creak as a cue to utterance-finality when a syllable was at least 30% creaky. Therefore, syllables in the present study had to be at least 30% creaky (though often were >50% creaky) to be binary coded as creaky (though, of course, Chilean Spanish listeners may respond differently to creak than Mexican Spanish listeners, a topic I return to in the Discussion section). This restriction meant that some tokens, such as the utterance-initial /e/ in Figure 1, were excluded.
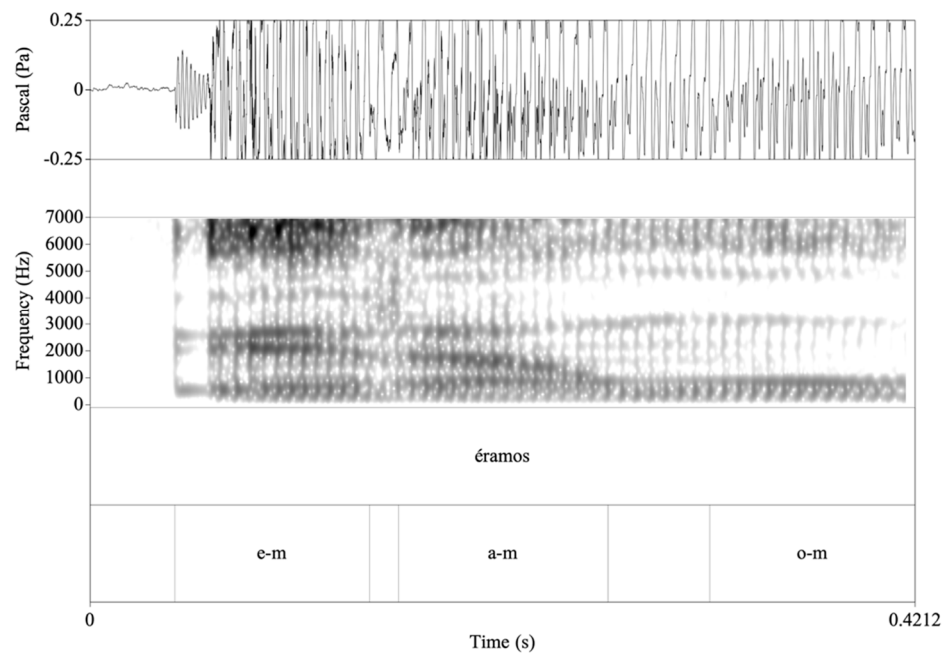
Figure 1. Waveform, spectrogram, and textgrid noting orthographic transcription and syllable-level coding (vowel identity and voice quality identity noted; "m" = "modal"). Speaker lp.m.3.05.

This utterance-initial vowel was produced with a small amount of glottalization but did not reach the 30% threshold to be coded as creaky. For this reason, this vowel was coded as modally voiced. In contrast, sometimes creak was readily apparent over the length of multiple syllables. For instance, in Figure 2, the speaker began creaking on the stressed /o/ vowel of the word "Condes" (Las Condes is a wealthy neighborhood in Santiago), and creak was present on the remaining syllables in the utterance.

2.2. Defining the Highly Creaky Utterances and Turns

From the larger dataset described above, I examined the 442 tokens of creaky voice that were not in utterance-initial or utterance-final position and selected utterances that had at least three creaky tokens within. Upon review of the data, I noted that nearly all the utterances within one speaker's turns (v.m.3.01) were highly creaky. This speaker was the oldest in the sample (75 years of age at the time of data collection), and his voice was also accompanied by "roughness" or an overall gravelly quality of voice (Hollien 1987). These are likely correlates of chronological age that make it difficult to separate creak used for discursive purposes from creak as a function of aging vocal folds. This speaker's highly creaky tokens were therefore excluded, leaving eighteen speakers' highly creaky tokens for further analysis (age range 21–58). Similar to the methods of Podesva (2013) and Lee (2015), these strict criteria yielded ninety-two "highly creaky" utterances for analysis produced by

eighteen speakers. The highly creaky utterances were produced by residents of both field sites and by both men and women.
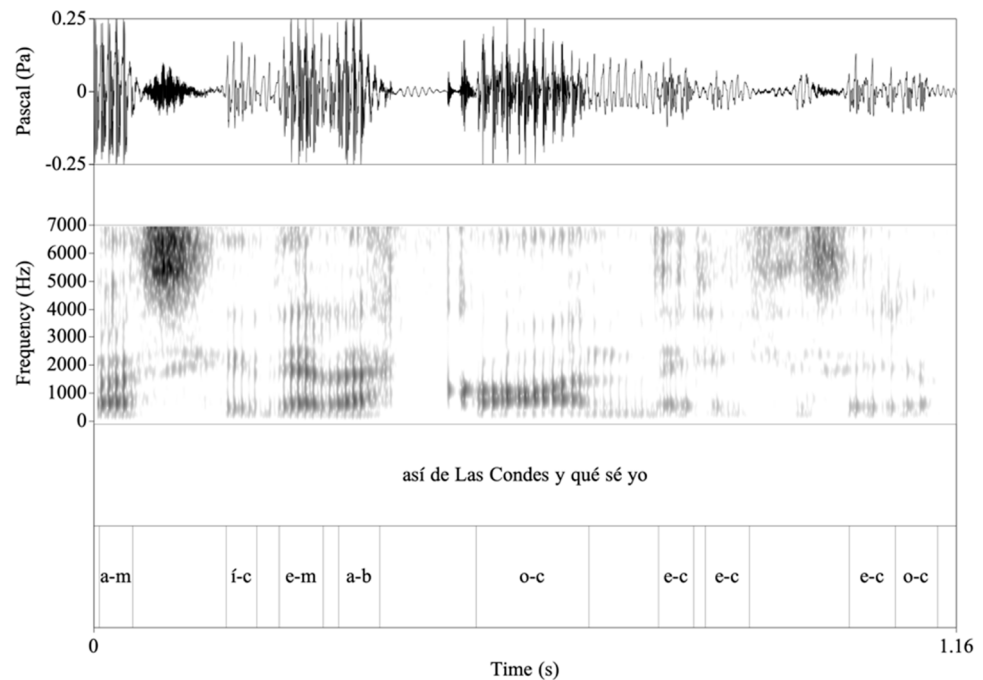
Figure 2. Waveform, spectrogram, and textgrid noting orthographic transcription and syllable-level coding (vowel phoneme and voice quality identity noted; m = "modal"; c = "creaky", b = "breathy"). Speaker lp.m.1.05.

Once I had identified the highly creaky utterances (henceforth HCUs), I turned to delineating the boundaries (beginning and end) of each of the turns that contained the HCUs. The onset of a highly creaky turn (HCT) was often the speaker's response to a question I had posed or a topic change of the speaker's own volition that began the DU. Two speakers produced two HCTs within the dataset while the other speakers produced one each, generating 20 HCTs for analysis. Each highly creaky turn had at least one HCU per turn, and one speaker had up to eleven HCUs within (mean = 3.98 HCUs per HCT). The turns were then categorized into discourse units following those established by Biber et al. (2021). The end of the turn coincided with syntactic, prosodic, and pragmatic completeness that indicated a transition-relevant place (Levinson and Torreira 2015).

3. Results and Discussion

In this section, I provide an overview of how creak is used within these data. I then examine each of the subtypes of the functions of creak using excerpts from the data and show several examples of how different functions may be used in tandem within single HCTs.

All 20 HCTs shared the same communicative purposes: all were primarily or secondarily focused on describing or explaining the past, describing or explaining time-neutral information, and sharing feelings and evaluations. If the primary focus was describing or explaining, the secondary focus of nearly all of them was sharing feelings and evaluations (in close alignment with the findings of Biber et al. (2021)). As described above, I posit that creak is used in two primary ways in these data: to organize discourse and to serve as tokens for alignment. In (1), I provide a brief explanation of each function.

1. Functions of creak present in the data

    a.   Creak used to organize discourse (39%)

      i.   At the end of turns, signaling a turn-yielding function (9/96 of the HCUs, 9%; 9/20 of the turns ended with several highly creaky syllables; 45% of the turns)[4]

      ii.   To signal a pivot between discourse units (DUs) (7/96 uses of HCUs; 7%)

      iii.   To signal a hedge or uncertainty or a site for self-repair (10/96 uses of HCUs; 10%)

      iv.   To indicate a parenthetical (or outside the discourse unit) joke, off-the-record comment, inner thought, and additional or preemptive provision of information (cf. Lee 2015) (11/96 HCUs; 11%)

   b.   Creak used to signal stance (61%)

      i.   Understanding: situations in which understanding may be relevant or at risk (30/96 HCUs; 31%)

      ii.   Affiliation: the recipient supports the action or stance being taken (28/96 HCUs; 29%)

3.1. Creak Used to Organize Discourse

Excerpts 1 through 7 detail the discourse-organizing functions of creak in these data. First, HCUs appeared at the end of HCTs, signaling a turn-yielding function. Of the 20 HCTs, HCUs ended 9 of them (45%). An example is included in Excerpt 1 below. In each of the excerpts, underlining indicates creaky voice, and arrows indicate HCUs. Breathy voice is indicated via italics, whisper voice via strikethrough, and falsetto voice in bold, and all non-creaky, non-modal phonation is described textually. In this excerpt, Marta was telling me about her relationship with her boyfriend and his son and noted that her boyfriend had just moved into her family home.

(1)   Marta describing a conversation between her and her boyfriend, who recently moved in with her; discourse unit: describe/explain the past (primary); sharing feelings and evaluations (secondary).

1 2	Yo le pregunté el otro día po le dije \<quote-past self\> Oye, ¿qué tú hai echado de menos a tus papás? \</quote\> // \<softer\> //
3 4 5	\<quote\> No, \</quote\> \<lengthening\> me dijo pero se le pusieron sus ojos llenos de \<emphasis\> lá**gri**mas \</emphasis, falsetto\> y a mí me dio pena y lo abracé po //
6 7 →	Lo abracé y me di*jo* \<breathy\>: \<quote\> Sí, sí igual los extraño pero me siento bien aquí contigo \</quote\> me dijo que él \<lengthening/\> //
8 9 → 10	Que él es **hombre** \<emphasis; falsetto\> me dijo \<quote\> *Yo soy hombre* \<breathy\>, a mí no me va a costar tanto como te hubiera costado a ti po \</quote\> me di*jo* \<whisper\>. //

1 2	I asked him the other day *po*[5] I said \<quote-past self\> \<softer\> Hey, have you been missing your parents? \</quote\> \</softer\> //
3 4 5	\<quote\> No \<lengthening\> \</quote\> he told me but his eyes filled up with \<emphasis\> tears \</emphasis; falsetto\> and I felt sad for him and I hugged him *po*.
6 7	I hugged him and he said to me \<quote\> yes, yes I do miss them but I feel good here with you \</quote\> he said that he \<lengthening\>
8 9 10	He's a \<emphasis\> man \</emphasis, falsetto\> he told me \<quote\> I'm a man, it's not going to be as hard for me as it would have been for you *po* \</quote\> he told me.

Utterance 4 (Lines 7–8) is one of only four HCUs in the dataset that comprise two sets of three-syllable stretches of creak in the same utterance. I focus on the final syllables of the utterance here. Of the last five syllables of the utterance (Line 8, starting with "ti po"), four are creaky. While the very final syllable in the utterance, and indeed the turn overall, is whispery (which is found in several turns throughout the dataset), I posit that these several syllables of creak in the turn-final position indicate that the turn is ending. These creaky syllables also co-occurred with syntactic completeness (Levinson and Torreira 2015) and pitch and intensity lowering.

Creak on utterance-final, turn-medial syllables is also prevalent. For instance, in this same turn, two of the three turn-medial utterances ended with one or two creaky syllables. A similar phenomenon is represented in Excerpt 2. In this excerpt, I had just asked the speaker (Carmen) if she was the "regalona", Chilean slang for the favorite, or the apple of the parent's eye.

(2) Carmen describing her relationships with her siblings; DU: sharing feelings and evaluations (primary); describing/explaining the past (secondary).

1		sí <lengthening/> no <rising intonation> no la regalona <emphasis> tamp<u>oco</u>
2		<emphasis/> <u>no no</u> <silencio/>
3		y con mi hermano ahí <u>no más</u>
4	→	no me llev<u>o bien con él</u> <silence/>
5		con mi her<u>mana</u>– //
6	→	<emphasis> más <emphasis/> todav<u>ía con mi hermana</u> //
7		porque resulta que mi herm<u>ana</u> eh <lengthening/> la casa donde vivo es donde vivo es
8		de mi <emphasis> mamá <emphasis> //
9		que <breathy> *en paz descanse* <breathy> /
10		<u>e</u>lla <emphasis> fallec<u>ió</u> <emphasis/>
11	→	y re<u>sulta que</u> <lengthening/>
12	→	la casa estaba en nom<u>bre de e*lla*</u> <breathy> ///
13		y qué es lo que pasa que ahora es mi <rising intonation > hermana la que <rising intonation; lengthening/>
14		busca problemas por la ca<u>s*a*</u> <whisper> //
15	→	que a ella le per<u>tenece</u> la <emphasis> ca<u>sa</u> <emphasis/> <lengthening/>
16	→	<u>que ella</u> <lengthening/>
17	→	que aquí <u>ella</u> «<u>Oh</u>» me sa<u>caba</u> varias veces en cara la casa /
18		[en]tonces le le digo yo que eso no es <emphasis> justo <emphasis; rising intonation>
19		porque ahora le pertenece a los <emphasis> tres </emphasis> po /
20		<u>an</u>tes no porque antes existía había una ley que era pa las pu<u>ras</u> <lengthening>
21		las que eran <emphasis> solte<u>ras</u> <emphasis/> /
22		pero ahora <emphasis> no <emphasis/> <u>porque ahora ya está todo</u> <emphasis>
23	→	<u>cambiado</u> <emphasis/> <u>ahora ya</u> ///

1	Yes <lengthening/> no <rising intonation> no, not the <emphasis> *regalona* <emphasis>
2	no no
3	Me and my brother more or less
4	I don't get along with him
5	with my sister--
6	even more with my sister

7 8	because it turns out that my sister eh \<lengthening> the house that I live in belongs to my \<emphasis> mom \<emphasis>
9	may she rest in peace
10	she passed away
11	and it turns out
12	the house was in her name
13	and what happens is that now it's my sister who is \<rising intonation; lengthening/>
14	looking for trouble over the house
15	that it belongs to her
16	that she
17	that here she \<quote> Ohh \<quote> she's thrown the house in my face several times
18 19	so I've told her that that's not fair because now it belongs to the \<emphasis> three \<emphasis> of us *po*
20	before no because before there was a law that it was for only
21	women who were \<emphasis> single \<emphasis/> /
22	but not anymore because now all that is changed now

Of the 23 syllables in the turn-final utterance, 14 of them are creaky. Additionally, of the 19 utterances within this turn, 11 are creaky on the final syllable. Perhaps, like Ogden (2001) describes for Finnish, short creak (creak on only one syllable) in utterance-final position in the middle of a larger project is a turn-holding strategy, while speakers may signal turn-ending with longer stretches of creak. It is also possible that this is simply the natural result of declination as air runs out toward the end of the utterance, as described by Podesva (2013).

The second use of creak to organize a turn is found at the adjunct of two discourse units that comprise differing communicative purposes. Seven of the 92 HCUs had this purpose. For instance, Excerpt 3 reveals how the communicative purpose of the speaker shifts within a single utterance. In this excerpt, after the speaker explains how she became involved with a neighborhood NGO, she shifts her communicative goal to reflect on the importance of the NGO's presence in the community and to offer her own perception of the community center.

(3) Mercedes describing the local community; DU: describing/explaining (time neutral) (primary); sharing feelings and evaluations (secondary)

1		Porque muchas familias //
2		Sobre todo en co*munas* \<breathy voice> *a*<u>*sí*</u> //
3		<u>E</u>xiste mucho lo que \<lengthening/> //
4		Violencia in<u>tra</u>familiar //
5		Mu*cho* mucho de*masia*<u>*do*</u> //
6		Eh mucha drogadicción //
7		Mucho alcohol //
8	→	Entonces la iniciativa de La Colo<u>ra</u>da a <u>mí</u> me pa<u>re</u>ce //
9	→	<u>Pero</u> \<emphasis> <u>excelen</u>*te* \<breathy voice/> \<emphasis/> \<lengthening/> //

1	Because many families
2	Especially in communities like this one

3	There exists a lot of
4	intrafamilial violence
5	A lot a lot too much
6	Eh a lot of drug addiction
7	A lot of alcohol
8	So the initiative of La Colorada to me seems
9	Just excellent

In Line 8, the speaker shifts from a time-neutral description of the community to an utterance that introduces the affective judgment in Line 9. In Line 8, the discourse marker "entonces" begins to mark the shift in communicative purpose, which is then doubly indicated by creaky voice later in the phrase. Additionally, the setup of the affective judgment is seen not only in the phrase "me parece" (*it seems to me*) but also in the grammatically optional unit of "a mí" (*to me*), essentially creating a phrase that doubly indicates that the speaker is about to offer her positive personal perception, set up in contrast to the negative portrayal of the community in the previous lines.

The third use of discourse-organizing creak indicates hedging or other uncertainty or signals self-repair. This use can be seen in Excerpts 4 and 5. For instance, Excerpt 4 below by Mauricio reengages a topic that we had been previously discussing (the socioeconomic differences present in my fieldwork neighborhoods) though I had just asked him an unrelated question: whether he had traveled outside of Chile.

(4) Mauricio describing travel experiences; DU: describe/explain (time neutral) (primary); sharing feelings and evaluations (secondary).

1		Eh, <u>no</u>
2		No he salido de ~~Chile~~ <whisper voice>
3		Tengo <lengthening> <u>a</u>migos sí que han <u>sa</u>lido mucho
4		<u>O</u> sea
5		Acá igual
6 7	→	Por ejemplo como te digo a<u>sí</u> de Las <emphasis> <u>Condes</u> <emphasis/> y qué sé yo
8		<u>Eh</u>
9		Sí me ha llamado harto la atención //
10		<u>E</u>so ha sido como una <lengthening> //
11		Como diferen<u>cia</u> //
12	→	<u>Así como</u> bastan<u>te</u> //
13		De estatus socioeconómi<u>co</u> <rising intonation> //
14		Y solamente me he dedicado a explorar acá lo que es <u>Chi</u>le no más //

1	Uh, no
2	I haven't left Chile
3	I do have friends that have traveled a lot
4	That is
5	here the same

6 7	For example like I was saying like from Las Condes and what do I know
8	Uh
9	That has really stood out to me
10	That's been a
11	like a difference
12	just like pretty . . .
13	about socioeconomic status
14	And I've dedicated myself to exploring here just within Chile

In Line 12, Mauricio is using creak to hedge or signal his uncertainty about how to formulate the remainder of this sentence most appropriately. In Chile, socioeconomic status is highly salient, but like in the United States, explicit talk about class can be taboo, and so the appropriate formulation eludes him in this line.

Another reformulation is seen in Excerpt 5 by Gustavo. I had just asked him which of his siblings he got along with best or with whom he was closest. In Line 3, Gustavo offers more detail about how he is closest to his brother, who lives in the south of Chile, whom he had identified in Line 2. He describes how he is like that brother in that they share many things, including a commitment to and affinity for something that he begins to identify at the end of Line 4. He cuts himself off, and in HCU Line 5, identifies that this thing they share is social action. Using creaky voice, Gustavo creates a repair from the previous line that he had cut off.

(5) Gustavo describing his relationships with his siblings; DU: sharing feelings and evaluations (primary); describe/explain the past (secondary).

1		Em
2		El que vive en el sur
3 4		Que yo comparto con él <emphasis> muchas <emphasis/> cosas como que a los dos nos gusta mucho la-
5	→	La acción social se podría decir //
6		Él también vivió en una población
7		Y, y
8		Y él estuvo trabajando en el Techo para Chile mucho tiempo
9		Después trabajó en <emphasis> Chiloé <emphasis> en el Techo pa Chi*le* <breathy>
10		Y, hace poquito se salió po //
11 12 13	→	Entonces como que los dos tenemos muchos temas en común siempre que nos juntamos como que salvamos el <emphasis> mundo <emphasis> por así decirle y creamos <emphasis> proyectos <emphasis> igual //
14		Nunca *ha*cemos <breathy> nada pero igual nos quedamos *ahí* <breathy/laughter> //
15		En la conversa //

1	Um
2	The one that lives in the south
3–4	I share a lot of things with him like both of us really like the
5	Social action you could say

6	He also lived in a low-income community
7	and, and
8	and he worked for Techo para Chile for a long time
9	After that he worked in Chiloé for the Techo para Chile
10	And, a little bit ago he got out *po*
11	So the two of us have a lot in common every time
12	we get together it's like we save the world for lack of a better way to say it and we
13	come up with projects too
14	We never do anything but it just stays there
15	In our conversation

Finally, the fourth category of creak used as a discourse organizer is similar to Lee's (2015) parenthetical grouping, shown in Excerpts 6 and 7. In this category, creaky voice is used to indicate off-the-record comments, jokes, and additional or preemptive provision of information. For instance, in Excerpt 6, Claudio is responding to my question about whether he gets along with his siblings and goes on to say much more about their present versus past relationship.

(6) Claudio responding to my question about whether he gets along with his siblings; *DU: describe/explain (time neutral); share feelings and opinions (secondary).*

1		¿con los, mis hermanos <rising intonation>? /
2		Eh s<u>í</u> /
3		<u>Sí</u> con mis hermanos /
4		<u>a</u>parte que estamos tod*os* <breathy voice> /
5		<emphasis> *des*parramados <emphasis> <breathy voice> /
6		tratamos de /
7		<u>esta</u>r <u>lo</u> <lengthening/> /
8		lo <emphasis> máximo uni<u>dos</u> </emphasis> /
9	→	no es como <emphasis> <u>antes</u> <emphasis> de /
10	→	<u>pu</u>cha a ver <u>cómo</u> le <u>pu</u>edo ex<u>plicar</u> /

1	With the, my siblings?
2	Eh yes
3	Yes with my siblings
4	Aside from the fact that we're all
5	scattered around
6	We try to
7	be the
8	most connected
9	it's not like before that
10	Jeez, let's see, how can I explain it to you

In Line 10, Claudio verbally expresses his uncertainty, stepping out of the frame of his response to my question about his siblings to make an off-the-record comment. He is frustrated at his inability to express himself exactly how he wants to (indicated by his

creaky production of the gentle expletive "pucha"), which he then follows with the phrase "a ver <u>cómo</u> le puedo expli<u>car</u>" (*how can I explain this to you*).

Additional or preemptive information was also included in Excerpt 7. I had just asked the speaker how she had met her boyfriend, and this HCT begins with her response.

(7) Carolina explaining how she met her boyfriend; DU: describe/explain the past (primary); sharing feelings and evaluations (secondary).

1		Eh <lengthening> mi polo<u>lo</u> <rising intonation/>
2 3	→	<u>En</u> el co<u>legio</u> en mi primer co**legio** <falsetto> yo estuve en <emphasis> <u>dos</u> <emphasis> co<u>legios</u> //
4		Y <lengthening> lo conocí
5	→	Cuando te<u>nía</u> <emphasis> do<u>ce</u> <emphasis> a~~ños~~ <whisper>
6		Polole**amos** <falsetto>
7		Típico pololeo de cabros chicos
8		De la <emphasis> mano <emphasis> no sé qué
9		Y terminamos porque su papá es
10	→	*S~~u~~* <whisper> papá es diplo<u>mático en</u>tonces viajan por muchos pa<u>í</u>*ses* <falsetto>
11		Pasaron <emphasis> doce a**ños** <emphasis> <falsetto>
12		<u>Que</u> no nos vimos más yo no supe de él
13		Ni Facebook ni nada <rising intonation>
14		Volvió <emphasis> este a**ño** <emphasis> <falsetto voice>
15		El vivió doce a<u>ños</u> en Armenia <rising intonation>
16		Y <lengthening> se vino <rising intonation>
17		Se le acabó la vi<u>sa</u> **creo** <falsetto>
18		Y <u>y</u> no po me agre**gó** <falsetto>
19		<u>A Facebook</u> <rising intonation>
20		Lo invité a un matrimonio <u>y</u> quedamos
21		Enamo<u>rados</u>
22		<u>Otra vez</u>
23	→	Y <lengthening> ahora estamos pololean~~do~~ *<whisper>*
	E:	Wuau, ¡es como una película!
1		Uh my boyfriend
2–3		In high school in my first high school I was in two high schools
4		And I met him
5		When I was twelve years old
6		We went out
7		A typical relationship between young kids
8		Holding hands and what have you
9		And we broke up because his dad is
10		His dad is a diplomat so they travel a lot
11		Twelve years went by
12		In which we didn't see each other at all and I didn't hear anything from him
13		Not even Facebook or anything

14		He came back this year
15		He'd spent twelve years in Armenia
16		And he came back
17		His visa expired I think
18		And and no *po* he added me
19		on Facebook
20		I invited him to a wedding and we fell
21		in love
22		again
23		and now we're back together
	E	Wow, it's like a movie!

In Utterance 2 (Lines 3 and 4), we see the first of two parenthetical additions of information. The speaker has begun the narrative of how she met her boyfriend but then deviates almost immediately to offer additional information regarding her primary school experience, specifically that she had been in one school but then changed to another (resulting in two different schools, as she mentions). I had asked her earlier what school she had gone to, and she had named one; here, she offers additional background information that is not immediately relevant to the current narrative. Another line of additional or background information is offered in Line 10 to explain why they had broken up: because her boyfriend's father was a diplomat, they had to leave Chile for his next posting. This line serves to indicate that had her boyfriend not been required to leave; they would likely have stayed together. The creak in Line 23, the final utterance of the turn, serves to indicate that the turn has come to an end, creating a transition-relevant place (like those shown in Excerpts 1 and 2). I responded appropriately by remarking that it was like a movie, and she replied that yes, many people tell her that their story is like a movie or fairytale.

The final example of parenthetical usage of creaky voice in this dataset is that of humor or jokes. In Excerpt 5 seen above, Gustavo describes how close he is with his brother due to their shared commitment to social action. In Lines 11–13, he says:

> "Entonces como que los dos tenemos muchos temas en común siempre que nos juntamos como que salvamos el <emphasis> mundo <emphasis> por así decirle y <u>creamos</u> <emphasis> <u>proyectos</u> <emphasis> <u>igual</u> //"

The use of creaky voice here follows a usage of irony in a hyperbolic sense (i.e., "salvamos el mundo" *we save the world*), and the creaky "creamos proyectos igual" (*we come up with plans, too*) conjures up a vision of the two brothers, affiliated in relationship and worldview, chatting and coming up with great plans. In the retelling, Gustavo downplays this project creation in a teasing fashion. This line also serves as the setup for the humorous, self-deprecating joke in Lines 14–15, in which he says "Nunca hacemos <u>na</u>da pero igual nos quedamos ahí" meaning *but we never actually carry out those plans*).

3.2. Creak Used to Invoke Alignment

I now turn to the use of creaky voice to invoke a convergent orientation between speaker and listener. Excerpts 8 through 10 reveal that creak is used to invoke alignment in two ways: to ensure understanding and to recruit affiliation on the part of the listener. Excerpt 8 exemplifies both types. Constanza lives with her parents, her sister, and her sister's two children and essentially acts as a co-parent to them. I asked specifically about whether she was close with her sister. In Excerpt 8, we first see creaky voice used to ensure understanding and later in the turn to seek affiliation.

(8) Constanza describing her relationship with her sister; DU: sharing feelings and evaluations (primary); describe/explain (time neutral) (secondary).

1	A mi \<emphasis\> herma*na* \<breathy voice\> \<emphasis/\> \<silencio\> //
2	Sí \<lengthening/\> //
3	No tengo problemas con ella //
4	De hecho, hay veces que parecemos \<emphasis\> matri*monio* \<breathy voice\> \</emphasis\>//
5	~~*Es verdad*~~ \<whisper and breathy voice\> //
6	Sí, es muy divertido cuando nos em ponemos a conversar \<rising intonation\> //
7 →	Y le digo *las* \<breathy, emphasis\> <u>cosas</u> \</emphasis\> //
8 →	y a veces <u>le echo</u> //
9	le digo un par de \<emphasis\> garabatos \<emphasis/\> y
10 →	<u>ella me los</u> \<emphasis\> devuelve \<emphasis\> y nos ponemos a \<emphasis\> pele*ar*
11	\<emphasis\>, el \<emphasis\> mano*tazo* \<emphasis\>, *pare-* \<breathy voice\> //
12	En serio que parecemos un matrimonio a vec<u>es</u> //
13	Es muy divertido //

1	With my sister
2	Yes
3	I don't have a problem with her
4	In fact, there are some times that we act like a married couple
5	It's true
6	Yes, it's really fun when we um start to chat
7	And I tell her things
8	Sometimes I toss
9–11	I throw out a couple of curse words, she throws them right back at me, and we start to fight, smack each other, we seem-
12	Seriously we seem like a married couple sometimes
13	It's really fun

In Line 4, the speaker makes an unexpected claim (that of sisters acting as a married couple). Though it is not recorded here, based on the response in Line 5 ("It's true!"), I must have responded in a nonverbal way that indicated disbelief or doubt, which could have included raised eyebrows or other markers of incredulity. Highly creaky Lines 7–9 are aimed at ensuring comprehension on my part via additional detail and support for her claim. As Clayman and Raymond (2021) indicate, this type of alignment token is present in utterances that are potentially problematic for recipients to grasp as intended (p. 300). Conversely, the affiliation use of creak is seen specifically in highly creaky Line 9. This utterance includes mention of several socially deviant behaviors, including cursing at one another and slapping each other. As Clayman and Raymond (2021) indicate, throughout this reporting of misdeeds, when the listener's affiliation may be seen as at risk, creaky voice as an alignment token "invokes support for the questionable action and often elicits confirmation" (p. 305). Interestingly, the slapping retell is overlaid not by creaky voice but by breathy voice. Perhaps, as other scholars have shown, non-modal voice quality may serve complementary purposes across the discourse (Podesva 2007, p. 487).

Other affiliation-seeking tokens are present throughout the dataset, such as above in Mauricio's Excerpt 4. In Line 3, Mauricio responded to my question about whether he had traveled outside the country by saying no, but that several of his friends had. He then uses two separate utterances, "O sea" (*I mean*) and "Acá igual" (*here the same*), to appropriately formulate his next statement, which begins with two separate discourse markers: "por

ejemplo" (*for example*) and "como te digo" (*like I was saying*) before the highly creaky portion of the utterance: "así de Las <emphasis> <u>Condes</u> <emphasis/> y qué sé yo" (*like from Las Condes and what do I know*). By using creaky voice in this utterance, Mauricio is not simply reporting where his friends are from but positioning Las Condes and its inhabitants as separate from himself, geographically but also socially, and inviting me to align with him in endorsement of this positioning. *Las Condes,* in this utterance, is emblematic of the moneyed social classes in Santiago (also comprising neighborhoods such as Vitacura, Lo Barnechea, and La Dehesa; (Rodríguez Vignoli 2007)). These neighborhoods are on the periphery of the city, difficult to access without a personal vehicle, and many of their residents are aligned with the political right (AS Chile 2020). Mauricio seeks my affiliation, my convergent orientation toward these neighborhoods, and perhaps negative valence or critical edge toward some of these characteristics of their inhabitants (Clayman and Raymond 2021, p. 303) in contrast with his own identity. I posit that creaky voice enables him to invoke this affiliation without explicitly stating it.

In the final two excerpts presented here, speakers Gonzalo and Mercedes use creak both to organize their discourse and to invoke alignment. Gonzalo uses creaky voice in Excerpt 9 to ensure understanding and to recruit affiliation for behavior that might be seen as controversial. At the beginning of the turn, I asked Gonzalo whether his parents are stricter with him than his younger female siblings and asked him about a specific example between Lines 8 and 9 or to expand upon his parents' hopes for him.

(9) Gonzalo responding to my question about whether his parents are stricter with him than with his younger (female) siblings; DU: sharing feelings and evaluations (primary); describing/explaining the past (secondary).

	E:	Y por ser el mayor, ¿tus papás son más estrictos contigo? <¿Más exigentes?>
1		<<u>Sí</u>.>
2		O sea <lengthening> <u>sí</u>
3		No sé si ahora tanto porque las otras ya son mayores de edad, pero
4		Antes, sí po a mí me llegaba el reto y después a la otra . . .
5		Más o menos no más <u>po ya</u> porque me habían retado a mí
6		Pero, sí aparte que soy el único hombre po en<u>to</u>nces
7		Yo creo que me ponen más . . .
8		<u>Más res</u>ponsabilidades
	E:	¿Y qué quieren para ti? ¿Quieren que tú seas ingeniero y esa cuestión? ¿O qué es lo que quieren para ti?
9		*Ø* <whisper> sea, cuando era chico siempre me presio<u>na</u>ron de estudiar
10		<u>U</u>na de las carreras principa<u>les</u> *po* <breathy>
	E:	¿Y cuáles son esas?
11	→	<u>No</u> sé co<u>mo</u> arquitectura, <u>ingeniería</u> comercial
12		Y yo como que no sabía mucho, no estaba ni ahí
13		Pero tenía que estudiar porque no me iban a dejar no estudiar
14	→	<u>Di</u>je, <quote> me conven<u>cí</u>, <u>ya</u> ingenie<u>ría</u> comercial <u>po</u>, si, <u>debe</u> ser lo mío po, si no
15		sé qué hacer <quote>
16		Y <u>chao</u>, <u>me</u> terminé, metiéndome a la-
17		A la U ahí
18	→	<u>Y</u> como era una escuela de negocios igual, no sé po era entreteni<u>do</u>, de <u>repente</u>, pero
19		después <u>ya</u> empecé a cha<u>tear</u> y los <u>ramos</u> y los ramos y los <u>ramos</u> y <u>no</u> <lengthening>

20	→	Ya no podía más
21		Por eso me, al final
22	→	Por eso to<u>mé</u> la decisión de irme \<lengthening\> <u>a</u> Mad<u>rid</u> *po* \<whisper\>
23	→	<u>Pa</u> cambiar un <u>poco</u> el aire
24	→	<u>Pa</u> seguir avanzando / /

	E	So because you're the oldest, do you think your parents are stricter with you? More demanding?
1		Yes
2		I mean yes
3		I don't know about now so much because the girls are over 18 but
4		before, yes *po* they always punished me and then to them
5		They only got punished a little because they'd already punished me
6		But yeah apart from that I'm the only boy so
7		I think they put more
8		more responsibilities [on me]
	E	And what do they want for you? Do they want you to be an engineer and all that? Or what do they want for you?
9		I mean, when I was younger they always pressured me to study
10		one of the primary careers *po*
	E	And which are those?
11		I don't know like architecture, commercial engineering
12		And since I didn't know that much, I didn't really care
13		But I had to study because they weren't going to let me not study
14 15		I said, I'm convinced, ok, commercial engineering, that must be my thing, I don't know what else to do
16		And bye, I eventually enrolled in the
17		in the university there
18 19		and since it was really a business school, I dunno, it was fun, sometimes, but after that I started to get fed up and the classes and the classes and the classes and I couldn't
20		I couldn't do it anymore
21		So then I, finally
22		So then I made the decision to go to Madrid *po*
23		For a change of scenery
24		To keep moving forward

In highly creaky Line 11, Gonzalo responds to my naïve question about what the primary degrees of study are in Chile. This line serves as a clarification or additional information but has the primary goal of enabling me to understand this relevant background information, leading to its categorization as an understanding-type alignment token. Its importance becomes even more apparent throughout the turn because, in Line 14, the speaker reports his own thought process that resulted in him undertaking the study of commercial engineering. This highly creaky utterance is also seeking understanding as to why he would embark on the study of a degree that he later realized he did not like. Lines 18, 19, and 20 are seeking affiliation, invoking the recipient's support for an ostensibly

"controversial" action or stance (Clayman and Raymond 2021, p. 303) of failing college courses. In Line 19, the speaker also repeats the same phrase ("los ramos" *the courses*) three times to indicate that the courses became overwhelming and contributed to his burnout. Repetition has been shown to be a formal expression of a speaker's evaluative stance, emphasizing the speaker's take on the relative importance of some narrative units (Labov and Waletsky 1967, p. 39). Line 20 is an explicit expression of the speaker's affective stance toward his career choice, which he had been building up to in the previous lines. Similarly, Lines 22–24 comprise the resolution of the speaker's narrative as he describes the choice he made to enroll in a university in Madrid. Overall, the creak in these lines enables the speaker to cast himself in a favorable light as he describes and emphasizes the strange and unusual character of his situation (Labov and Waletsky 1967, p. 34).

Finally, several different functions of creak as well as a critical examination of my role as the interlocutor, are present in Excerpt 10. In Line 5, as Mercedes is retelling the story of how her husband abandoned her, she jokingly downplays being left with only CLP 2000 (approximately USD 4) in her pocket by saying that it was all right because her mother treated her to lunch. She uses this joke to downplay the gravity of what had happened to her and ends this utterance with several creaky syllables. I hypothesize that the creaky voice here was not necessarily used to indicate a transition-relevant place nor simply as a joke-telling parenthetical use. Rather, I posit that the creaky syllables and downplaying jokes allow the speaker, who is momentarily portraying herself in an unfavorable light, both to acknowledge the delicateness of the situation and to try to remedy it (Haakana 2001, p. 213). The speaker may have been looking for some sort of indicator (a smile, a laugh, etc.) that her joke had been understood and that the remedy had served its purpose. However, I was expecting a continuation of the story rather than a humorous utterance and was not prepared to respond appropriately. While I do not have video evidence of this misstep, given her discourse marker in Line 7 and the hedge at the beginning of Line 8, I believe that this was a missed opportunity to respond and empathize in an appropriate way with the affiliation-seeking of the creaky, downplaying joke.

(10) Mercedes explaining the ending of her relationship with her children's father; DU: describing/explaining the past (primary); sharing feelings and evaluations (secondary).

1		Eh \<silencio/> //
2		Se *las* \<breathy, emphasis> llev*ó* \<breathy voice> \</emphasis> //
3		Él-- //
4	→	Yo me quedé con dos mil <u>pesos en el</u> \<emphasis> <u>bolsi</u>*llo* \<breathy, emphasis/>//
5 6	→	Pero mi mamá me invitaba a \<emphasis> almorzar \</emphasis> así que <u>no era co</u><u>mo</u> \<silencio/> //
7		Bueno //
8		Eh yo dije \<quote> bueno ni impor*ta* \<breathy> //
9		Cuando regrese con *las niñas* \<breathy; rising intonation> //
10		Eh me va a dar para <u>te</u>ner yo \<silencio/> \<quote/> //
11	→	No <u>fue así</u> //
12		Llegó con las niñas //
13	→	Manejando cur<u>ado y todo</u> //
14	→	Y me dijo \<quote> yo de aquí en adel<u>ante no te doy</u> //
15		\<emphasis> un pe*s*o̶ \<whisper voice> \<emphasis/> //
16		Arrégla<u>te</u>la <u>co</u>mo \<emphasis> pue*das* \<breathy voice>\<emphasis/> //
17		Me deman<u>das</u> //

18		Y ahí cuando– //
19		Salga el //
20	→	Eh <lengthening> lo que yo te tengo que dar **para las niñas** <falsetto voice/> //
21	→	Ahí conversamos <quote/> //

1	Uh
2	He left with them
3	He
4	I was left with two thousand pesos in my pocket
5	But my mom invited me to lunch so it wasn't like
6	Well
7	Uh I said well it doesn't matter
8	When he gets back with the girls
9	He'll give me something so that I have . . .
10	That didn't happen
11	He arrived with the girls
12	driving drunk and everything
13	and he said to me from now on I'm not giving you
14	one peso
15	Figure it out as you're able
16	Sue me
17	And then when
18	It arrives
19	Uh, what I have to give you for the girls
20	Then we'll talk

This excerpt also boasts several instances of creak used to organize discourse as well as additional alignment tokens. First, in Line 4, in addition to communicating information, the creaky voice enables the speaker to not only assert that what her ex-husband did was deplorable but also invites me to agree or align with this perspective. Other indicators of this alignment seeking are present throughout the turn. For instance, Lines 8–10 are reported speech from her own past self, optimistically expecting that upon her husband's return, he will see the error of his ways. Line 11 is a creaky discourse-organizing marker, shifting the communicative purpose from optimistic expectation to reporting information from the past. It also has a dramatic effect of indicating an action contrary to the speaker's (and therefore also the hearer's) expectations.

Similarly, highly creaky Line 13 is also seeking affiliation. Mercedes reports that her husband was driving drunk with her daughters in the car. In this reporting of behavior in breach of societal norms, the creaky voice invokes support for her position as critical of her husband, who has carried out this questionable and dangerous action (cf. Clayman and Raymond 2021, p. 303). This enables the speaker to position herself as reasonable and empathetic in comparison to her husband, who was unnecessarily cruel and careless. His reported cruelty is again confirmed in Line 14, in which he says he will not give her a single *peso* more. The reported speech of her husband (Lines 14–21) contributes to this distancing and negative portrayal, enabling the speaker to adopt a different voice from her own (cf. Stivers 2008, p. 39).

Line 20 is a highly creaky discourse organizer, enabling the speaker to repair the thought she started in Line 18. Perhaps she could not recall the exact wording of her husband's quote. However, it is more likely that she was not quoting verbatim but rather expressing his general sentiment to achieve a rough approximation of his positioning (cf. Clayman and Raymond 2021, p. 300). Lines 18 and 19 are both syntactically incomplete, and the speaker starts Line 20 with a filled pause "eh", signaling that she was still searching for the most appropriate way to express the term or idea she was looking for (something like the court-ordered child support amount). Line 21 rounds out the turn, creating a transition-relevant place via creaky voice, syntactic completeness, and the end of the reported speech.

Via the analysis presented here, I have identified two major functions of creak in these Chilean Spanish data: creak is used to organize discourse and to invoke alignment. I have shown that highly creaky utterances are used to organize discourse in four ways: at the end of larger utterances to signal a turn-yielding function; to indicate a pivot between discourse units (DUs) or communicative purposes; to signal a hedge or uncertainty or a site for self-repair; and to indicate a parenthetical (outside the discourse unit) joke, off-the-record comment, or additional or preemptive provision of information. These discourse-organizational uses account for approximately 40% of the highly creaky utterances in the dataset. The first finding aligns with previous work in English (Henton and Bladon 1988; Podesva 2013; Lee 2015) and Spanish (Garellek and Keating 2015; González et al. 2022) that shows that creak coincides with higher-order prosodic final positions. It also suggests that in Chilean Spanish, creaky voice may signal a turn-yielding function as in Finnish (Ogden 2001). Previous studies have identified that prosody plays a role in turn-yielding in addition to lexico-syntactic and pragmatic cues. For instance, Bögels and Torreira (2015) tested Dutch listeners' responses to the questions "Are you a student?" and "Are you a student at Radboud University?" to see whether they attended to lexico-syntactic cues or these plus prosodic cues to indicate turn-finality. The authors found that increases in both pitch and duration significantly contributed to quicker and more accurate turn-prediction. However, these scholars did not examine voice quality in their analyses. In a corpus-based examination of American English, Argentine Spanish, and Slovak, Heldner et al. (2019) found that the presence of creak is more likely in a turn-final, transition-relevant place (i.e., a transition where the next speaker starts talking following a short silence) than in turn-holding or backchannel positions. These findings point to the possibility of a more robust relationship between creaky voice and turn-finality than previously considered, and the results of the present study offer preliminary evidence for creaky voice as a turn-yielding cue. However, this brief review of previous studies draws attention to a body of work that has aimed to empirically define turn-ending. In contrast, the present study has taken a broader view of a turn as a recognizably complete utterance within a given context (Hoey and Kendrick 2017). To address this limitation, it would be helpful to examine the prosodic correlates of turn-finality more closely, such as voice quality, duration, and *f0*, and operationalize the lexico-syntactic features of turn-finality. Moreover, these judgments of turn-finality have not accounted for how Chilean Spanish listeners themselves interpret turn-ending and whether they attend to creak in making these judgments. Further perceptual support for this claim connecting highly creaky utterances with turn-finality is required.

Regarding the second organizational function, to my knowledge, this is the first study that has associated creak with discursive or communicative purpose-based shifting. As I have shown here, participants described events and people in the past and present and shared their feelings and opinions throughout the conversations, using creak to indicate a shift in communicative purpose. While this use is categorized as a discourse organizational tool, creak's co-occurrence with changes in communicative purpose also exhibits a stance-taking function: it is up to participants to decide who and what they will talk about and how they will express their personal stances and evaluations toward these chosen topics of conversation. Speakers are, therefore, agents in determining not only what they say but how they say it. This finding aligns with previous work (Grieser 2019; Kiesling 2009) that

has shown how speakers incorporate multiple linguistic tools to style different parts of themselves in their quest to make social meaning, and with recommendations to focus on the contribution of the listener to the meaning-making process (Soukup 2018).

Creaky voice used to indicate hedging, uncertainty, or sites for self-repair were also found in Lee's (2015) English-speaking data, as were examples of parenthetical or preemptive provision of information. This may indicate a cross-linguistic tendency that should be verified further using data from other languages. Creak used in tandem with hedging or uncertainty may also be indicative of a type of embodied cognition (Lakoff and Johnson 1999; Fauconnier and Turner 2002; Gibbs 2003) which connects the thought process and bodily experience. That is, speakers physically represent their current inability to complete their thought via a bodily expression of creak, which studies have shown to be accompanied by lower pitch and intensity (Shayan et al. 2011; Lefkowitz and Sicoli 2007), perhaps indicative of communication directed internally toward the self rather than outwardly toward the listener.

4. Conclusions

In this paper, I have shown how speakers of Chilean Spanish use creaky voice to organize their discourse and to position themselves in interaction with an unfamiliar interlocutor. Creak is primarily used in these data as an alignment token that invokes the listener's understanding or affiliation. Understanding may be at risk due to the speaker's formulation of their thoughts (often accompanied by false starts or hesitations) or the listener's non-local positioning (as an L2 speaker of Chilean Spanish and not a member of the local community). Affiliation may be at risk due to the speaker's retelling of socially risky behavior or stance or because of their positioning regarding other actions or others' perspectives. Similarities between the results of this study and Clayman and Raymond's (2021) analysis of the particle *you know* confirm that it is not creaky voice per se that can act as an alignment token but rather that the co-occurrence of creaky voice with lexical, syntactic, and prosodic features within a particular setting and with a particular sociocultural backdrop can invoke a convergent orientation between speaker and listener.

This analysis is one of the first of its kind within Chilean Spanish sociolinguistics. It has allowed for an examination of communication that eschews macrosocial demographic groupings typical of the first wave of sociolinguistics and centers stance as a resource for individual action (Jaffe 2009, p. 20). This focus on speakers' agency, affiliation, and their broader discursive purposes may inform other examinations of voice quality as an interactional resource, particularly given its focus on a non-English language.

Additional research in this area, including analyses of speakers' perception of the meanings of creaky voice, could contribute to determining whether this may be an indexical change in progress (Hall-Lew et al. 2021a), but as it stands, this study has shown that creaky voice can help construct meaning in situated interactions, serving as a resource for taking stances and making social moves (Hall-Lew et al. 2021b, p. 1).

Notes

[1] I align with Briggs (1986), Speer (2002), and Van den Berg et al. (2003), who advocate for treating the sociolinguistic interview as a conversation. Briggs (1986) specifically recommends critical analysis of the interview to examine points at which the interviewer and interviewee have misunderstood each other, as well as exploring the "communicative roots" of the interview to examine its norms with attention paid to the speaker's milieu as well as the discursive meaning being made.

[2] The age classifications are based on the year of birth related to the Pinochet dictatorship, aligning with Delforge's (2009) approach to relying on important historical and sociopolitical events to delineate age groups. The youngest speakers were at least 18 and were born post-dictatorship (so were 25 or younger at the time of data collection), 26–41-year-olds were born during the dictatorship, and 42+ were born prior to the 1973 coup.

[3] The EMIS system created by Sadowsky (2021) determines socioeconomic stratification based on the level of education and the occupation of the individual (or, in the case of younger speakers in the present study, the individual's parents). These groups are Chile-specific and, according to Sadowsky (2021), are approximately comparable in English to extreme upper (A), upper (B), upper middle (Ca), lower middle (Cb), lower (D), and extreme lower (E).

[4] There were a total of 92 unique highly creaky utterances, but four of them had two groupings of three or more highly creaky syllables: one at the end of the utterance (co-occurring with the end of the turn) and an additional group of three or more highly creaky syllables not in utterance-final position. I only counted an utterance twice if it showed these specific characteristics, making the denominator for all proportions 96.

[5] "Po" is a Chilean Spanish discourse marker deriving from the standard Spanish adverb "pues". It is used frequently in informal Chilean Spanish, often to underscore a point, signifying something like 'certainly' (Makihara 2005). Pragmatic cross-dialectal work has indicated that *pues* as a discourse marker is not grammatically essential, but leaving it out can affect the force of the utterance and/or the discourse cohesion (Fuentes-Rodríguez et al. 2016). These authors also indicate that *po(h)* used in Chilean Spanish may have multiple functions, including intensification and closing, suggesting that its meanings may differ based on utterance position and pragmatic value. For this reason, I have chosen not to gloss discursive *po* in the transcriptions above.

References

Acton, Eric K. 2021. Pragmatics and the Third Wave: The Social Meaning of Definites. In *Social Meaning and Linguistic Variation: Theorizing the Third Wave*. Edited by Lauren Hall-Lew, Emma Moore and Robert Podesva. Cambridge: Cambridge University Press, pp. 105–26. Available online: https://www.emich.edu/english/documents/faculty/pragmatics_and_the_third_wave.pdf (accessed on 22 July 2022).

AS Chile. 2020. Resultados Plebiscito Nacional 2020: Mapa por comunas de Santiago de Chile. *Diario AS*, October 26. sect. Actualidad. Available online: https://chile.as.com/chile/2020/10/26/actualidad/1603716098_944619.html (accessed on 12 December 2022).

Biber, Douglas, Jesse Egbert, Daniel Keller, and Stacey Wizner. 2021. Towards a Taxonomy of Conversational Discourse Types: An Empirical Corpus-Based Analysis. *Journal of Pragmatics* 171: 20–35. [CrossRef]

Bolyanatz, Mariška. 2023. Non-modal voice quality in Chilean Spanish. *Spanish in Context*, forthcoming.

Bögels, Sara, and Francisco Torreira. 2015. Listeners Use Intonational Phrase Boundaries to Project Turn Ends in Spoken Interaction. *Journal of Phonetics* 52: 46–57. [CrossRef]

Briggs, Charles L. 1986. *Learning How to Ask: A Sociolinguistic Appraisal of the Role of the Interview in Social Science Research*. Cambridge: Cambridge University Press.

Chafe, Wallace. 1993. Prosodic and Functional Units of Language. In *Talking Data: Transcription and Coding in Discourse Research*. Edited by Jane Edwards and Martin Lampert. Hillsdale: Lawrence Erlbaum Associates, Inc., pp. 33–44.

Clayman, Steven E. 2012. Turn-Constructional Units and the Transition-Relevance Place. In *The Handbook of Conversation Analysis*. Edited by Jack Sidnell and Tanya Stivers. Chichester: Wiley-Blackwell, pp. 150–66.

Clayman, Steven E., and Chase Wesley Raymond. 2021. 'You Know' as Invoking Alignment: A Generic Resource for Emerging Problems of Understanding and Affiliation. *Journal of Pragmatics* 182: 293–309. [CrossRef]

Crowhurst, Megan J. 2018. The Influence of Varying Vowel Phonation and Duration on Rhythmic Grouping Biases among Spanish and English Speakers. *Journal of Phonetics* 66: 82–99. [CrossRef]

Dallaston, Katherine, and Gerard Docherty. 2020. The Quantitative Prevalence of Creaky Voice (Vocal Fry) in Varieties of English: A Systematic Review of the Literature. *PLoS ONE* 15: e0229960. [CrossRef] [PubMed]

Davidson, Lisa. 2021. The Versatility of Creaky Phonation: Segmental, Prosodic, and Sociolinguistic Uses in the World's Languages. *Wiley Interdisciplinary Reviews Cognitive Science Cognitive Science* 12: e1547. [CrossRef] [PubMed]

Delforge, Ann Marie. 2009. The Rise and Fall of Unstressed Vowel Reduction in the Spanish of Cusco, Peru: A Sociophonetic Study. Ph.D. thesis, University of California, Davis, CA, USA.

Dilley, Laura, Stefanie Shattuck-Hufnagel, and Mari Ostendorf. 1996. Glottalization of Word-Initial Vowels as a Function of Prosodic Structure. *Journal of Phonetics* 24: 423–44. [CrossRef]

Du Bois, John. 2007. The Stance Triangle. In *Stancetaking in Discourse: Subjectivity, Evaluation, Interaction*. Amsterdam and Philadelphia: John Benjamins Publishing Company, pp. 139–82. Available online: http://dubois.faculty.linguistics.ucsb.edu/DuBois_2007_Stance_Triangle_M.pdf (accessed on 8 August 2022).

Du Bois, John, Stephan Schuetze-Coburn, Susanna Cumming, and Danae Paolillo. 1993. Outline of Discourse Transcription. In *Talking Data: Transcription and Coding in Discourse Research*. Edited by Jane Edwards and Martin Lampert. Hillsdale: Lawrence Erlbaum Associates, Inc.

Duranti, Alessandro. 1986. The Audience as Co-Author: An Introduction. *Text* 6: 239–47. [CrossRef]

Eckert, Penelope. 2012. Three Waves of Variation Study: The Emergence of Meaning in the Study of Sociolinguistic Variation. *Annual Review of Anthropology* 41: 87–100. [CrossRef]

Esposito, Christina. 2003. Santa Ana Del Valle Zapotec Phonation. Master's thesis, UCLA, Los Angeles, CA, USA.

Esposito, Christina, and Sameer ud Dowla Khan. 2012. Contrastive Breathiness across Consonants and Vowels: A Comparative Study of Gujarati and White Hmong. *Journal of the International Phonetic Association* 42: 123–43. [CrossRef]

Fauconnier, Gilles, and Mark Turner. 2002. *The Way We Think: Conceptual Blending and the Mind's Hidden Complexities*. New York: Basic Books.

Fuentes-Rodríguez, Catalina, María Elena Placencia, and María Palma-Fahey. 2016. Regional Pragmatic Variation in the Use of the Discourse Marker Pues in Informal Talk among University Students in Quito (Ecuador), Santiago (Chile) and Seville (Spain). *Journal of Pragmatics* 97: 74–92. [CrossRef]

Garellek, Marc. 2014. Voice Quality Strengthening and Glottalization. *Journal of Phonetics* 45: 106–13. [CrossRef]

Garellek, Marc, and Patricia Keating. 2011. The Acoustic Consequences of Phonation and Tone Interactions in Jalapa Mazatec. *Journal of the International Phonetic Association* 41: 185–205. [CrossRef]

Garellek, Marc, and Patricia A Keating. 2015. Phrase-Final Creak: Articulation, Acoustics, and Distribution. Presented at the 2015 Annual Meeting of the Linguistic Society of America, Portland, OR, USA, January 8–11. Available online: https://idiom.ucsd.edu/~mgarellek/files/Garellek_Keating_2015_LSA.pdf (accessed on 22 June 2022).

Gibbs, Raymond W. 2003. Embodied Experience and Linguistic Meaning. *Brain and Language* 84: 1–15. [CrossRef] [PubMed]

Goffman, Erving. 1974. *Frame Analysis: An Essay on the Organization of Experience*. Cambridge: Harvard University Press.

González, Carolina, Christine Weissglass, and Daniel Bates. 2022. Creaky Voice and Prosodic Boundaries in Spanish: An Acoustic Study. *Studies in Hispanic and Lusophone Linguistics* 15: 33–65. [CrossRef]

Grieser, Jessica A. 2019. Investigating Topic-Based Style Shifting in the Classic Sociolinguistic Interview. *American Speech* 94: 54–71. [CrossRef]

Haakana, Markku. 2001. Laughter as a Patient's Resource: Dealing with Delicate Aspects of Medical Interaction. *Text—Interdisciplinary Journal for the Study of Discourse* 21: 187–219. [CrossRef]

Hall-Lew, Lauren, Amanda Cardoso, and Emma Davies. 2021a. Social Meaning and Sound Change. In *Social Meaning and Linguistic Variation*, 1st ed. Edited by Lauren Hall-Lew, Emma Moore and Robert Podesva. Cambridge: Cambridge University Press, pp. 27–53. [CrossRef]

Hall-Lew, Lauren, Emma Moore, and Robert Podesva. 2021b. Social Meaning and Linguistic Variation: Theoretical Foundations. In *Social Meaning and Linguistic Variation: Theorizing the Third Wave*. Edited by Lauren Hall-Lew, Emma Moore and Robert Podesva. Cambridge: Cambridge University Press, pp. 1–24.

Heldner, Mattias, Marcin Włodarczak, Štefan Beňuš, and Agustín Gravano. 2019. Voice Quality as a Turn-Taking Cue. In *Interspeech 2019*. Graz: ISCA, pp. 4165–69. [CrossRef]

Henton, Caroline, and Anthony Bladon. 1988. Creak as a Sociophonetic Marker. In *Language, Speech, and Mind: Studies in Honour of Victoria A. Fromkin*. Edited by Larry M Hyman and Charles Li. London: Routledge, pp. 3–29.

Hildebrand-Edgar, Nicole. 2016. Creaky Voice: An Interactional Resource for Indexing Authority. Master's thesis, University of Victoria, Victoria, BC, Canada. Available online: http://dspace.library.uvic.ca/bitstream/handle/1828/7437/Hildebrand-Edgar_Nicole_MA_2016.pdf?sequence=1&isAllowed=y (accessed on 8 September 2022).

Hoey, Elliot M., and Kobin H. Kendrick. 2017. Conversation Analysis. In *Research Methods in Psycholinguistics: A Practical Guide*. Edited by Annette de Groot and Peter Hagoort. Hoboken: Wiley Blackwell, pp. 151–73. Available online: https://pure.mpg.de/rest/items/item_2328034/component/file_2328033/content (accessed on 8 September 2022).

Hollien, Harry. 1987. 'Old Voices': What Do We Really Know about Them? *Journal of Voice* 1: 2–17. [CrossRef]

Jaffe, Alexandra M. 2009. Introduction: The Sociolinguistics of Stance. In *Stance: Sociolinguistic Perspectives*. Edited by Alexandra M. Jaffe. Oxford Studies in Sociolinguistics. Oxford and New York: Oxford University Press, pp. 1–28.

Keating, Patricia, Marc Garellek, and Jody Kreiman. 2015. Acoustic Properties of Different Kinds of Creaky Voice. Paper presented at ICPhS 2015, Glasgow, UK, August 10–14.

Kiesling, Scott. 2009. Style as Stance: Stance as the Explanation for Patterns of Sociolinguistic Variation. In *Stance: Sociolinguistic Perspectives*. Edited by Alexandra Jaffe. Oxford: Oxford University Press, pp. 171–94.

Labov, William, and Joshua Waletsky. 1967. Narrative Analysis: Oral Versions of Personal Experience. In *Essays on the Verbal and Visual Arts: Proceedings of the 1966 Annual Spring Meeting of the American Ethnological Society*. Edited by June Helm. Seattle and London: University of Washington Press, pp. 12–44.

Lakoff, George, and Mark Johnson. 1999. *Philosophy in the Flesh: The Embodied Mind and Its Challenges to Western Thought*. Chicago: University of Chicago Press. Available online: https://www.degruyter.com/database/COGBIB/entry/cogbib.7098/html?lang=en (accessed on 15 December 2022).

Lee, Sinae. 2015. Creaky Voice as a Phonational Device Marking Parenthetical Segments in Talk. *Journal of Sociolinguistics* 19: 275–302. [CrossRef]

Lefkowitz, Daniel, and Mark A. Sicoli. 2007. Creaky Voice: Constructions of Gender and Authority in American English Conversation. Panel on Phonation, Voice, and Style presented at the American Anthropological Association, November 28. Available online: https://anthropology.as.virginia.edu/sites/anthropology.as.virginia.edu/files/lefkowitz.sicoli_creaky.voice.constructions.of.gender.and.authority.pdf (accessed on 22 June 2022).

Levinson, Stephen C., and Francisco Torreira. 2015. Timing in Turn-Taking and Its Implications for Processing Models of Language. *Frontiers in Psychology* 6. [CrossRef] [PubMed]

Makihara, Miki. 2005. Rapa Nui Ways of Speaking Spanish: Language Shift and Socialization on Easter Island. *Language in Society* 34: 727–62. [CrossRef]

Ogden, Richard. 2001. Turn Transition, Creak and Glottal Stop in Finnish Talk-in-Interaction. *Journal of the International Phonetic Association* 31: 139–52. [CrossRef]

Podesva, Robert. 2007. Phonation Type as a Stylistic Variable: The Use of Falsetto in Constructing a Persona. *Journal of Sociolinguistics* 11: 478–504. [CrossRef]

Podesva, Robert. 2013. Gender and the Social Meaning of Non-Modal Phonation Type. *Proceedings of the Annual Meeting of the Berkeley Linguistics Society* 37: 427–48. [CrossRef]

Podesva, Robert, and Patrick Callier. 2015. Voice Quality and Identity. *Annual Review of Applied Linguistics* 35: 173–94. [CrossRef]

Redi, Laura, and Stefanie Shattuck-Hufnagel. 2001. Variation in the Realization of Glottalization in Normal Speakers. *Journal of Phonetics* 29: 407–29. [CrossRef]

Rodríguez Vignoli, Jorge. 2007. Segregación residencial, migración y movilidad espacial. El caso de Santiago de Chile. *Cadernos Metrópole.* 17: 135–68.

Sacks, Harvey, Emanuel Schegloff, and Gail Jefferson. 1978. A Simplest Systematics for the Organization of Turn Taking for Conversation. In *Studies in the Organization of Conversational Interaction*. Edited by Jim Schenkein. Cambridge: Academic Press, pp. 7–55. [CrossRef]

Sadowsky, Scott. 2021. EMIS: Sistema de estratificación socioeconómica para la investigación lingüística. In *Chilean Spanish Linguistics: Studies on Variation, Innovation, Contact, and Identity*. Edited by Brandon M. A. Rogers and Mauricio A. Figueroa Candia. Wilmington: Vernon Press, pp. 367–96.

Schegloff, Emanuel, Elinor Ochs, and Sandra Thompson. 1996. Introduction. In *Interaction and Grammar*. Edited by Elinor Ochs, Emanuel Schegloff and Sandra Thompson. Studies in Interactional Sociolinguistics. Cambridge: Cambridge University Press, pp. 1–51. Available online: http://www.sscnet.ucla.edu/anthro/faculty/ochs/articles/Ochs_Schegloff_Thompson_1996_Introduction.pdf (accessed on 6 February 2023).

Selting, Margret. 2000. The Construction of Units in Conversational Talk. *Language in Society* 29: 477–517. [CrossRef]

Shayan, Shakila, Ozge Ozturk, and Mark A. Sicoli. 2011. The Thickness of Pitch: Crossmodal Metaphors in Farsi, Turkish, and Zapotec. *The Senses and Society* 6: 96–105. [CrossRef]

Sicoli, Mark A. 2010. Shifting Voices with Participant Roles: Voice Qualities and Speech Registers in Mesoamerica. *Language in Society* 39: 521–53. [CrossRef]

Sidnell, Jack. 2013. Basic Conversation Analytical Methods. In *The Handbook of Conversation Analysis*. Edited by Jack Sidnell and Tanya Stivers. Chichester: Wiley-Blackwell, pp. 77–99.

Slifka, Janet. 2006. Some Physiological Correlates to Regular and Irregular Phonation at the End of an Utterance. *Journal of Voice* 20: 171–86. [CrossRef] [PubMed]

Soukup, Barbara. 2018. Contextualizing the Third Wave in Variationist Sociolinguistics: On Penelope Eckert's (2018) Meaning and Linguistic Variation. *Vienna English Working Papers* 27: 51–67.

Speer, Susan A. 2002. 'Natural' and 'Contrived' Data: A Sustainable Distinction? *Discourse Studies* 4: 511–25. [CrossRef]

Stivers, Tanya. 2008. Stance, Alignment, and Affiliation During Storytelling: When Nodding Is a Token of Affiliation. *Research on Language & Social Interaction* 41: 31–57. [CrossRef]

Torres, Peter Joseph, Stephen Gresham Henry, and Vaidehi Ramanathan. 2020. Let's Talk about Pain and Opioids: Low Pitch and Creak in Medical Consultations. *Discourse Studies* 22: 174–204. [CrossRef]

Van den Berg, Harry, Margaret Wetherell, and Hanneke Houtkoop-Steenstra. 2003. *Analyzing Race Talk: Multidisciplinary Perspectives on the Research Interview*. Cambridge: Cambridge University Press.

Wilce, James M. 1997. Discourse, Power, and the Diagnosis of Weakness: Encountering Practitioners in Bangladesh. *Medical Anthropology Quarterly* 11: 352–74. [CrossRef]

Wolk, Lesley, Nassima B. Abdelli-Beruh, and Dianne Slavin. 2012. Habitual Use of Vocal Fry in Young Adult Female Speakers. *Journal of Voice: Official Journal of the Voice Foundation* 26: e111–e116. [CrossRef]

Yuasa, Ikuko. 2010. Creaky Voice: A New Feminine Voice Quality for Young Urban-Oriented Upwardly Mobile American Women? *American Speech* 85: 315–37. [CrossRef]

Article

Correntino Spanish Memes and the Enregisterment of Argentine Guarani Loanwords

Justin Pinta

Department of Classical and Modern Languages and Literatures, Mississippi State University, Starkville, MS 39759, USA; jpinta@cmll.msstate.edu

Abstract: The intense contact between Guarani and Spanish in the Argentine province of Corrientes has produced a wide array of mutual contact effects, the most visible being widespread borrowing in both directions. This article examines a previously unreported feature of Argentine Guarani loans in Correntino Spanish: the social value they have acquired. Building on the growing body of work in sociolinguistics on internet memes, which are sites of phenomena rich in sociolinguistic value, an analysis is provided of Argentine Guarani loans in Correntino Spanish using an original corpus of memes collected from Correntino Instagram pages. Such memes, whose intended audience is monolingual, are a valuable source of Correntino Spanish features, which are used for humorous, ironic, or nostalgic effect. Via analysis of the relationship between these loans and the kinds of memes they appear in, I show that these loans have undergone enregisterment, i.e., they have taken on additional social meaning that allows them to index a complex variety of ideological stances toward Correntino social phenomena and character types. The results of this process evidence the fact that language contact, as an engine of variation, creates fertile ground for the emergence of social meaning and that memes are a productive and promising window into the (re)creation and evolution of such meaning.

Keywords: enregisterment; indexicality; character types; Spanish; Guarani; loanwords; memes

1. Introduction

The scholarly treatment of language variation has evolved since variationist analysis began in earnest in the mid-twentieth century. Variationist studies are commonly grouped into three "waves" (Eckert 2012), delineated by differing theoretical treatment of the emergence of social meaning in variation and shifting approaches to speaker agency. As Hall-Lew et al. (2021, p. 1) summarize:

> Traditional first- and second-wave approaches have treated variation as a window into language change (Weinreich et al. 1968), and have examined the stratification of variation according to both macrosocial categories (e.g., gender, class) and locally significant categories (e.g., jocks, burnouts). Proponents of a third-wave approach focus on linguistic variation as a resource for taking stances, making social moves, and constructing identity.

Approaches in the third (and current) wave, by virtue of their emphasis on the role variation plays in the construction of social categories as opposed to being mere by-products of them, offer important insight into the use of contact features in Correntino Spanish, a variety of Spanish unique to the Argentine province of Corrientes. Correntino Spanish is notable among Spanish varieties for the myriad ways in which it has affected and been affected by Argentine Guarani[1] (Cerno 2017; Dietrich 2002), a variety of Guarani unique to Corrientes (closely related to but historically and linguistically differentiated from Paraguayan Guarani). Due to a complex mixture of sociohistorical and ideological factors, Correntinos find themselves situated geographically, culturally, and linguistically

between Paraguay and Buenos Aires, the latter being the perceived core of Argentine culture throughout Latin America. The linguistic hallmarks of the province—Spanish-influenced Guarani and Guarani-influenced Spanish—are in many ways maligned by non-Correntinos and Correntinos alike, with Argentine Guarani characterized as an illegitimate variety of Guarani, not as "pure" as the variety spoken in Paraguay, and Correntino Spanish similarly characterized as an "impure" variety of Spanish (Pinta 2022). This overtly linguistic discrimination of Corrientes and its inhabitants, discursively justified on grounds of language purism, is inevitably linked to the social characteristics of the province, which is seen by many non-Correntino Argentines—in particular by *porteños*, i.e., citizens of Buenos Aires—as a socially backwards province characterized by poverty and superstition. An understanding of the social environment of Correntino Spanish (and Argentine Guarani) helps elucidate the linguistic phenomena that make the province unique, phenomena drawn from a diverse and rich pool of linguistic resources which are recruited by speakers in myriad ways to achieve myriad social ends.[2]

While the list of linguistic features that make Correntino Spanish unique as a variety goes beyond contact effects, the influence of Guarani on this variety is uncontroversially the most important of its delineating characteristics, some but not all of which are shared by Paraguayan Spanish and very few of which are shared by other Spanish varieties spoken in Argentina. Guarani-source contact effects are found in all areas of the grammar, the most visible, of course, being the lexicon. In this article, I examine a previously unreported feature of Argentine Guarani loans in Correntino Spanish: the social value they have acquired. Via analysis of an original corpus of internet memes collected from Correntino Instagram pages, I draw connections between the use of Guarani loans and the thematic patterns of the memes they appear in. I show that Guarani loans have undergone enregisterment (Agha 2003), i.e., that they have taken on additional social meaning which allows them to index a complex variety of ideological stances toward Correntino social phenomena and character types. This study therefore not only addresses social meaning in the context of third-wave sociolinguistic analysis generally but also foregrounds the role that contact features play in the accrual and deployment of social meaning in communities whose linguistic repertoires are the products of intense language contact.

Such social meaning is malleable and subject to context-specific interpretation. In Correntino internet memes, Guarani loans may in one meme be used in a fond, nostalgic manner while in another be used in the service of self-deprecating humor. They can be differentially associated with the romanticized ideal of the Correntino gaucho, i.e., the hard-working family man who is a skilled horse rider, expert in agricultural matters, master of the outdoors, and dependable head of household, as well as the Correntino womanizer, i.e., the smooth-talking, often heavy-drinking bender of the truth who is trusted by neither his romantic partner nor close friends. In their discussion of the connection between linguistic forms and social meaning, Hall-Lew et al. (2021, p. 5) note that "[t]he social meaning(s) that listeners arrive at, however vaguely, can only be determined in the moment of use, dependent on the particular ideologies made relevant in context". This notion will prove crucial to an understanding of the varied and sometimes contradictory types of social meaning which Guarani loans are linked to, as will be further developed below.

This article is structured as follows. Section 2 provides the background necessary to situate the study, addressing both its theoretical grounding and internet memes as objects of inquiry. In Section 3, methodological details are provided, including the composition of the meme corpus on which this study is based and the analytical approach adopted here. Section 4 contains a thematic analysis of the memes in the corpus, followed by Section 5, in which the phenomena in the corpus are linked to the previous literature discussed in Section 2, demonstrating how these patterns inform our understanding of the social value of Guarani loans in Correntino Spanish generally. Finally, the article concludes with summarizing remarks in Section 6.

2. Background

2.1. Enregisterment, Indexical Fields, and Character Types

The nature of the relationship between linguistic variables, speech styles, and social categories has been at the heart of much of the work within third-wave sociolinguistics and a crucial theoretical framework used to understand the mechanism by which linguistic variables become meaningful social objects has been that of enregisterment. First used and defined by Agha (2003, p. 231) as "processes through which a linguistic repertoire becomes differentiable within a language as a socially recognized register of forms", the process of enregisterment produces an identifiable relationship between linguistic features and social categories. The acquired social value of such features in turn demarcates them in important ways from other features that are part of the same linguistic system. The attribution of social value to a speech variety is contingent on the enregisterment of a subset of linguistic features of that variety, with the members of the subset being a class subject to constant reinterpretation and, over time, addition or subtraction. Agha (2003, p. 232) notes that such cultural value "is not a static property of things or people but a precipitate of sociohistorically locatable practices, including discursive practices, which imbue cultural forms with recognizable sign-values and bring these values into circulation along identifiable trajectories in social space". As will be shown later, Correntino Spanish memes provide a unique and relatively novel window into the ways in which linguistic features that have come to acquire such social value are used and understood.

The theory of enregisterment has been applied to a number of sociolinguistic case studies since its introduction nearly 20 years ago. (For a variety of examples, see the 2009 special issue of *American Speech*, vol. 84, no. 2.) Zhang (2021) discusses the enregisterment of Cosmopolitan Mandarin, a novel style of Mandarin that is shown to participate in the creation of new social distinctions as opposed to being a mere product of them. In her analysis of the enregisterment of the distinguishing features of this variety, Zhang (2021, p. 281) notes "that the emergence of [Cosmopolitan Mandarin] is a contested process fraught with indexical instability and multiplicity, a process whereby [Cosmopolitan Mandarin] and its constitutive elements are imbued with varied and often conflicting indexical values". This resulting set of conflicting values mirrors what occurs in Corrientes, as will be further elaborated on below.

Another well-known application of enregisterment is its use by Johnstone et al. (2006) in their analysis of a set of linguistic features unique to the variety of North American English spoken in Pittsburgh, Pennsylvania. They provide a classification of variable indexical meaning via analysis that combines a taxonomy of social meaning proposed by Labov (1972) with the orders of indexicality proposed by Silverstein (2003).[3] They show that some features of "Pittsburghese" have gone through a three-step process: (1) acquiring first-order indexicality (i.e., becoming an "indicator"), by which they are regionally correlated with a group of speakers but not overtly recognized as such; (2) acquiring second-order indexicality (i.e., becoming a "marker"), by which speakers become aware of a connection between such features and locality, therefore allowing the features to become available for social work; and (3) acquiring third-order indexicality (i.e., becoming a "stereotype"). Once features acquire third-order indexicality, the conditions are created for speakers (and even non-speakers) of a given variety to "use regional forms drawn from highly codified lists to perform local identity, often in ironic, semiserious ways" (Johnstone et al. 2006, p. 83). Memes, as discussed below, provide a clear window into both the recruitment of features for such use and the end result of the set of social meanings they can acquire.

Such a set of meanings, which speakers draw from and alter according to their own ideological stances, constitute what Eckert (2008) calls an "indexical field", or a non-static constellation of meanings which are linked ideologically. Building on the influential work of Silverstein (2003) on indexicality and indexical orders, Eckert proposes indexical fields as a way of accounting for constantly shifting social meanings associated with a synchronically static set of linguistic features, noting that indexical fields are "fluid, and each new activation has the potential to change the field by building on ideological connections"

(Eckert 2008, p. 454). Experimental work by Campbell-Kibler (2007) demonstrates this fluid nature of indexical fields by showing that listener evaluations of the social meaning of enregistered features—in the case of this study, the rendering of the morpheme *-ing* in US English as either [ɪn] or [ɪŋ]—is malleable according to context. She notes that the use of the former variant "influences the perception of education across multiple speakers on the one hand, and in a subset of the speakers also affects an accent associated with lack of education", going on to stress that such seemingly paradoxical interconnections "are a fundamental aspect of sociolinguistic variation" (Campbell-Kibler 2007, p. 55).

The variation in interpretation of the social value of enregistered features further allows for the same features to be associated with differing speech styles, and subsequently, differing stereotypes of who uses such features. This phenomenon has been addressed in work on "character types"—sometimes used synonymously with "characterological figures" (Agha 2007, p. 177) or "personae" (Drager et al. 2021, p. 176)[4]—i.e., "stereotypical, reified figure[s] such as the California 'Valley girl,' often circulated via mass media" (Starr 2021, pp. 315–16). A crucial component of character types, as noted by Agha (2007, p. 177), is that they are "performable through a semiotic display or enactment (such as an utterance)". That is, one end product of the process of enregisterment is that a set of enregistered features can be drawn on in the portrayal of stereotypical figures in ways that are locally meaningful. These representations are linked to larger societal characteristics that manifest themselves in individualized representations. As Eckert (2016, p. 75) notes (using "personae" in lieu of "character types"): "The attention to personae shifts the focus away from the social aggregate to individuals as they move through identities and situations. However, it does not amount to a study of the individual, but of the structure within which individuals find and make meaning". Internet memes are a productive medium for the reproduction of such character types, which are construed via use of enregistered features, as further discussed below.

2.2. Memes as Objects of Analysis

Recent years have seen the emergence of internet memes as objects of sociolinguistic interest and analysis, in line with a growing awareness of memes as sites of phenomena rich in sociolinguistic value. Such value emerges both from the linguistic features that surface in these contexts but not in formal writing contexts and from the content of the memes themselves, which is often a valuable source of sociocultural references. A comparison could be made to the value of the colloquial Latin graffiti at Pompeii, which provides a source of linguistic and sociocultural information not available in more formal texts (Jones 2016). In the same way, memes can provide a space for colloquial, informal linguistic features—precisely the kinds of features that are so often imbued with social meaning—which might not appear in other written contexts.

The term "meme" was first coined by Dawkins (1976), who used it within the context of evolutionary biology, contrasting memes with genes.[5] Taking Dawkins as a starting point, Davison (2012, p. 122) discusses the transition to the modern meaning of "meme", ultimately providing his own definition:

> In Dawkins's original framing, memes described any cultural idea or behavior. Fashion, language, religion, sports—all of these are memes. Today, though, the term "meme"—or specifically "Internet meme"—has a new, colloquial meaning. While memes themselves have been the subject of entire books, modern Internet memes lack even an accurate definition. There are numerous online sources (Wikipedia, Urban Dictionary, Know Your Meme, Encyclopedia Dramatica) that describe Internet memes as the public perceives them, but none does so in an academically rigorous way. Given this, I have found the following new definition to be useful in the consideration of Internet memes specifically: *An Internet meme is a piece of culture, typically a joke, which gains influence through online transmission.*

To narrow the definition even further, internet memes are understood in popular culture as being images (or sometimes short videos, gifs, or chunks of text), often with accompanying captions, and this is the sense in which "meme" is used here. An example of a meme from the corpus (whose composition will be discussed in detail in Section 3.1) is given in Figure 1, with an English translation provided immediately below it. Two Guarani loans are found in this meme: *caú* 'drunk' (from Guarani *ka'u*) and *que*, an intensifier particle used in imperative contexts (from the Guarani form *ke*, whose phonetic realization and function are the same as the Spanish borrowed form).

> Si hago alguna pavada estando caú no te vayas que a reír
> @memelandiacorrentino
> Tu amigo

-If I do something stupid while drunk don't go laughing
-Your friend

Figure 1. Example of a meme from the Correntino Spanish meme corpus containing the Guarani loans *caú* 'drunk' and *que*, an imperative intensifier particle.[6]

The key consideration for this analysis is not what does and does not count as a meme[7] but rather that a central property of memes is their sociocultural relevance, often to narrow or relatively small social groups. Such memes must necessarily be created by a member of such a social group (or someone with a thorough understanding of the group norms and values) and are likely to only be fully understood or appreciated by a member of the same group. For instance, while the general content of the meme provided in Figure 1 is easily interpretable by anyone, there is an additional layer of content that grounds the meme locally: the man laughing in the photo is the current governor of Corrientes, Gustavo Valdés.

As Shifman (2011, p. 188–89) points out, "[O]nly memes suited to their socio-cultural environment will spread successfully; the others will become extinct...While some memes are global, others are more culture specific, shaping collective actions and mindsets". In bringing up the notion of a meme's success, Shifman refers to the propensity of a meme to be shared and to spread, either in its verbatim form or in an evolved form in which someone else reproduces the template and aesthetic but changes the joke. The ability of a meme to evolve, and the agency of the individual in the reinterpretation and redesign of a meme, is a crucial factor in a meme's "virality", or ability to spread quickly (Varis and Blommaert 2015), but a meme's sociocultural relevance is ultimately the most important factor in determining its success. Knobel and Lankshear (2007, p. 209), in an early analysis of memes, comment on this as well, noting that one of the most important factors in the success of a meme is that it have "[a] rich kind of intertextuality, such as wry cross-references to different everyday and popular culture events, icons or phenomena".

This "rich kind of intertextuality" can manifest itself culturally, linguistically, or both. For this reason, memes are often treated as a unique genre, with Wiggins (2019, p. 40) clarifying, "[A] meme, viewed as a genre, is not simply a formula followed by humans to

communicate, but represents a complex system of social motivations and cultural activity that is both a result of communication and impetus for that communication". It is precisely this property of memes that has drawn the attention of sociolinguists, combined with the fact that the common tools of linguistic analysis are extremely well suited for the analysis of memes (Cochrane et al. 2022).

Meme analysis in sociolinguistics, a subset category of the larger body of work addressing language in computer-mediated contexts (e.g., see Squires 2016), has proven a valuable approach in a variety of cases. Ndoci (forthcoming), for instance, via analysis of Greek memes, demonstrates how the construction of the Albanian immigrant stereotype is propagated via memes that utilize stereotyped features of the L2 Greek of Albanians. Members of Greek society without direct interaction with Albanian immigrants thus become aware of a link between these features and social notions of Albanians in Greece, subsequently further propagating harmful stereotypes via the resulting mock language (Hill 2007); memes provide a vehicle for such a phenomenon, which would not be possible otherwise among Greeks without direct exposure to such linguistic features.

Other studies have looked at how specific memes have accrued complex social meaning. Aslan and Vásquez (2018) analyze the "Cash me ousside/howbow dah" meme—a meme based on an utterance produced by a teenage girl on an American talk show in 2016—and the ways in which it indexes a variety of differing social meanings and categories. Drawing on the notion of "citizen sociolinguistics" (Rymes and Leone 2014), an approach that focalizes the participation of non-linguists in sociolinguistic exploration, they analyze metalinguistic commentary on YouTube videos to show how race, region, education, and class converge in the interpretation and reproduction of this meme, noting "that these categories overlap in complex, and not always predictable, configurations" (Aslan and Vásquez 2018, p. 406). Similarly, Procházka (2019) demonstrates how memes index a complicated mixture of ideologies via analysis of the "countryball" meme comics, which soared in popularity during the European migrant crisis. Countryball memes involve simplistic drawings of ball-shaped characters whose color patterns resemble the flags of different countries and whose linguistic and behavioral characteristics are satirized representations of sociocultural stereotypes associated with the countries they represent. The use and interpretation of these memes, located in the context of anti-immigrant and nationalist ideologies, resist a simple good-bad binary and evidence more complicated and overlapping ideological positions that are highly individualized. The analysis provides an important case study of memes as valuable and informative sociolinguistic objects of modern inquiry, concluding that "memes are not a mere product of participatory culture, but rather a powerful instigator of technosocial and often heteroglossic practices that co-organize social life in the new polycentric collectivities appearing on social media" (Procházka 2019, p. 717).

Drawing on these studies and the growing body of literature on memes in general, this study provides a novel analysis of the use of Argentine Guarani loans in Correntino Spanish memes. The uniqueness of Correntino Spanish combined with the particularly strong sense of Correntino identity allows for a sociolinguistic context that produces a complicated array of phenomena. As discussed in these previous analyses of memes, the linguistic features in question here, i.e., Argentine Guarani loans, are used in the service of a variety of seemingly conflicting ideological stances, themselves rooted in the contrast between local identity and the larger categories of *porteño* and Paraguayan identities. Correntino identity is forged simultaneously by drawing on aspects of these larger identities and by reinterpreting those aspects in opposition to them. Memes provide an easily accessible and easily malleable resource by which to produce and reproduce ideological stances about what is and is not "Correntinoness". To demonstrate this, a unique corpus of Correntino Spanish memes was created, the composition of which is discussed in the following section.

3. Methodology

3.1. Meme Corpus

The data informing this study come from a corpus of 409 unique memes, all in Correntino Spanish, compiled from 2020 to 2022. The criterion for inclusion in the corpus was whether or not a given meme contained at least one Guarani loanword. However, given that the goal in creating the corpus was to focus narrowly on Correntino Spanish, all Guarani loanwords that have entered Argentine Spanish broadly conceived were left out. The criterion for establishing this categorization was made on the basis of the geographic diffusion of a given loan. Guarani loans that are common in the urban variety of Spanish spoken in Buenos Aires, such as loans for animal names (e.g., *yacaré* 'caiman'), placenames (e.g., *Iguazú*, the name for Iguazu Falls), and other various loans (e.g., *tereré*, yerba mate drunk cold), were excluded from the corpus, as they are not emblematic of local speech and thus not available for explicit connection with Correntino culture and identity in the same way. The assessment of whether or not a given Guarani-source form is common in Buenos Aires was made on the basis of my own high degree of familiarity with that variety and, in cases of doubt, in consultation with native speakers of non-Correntino Argentine Spanish varieties.

The sources of the memes are 11 Instagram pages that are run by Correntinos and intended for Correntinos; these pages were chosen via an in-depth survey of Correntino Instagram pages, and all pages that were thematically locally oriented were included. These pages, along with their Instagram handle,[8] number of followers, and the location out of which each page is run, are provided in Table 1. Additionally, a map of the province is provided in Figure 2, in which each of the locations in Table 1 is indicated. Some pages contributed more memes to the corpus than others, as some were much more prolific and active in posting than others. The number of memes from each specific source page is found in Figure 3.

The process of "collecting" a meme involved manually identifying it as containing a loan, screenshotting it, saving it, categorizing it by page of origin, and entering it into a database in Excel. Each meme was coded for the amount of loans it contained, what the loans were, whether the loans were morphosyntactic or lexical, and whether or not the content of the meme made explicit reference to Correntino identity or culture (discussed further in Section 3.2). Specific URL sources for each of the memes reproduced in this article are given in a footnote at the bottom of the page on which the meme appears, alongside the name of the account from which the meme was taken.

Table 1. The 11 Instagram pages from which the 409 memes comprising the Correntino Spanish meme corpus were sourced, along with their Instagram handles, the number of followers as of 27 October 2022, and the city/town out of which the page is run.

Page Name	Page Handle	Followers	Location
Memes Bella Vissta	@memesbellavissta	*defunct*	Bella Vista
El Guriok	@elguriok	10.9 K	Corrientes City
El Miami Correntino	@elmiamicorrentino	56.4 K	Corrientes City
Memelandia Correntino	@memelandiacorrentino	89.2 K	Corrientes City
Típico de Correntino	@tipicodecorrentino	30 K	Corrientes City
Libres Yorkk	@libresyorkk	106 K	Paso de los Libres
Un Correntino Dice	@uncorrentinodice	1149	*unspecified*
Tavetayape	@tavetayape	1288	Yapeyú
Angaú Memes	@angaumemes.ok	1012	Paso de la Patria
Meme Amela	@meme_amela	2444	Corrientes City
Las Correntinas Dicen	@lascorrentinasdicen	1797	Corrientes City

Figure 2. Map of Corrientes indicating the location of the cities/towns out of which the Instagram pages in Table 1 are run.[9]
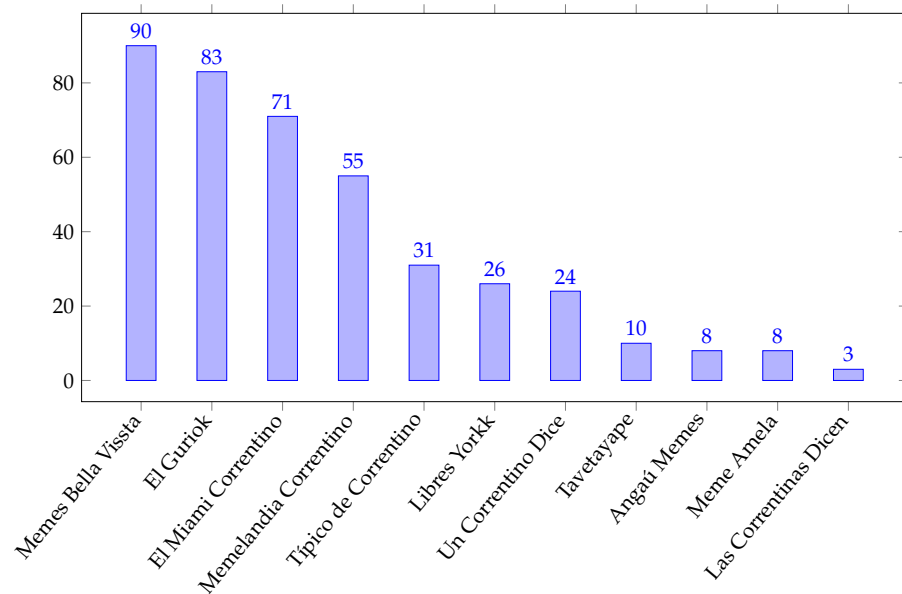
Figure 3. Number of memes taken from each individual Instagram source page (total: 409).

It is worth noting that the corpus here is not exhaustive in the sense of representing each and every meme across all 11 Instagram pages that contains a Guarani loan. Given that memes were collected manually, fully exhausting all 11 pages was beyond the scope of this study (for example, Memelandia Correntino alone contains well over 14,000 memes). However, many of the pages with smaller meme counts were fully exhausted, and in the cases of pages such as Memelandia Correntino, memes were not cherry-picked, as each

collecting "session" consisted of searching an uninterrupted stretch of memes from a given page. This strategy mirrors the common approach in spoken speech analysis in which an uninterrupted stretch of speech is examined in order to extract all occurrences of a given variable. In such a stretch, all memes that contained a Guarani loan (other than those Guarani loans that have entered the general Argentine Spanish lexicon, as discussed above) were collected. Generally speaking, most memes from these pages do not contain Guarani loans, and memes without loans were ignored.[10]

Various memes contained more than one loan. Given this, the corpus contains 409 memes but 498 loans. Within these 498 loans, a basic division can be made between grammatical loans (i.e., borrowed morphemes, such as epistemic particles or evidentials, whose function is morphosyntactic) and lexical loans (i.e., borrowed nouns and adjectives). Of the 498 loans, 268 (53.8%) are grammatical and 230 (46.2%) are lexical.

The lexical loans from Guarani in Correntino Spanish are numerous, and the corpus reflects this, with a wide variety of loans being represented from various semantic fields. The grammatical loans, however, are more restricted, with 15 total being registered in the corpus. The number of occurrences of each is given in Figure 4, followed by a description of the Guarani source form and the meaning of the resulting Spanish loan in Table 2.
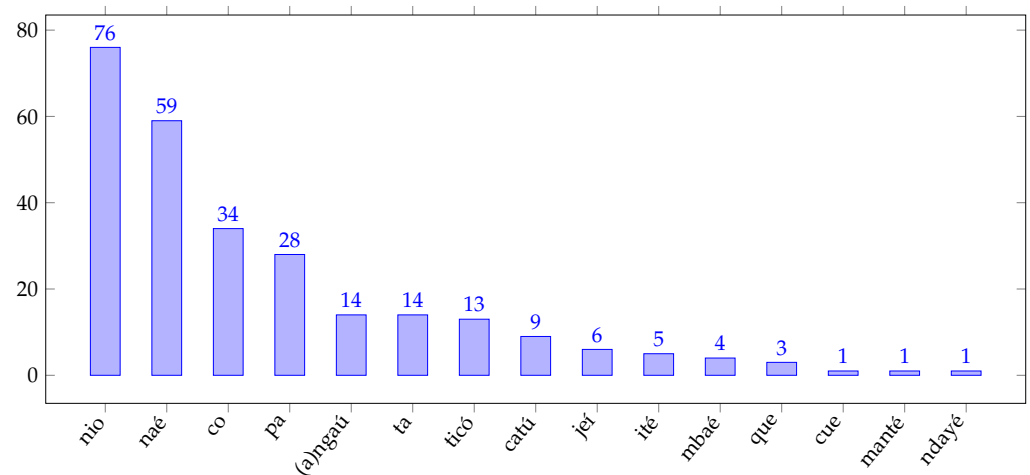

Figure 4. The number of occurrences of each of the 15 grammatical loans registered in the corpus. The Guarani source forms for each can be found in Table 2.

With regard to the meaning of each grammatical loan, it is important to mention that the precise semantics of these loans is highly nuanced and contextually specific. No studies have provided an in-depth description of their semantics or pragmatics, and careful work is needed to provide a satisfactory account of how they are used synchronically and how they have evolved diachronically. Most of them strongly resemble the subset of discourse particles often called "modal particles" in other languages. These particles, best known from Germanic languages (in particular German, e.g., *ja*, *doch*, *halt*, etc.), are uninflected words usually confined to conversational or informal contexts that make subtle pragmatic contributions to a given utterance.[11] Karagjosova (2004, p. 20) notes,

> The meanings of [modal particles] are abstract and difficult to generalise over their particular contexts of occurrence and are usually explicated in terms of the propositional attitudes of their speaker. [Modal particles] can be omitted without influencing the truth value of their carrier sentence. They are typically used in dialogues and as such are primarily oriented towards the addressee and his preceding utterance.

Waltereit's (2001, p. 1392) description of modal particles addresses another commonly reported characteristic, that of their elusive meaning:

[Modal particles] usually cannot carry stress, they cannot be coordinated, they cannot by themselves form a sentence, and their scope ranges over the entire sentence. Besides being identifiable on formal grounds alone, they are also clearly felt to be in some way semantically and pragmatically homogeneous (an intuition which is already reflected in the term *modal* particle or the more traditional German term *Abtönungspartikel*), although their semantics and pragmatics seem to be particularly elusive and difficult to grasp.

These morphosyntactic descriptions and the observation that the meaning of such forms is "particularly elusive and difficult to grasp" apply to most of these particles in Correntino Spanish as well. The source forms for many of such loans are epistemic particles in Guarani, and in many cases the resulting loans have semantically drifted from the original Guarani form[12] and have come to acquire multifunctionality (Wiltschko et al. 2018), i.e., differing pragmatic interpretations depending on the discourse context. For example, the interpretation of *angaú*, which can mean "fake" or "supposedly" (see Table 2), depends on the pragmatic context, e.g., *Es de buena marca angaú* could be variably translated as "It's supposed to be a good brand (and may be)" or "It's supposedly a good brand (but absolutely isn't)".

Table 2. All 15 grammatical loans found in the corpus, by order of frequency (see Figure 4), along with the Guarani source form and the general meaning of the resulting Spanish form.

Spanish Form	Guarani Source Form	English Gloss
nio	*ningo*	intensifier morpheme
naé	*ndaje*	intensifier morpheme
co	*ko*	intensifier morpheme
pa	*pa*	interrogative morpheme
(a)ngaú	*nga'u*	'supposedly', 'fake'
ta	*ta*	interrogative morpheme
ticó	*tiko*	interrogative morpheme
catú	*katu*	intensifier morpheme
ité	*ite*	intensifier morpheme
jeí	*he'i*	'supposedly'
mbaé	*mba'e*	'maybe', intensifier morpheme
que	*ke*	intensifier morpheme (imperative)
cue	*kue*	'former', 'ex-'
manté	*mante*	'only'
ndayé	*ndaje*	intensifier morpheme

The glosses of the resulting Spanish forms in Table 2 are given in the most general terms possible and are based on my familiarity with these forms in spoken Correntino Spanish in conjunction with descriptions provided to me by native speakers. There are notably various particles listed as "intensifier morphemes" (most of which come from Guarani epistemic particles) and "interrogative morphemes", but this should not be interpreted as though the forms within these classes are interchangeable, as the syntactic properties of the loans are not identical, and they are no doubt differentiated semantically in subtle ways which have not yet been described.[13] Interestingly, far more is known about the semantics of the source forms in Argentine Guarani (see Cerno 2013) than the resulting Correntino Spanish loanwords. The descriptions given here in Table 2 are a loose guide to the interpretation of these particles and not intended to be a satisfactory account of their semantics.

With regard to the presentation of memes here, each meme is accompanied by a line-for-line English translation found immediately below the meme. The translations here, all my own, prioritize functional equivalence over formal equivalence and accordingly aim to render the content in an idiomatic way. The nature of many of the grammatical loans makes their translation into English a challenging task, as there is rarely a one-to-one lexical correspondence between a given form and an English equivalent. As such, their rendering in English can vary from context to context. Subcaptions below each individual meme

indicate the Guarani loans found within it; grammatical loans can be cross-referenced with Table 2.

Methodologically, there is no distinction made here between the text in the image of the meme and the caption immediately below the meme, and thus loans in the meme itself were counted in addition to loans in the meme caption. This is due to the fact that meme captions are often as important as the content of the meme itself, representing an integral, sometimes crucial source of the humor of the overall entity. (Some memes, in fact, are uninterpretable when isolated from their caption). Meme captions are common sites for the meme creator to write using their own voice, adding commentary or a response to the content of the meme, often in the first person. Two examples of memes in which the loan is found within the caption are found in Figure 5.

-That's just the way it is
-Your lies—Me
[Caption]: A true idiot

(**a**) Use of *ité*

-It's cold outside, put on a sweater
-A sweeeeeater
[Caption]: Really cold

(**b**) Use of *naé*

Figure 5. Examples of memes in which the loan is found in the image caption as opposed to the meme itself.[14]

3.2. Theoretical Approach

My theoretical approach to the analysis of the meme corpus is rooted in grounded theory (Glaser and Strauss 1967; Strauss and Corbin 1990). I center my approach on the discussion of grounded-theory text analyses given by Bernard (2006, p. 492–97), which focuses on how to thematically synthesize a given text, grounding the resulting thematic generalizations in careful evaluation of the data. Memes by their very nature lend themselves well to such an approach, given their relative thematic simplicity compared to much lengthier texts such as interview transcriptions. The overall goal here is to identify the commonly recurring themes in the memes in which Argentine Guarani loanwords appear in order to reveal relationships between the use of these loans and particular social themes, in turn offering insight into the social value of these highly visible contact features.

To investigate this connection between Guarani loans and social meaning, each meme in the corpus was coded according to whether or not the content was in some way Correntino-specific. Ten criteria were used to justify counting a meme as "Correntino-specific". Memes were evaluated on whether they contained the following:

1. Overt reference to Corrientes, for any reason at all;
2. Reference to food or drink typical of the province (e.g., *mate*, *torta parrilla*, *torta frita*, *guiso*, *asado*, etc.);
3. Reference to flora, fauna, or ecological characteristics typical of the province (e.g., capybaras, caimans, yellow anacondas, the wetlands, etc.);
4. Reference to local cultural belief systems or entities (e.g., Correntino *payé*,[15] the *pombero*,[16] etc.);
5. Reference to local cultural customs involving song or dance (e.g., *chamamé*);
6. Reference to rural life in the countryside (e.g., agriculture, horses, the summer heat, drought, mosquitoes, etc.);
7. Reference to the provincial rivalry between Corrientes and Chaco;
8. Reference to the perceived status of Corrientes as less developed than other Argentine provinces;
9. Reference to Correntino politics;
10. Reference to the local stereotype of Correntino machismo, generally conceived[17] (e.g., that Correntino men have various romantic partners behind the back of their wife or girlfriend, that Correntino men are known for making advances on the female partners of their own close male friends, etc. Memes concerning romantic relationships in a general sense did not meet this criterion).

All memes in the corpus that did not meet these thematic criteria were counted as general, i.e., not Correntino-specific. Thematic patterns were common among this category of memes as well, with many of them having content that referred to schoolwork, Argentine culture broadly conceived (commonly about Argentine national politics), the COVID-19 pandemic, and general current events. Two examples of memes that fall into this general category are given in Figure 6.

-What did you do in quarantine my friend?
-Me: Was I supposed to do something?

(**a**) Use of *ticó*

-Hey are you taking any exams?
-Yeah yeah but in February
-We're already in February

(**b**) Use of *nio*
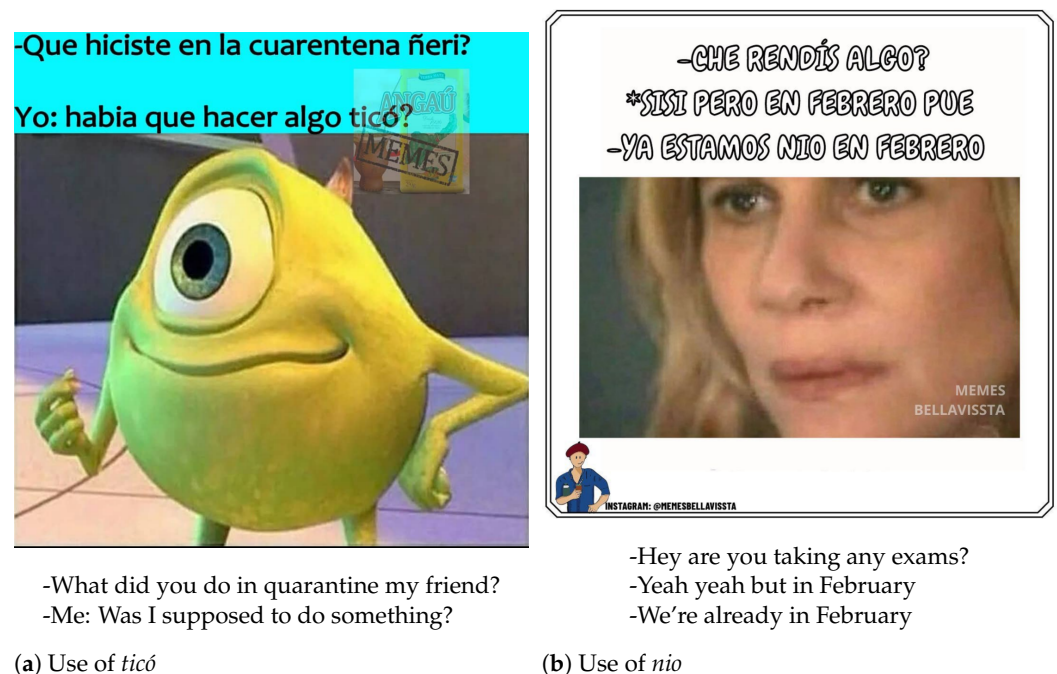
Figure 6. Memes whose content is not Correntino-specific but still contain Guarani loans.[18]

4. Analysis

The memes that constitute the corpus discussed here are thematically often concerned with local issues of Correntino culture and current events. These memes serve as a niche for local content, as larger, more popular meme pages exist that focus on Argentina or Latin America more broadly. The role of these Correntino pages is therefore to replicate

the patterns employed by meme creators in general, but in such a way as to provide localized content which is specifically aimed at Correntinos. Accordingly, and crucially, these memes consistently employ Correntino Spanish. There is no Guarani codeswitching in these memes, as their intended audience is unquestionably largely monolingual. All loans found in these memes, be they grammatical or lexical, are perfectly comprehensible by monolingual speakers of Correntino Spanish (but would not be comprehensible by speakers in, for example, Buenos Aires). The most common location for the production of these memes, as previously shown in Table 1, is Corrientes City, which is a virtually entirely monolingual space. As seen in Figure 2, even the memes that come from outside the provincial capital all come from population centers on the geographic periphery of the province, i.e., regions where Guarani has limited presence (Cerno 2013; Pinta 2022).

In Corrientes as elsewhere, internet culture is seen as aligned with modernity, technology, and urban lifestyles, and the content of the memes directly reflects this. Argentine Guarani is often portrayed as being in conflict with such things, being largely associated with tradition, agriculture, and rural lifestyles (Pinta 2022), a depiction of indigenous languages which is found throughout Latin America (Canessa 2012; French 2000). While there is a growing kind of cultural nostalgia in urban areas such as Corrientes City for things associated with the *campo* ('countryside') as being emblematic of the province in general, and while such topics do surface in the memes produced there, the use of Argentine Guarani in an overt way, even in a codeswitching context, is simply not an available linguistic resource for the younger, monolingual speakers of Correntino Spanish who are producing and consuming these memes.[19]

Given their locally focused nature, the memes make frequent use, occasionally in metalinguistic ways, of typically Correntino phonological, morphosyntactic, and lexical features. Prior to getting to the heart of the analysis here—lexical features, i.e., Guarani loans—it is worth looking at a phonological feature typical of Correntino Spanish that appears in Correntino memes: the use of the locally emblematic assibilated rhotic. While the majority of the features which delineate Correntino Spanish as a unique variety are Guarani in origin, a few are shared by other varieties of Latin American Spanish outside situations of Guarani contact, and this rhotic phenomenon is an example of this latter group. (It is, of course, interpreted in Corrientes as a salient phonological marker of local Spanish nonetheless.) Consideration of this feature is valuable as a prelude to the discussion of Guarani loans, as it serves as an excellent example of the ways in which local features are recruited in memes; the assibilated rhotic is found in metalinguistic contexts, affectionate contexts, and contexts of self-deprecating humor, and this thematic variety mirrors similar patterns which will be explored below in memes containing Guarani loans.

Correntino Spanish, like nearly all Spanish dialects, has two rhotic phonemes; in careful speech or formal contexts these are identical to the standard Spanish rhotics: the tap /ɾ/ and the trill /r/. However, in informal contexts in Correntino Spanish, the trill is assibiliated, leading to a realization closer to [ʐ] or [ʒ]. (Similar realizations in other parts of the Spanish-speaking world, notably in the Andes, are often represented as [r̝].) This is represented orthographically by Correntinos, wishing to emphasize local pronunciation in written form, with *sh*. This written form is commonly employed in memes, often in ironic or humorous ways, and examples are seen in Figures 7 and 8.[20]

It is useful here to contrast the meme in Figure 7a with those in Figure 8. In Figure 7a, we see a man at work, obviously busy, responding to a text message from an individual we are to assume is his girlfriend. She asks if he is busy, and he replies that he is never too busy for her. Here we see the assibilated rhotic in the pet name *reina* 'queen,' written as *sheina*; this use is less a direct link to an aspect of Correntino culture and more a feature of the affectionate pet name being used. In speaking to a romantic partner, an informal register is assumed, and the use of the assibilated rhotic is perfectly appropriate in such a context. In the memes in Figure 8, we do not see quotations as in Figure 7a, but rather these memes are meant to directly communicate with the reader. In Figure 8a, we see what look like military boots but are actually "Croc Martens", i.e., boots made to resemble leather Doc Martens

brand boots that are in reality Crocs, the brand known for making inexpensive clogs made from a material similar to plastic or rubber. In Figure 8b, we see an image of someone ringing an electric doorbell with a video camera which is juxtaposed with a barking dog; the source of humor is the notion that "in other houses" doorbells are modern while in "Coshientes" the doorbell is a barking dog.

-Are you busy?
-For you never baby, what's up?

(**a**) "sheina" in place of *reina* ('queen')

When someone tells me that they can't understand what I say with "r"

(**b**) Metalinguistic reference

Figure 7. Examples of assibiliated rhotics being represented orthographically as *sh* in a romantic pet name (**a**) and referenced metalinguistically (**b**) in Correntino Spanish memes.[21]

When you're a police officer in Corrientes

(**a**) "Coshientes" in place of *Corrientes*

The doorbell in other houses / in Corrientes

(**b**) "Coshientes" in place of *Corrientes*

Figure 8. Examples of assibiliated rhotics being represented orthographically as *sh* in contexts of self-deprecating humor in Correntino Spanish memes.[22]

Thus, in Figure 7a, we see the assibilated rhotic as representative of an affectionate or intimate way of speaking, while in Figures 8a,b it is used to humorously portray Corrientes in a self-deprecating way. Although the writing of *Corrientes* as *Coshientes* can also take place in affectionate contexts which evoke local pride, the examples in Figure 8 illustrate the

more common self-deprecating contexts, which humorously address what are perceived as societal or cultural shortcomings of the province. The latter contexts are far more common in the corpus than the former, and the humor in these cases is found in indirect references to the perception that Corrientes are less advanced or less modern than other regions of Argentina. Such humor is, in my experience, extremely common among locals in Corrientes. Jokes made by them about them are common, but crucially such jokes are only considered appropriate if made by locals. A joke about Corrientes made by an Argentine from another province would nearly certainly be seen as insulting in Corrientes. This is bound up with the use of the assibilated rhotic as a phonological feature that is emblematic of the province—only a Correntino could be saying *en Coshientes* in these examples. Use of this feature combined with the larger context of these memes as clearly coming from Correntino meme pages clearly communicates this as self-aware and self-deprecating—and therefore humorous as opposed to insulting.

Returning to the corpus under discussion here, we frequently see the same kind of humorous connections between "Correntinoness" and linguistic features at play but via lexical items in lieu of phonological features. To take the most obvious class of examples, various memes make metalinguistic reference to Guarani loans, and examples of this are provided in Figures 9 and 10.

‘danger’—‘careful’
‘chaque’

Me trying to talk without saying *naé*, *nio*, *nomateígo*, *pue*, *angaú*

(**a**) Use of *chaque*

(**b**) Use of *naé*, *nio*, and *angaú*

Figure 9. Memes making metalinguistic reference to Guarani loans.[23]

In Figure 9a, we see the Guarani loan *chaque*, meaning ‘[be] careful’ or ‘look out,’ being used in a meme which takes imagery from an advertisement for the 2021 film *Godzilla vs. Kong*. The standard Spanish forms *peligro* ‘danger’ and *cuidado* ‘[be] careful’ are seen fighting, overlaid on images of Godzilla and King Kong, only to be chased away by *chaque* with a baseball bat, which is overlaid on an image of a dog (which is itself a reproduction of the popular "Cheems" meme, in turn an iteration of the "Doge" meme). The implication of this meme—that *chaque* is the superior of the three terms, as the others are seen fleeing from it while it chases them away with a baseball bat—is relatively straight-forward.

In Figures 9b and 10, we see variations of a recurring theme in the corpus: that of Correntinos being unable to avoid using Guarani loans, in particular the more frequent grammatical loans. In Figure 9b, we see the cartoon character SpongeBob SquarePants in two stages of increasing physical discomfort. The caption "Me trying to talk without saying *naé*, *nio*, *nomateígo*,[24] *pue*,[25] *angaú*", when linked with this visual, communicates the idea that to attempt to speak without using these words is a source of visceral discomfort.

Similarly, in the two-part meme in Figure 10 we see a cartoon character (also from SpongeBob SquarePants) holding a drink, the contents of which are shown to be a series of emblematically Correntino words and expressions, most of them loans from Guarani. The second part of the meme shows the character finishing the drink in one motion. Notably, there is agency in this meme that is not clear in Figure 9b; these lexical forms are deliberately "ingested" in this case, presented as something the speaker is not resisting but rather desires.

-Me
-Saying *we, oló, así naé, allá, oatatá,
ticó, qué pa, así co nomás, nio*

Figure 10. A two-part meme metalinguistically referencing various Guarani loans, i.e., *naé, oatatá,
ticó, pa, co,* and *nio.*[26]

The memes in Figures 9 and 10 demonstrate Correntinos' awareness of a host of lexical items that characterize Correntino Spanish. While some of the forms of non-Guarani origin are not unique to Correntino Spanish (e.g., *pue, we*[27]), the borrowed forms are in general not found in the Spanish of other Argentine provinces,[28] which allows for their recruitment in metalinguistic ways.

While the metalinguistic use of Guarani loans is telling, the vast majority of the memes that make use of such loans do not do so metalinguistically. Nonetheless, throughout the corpus Guarani loans appear in memes whose content is explicitly or implicitly Correntino in nature. Examples are provided in Figures 11 and 12.

In Figure 11a, we see reference to a common local cultural trope of Corrientes being a place where gossip is common and spreads quickly. The use of *catú* here plays the double role of grammatically underscoring the fact that "everything is known" and authenticating the phrase as Correntino, grounding it in a commonly shared cultural notion. This is further emphasized by the aesthetic nature of the meme, which is made to resemble the aesthetic used by the provincial government in tourism campaigns and public service announcements. The last of the three phrases—*Corrientes somos todos* 'We are all Corrientes'—is the slogan used by the provincial government in such contexts, and the ironic use here adds a further layer of humor coming as it does after the point that everyone knows everything about one another in Corrientes.

In Figure 11b, we see a scene both common and beloved throughout the province, that of *torta parrilla* (a large piece of bread cooked on a grill) and *yerba mate*, being described via a classically Correntino linguistic construction. *De ma*, a reduced form of *de más* (often written *demá*), is a common intensifier in both Correntino Spanish and Argentine Guarani.[29] This used alongside the Guarani-origin intensifier *nio*, together describing *rico* 'delicious,' yields an English translation along the lines of "absolutely delicious". The recruitment of
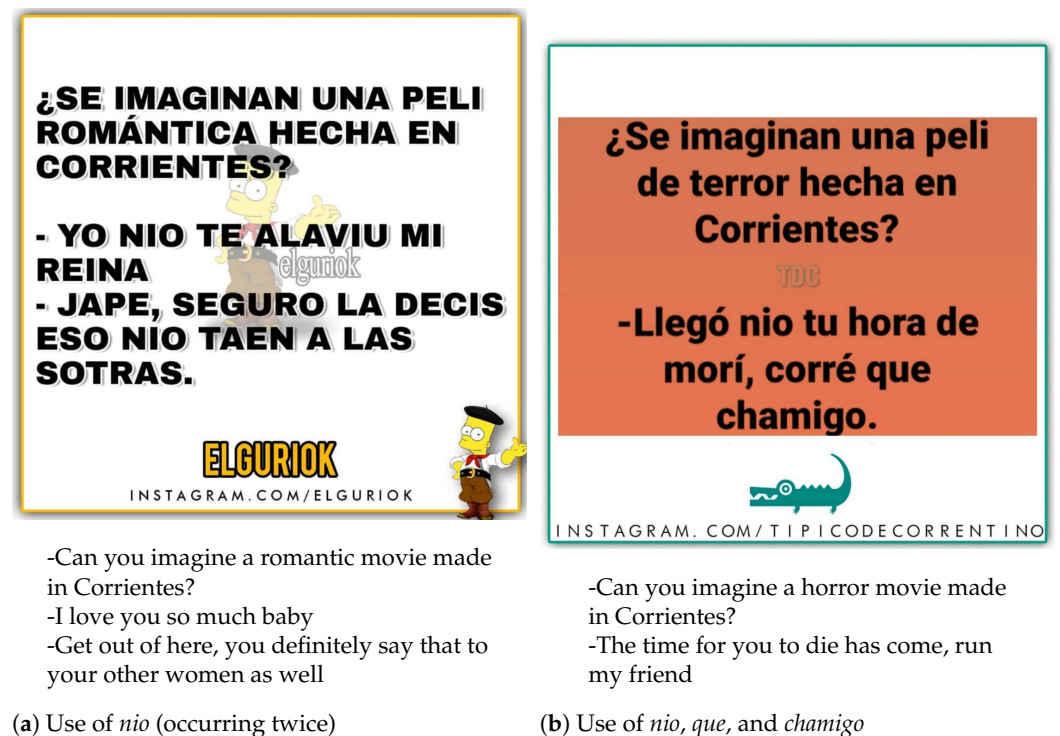
localized linguistic features to describe a localized culinary custom as delicious underscores both as emblematic of the province.

-What happens between two people, stays between two people
-But we're in Corrientes and here absolutely everything is known, and if it isn't known, it's made up
-We are all Corrientes

(**a**) Use of *catú*

Absolutely delicious

(**b**) Use of *nio*

Figure 11. Memes whose content is Correntino-specific and contains Guarani loans.[30]

-Can you imagine a romantic movie made in Corrientes?
-I love you so much baby
-Get out of here, you definitely say that to your other women as well

(**a**) Use of *nio* (occurring twice)

-Can you imagine a horror movie made in Corrientes?
-The time for you to die has come, run my friend

(**b**) Use of *nio*, *que*, and *chamigo*

Figure 12. Memes whose content is Correntino-specific and contains Guarani loans. In these cases, both make reference to what a Correntino-made movie of a particular genre would be like.[31]

The memes in Figure 12 are versions of the common theme in Correntino memes of imagining what an authentically Correntino version of a given media genre would look

like—in this case, what a horror movie and a romantic movie made in Corrientes would be like. Figure 12a is rich in Correntino phonological, morphosyntactic, and lexical features; the Guarani loan *nio* appears twice, and phonological features common to Correntino Spanish are represented orthographically (i.e., reductions such as *taén* from *también* 'also,' *jape* from *rajá pue(s)* 'get out of here'). The "loan" *alaviu* is a Hispanicized rendition of the English phrase *I love you*,[32] clearly used here as symbolic of Hollywood romantic movies. The humor is found in imagining the use of such emblematically Correntino features in a movie, and these features complement the non-linguistic Correntino content of the meme as well—the local stereotype of Correntino men as smooth talkers who have more than one romantic partner. The recruitment of Guarani-origin features for similar purposes is visible in Figure 12b as well. The loans *nio* and *que* (from the Guarani morpheme *ke*, used to strengthen imperatives) complement the deleted coda *r* in *morir*, written as *morí*,[33] and the ubiquitous *chamigo* ('my friend,' a combination of the Guarani possessive *che* 'my' and Spanish *amigo* 'friend').

As evidenced in Figure 12, a variety of linguistic resources can be utilized to ground a meme as authentically Correntino, and the corpus is full of such occurrences. While there exist various linguistic features seen as emblematic of Correntino Spanish that are not Guarani in origin (e.g., *pue*, *we*, *demá*, etc.),[34] they constitute exceptions, if important ones, to the generalization that Guarani-origin features are the defining characteristics of Correntino Spanish as a distinct variety not only within Argentina as a whole but within the Argentine northeast more narrowly. Memes whose purpose is to evoke Correntino ways of speaking, whether metalinguistically or not, inevitably rely heavily on Guarani-origin linguistic material to achieve this end.

Returning to the categorization criteria provided in Section 3.2, 241 of the 409 memes (58.9%) contain Correntino-specific content, while 168 (41.1%) contain general content. Of the 168 memes without reference to Correntino cultural phenomena, many of them still concern everyday aspects of Correntino life but not aspects unique to Corrientes. For example, many of the memes reference the devaluation of the Argentine peso.[35] This devaluation has had a profound impact on the less economically prosperous provinces of Argentina, with Corrientes being one of the most affected. While this economic reality is in many ways a central characteristic of Correntino life at present, memes addressing such a reality are being made throughout the country, meaning such memes in the corpus cannot be considered specific to Corrientes. Despite these and similar memes being classified as general, the corpus still contains more Correntino-specific content than not, a pattern that is indicative of a link between Guarani loans and Correntino culture generally.

5. Argentine Guarani Loans as Enregistered Features

Argentine Guarani loans, representing the core of a constellation of linguistic features which constitute and demarcate Correntino Spanish as a unique Spanish variety, have come to acquire a variety of indexical values, some in conflict and paradoxical, mirroring the cases of Cosmopolitan Mandarin (Zhang 2021) and American English *-ing* (Campbell-Kibler 2007), as discussed in Section 2.1. The simultaneous indexing of positive and negative values associated with Corrientes and Correntinos evidences the complexity and malleability of the social meaning of these forms, which varies via context and can change over time.

As described by Johnstone et al. (2006, p. 83) in their discussion of the enregisterment of particular features of Pittsburgh English, the conditions have formed in which Correntinos can "use regional forms drawn from highly codified lists to perform local identity, often in ironic, semiserious ways", and we can account for this by returning to the previously discussed notion of character type. The varied social meanings of Guarani loans are used to evoke various Correntino character types that serve to amplify these meanings and make them increasingly visible. The stereotype of the smooth-talking, womanizing Correntino man is one such character type. Memes that draw on this stereotype are seen in Figure 13.
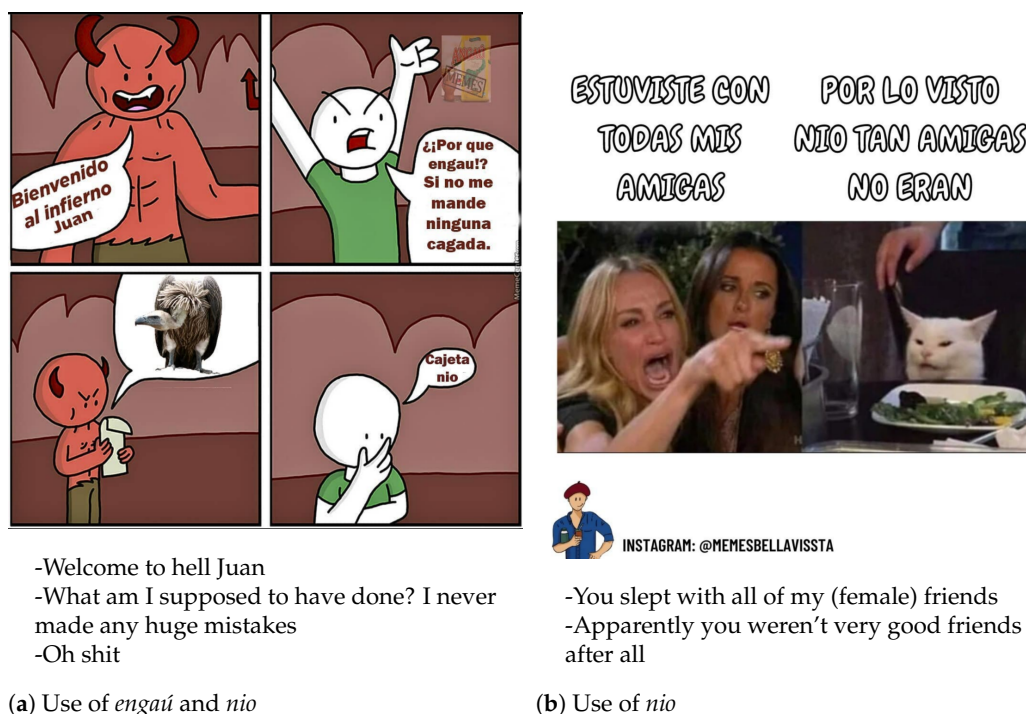
-Welcome to hell Juan
-What am I supposed to have done? I never made any huge mistakes
-Oh shit

(**a**) Use of *engaú* and *nio*

-You slept with all of my (female) friends
-Apparently you weren't very good friends after all

(**b**) Use of *nio*

Figure 13. Memes illustrating the character type of the womanizing Correntino man.[36]

In Figure 13a, we see a man being welcomed to hell by the devil. When the man asks why he is in hell, the devil looks at a picture of a vulture. The vulture here is symbolic in the same way as the *chajá* mentioned earlier; i.e., *buitre* 'vulture' is a common designation for a man who seduces the girlfriend or wife of another man. Upon realizing that is why he is there, the man no longer protests, accepting both his guilt and that that is a legitimate reason for his punishment. In Figure 13b, we see a local version of the common woman-yelling-at-a-cat meme in which a woman accuses someone of having slept with all of her friends. The accused here, a cat who is understood to stand in for her male boyfriend or love interest, responds not by defending himself but rather by saying they must not have been very good friends after all. In choosing to not deny the accusation, the implication of his guilt is clear. These memes both provide a window into a common characteristic of the womanizing Correntino character type, that of a lack of sympathy or sense of guilt, of an almost brazen disregard for the consequences of his actions. The use of the Guarani loans *engaú*[37] and *nio* authenticate this meme as linguistically Correntino and align it with the local stereotype of the *buitre/chajá*.

The womanizer character type directly contrasts with another common character type, that of the morally upright, hardworking Correntino gaucho. In fact, this character type is so common that the watermark of the meme page Memelandia Correntino, i.e., the symbol used on many of the memes to prevent others from passing them off as their own, is a gaucho whose name is given as Don Vicente. Occasionally, serious posts of Don Vicente are made, such as each April 30th, the local "day of the rural worker", which celebrates those Correntinos who work in various agricultural capacities throughout the province. Although the watermark is usually made semi-transparent, in one meme in the corpus, seen in Figure 14, we can see it clearly, adjacent to the main content to which it is thematically unrelated.

The character type of the hardworking Correntino gaucho takes various shapes in Correntino memes, but the uniting traits are honesty, service to the community, and a commitment to hard work. Two memes, different in aim but both drawing on this character type, are given in Figure 15.

-What are you doing here?
-You said online that you're making *mbaypú*.

Figure 14. A meme illustrating the water mark of the meme page Memelandia Correntino. The loan *mbaypú* occurs, from the Guarani source form *mbaipy*, a common local dish similar to polenta that is traditionally made from corn flour.[38]

-Excellent Correntino

(**a**) Use of *nio*

-For 1,000 pesos I'll do whatever you want
-Wonderful, let's go herd cattle from the island, I'm without a farmhand

(**b**) Use of *nio*

Figure 15. Memes illustrating the character type of the hardworking Correntino gaucho.[39]

In Figure 15a we see a man dressed in prototypically rural attire who, wearing a mask out of respect for local anti-COVID measures, is at a local voting center on voting day. The simple description "excellent Correntino" is a relatively rare instance of a "meme" without irony or humor of any kind. It is meant as a self-serious homage to the idealized Correntino gaucho, a classically attired, upright community member doing his civic duty. In Figure 15b we see a different representation of the same character type, involving humor more reminiscent of the other memes in the corpus. A presumed sex worker is speaking to a man through his car window, offering any service for 1000 pesos. He responds happily, saying that he is in need of a farmhand to help him herd cattle. In addition to the obvious irony, this meme is understood to be reflective of the idealized Correntino gaucho, whose loyalty to his work and moral fortitude are his highest priority. In both memes in

Figure 15, we again see the use of Guarani loans (and *de ma* in the case of Figure 15b) as sociolinguistically grounding them in a local way, offering an additional layer of social meaning beyond what would be found in a generalized meme whose target audience was all of Argentina or Latin America generally.

As evidenced by the linguistic and thematic patterns in Correntino memes, Guarani loanwords in Correntino Spanish have come to acquire social meaning that allows for uses that cannot be solely accounted for by their semantics. They have undergone enregisterment, the result of which has afforded them locally specific cultural value. Just as a collection of features of Pittsburgh English, as discussed by Johnstone et al. (2006), underwent enregisterment processes that took them through the *indicator → marker → stereotype* evolutionary pathway, the same processes have resulted in the accrual of social meaning in Guarani loans in Correntino Spanish to the point of allowing their recruitment for use in local stereotypes, as evidenced by the character types visible in Correntino memes.

Correntino Spanish features are simultaneously associated with the romanticized ideal of the Correntino gaucho as well as the Correntino womanizer. The use of Guarani loans in the representation of such character types goes well beyond demarcating Correntino Spanish as a mere geographic variety; Podesva (2011, p. 9) notes that character types crucially "index much more than simply the regions from which such characters originate. They additionally index cultural values". The malleable indexical values of the features used to evoke Correntino character types in turn create differing character types themselves, which are part of the process of production and reproduction of such cultural values. As Eckert (2008, p. 464) notes, "[t]he use of a variable is not simply an invocation of a preexisting indexical value but an indexical claim which may either invoke a pre-existing value or stake a claim to a new value". Such properties, whether at the level of a particular feature, the constellation of features, or the character types such features are linked to, allow for rapid evolution of social meaning, something which memes as a genre are in a perfect position to respond to and even assist in.

The agency of individuals in the process of interpreting and reproducing the social meaning of a set of forms was something commented on by Agha (2003, p. 242) in his pioneering work on enregisterment, in which he notes,

> [I]t is not my purpose to assert that public sphere representations (such as the 'mass media' depictions discussed earlier) determine individual views, or anything of the sort. Contemporary mass media depictions are themselves the products of individuals caught up in larger historical processes; and the 'uptake' of such messages by audiences involve processes of evaluative response that permit many degrees of freedom. I am concerned rather with the ways in which these representations expand the social domain of individuals acquainted with register stereotypes, and allow individuals, once aware of them, to respond to their characterological value in various ways, aligning their own self-images with them in some cases, transforming them in others through their own metasemiotic work.

Memes function as a "public sphere representation" par excellence in that they are easily accessible, easily transformed according to an individual meme creator's intention, and rapidly disseminable in a modified version. Correntino memes, with a host of linguistic features available to ground them as Correntino, are free to be used to index a complex variety of shifting social values and to portray an evolving set of character types. In fact, within the broader world of memes, originality via modification is in essence the defining characteristic of a good meme. The Correntino memes that do the best, i.e., which garner the greatest number of likes and comments, are commonly not memes created out of thin air but are rather locally tailored instantiations of meme templates, trending throughout the internet at large or more narrowly throughout Latin American or Argentine corners of it, which are then cleverly modified to make them Correntino. In this way, memes as a genre encourage a kind of quickened evolution of the form and meaning of the memes themselves, which in turn creates optimal conditions for subtle shifting of the accrued social meaning of the linguistic features that they employ.

The use of Guarani loans in Correntino memes in the service of reproducing stereo-types demonstrates that they are above the level of social awareness, and they are accordingly available for explicit social commentary. An illustrative example is found in Figure 12a, the meme that imagines what a romantic movie made in Corrientes would be like, with the proposed conversation between two characters reproduced here:

-Yo nio te alaviu mi reina 'I love you so much baby'
-Jape, seguro la decís eso nio taén a las sotras 'Get out of here, you definitely say that to your other women as well'

The double use of *nio*—the most common of the grammatical loans in the corpus (see Figure 4)—is a central part of the amalgamation of Correntino features here. These features are part of a complicated network of meanings and stances that allow the meme to work; they are predicated on the reader's understanding of, at minimum: (1) the character type of the Correntino womanizer, (2) the lack of movies in which Correntino Spanish is heard (and by consequence the perception of a serious movie in Correntino Spanish as ironic), (3) the status of English as a common language of romantic films, and (4) the social value held by the linguistic features found in this conversation. This meme simply would not work in standard Spanish—its comedic value is entirely dependent on an irony that is only possible given the social value of the linguistic features used here. Only enregistered features are available for such work. Only features which have come to acquire locally grounded social meaning can be used effectively given that, in the case of memes, other non-linguistic characteristics which might be available to do such work in an embodied interaction (e.g., posture, gesture, clothes, etc.) are unavailable. The task of communicating "Correntinoness" here falls exclusively on linguistic features.

I echo Babel (2011, p. 56) in noting that "the theory of enregisterment provides a useful tool for interpreting the effects of language contact", given that, as in the case of Argentine Guarani loanwords in Correntino Spanish, language contact is an engine of language variation, and language variation and social meaning are inextricable, as third-wave (and, indeed, first- and second-wave) sociolinguistic research has carefully demonstrated. Correntino Spanish, as a subvariety of a macrovariety often labeled "Argentine Spanish", shares a wealth of linguistic features with Spanish speakers of other Argentine provinces such as Buenos Aires, Córdoba, Santiago del Estero, and so on. However, the array of Guarani contact features found in Correntino Spanish are visibly absent in these other closely related Argentine varieties, which allows for their recruitment as local, Correntino emblems. Notably, this can occur whether or not speakers are aware of their status as contact effects in the diachronic sense—that they are recognized as Correntino is all that is necessary for them to become enregistered in this way. The close contact Correntino Spanish has had with Argentine Guarani has endowed it with a particularly rich and diverse set of features which, although not the only enregistered features in Correntino Spanish, make an excellent case study for the targeting of contact-induced variation by the process of enregisterment.

6. Conclusions

This article has aimed to illuminate a previously unanalyzed aspect of the complex situation of language contact in the Argentine province of Corrientes: the social value of Argentine Guarani loanwords in Correntino Spanish. Via analysis of a unique corpus of internet memes, Guarani loans in Correntino Spanish are shown to have undergone enregisterment; i.e., they have become used and understood by speakers in ways that link them to a web of Correntino cultural characteristics and "Correntinoness" generally. Linguistic emblems of "Correntinoness" are simultaneously used to portray Correntino pride and in comedic self-deprecation; they are used to evoke positive and negative Correntino character types. Such differing values are contextually dependent, with the "hearer" (or in the case of memes, the reader) playing an active role in their emergence. These values are linked to larger social space; as Eckert (2008, p. 455) notes, "we construct a social landscape through

the segmentation of the social terrain, and we construct a linguistic landscape through a segmentation of the linguistic practices in that terrain". The Correntino linguistic landscape, being in many ways defined by contact features, provides speakers with a varied set of unique features with which to make connections to larger social phenomena.

Analysis of the enregisterment of such contact features not only allows for deeper understanding of the language contact situation in Corrientes generally, but also supports the notion, advanced by Babel (2011), that the theory of enregisterment provides a nuanced approach to situations of intense language contact—situations in which a complete account of language mixture requires not only an understanding of the linguistic mechanisms involved but also an understanding of the social value of the resulting contact features and the ways in which speakers exploit such value. By drawing on a complicated indexical field, whose social values are rooted in ideologies surrounding complex and paradoxical elements of Correntino identity, speakers deploy Guarani-origin contact features to achieve social ends, something only possible via the process of enregisterment. The ways in which such social ends are achieved are visible in Correntino internet memes, which are grounded in local perceptions, values, and stereotypes. The continued popularity of memes in Corrientes and the use of regionally grounded forms within those memes will no doubt both reflect and foster continued evolution of the social value of Guarani loans in Correntino Spanish.

While the use of internet memes as a primary object of analysis is not without limitations—internet memes, after all, are not naturalistic speech—the kinds of linguistic resources which meme creators employ in the act of creation/modification must necessarily be authentic for a meme to have the desired effect, particularly as it pertains to locally specific humor or irony. This article thus provides a further example of memes as windows into the kinds of sociolinguistic phenomena that can further our understanding of how linguistic features—in this case, contact features—become meaningful social objects.

Notes

[1] The most common designation for this variety of Guarani is the Spanish term *guaraní correntino*. In English, I opt for use of the term "Argentine Guarani" (as opposed to "Correntino Guarani" or "Correntinean Guarani") for reasons outlined in Pinta 2022, pp. 6–10.

[2] Throughout this article I will use the unspecified term "Guarani" as synonymous and interchangeable with the more precise "Argentine Guarani", the variety that is in contact with Correntino Spanish. Argentine Guarani is treated by most scholars as a mere geographical extension of Paraguayan Guarani into Argentina—and therefore "Paraguayan Guarani" is commonly described as being spoken in Corrientes—despite the historical and linguistic differences between them that hinder mutual intelligibility and the societal differences that have produced vastly different sociolinguistic realities for speakers of the two varieties (Pinta 2022). It is, of course, perfectly acceptable to talk about the influence of "Guarani" on "Spanish" in the context of northern Argentina and Paraguay as a general region, abstracting over dialectal differences; however, the source of the Guarani loans in the Spanish spoken in Corrientes is, in some cases, demonstrably Argentine Guarani and not Paraguayan Guarani. Accordingly, all references to "Guarani" here are to be read as implying "Argentine Guarani", and in cases where ambiguity is possible or precision is necessary then appropriately precise terms will be used.

[3] For further discussion of the relationship between these two approaches, see Eckert 2008, pp. 463–564.

4   Some scholars prefer a more fine-grained distinction between these and related categories, e.g., see Moore and Podesva 2009.

5   For a detailed treatment of the history of the term and concept, see Wiggins 2019, pp. 1–20.

6   Source: Memelandia Correntino, https://www.instagram.com/p/CZnjEPuFQar (accessed on 14 June 2023).

7   In fact, some of the memes in the corpus informing this article might fall outside a narrow definition of what a meme is, given that images are often seen as necessary components of memes in the prototypical sense, and some memes in the corpus are text unaccompanied by imagery.

8   Handles provide the web address of each individual page; by replacing the "X" in `https://www.instagram.com/X` with a given handle (minus the @) the address can be recovered, e.g., https://www.instagram.com/elmiamicorrentino (accessed on 14 June 2023). All pages are live as of 14 June 2023 with the exception of Memes Bella Vissta, which has been removed from Instagram for unknown reasons.

9   The map in Figure 2 was taken from Wikimedia Commons (https://commons.wikimedia.org/wiki/File:San_Miguel_(Provinci a_de_Corrientes_-_Argentina).svg, accessed on 14 June 2023). It is reproduced here under the terms of the Creative Commons Attribution-Share Alike 3.0 Unported license and has been altered for the purposes of this article.

10 Exceptions to this include some memes which either directly reference Correntino culture in meaningful ways or memes which make metalinguistic reference to Correntino Spanish phonological features. These additional memes do not figure into the 409-meme corpus referenced here (unless they also happen to contain a loan), but some of them will be referenced below.

11 For a list of the properties of modal particles that justify them as a specific class, see Abraham 1991, pp. 4–5.

12 See Schwenter and Waltereit 2010 for discussion of this kind of historical change in additive particles.

13 See Marco and Arguedas 2021 for discussion of intensification and the complexity of defining it.

14 Source (a): Típico de Correntino, https://www.instagram.com/p/CF2GiLEnrIi (accessed on 14 June 2023).
    Source (b): Angaú Memes, https://www.instagram.com/p/CCqWrU9HVokma6pAzyLpNmS7hb-5zrOTrzdC8c0 (accessed on 14 June 2023; this source URL only works for Instagram users who are followers of this page).

15 Itself a loan whose source is the Guarani form *paje*, it is described as local charm or magic, in both positive and negative contexts.

16 A mythological figure of Guarani origin who is said to live in rural regions. Alternatively described in positive and negative lights, e.g., as a protector of local wildlife or as a troublemaker, belief in the *pombero* is common among the rural inhabitants of the province, who both fear and respect him. During fieldwork in the rural interior of the province in 2017 and 2018, various locals recounted to me stories of having seen the *pombero*, and one individual very earnestly warned me of him.

17 While it could be argued that this kind of generalized machismo is not locally specific to Corrientes (and is perceived as common throughout Argentina, or Latin America in general), this category deserves inclusion here due to a variety of Correntino-specific lexical items which are used to describe such phenomena in a locally specific way. Many of these terms are Guarani in origin, e.g., *chajá*, from Guarani *chahã*, the name for the Southern screamer (Chauna torquata), a common bird species native to the province, which in turn came to be used for a man who seduces the wife or girlfriend of another man.

18 Source (a): Angaú Memes, https://www.instagram.com/p/CCXaY8FDEXmt07l36LPhwN3k25rnBdgoMYhqNE0 (accessed on 14 June 2023; this source URL only works for Instagram users who are followers of this page). The source page of the meme in (b), Memes Bella Vista, has been removed from Instagram, and accordingly the source URL is no longer available.

19 To clarify further, "the use of Argentine Guarani in an overt way" refers to proficiency of some kind in the language itself and does not include the use of Guarani loanwords in Spanish. Just as the use of *raccoon* or *persimmon* by an English speaker speaking North American English does not imply knowledge of Powhatan, the use of the Guarani loans discussed in this article by a Spanish speaker speaking Correntino Spanish does not imply knowledge of Guarani.

20 It is worth mentioning that while these memes come from Instagram pages used as sources for the corpus (those from Figure 7 are from Memes Bella Vista, and those from Figure 8 are from El Miami Correntino), they are not part of the 409 memes that constitute the corpus given that they do not contain Guarani loans.

21 The source page of both memes in Figure 7, Memes Bella Vista, has been removed from Instagram, and accordingly the source URLs are no longer available.

22 Source (a): El Miami Correntino, https://www.instagram.com/p/CJiy0HwnuKS (accessed on 14 June 2023).
    Source (b): El Miami Correntino, https://www.instagram.com/p/CIbPSpAHKo2 (accessed on 14 June 2023).

23 Source (a): El Guriok, https://www.instagram.com/p/CLaxD1KH8J_ (accessed on 14 June 2023).
    Source (b): El Guriok, https://www.instagram.com/p/CMsgltSnxAa (accessed on 14 June 2023).

24 *Nomás te digo* 'I'm just saying' but written as it is frequently pronounced, including the deletion of various consonants, i.e., [no.ma.t̪e.ˈi.ɣo] instead of [no.mas.t̪e.ˈði.ɣo] in careful speech.

25 The Correntino realization of the highly frequent form *pues*, a discourse marker often found phrase-finally in Correntino Spanish.

26 Source: El Miami Correntino, https://www.instagram.com/p/CLrpCZ_HnVe (accessed on 14 June 2023).

27 A common phrase-initial discourse marker.

28 It should be noted that some of these loans occur in Paraguayan Spanish as well. However, others—e.g., the interrogative marker *ta* (Cerno 2013, p. 227)—are specifically Argentine Guarani in origin (i.e., their source form is nonexistent in Paraguayan Guarani)

and therefore are unique to Correntino Spanish. The fact that some Guarani loans are shared between Correntino Spanish and Paraguayan Spanish should not be interpreted to mean that these two varieties of Spanish are largely the same. Various features, in particular phonological but also morphosyntactic and lexical, demarcate Paraguayan Spanish and Correntino Spanish as being clearly and immediately recognizably distinct Spanish varieties (for discussion along these lines specifically regarding contact features, see Estigarribia et al. 2023).

29 The original etymological source of this form, while unclear, may be the Portuguese *demais* 'much, too much,' which patterns in many ways like the forms in Argentine Guarani and Correntino Spanish (Leonardo Cerno, personal communication, 8 December 2022).

30 Source (a): Memelandia Correntino, https://www.instagram.com/p/CavIyqKl3oM (accessed on 14 June 2023).
Source (b): Un Correntino Dice, https://www.instagram.com/p/CM703fTswy0 (accessed on 14 June 2023).

31 Source (a): El Guriok, https://www.instagram.com/p/CKO-ZLqH1Do (accessed on 14 June 2023).
Source (b): Típico de Correntino, https://www.instagram.com/p/CJCM6SXHZWg (accessed on 14 June 2023).

32 It is interestingly treated as just a verb here as opposed to a verb phrase, replacing the Spanish verb form but not the Spanish pronouns *yo* 'I' or *te* 'you'.

33 Coda deletion, a common feature of Correntino Spanish, is likely also a contact effect (or at the very least enforced by contact effects), given that Guarani phonology natively prohibits syllable codas (Cerno 2013). Many Spanish loans in Guarani whose source forms contained codas were repaired via coda deletion during the process of loanword adaptation (Pinta and Smith 2017), leading to forms that phonologically pattern with *morí* in this case.

34 In my experience, many Correntino Spanish speakers, whether monolingual or bilingual, do not seem to be aware of the source language of the discourse markers in this variety, presumably due to the fact that Spanish-origin forms have been borrowed into Argentine Guarani and vice versa, thus making the source of a given form unclear. I have heard speakers claim that some Guarani-origin particles come from Spanish, e.g., *nio*, and that some particles native to Spanish are in fact borrowed from Guarani, e.g., *pue(s)*. This is perhaps also attributable to the relatively similar phonologies of the two languages (see (Pinta and Smith 2017) and the references cited therein).

35 The exchange rate between the Argentine peso and the US dollar on my first visit to Argentina, in May of 2007, was roughly 3:1; the current official exchange rate as of October of 2022, 15 years later, is approximately 156:1.

36 Source (a): Angaú Memes, https://www.instagram.com/p/CB5_mucjC2S3b1egCbq1JG1BMzrv6xNXylAsp00 (accessed on 14 June 2023; this source URL only works for Instagram users who are followers of this page). The source page of the meme in (b), Memes Bella Vista, has been removed from Instagram, and accordingly the source URL is no longer available.

37 This is a common variant of the previously given form *angaú*; they are semantically identical. As indicated in Table 2, the source form (for both) is the Guarani form *nga'u*, and the Spanish forms *angaú/engaú* differ in the epenthetic vowel used to resolve the /ᵑg/ onset, which does not natively occur in Spanish.

38 Source: Memelandia Correntino, https://www.instagram.com/p/CMqPzByFn4a (accessed on 14 June 2023).

39 Source (a): Tavetayape, https://www.instagram.com/p/CTupPgOrw36 (accessed on 14 June 2023).
Source (b): Memelandia Correntino, https://www.instagram.com/p/CLxfNSLlm1hy_KTULPshgMnsvLQq90dx6x9vog0 (accessed on 14 June 2023).

References

Abraham, Werner. 1991. Introduction. In *Discourse Particles*. Edited by Werner Abraham. Amsterdam: John Benjamins, pp. 1–10.

Agha, Asif. 2003. The social life of cultural value. *Language and Communication* 23: 231–73. [CrossRef]

Agha, Asif. 2007. *Language and Social Relations*. New York: Cambridge University Press.

Aslan, Erhan, and Camilla Vásquez. 2018. 'Cash me ousside': A citizen sociolinguistic analysis of online metalinguistic commentary. *Journal of Sociolinguistics* 22: 406–31. [CrossRef]

Babel, Anna M. 2011. Why Don't All Contact Features Act Alike? Contact Features as Enregistered Features. *Journal of Language Contact* 4: 56–91. [CrossRef]

Bernard, H. Russell. 2006. *Research Methods in Anthropology: Qualitative and Quantitative Approaches*, 4th ed. New York: Rowman & Littlefield Publishers.

Campbell-Kibler, Kathryn. 2007. Accent, (ING) and the social logic of listener perceptions. *American Speech* 82: 32–64. [CrossRef]

Canessa, Andrew. 2012. *Intimate Indigeneities: Race, Sex, and History in the Small Spaces of Andean Life*. Durham: Duke University Press.

Cerno, Leonardo. 2013. *El Guaraní Correntino: Fonología, Gramática, Textos*. Frankfurt am Main: Peter Lang.

Cerno, Leonardo. 2017. Aspects of Dialectal Diversification of Guarani in Paraguay and Corrientes: Contact between Two Given Languages in Different Settings. In *Guarani Linguistics in the 21st Century*. Edited by Bruno Estigarribia and Justin Pinta. Leiden: Brill Publishers, pp. 348–78.

Cochrane, Leslie, Alexandra Johnson, Aubrey Lay, and Ginny Helmandollar. 2022. "One does not simply categorize a meme": A dual classification system for visual-textual internet memes. *Proceedings of the Linguistic Society of America 7*: 5260. [CrossRef]

Davison, Patrick. 2012. The Language of Internet Memes. In *The Social Media Reader*. Edited by Michael Mandiberg. New York: New York University Press, pp. 120–34.

Dawkins, Richard. 1976. *The Selfish Gene*. Oxford: Oxford University Press.

Dietrich, Wolf. 2002. Guaraní criollo y guaraní étnico en Paraguay, Argentina y Brasil. In *Current Studies on South American Languages*. Mily Crevels, Simon van de Kerke, Sergio Meira, and Hein van der Voort. Leiden: Leiden University, pp. 31–41.

Drager, Katie, Kate Hardeman-Guthrie, Rachel Schutz, and Ivan Chik. 2021. Perceptions of Style: A Focus on Fundamental Frequency and Perceived Social Characteristics. In *Social Meaning and Linguistic Variation: Theorizing the Third Wave*. Edited by Lauren Hall-Lew, Emma Moore, and Robert J. Podesva. Cambridge: Cambridge University Press, pp. 176–201.

Eckert, Penelope. 2008. Variation and the indexical field. *Journal of Sociolinguistics* 12: 453–76. [CrossRef]

Eckert, Penelope. 2012. Three Waves of Variation Study: The Emergence of Meaning in the Study of Sociolinguistic Variation. *Annual Review of Anthropology* 41: 87–100. [CrossRef]

Eckert, Penelope. 2016. Variation, meaning and social change. In *Sociolinguistics: Theoretical Debates*. Edited by Nikolas Coupland. Cambridge: Cambridge University Press, pp. 68–85.

Estigarribia, Bruno, Justin Pinta, and Ernesto López Almada. 2023. Guarani influence on Spanish in contact situations: A comparison between Paraguayan and Correntino Spanish. In *Mutual Influence in Situations of Spanish Language Contact in the Americas*. Edited by Mark Waltermire and Kathryn Bove. Abingdon: Routledge, pp. 134–57.

French, Brigittine M. 2000. The Symbolic Capital of Social Identities: The Genre of Bargaining in an Urban Guatemalan Market. *Journal of Linguistic Anthropology* 10: 155–89. [CrossRef]

Glaser, Barney G., and Anselm L. Strauss. 1967. *The Discovery of Grounded Theory: Strategies for Qualitative Research*. New York: Aldine.

Hall-Lew, Lauren, Emma Moore, and Robert J. Podesva. 2021. Social Meaning and Linguistic Variation: Theoretical Foundations. In *Social Meaning and Linguistic Variation: Theorizing the Third Wave*. Edited by Lauren Hall-Lew, Emma Moore and Robert J. Podesva. Cambridge: Cambridge University Press, pp. 1–23.

Hill, Jane H. 2007. Mock Spanish: A Site for the Indexical Reproduction of Racism in American English. In *Race, Ethnicity, and Gender: Selected Readings*. Edited by Joseph F. Healey and Eileen O'Brien. Thousand Oaks: Sage Publications, Inc., pp. 270–84.

Johnstone, Barbara, Jennifer Andrus, and Andrew E. Danielson. 2006. Mobility, Indexicality, and the Enregisterment of "Pittsburghese". *Journal of English Linguistics* 34: 77–104. [CrossRef]

Jones, Taylor. 2016. Tweets as graffiti: What the reconstruction of Vulgar Latin can tell us about Black Twitter. In *English in Computer-Mediated Communication: Variation, Representation, and Change*. Lauren Squires. Berlin: Mouton De Gruyter, pp. 43–67.

Karagjosova, Elena. 2004. *The Meaning and Function of German Modal Particles*. Saarbrücken: Saarland University dissertation.

Knobel, Michele, and Colin Lankshear. 2007. Online Memes, Affinities, and Cultural Production. In *A New Literacies Sampler*. Edited by Michele Knobel and Colin Lankshear. New York: Peter Lang, pp. 199–227.

Labov, William. 1972. On the Mechanism of Linguistic Change. In *Sociolinguistic Patterns*. Philadelphia: University of Pennsylvania Press, pp. 160–82.

Marco, Marta Albelda, and María Estellés Arguedas. 2021. De nuevo sobre la intensificación pragmática: Revisión y propuesta. *Estudios Románicos* 30: 15–37.

Moore, Emma, and Robert J. Podesva. 2009. Style, indexicality, and the social meaning of tag questions. *Language in Society* 38: 447–85. [CrossRef]

Ndoci, Rexhina. Forthcoming. The Linguistic Construction of Migrant Ethnic Identity: Names, Name-Calling, and Internet Memes at the Service of (de)Racialization. Doctoral Dissertation. The Ohio State University, Columbus, OH, USA.

Pinta, Justin. 2022. *Minority Language Resilience in Corrientes, Argentina: Argentine Guarani and Spanish in Contact*. Columbus: The Ohio State University.

Pinta, Justin, and Jennifer L. Smith. 2017. Spanish Loans and Evidence for Stratification in the Guarani Lexicon. In *Guarani Linguistics in the 21st Century*. Edited by Bruno Estigarribia and Justin Pinta. Leiden: Brill Publishers, pp. 285–314.

Podesva, Robert J. 2011. The California Vowel Shift and Gay Identity. *American Speech* 86: 32–51. [CrossRef]

Procházka, Ondřej. 2019. Chronotopic representations as an effect of individuation: The case of the European migrant crisis. *Language in Society* 49: 717–43. [CrossRef]

Rymes, Betsy, and Andrea R. Leone. 2014. Citizen sociolinguistics: A new media methodology for understanding language and social life. *Working Papers in Educational Linguistics* 29: 25–43.

Schwenter, Scott A., and Richard Waltereit. 2010. Presupposition accommodation and language change. In *Subjectification, Intersubjectification and Grammaticalization*. Edited by Kristin Davidse, Lieven Vandelanotte and Hubert Cuyckens. Berlin: De Gruyter Mouton, pp. 75–102.

Shifman, Limor. 2011. An anatomy of a YouTube meme. *New Media & Society* 14: 187–203.

Silverstein, Michael. 2003. Indexical order and the dialectics of sociolinguistic life. *Language and Communication* 23: 193–229. [CrossRef]

Squires, Lauren, ed. 2016. *English in Computer-Mediated Communication: Variation, Representation, and Change*. Berlin: Mouton De Gruyter.

Starr, Rebecca Lurie. 2021. Changing Language, Changing Character Types. In *Social Meaning and Linguistic Variation: Theorizing the Third Wave*. Edited by Lauren Hall-Lew, Emma Moore, and Robert J. Podesva. Cambridge: Cambridge University Press, pp. 315–37.

Strauss, Anselm, and Juliet Corbin. 1990. *Basics of Qualitative Research: Grounded Theory Procedures and Techniques*. Thousand Oaks: Sage Publications, Inc.

Varis, Piia, and Jan Blommaert. 2015. Conviviality and collectives on social media: Virality, memes, and new social structures. *Multilingual Margins* 2: 31–45. [CrossRef]

Waltereit, Richard. 2001. Modal particles and their functional equivalents: A speech-act-theoretic approach. *Journal of Pragmatics* 33: 1391–417. [CrossRef]

Weinreich, Uriel, William Labov, and Marvin Herzog. 1968. Empirical foundations for a theory of language change. In *Directions for Historical Linguistics: A Symposium*. Edited by Winfred P. Lehmann and Yakov Malkiel. Austin: University of Texas Press, pp. 95–188.

Wiggins, Bradley E. 2019. *The Discursive Power of Memes in Digital Culture: Ideology, Semiotics, and Intertextuality*. New York: Routledge.

Wiltschko, Martina, Derek Denis, and Alexandra D'Arcy. 2018. Deconstructing variation in pragmatic function: A transdisciplinary case study. *Language in Society* 47: 569–99. [CrossRef]

Zhang, Qing. 2021. Emergence of social meaning in sociolinguistic change. In *Social Meaning and Linguistic Variation: Theorizing the Third Wave*. Edited by Lauren Hall-Lew, Emma Moore, and Robert J. Podesva. Cambridge: Cambridge University Press, pp. 267–91.

What Does It Meme? English–Spanish Codeswitching and Enregisterment in Virtual Social Space

Kendra V. Dickinson

Department of Spanish and Portuguese, Rutgers University, New Brunswick, NJ 08901, USA;
kendra.dickinson@rutgers.edu

Abstract: This project investigates English–Spanish codeswitching in internet memes posted to the Facebook page, We are mitú (mitú), and analyzes how lexical insertions and quotatives contribute to the enregisterment of linguistic patterns and the construction of collective identity among U.S. Latinx millennials in virtual social spaces. Data include instances of lexical insertion (n = 280) and quotative mixed codes (n = 114) drawn from a collected corpus of 765 image–text memes. The most frequent lexical insertions included food items (e.g., *elote* and *pozole*), kinship terms (e.g., *abuelita* and *tía*), and culturally specific artifacts or practices (e.g., *quinceañera* and *lotería*), which reflect biculturalism and rely on a shared set of references for the construction of a group identity. Additionally, the quotatives in the data construct Spanish-speaking characterological figures that enregister a particular brand of U.S. Latinx millennial identity that includes being bilingual, having Spanish-speaking parents, and having strong ties to Latinx culture. Overall, this work highlights not only internet memes as a vehicle for enregisterment, but also, and more importantly, how the language resources employed within them work to enregister linguistic and cultural norms of U.S. Latinx millennials, and thereby, play a role in identity construction in virtual social spaces.

Keywords: enregisterment; identity; memes; Spanish

1. Introduction

The combined use of English and Spanish on social media in the United States represents the confluence of two phenomena: (1) the coexistence of Spanish and English in the U.S. (Ortman and Shin 2011; U.S. Census Bureau 2015) and (2) the widespread use of social media for virtual community interaction. This project is situated at the intersection of these two phenomena, investigating the use of English–Spanish mixed code in image–text internet memes (henceforth referred to simply as "memes"). The goal of this paper is to analyze the ways in which the language resources employed in these memes can work to enregister linguistic and cultural norms of U.S. Latinxs, and thus, play a role in identity construction.

This project focuses on the Facebook page, We Are mitú (mitú),[1] a U.S.-based Latinx[2] online media organization, and analyzes how codeswitching practices within the memes posted by the page both illustrate Spanish–English contact in the U.S. and work to construct U.S. Latinx millennial identity.[3] This site was chosen for several reasons. First, it represents a large repository of bilingual meme content in a single location. Second, mitú describes itself as "100% American and 100% Latino", inspired by its community "to create authentic, culturally relevant stories", and it frequently posts videos, articles, and humorous content in English and Spanish. This express goal of engaging with U.S. Latinx experiences makes it an interesting site for analyzing the ways in which language can be used to construct identities. Third, as of December 2019, the group's Facebook page had over four million followers. While it is not possible to ascertain the number of individuals who regularly consume the media posted to the page, the sheer number of followers reflects its popularity and potential for reaching a large audience. Overall, this page offers a unique opportunity

for the investigation of the enregisterment of linguistic forms in memes targeted to U.S. Latinx individuals.

The target audience of the page is listed as internet users between the ages of 18 and 44. In a 2015 interview, the mitú cofounder, Beatriz Acevedo, noted that one of the goals of the organization was to be the "voice of the millennial generation".[4,5] In this vein, much of the page's content contains overt acknowledgement of the age of its intended audience. For example, the post in Figure 1 claims that "we", referring to both the poster and the intended audience, are old enough to be spending time looking for shopping deals (*las ofertas*). Furthermore, the post includes the hashtag *#señora*, a term used in Spanish to refer to either a married or an older woman. Additionally, many cultural references within these posts relate to things that were popular among adolescents in the early 2000s, when millennials were coming of age. For example, Figures 2 and 3 refer to the Mexican television show, *Rebelde*, which aired during this time of millennial adolescence.

Figure 1. Generational reference on mitú.

Figure 2. *Rebelde* reference #1 on mitú.

Figure 3. *Rebelde* reference #2 on mitú.

In this paper, I use both quantitative and qualitative analyses to illustrate how the communicative functions of the lexical and quotative codeswitching patterns employed in

these memes are not only a reflection of U.S. Latinx millennials' day-to-day experiences, but also work to construct and enregister an ethnolinguistic identity in a dynamic internet community. More broadly, this work highlights the role of the internet as a sociolinguistic space, the recruitment of image–text memes as a vehicle for the transmission of meaning, and how the linguistic resources employed within these memes can enregister linguistic practices.

2. Background

The present work draws on theoretical insights from a number of related fields, including previous research on codeswitching, computer-mediated communication, social theory, and third-wave sociolinguistics. In what follows, I synthesize some of the important findings from these fields that inform the analysis of linguistic practices in the memes under consideration.

Memes, in the broadest sense, are units of cultural transmission that are replicated and reproduced in human communities through imitation, as defined by evolutionary biologist and social theorist, Richard Dawkins (Dawkins 1976, p. 205). According to Blackmore (1999, p. 7), "Everything that is passed from person to person in this way is a meme", including skills, ideas, and elements of language. Memes and their transmission have greatly increased with the rise of the internet as means a communication (Blackmore 1999), as the internet provides an ideal platform for the exchange of ideas across space and time.

In present-day popular culture, the term "meme" refers to a cultural or community idea that is replicated and mutated on internet-based platforms. According to Lankshear and Knobel (2007, p. 202), the term "meme" is used among internet users in "describing the rapid uptake and spread of a particular idea presented as a written text, image, language 'move,' or some other unit of cultural 'stuff.'" While it is still an emerging area of analysis, various authors have endeavored to identify the shared characteristics of internet memes. Zappavigna (2012) describes internet memes as a display of interpersonal, semiotic bonding used by in-group members to signal solidarity. Memes continually evoke a shared set of references that enhance a particular group of users' sense of unity (Baym 1995; Miltner 2014) by tapping "into shared popular culture experiences and practices" (Lankshear and Knobel 2007, p. 207). According to Zappavigna, memes are a kind of "inside joke", where the shared ideas of a particular community, which are not necessarily humorous in isolation, are combined in novel and surprising ways. Furthermore, Miltner (2014) discusses how particular referential knowledge is required to "get the joke" of a particular meme, creating a sort of "communal wall" between those who understand the joke and those who do not, making memes part of an interconnected, self-referential network of meaning bound to the context of their creation and dissemination.

The very nature of internet memes facilitates processes of enregisterment, or the process by which linguistic forms become ideologically associated with social identities (Agha 2003). In later work, Agha (2005, p. 38) elaborates on enregisterment as "processes whereby distinct forms of speech come to be socially recognized (or enregistered) as indexical of speaker attributes by a population of language users", such as "Japanese women's language" (Inoue 2003) or "internet language" (Squires 2010). In Section 4, I apply this framework to show how the repetitive nature of internet memes can facilitate enregisterment in mitú's content.

Internet memes can also represent performances that reinforce existing links between linguistic features and particular identities. For example, Shifman (2013) analyzes the popular 2007 video "Leave Britney Alone", in which an internet user tearfully and forcefully defends pop icon, Britney Spears. The author describes how the resulting memes emphasized the communicative strategies and codes used by the video's creator, such as high emotionality and yelling. Additionally, in their analysis of the "It Gets Better" internet movement supporting LGBTQ youths, Gal et al. (2016, p. 1710) conceptualize the memes in the campaign as "performative acts", designed either to persuade an audience or to construct collective identity and community norms.

The representations and performances that link language forms and social identities are often referred to as "characterological figures", or stereotypical personae that can be linked to speech (Agha 2006, p. 177). For example, in Johnstone's (2017) analysis of talking plush dolls meant to represent people from Pittsburgh, she finds that the speech attributed to the dolls both presupposes and helps to create the characterological figure of a "Yinzer", i.e., a working-class person from Pittsburgh. In other words, the linguistic characteristics and appearance of the dolls work to re-enregister an already enregistered variety by reinforcing the link between communicative style and the local working class. The fact that such characterological figures can be crafted by companies and media organizations and can implicitly promote the processes of enregisterment via their broadcasting to a particular audience are important for the current study.

The relationship between form and meaning can be different depending on who is listening and who the players are in a given interaction. For example, Johnstone (2011) analyzes enregisterment processes through a character in a radio DJ skit whose speech patterns are used to perform several identities, including a mother, a working-class person, and a Yinzer. The author notes that Pittsburgh-specific linguistic features are more likely to be salient to those listeners familiar with the local vernacular, while other features invoke more broad-reaching cultural schemas of motherhood to a more general audience. In this way, enregisterment does not create fixed associations between language forms and social meaning, but rather it is a process of "constituting possibilities" for the creation of varied sets of meaning for both speakers and hearers in a given interaction (Goebel 2007, p. 523).

The potential for characterological figures to perform multiple identities and to project diverse sets of meanings depending on the context and the listener is particularly relevant to the enregisterment of language on the internet, where messages are frequently repeated and altered and can be observed by a vast number of people. In his study of Jamaican Creole and English use on the internet, Hinrichs (2006, p. 110) describes switches between the two languages as explicit performances of "a typical Jamaican in conversation". Through this performance of "Jamaicanness", in-group members recruit stylistic, sociopragmatic resources as part of identity construction to position themselves in relation to the content of the discourse. Regarding Spanish in particular, in her study of the use of Mexican Spanish by transnational second-generation Mexican bilinguals on Facebook, Christiansen (2015a) found that individuals use multiple linguistic strategies to index their identity and counter-identity as *ranchero* (a social group associated with rural life in Mexico). The author interprets group members' alternation between embracing and distancing themselves from *ranchero* culture as a performance that is "reminiscent of their collective past but fitted to contemporary U.S. Mexican culture" (Christiansen 2015a, p. 699). In another study of the same online *ranchero* community, Christiansen (2015b) found that individuals use their perceived degree of "centrality" to *ranchero* culture, social networks, and the Spanish language to position themselves as more or less Mexican in an online environment.

Similarly, the memes on the mitú page often use characterological figures to construct and reinforce a particular kind of U.S. Latinx millennial identity, which may or may not reflect the real lives of the page's followers. For some followers, the memes can reinforce and re-enregister their own realities. For those with divergent experiences, the meme content can work to enregister the linguistic and cultural patterns they display as part of a generalized U.S. Latinx millennial experience. In both cases, the content of the memes can work to construct a particular U.S. Latinx millennial identity, relying on the perception of shared community experiences and understanding. Previous work on identity building in mitú's content has shown that this perception of shared experiences is a critical element of the organization's content. Gutiérrez (2021, p. 80) describes the cultural and language practices in mitú content videos on the Pero Like channel as "emblematic of a generation, elucidating how Latinx millennials come to mediate panethnicity while creating new forms of Latinx media representation". Additionally, the author notes how the use of characters such as the *abuela* ('grandmother') by the content creator, Jenny Lorenzo, based on her own grandmother, are meant to exemplify a sort of universal Latina immigrant

grandmother figure (p. 90) and represent the Latinx millennial experience as well as generational differences.

Given that the present study focuses on how codeswitching is deployed within the memes posted to mitú, it is critical to establish what is meant by the term. There has been a debate in linguistics about codeswitching typologies and what exactly constitutes true codeswitching. Nevertheless, it is generally agreed that the terms *intersentential* and *intrasentential codeswitching* refer to complete switches from one language or variety to another between and within sentence boundaries, respectively (Poplack 1980; MacSwan 2014; inter alia). A more contested type involves the use of a lexical item from one language within a different matrix language. The literature has used a variety of terminologies to describe this, including *insertion* and *borrowing* (Poplack 1980; MacSwan 2014; Muysken 2000; Woolford 1983; inter alia), though these phenomena have often been differentiated from other types of language mixing and have traditionally been classified as falling outside of the umbrella of codeswitching (Poplack 1993, pp. 255–56, 279). However, Muysken (2000, p. 69) suggests that, while there may be formal, structural differences between borrowing and codeswitching, they can fulfill similar discursive functions, such as the symbolic role of marking multicultural identity. Backus (2015) also follows this assertion, stating that, while borrowings and codeswitching may differ diachronically, their synchronic function and sociopragmatic indexicality are often the same. Therefore, *codeswitching* is used in the present work as a pre-theoretical term that encompasses all of the aforementioned types of language mixing, including all instances of Spanish and English being used together in the same discursive context. Following Androutsopoulos (2015), multilingualism can be better examined when considering "any discourse that draws on resources associated with more than one language". Additionally, so-called hybrid linguistic forms (e.g., *stressitos*, an English word with a Spanish diminutive suffix) can be included in this umbrella, going beyond a simple mixture of first and second languages (Wei 2018), and can play a role in identity construction (Wei 2011; Wei and Lin 2019). Importantly, this approach to codeswitching allows more robust analysis in that it permits us to observe all possible cases of multilingual discourse rather than being unnecessarily restrictive.

The social and linguistic processes that motivate codeswitching are critical to the present analysis. Previous research has explored this from several perspectives, including functional motivators, language dominance and shifts, lexical and structural mutual influences, and discursive sociopragmatic functions. For example, Muysken (1997, p. 364) suggests that language dominance is reflected in the direction of lexical insertion; first-generation immigrants tend to insert words from the host country into their native language, while subsequent generations tend to insert words from their family language into the language of the host country. Relatedly, Bentahila and Davies (1992) found that speakers of later generations in a host country tend to switch smaller constituents than earlier generations do. These patterns of smaller, often single-word codeswitches in Spanish are frequent in the memes posted to mitú, reflecting the sociodemographic reality for 65% of Latinx millennials, a critical target audience for the page, who are U.S.-born (Patten 2016).

Additionally, previous studies have shown that words that are switched from one language into another tend to be highly specific in meaning or may have differing connotations between the two languages (Backus 2001; Myers-Scotton and Jake 1995), such that biculturalism is often the motivation for lexical switches (Montes-Alcalá 2007). De Fina (2007) shows an example of this in her analysis of the maintenance of Italian food terms in an immigrant community as a means of the construction of an ethnic identity. Similarly, Gutiérrez (2017) discusses the use of food symbolism in selfies taken by Latinx individuals, which employ culturally specific dishes such as tacos and maize as identity representations. A similar trend will be shown in the data from mitú, whereby the content of lexical switches reflects patterns of biculturalism and culturally specific meanings.

3. Materials and Methods

To determine how codeswitching works to enregister linguistic and cultural norms of U.S. Latinx millennials in virtual social space, data were collected from the Facebook page, We are mitú. On Facebook, pages are places where businesses, brands, musicians, organizations, and other groups can connect with fans or customers by posting content that can appear on individual followers' timelines. Facebook users who follow the mitú page can receive regular updates about the content posted by the media organization.[6] Image–text memes were identified from among this content as photos that consisted of some combination of an image and saying and/or phrase, excluding advertisements (Figure 4), simple photos (Figure 5), or purely informational posts (Figure 6). This is not to say that these images do not portray cultural content and symbolism, but they do not fit the definition of a meme as adopted in this project.

Figure 4. Advertisement on mitú.

Figure 5. Photo post on mitú.
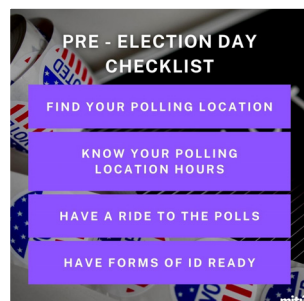
Figure 6. Informational post on mitú.

All the memes that were posted by the page over the two-month period from 1 September to 31 October 2019 were collected.[7] This was achieved by visiting the page each day and downloading all the memes that had been published since the previous day's collection. Though the two-month period is limiting, I began this project with the goal of obtaining a large enough sample of memes for analysis. Having collected over 700 memes as the end of October neared, I elected to finalize and analyze the corpus I had gathered.

The only items that were excluded were those that were reposted from previous dates already included in the corpus to avoid repetition. All the collected memes were given a unique ID tag and were transcribed into plain text. Then, each meme was examined to see if it contained mixed code. Those that did not include multiple languages were classified as monolingual and were coded as either monolingual English or monolingual Spanish.

After determining that a given meme included mixed code, the switch(es) within it were classified as intersentential, intrasentential, lexical, or quotative/attributive. The present work focuses only on the latter two types because they account for the majority of the data and most clearly reflect the processes of enregisterment discussed in Section 4. Fixed phrases, defined as instantiations of language use that are prefabricated, formulaic, and highly conventional (Kecskes et al. 2018), were excluded from the present analysis. These included proper names, such as the soda brand name, *Jarritos*, song lyrics, and idioms.

As defined in the present study, the term *lexical switch* refers to the insertion of elements from one language into an identifiable matrix base in another language, as exemplified in Figure 7. For lexical switches, the specific word being switched and its syntactic category were also recorded. The latter designation included the following categories: *noun*, *adjective*, *verb*, *pronoun*, or *other*, which included cases such as onomatopoeia and bound morphemes.


Use of the word *dulces* ('sweets') in an English sentence.

Figure 7. Example of lexical insertion of a Spanish word in an English matrix.

Additionally, the terms *quotative switch* and *attributive switch* both refer to instances where words are credited to a specific speaker, with the former referring to situations where the language appears as a direct quote in quotation marks (Figure 8) and the latter referring to cases where the language is attributed to another speaker (Figure 9), for example, in a subordinate clause. Both types exhibited similar patterns in the present data set, and therefore, were collapsed into a single *quotative/attributive* classification. After determining the type of switch(es) included in each meme, the direction of the switch was categorized as either English-to-Spanish or Spanish-to-English.

For quotative/attributive switches, the identity of the imagined speaker, or the person that the words were being attributed to, was recorded, e.g., *aunts* (Figure 10), *mom* (Figure 11), and *me* and *friend* (Figure 12). All the speaker classifications were determined from context and recorded as they appeared in the meme, e.g., "your mom". Then, speakers were grouped into larger categories such as "mother".

In summary, all the memes were classified according to the switch type, lexical item, and imagined speaker following the procedures described in this section. The following section presents the results based on this coding and shows how these factors play a role in the construction of Latinx identity and the enregisterment of linguistic norms in the online community of mitú.
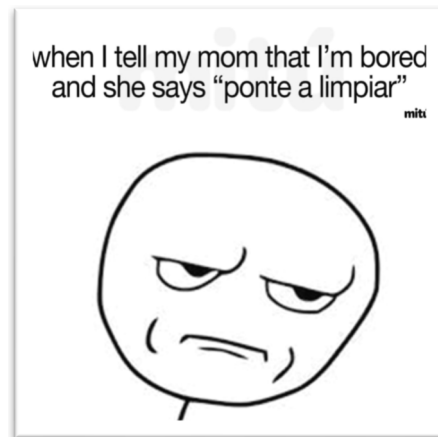
Direct quote: "get to cleaning".

Figure 8. Example of a Spanish quote in an English matrix attributed to "mother".

Attribution of the idea that the "dog behaves
better" than the child.

Figure 9. Example of a Spanish speech attributed to "mother".

Phrase: " Good morning and God bless you
all, attributed to aunts on Facebook.

Figure 10. Meme with speaker "aunts".

Quote from "mother" saying "You pay rent
or get to work" to the family dog.

Figure 11. Meme with speaker "mother".

"Cuando Way" street sign as play on words of
"¿Cuándo wey/güey?", which means "When, dude?" in
English.

Figure 12. Meme with imaged speaker "self" and "friend".

4. Results and Discussion[8]

4.1. Overview

I begin by presenting a description of the larger corpus that the memes under analysis are drawn from in order to highlight general trends and to justify my focus on a subset of the data. The corpus is made up of 765 image–text memes. Of these, 35.2% (n = 268) are monolingual, of which 84.0% (n = 225) are in English and 16.0% (n = 43) are in Spanish. Additionally, 3.1% (n = 24) of the memes contained Spanish-origin words that have been lexicalized in English, and therefore, could not be definitively attributed to either language (e.g., tequila, taco, and piñata), and 7.1% (n = 54) included fixed phrases (e.g., song lyrics and brand names), both of which were excluded from the quantitative analysis. The remaining 54.8% (n = 419) of the memes included a total of 526 instantiations of English–Spanish mixed code, including some cases in which there were multiple codeswitches per meme. Of these, 6.8% (n = 36) are switches from Spanish to English, 92.8% (n = 488) are switches from English to Spanish, and 0.4% (n = 2) are hybrid English–Spanish words, specifically *stress-itos* (English *stress* plus a Spanish diminutive suffix) and *chikibaby* (Spanish *chiki*, from *chiquito* 'small,' plus English *baby*). What follows is an analysis of a subset of the total 526 instances of codeswitching, including Spanish lexical insertions and Spanish

quotative/attributive switches, which make up 51.3% (n = 270) and 21.7% (n = 114) of the instances of mixed code in the corpus, respectively.

The primacy of mixed code in the corpus reflects the critical roles of both English and Spanish in the enregisterment of U.S. Latinx linguistic patterns. In particular, the collected memes reflect Krogstad and Gonzalez-Barrera's (2015) findings that 70% of participants of approximately millennial age reported using "Spanglish". These previous findings, in combination with the frequent use of English and Spanish within a single meme in the present data, illustrate the processes of the enregisterment of these linguistic patterns in virtual social spaces. The following sections will detail the ways in which the patterns of language use in these memes enregister mixed code, cultural artifacts, and Spanish-speaking parents as part of the U.S. Latinx millennial experience, offering the possibility for the construction of a collective identity.

4.2. Lexical Switches

A total of 280 lexical switches were recorded. Of these, 96.4% (n = 270) were Spanish words within an English matrix, with just 3.6% (n = 10) representing English words within a Spanish matrix. A total of 136 unique words are represented in the data, as some words are repeated in multiple memes. The vast majority of the lexical items were nouns, as shown in Table 1.

Table 1. Distribution of Spanish insertions by lexical category.

Lexical Category	Count	% of Total
noun	241	89.3%
adjective	13	4.8%
pronoun	9	3.3%
verb	4	1.5%
other	3	1.1%

As previously noted, other research has shown that younger generations tend to switch smaller linguistic elements of the language of their family's country of origin into the majority language (Bentahila and Davies 1992; Muysken 1997), and these types of switches are frequent in U.S. bilingual media (Mahootian 2005). This pattern holds true in the present data. As noted above, 51.3% (n = 270) of all instantiations of codeswitching in the data were insertions of Spanish lexical items into English matrix clauses. In addition, the prevalence of lexical insertions in these data, namely nominal (89.3%) and adjectival (4.8%) insertions, aligns with previous findings on the lexical categories that are most often inserted (Poplack 1980).

Figure 13 provides a visual representation of the distribution of the different words that constituted the lexical switches found in these memes. In this word cloud, the size of the word corresponds to its frequency of occurrence, i.e., words that are represented in larger text occurred more often than words that are represented in smaller text. For example, *chisme* ('gossip') occurred as a lexical insertion fourteen times, whereas *padrinos* ('godparents'), as shown in the lower left corner of the image, occurred only twice. As can be observed in the image, many of the more frequent words represent food items (e.g., *elote*, a grilled corn dish typical of Mexico, and *lomo saltado*, a dish made with meat and French fries typical in Peru), kinship terms (e.g., *abuela* 'grandmother' and *tía* 'aunt'), and culturally specific artifacts, traditions, or practices (e.g., *quinceañera*, a traditional 15th birthday celebration for girls, and *lotería*, a card-based game of chance typical in Mexico).

Notably, of the 270 Spanish lexical insertions, 29.6% (n = 80) are food items (Table 2), 35.2% (n = 95) are family or social terms (Table 3), 10.0% (n = 27) are other cultural items (Table 4), and 25.2% (n = 68) belong to other categories.

Figure 13. Word cloud of lexical switches.

Table 2. Lexical insertions related to food.

Word	Gloss	Count (n = 80)
agua(s) fresca(s)	beverage made of blended fruit, sugar and water	2
aguachiles	dish made from shrimp and vegetables; typical of Mexico	1
cafecito	from *café* ('coffee') with diminutive suffix *-ito*	4
caldo	broth	4
carne (asada)	(grilled) meat	4
churro	dessert made from fried dough	1
dulces	sweets	1
elote	grilled corn dish; typical of Mexico	5
elotero	seller of *elote*	2
empanadas	pastry turnover with savory filling	3
flan	sweet custard dessert	1
frijoles	beans	4
horchata	rice beverage	2
chile	hot pepper	1
limón(es)	lime(s)	2
lomo saltado	dish made with meat and French fries; typical of Peru	1
marisco	shellfish	1
mazapán	candy made from ground nuts, sugar, and honey	1
menudo	soup made with cow's stomach; typical of Mexico	1
nopal	edible cactus	1
paleta	ice pop made with frozen fruit	3
pan (con pavo)	bread (with turkey)	2
pan dulce, concha	sweet bread; typical of Mexico	4
plato	dish, plate	1
pozole	traditional Mexican stew made with hominy and meat	5
pupusa(s)	corn flour flatbread with filling; typical of El Salvador, Honduras	2

Table 2. *Cont.*

Word	Gloss	Count (n = 80)
pupusería	locale that sells *pupusas*	1
raspado	shaved ice	3
sandía	watermelon	2
sopa	soup	3
tamal(es), tamalito	tamale(s), with diminutive suffix *-ito*	4
taqueria	locale that sells tacos	2
taquero(s)	seller of tacos	4
tres leches	sponge cake made with evaporated, condensed, and whole milk	1
trompo al pastor	spit-grilled pork; typical of Mexico	1

Table 3. Lexical insertions related to family and society.

Word	Gloss	Count (n = 95)
abuela(s), abuelita(s)	grandmother(s); with diminutive suffix *-ita(s)*	9
ama	from *ama de casa* ('housewife')	1
amiguito(s)	a romantic partner that is not necessarily serious or exclusive	3
chisme	gossip	14
chimoso/a(s)	gossipy	7
comadre	close female friend, sister	3
compa(s)	from *compañera/o* ('comrade,' 'companion')	2
fulanito	from *fulano* ('so-and-so,' 'some guy') with diminutive suffix *-ito*	2
hermanito	from *hermano* ('brother') with diminutive suffix *-ito*	2
mamá	mom	4
mija(s), mijitas	term of endearment meaning sweetheart or dear	3
padre	father	1
padrinos	godparents	2
primo(s)	cousin(s)	8
señora	married or older woman	11
sobrinos	nieces and nephews	1
suegra	mother-in-law	4
tía(s)	aunt(s)	16
tío	uncle	2

Table 4. Lexical items related to culture.

Word	Gloss	Count (n = 27)
bautismo	baptism	1
brujería	witchcraft	3
cobija (de San Marcos)	a thick blanket decorated with images; from Mexico	5
El Cucuy	a mythical monster	1
la chancla	sandal or flip-flop used to discipline children	1
La Llorona	a mythical ghost of a weeping woman	2
lotería	card-based game of chance; typical of Mexico	2
mal de ojo	evil eye; thought to cause illness or misfortune	1
misa	Catholic mass	1
novelas	from *telenovelas*, Spanish-language soap operas	1
Quince (añera)	traditional 15th birthday celebration for girls	6
rancho	ranch; land where a family lives and raises animals or farms	2
rosario	rosary	1

The trends found in previous work related to cultural symbolism and bilculturalism have also been illustrated by the present data (De Fina 2007; Gutiérrez 2017; Montes-Alcalá

2007) For example, many of the lexical switches involved traditional food items of Latinx cultural groups, especially food traditional to Mexican–American communities, including *agua fresca*, *carne asada*, *elote*, *horchata*, *pozole*, and *tamales*, among others. Not only can these examples be considered from a utilitarian point of view, given that they have no direct translations in English, but their use invokes a shared set of references that are specific to in-group experiences and cultural practices.[9]

Relatedly, many familial and kinship terms appear as Spanish words inserted into English clauses in these data, including *abuela* ('grandmother'), *primo* ('cousin'), *tía* ('aunt'), *tío* ('uncle'), *padre* ('father'), *sobrino* ('nephew'), and *suegra* ('mother-in-law'). These words provide a clear example that supports Backus' (2001) proposal that highly specific words, as defined by the norms of the particular community they are used in, are more likely to be inserted. In some cases, as with the previously described food items, Spanish words are used where part of the meaning would be lost in translation. For example, the literal translation of *amiguito* would be something like 'little friend' in English, but in the memes analyzed here, it is used to talk about a romantic partner that is not necessarily serious or exclusive. This recruitment of familial and kinship terms in Spanish instead of using the corresponding terms in English suggests differences in the sociopragmatic meaning of the English and Spanish words. Therefore, upon closer examination, the words that are selected for lexical switches in these memes are not only highly specific and difficult to translate into English, but also reflect cultural practices that are closely connected to identity and are often meant to evoke meaning beyond the semantic meaning of the word (Kosoff 2014; Mahootian 2005).

4.3. Quotative/Attributive Switches

The corpus contained 153 instantiations of language attributed to a speaker in three general configurations: a Spanish quote within an English matrix (Figure 14), an English quote within an English matrix (Figure 15), and a Spanish quote within a Spanish matrix (Figure 16).

The majority of the quotative/attributive switches in the corpus consisted of Spanish quotes embedded within English matrix clauses (74.5%, n = 114), which are the focus of the present analysis. The remaining 25.5% (n = 39) of the memes based on the quotative template were monolingual, broken down into English quotes within English matrix clauses (19.0%, n = 29), and Spanish quotes within Spanish matrix clauses (6.5%, n = 10). There were no instances of English quotatives embedded in Spanish matrix clauses in the present data. These patterns are visualized in Figure 17.
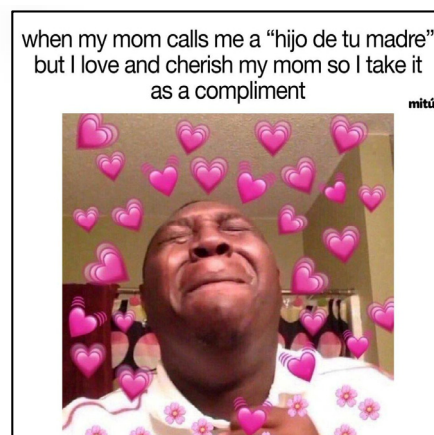
Phrase "son of your mother" attributed to "mother".

Figure 14. Example of a Spanish quote in an English matrix.

Phrase "Who I am to judge" attributed to "self".

Figure 15. Example of an English quote in an English matrix.

Phrases "You too" and "Likewise" without specific attribution.

Figure 16. An example of a Spanish quote in a Spanish matrix.

The corpus overall shows a strong tendency towards the use of quotatives in Spanish within English matrix clauses. This pattern can be considered reflective of the linguistic environments of many U.S. Latinx millennials (i.e., being surrounded by people who speak Spanish) and a desire to maintain quotes in their original language (Gumperz 1982; Halim and Maros 2014). Additionally, this tendency can work to enregister established patterns of language contact, intergenerational language shift, and the use of Spanglish among millennials (Krogstad and Gonzalez-Barrera 2015).

The imagined speakers to whom these English-to-Spanish quotative switches are attributed also follow a distinct pattern, as shown in Figure 18. The vast majority of Spanish quotative switches were attributed to older family members. A total of 42.1% (n = 48) of Spanish quotes were attributed to a "mother" figure, 16.7% (n = 19) were attributed to either "father" or "parents", and 12.3% (n = 14) were attributed to other older family members (e.g., *tía(s)*, *suegra*, or *abuela*) or community members (e.g., *señora*). Together, the quotes attributed to the aforementioned speakers account for over 70% of the quotative data. This reveals a pattern within these memes of imagined Latinx millennials that share the common ground of a close-knit community in which relatives are not only essential figures, but also speak Spanish. Another 14% (n = 16) were attributed to the "self", and the remaining 14.9% (n = 17) were attributed to non-specific individuals, Latinxs in general, other family members (e.g., *primos* 'cousins'), significant others, or friends.
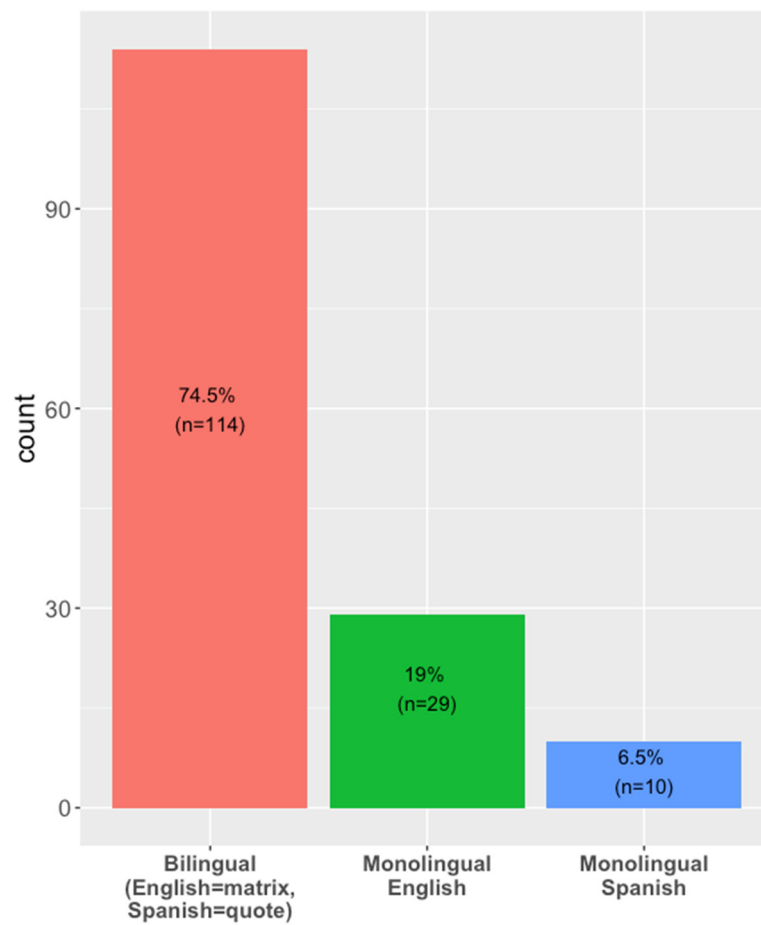
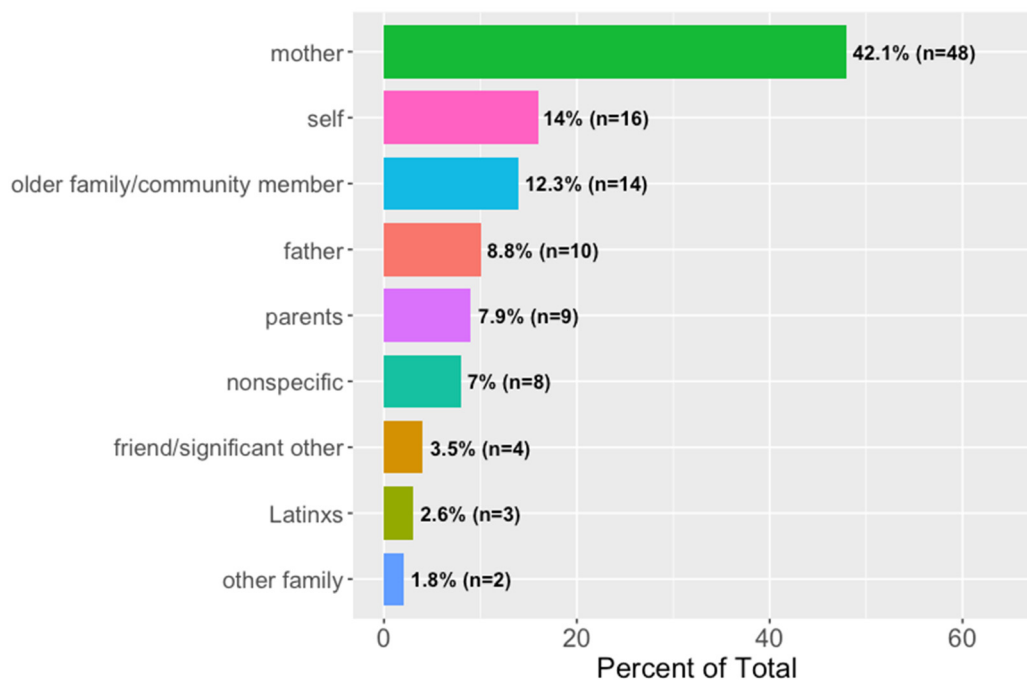**Figure 17.** Distribution of quotative switch types.

Figure 18. Distribution of imagined speakers of Spanish quotatives in English clauses.

Figure 19 contextualizes these findings by showing the rates of quotes attributed to these imagined speakers across the different quotative templates. First, the quotes

attributed to "parents", "father", and other family/community members occur categorically in Spanish. Second, the quotes attributed to the "mother" figure occur in Spanish 90.6% (n = 48) of the time and in English 9.4% (n = 5) of the time. Additionally, the figure of the "self" is depicted as frequently speaking both English (32.2%, n = 10) and Spanish (67.8%, n = 21). This reveals a tendency not only to attribute speech to the aforementioned figures, but also to attribute different language patterns to them. Finally, "other people" (e.g., Uber driver or teacher), "other family member" (e.g., cousins or aunts), "friends/significant others", and "Latinxs", in general, occur at higher rates in English quotatives.
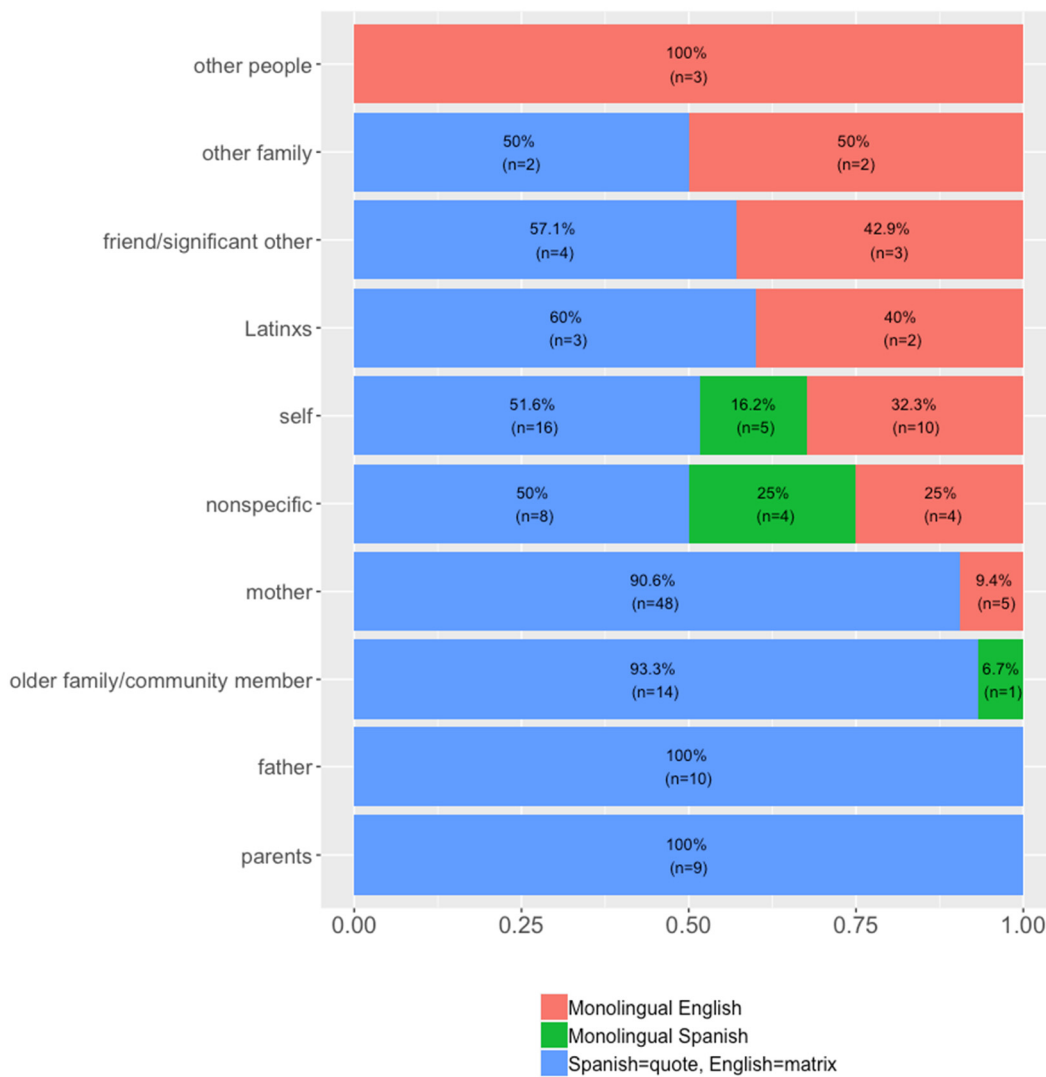

Figure 19. Distribution of quotative templates according to imagined by speaker.

While the instances of quotative mixed code in these memes may be, in part, reflective of the documented process of intergenerational language shifts, their messages also rely on more nuanced sociopragmatic and symbolic meanings. A wealth of previous studies have found that particular linguistic forms, in this case, codeswitching, can be used in internet communication to convey in-group membership and construct group identity (Androutsopoulos 2006; Dorleijn and Nortier 2009; Zappavigna 2012; inter alia). In the case of the present data, the expression of culturally specific lexical items and Spanish quotes attributed to older family members works to establish linguistic and cultural patterns that communicate the sociopragmatic meaning of being in an environment where Spanish is used, constructing a collective imagination of the U.S. Latinx millennial experience.

The communicative capacity of these memes relies on the construction of characterological figures and cultural symbolism to ideologically link the linguistic features employed within them to abstract social categories in a process of enregisterment (Agha 2003). The creation of these figures relies on the shared experiences of U.S. Latinx millennials, as previously described, but also uses the memetic template to create patterns of expression that can become enregistered through repetition and humor. One such characterological figure employed in these memes is the "mother", who is constructed as a figure whose primary language of expression is Spanish. In addition, the content of the memes where language is attributed to the "mother" figure follow specific patterns, in particular, the frequency of activities related to domesticity and raising children. For example, Figure 20 depicts the figure of the "mother" reprimanding a child for misbehavior, and Figure 21 shows the "mother" telling a child to contribute to a group clean up. Figure 22 depicts an example of a "mother" questioning her child's behavior at different life stages; first, at a younger age, the child is told to go and play outside, and second, at an older age, the child is told that they spend too much time out and about.

Phrase "Calm down or I'll calm you down"
attributed to "mother".

Figure 20. Meme depicting a mother reprimanding child.

Phrase "help clean up" attributed to
"mother".

Figure 21. Meme depicting a mother telling child to do chores.

Phrases "Go play outside" and "You are always
in the streets" attributed to "mother".

Figure 22. Meme depicting a mother questioning a child's behavior.

As in Johnstone's (2011) description of the characterological figures in radio skits, the depictions of motherhood displayed in these memes have the potential to enregister language forms within more than one cultural schema, because what is salient depends on the listener or, in this case, the internet audience. For example, for U.S. Latinx millennials who grew up with Spanish-speaking mothers, this pattern can serve to further enregister the relationship between use of the Spanish language and motherhood, not unlike the *abuela* figure described in Gutiérrez (2021). More broadly, even for those audience members for whom there is not a strong personal connection between speaking Spanish and motherhood, the content of the memes themselves may enregister certain practices stereotypically associated with motherhood,[10] such as domestic activities and monitoring children's behavior. In this way, both *what* is depicted in these memes and *how* it is depicted play a role in (re)enregistering the "mother" figure for U.S. Latinx millennials.

The present data show similar patterns, albeit not as frequently, related to the construction of the characterological figure of the "father". As with the "mother" figure, the "father" is characterized as speaking Spanish. Regarding the content of the memes related to the "father", they often depict stereotypically masculine behaviors, such as protectiveness, fixing things, and a hard emotional exterior that can be softened. For example, Figure 23 portrays the figure of the "father" secretly supervising his teen or adult female offspring, who has a male friend in their room, and Figure 24 shows a "father" figure telling their child to replace the oil in their car. Lastly, Figure 25 depicts a scenario in which the "father" figure sternly states that he does not want a dog, but later falls in love with it. As with the characterological figure of the "mother" described above, the repeated depiction of "father" as using Spanish can enregister or re-enregister these language forms, depending on the audience.

Phrase "You should leave the door open"
attributed to "father".

Figure 23. Meme depicting a father supervising child.

Phrase "Put oil in the car" attributed to "father".

Figure 24. Meme depicting a father in relation to chores.

Phrase "I don't want dogs in this house" attributed to "father".

Figure 25. Meme depicting a father's changing attitude towards a family dog.

This frequent depiction of the Spanish-speaking "mother" and "father" works to construct the U.S. Latinx millennial as having Spanish-speaking parents. This is supported by representations such as Figure 26, where both parents are portrayed as speaking Spanish, and the continued reiteration of this pattern in successive memes enregisters the relationship between these figures and the use of the Spanish language. In addition to the use of Spanish, memes such as Figure 26, which depict the imagined speakers as having large extended families, function to further construct strong family ties as part of the U.S. Latinx millennial identity. These memes represent several different behaviors that are attributed specifically to the parents of Latinx millennials, including referring to particular aesthetics (Figure 27), engagement in cultural activities (Figure 28), and sayings (Figure 29).
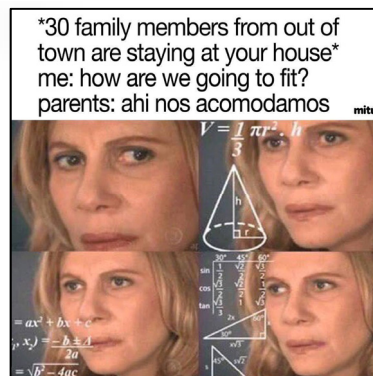
Phrase "we'll make ourselves comfortable" attributed to "parents".

Figure 26. Meme depicting parents using Spanish.

Image of bobblehead dogs used as dashboard
decorations.

Figure 27. Meme depicting parents' aesthetics.
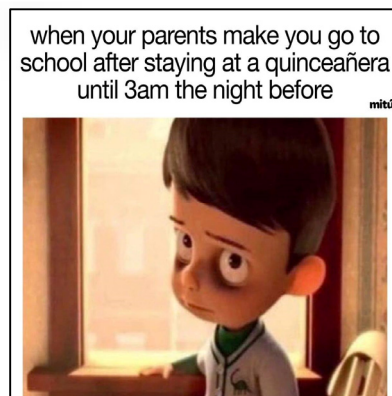
Reference to *quinceañera*, a traditional
15th birthday celebration for girls.

Figure 28. Meme depicting cultural activities.

Phrase "share with your brother" at-
tributed to "parents".

Figure 29. Meme depicting parents' sayings.

Finally, the last characterological figure salient in these data is the "self", which is portrayed as speaking both English and Spanish, constructing the Latinx millennial as someone who is bilingual. For example, the "self" figure is depicted using English in Figure 30,[11] using Spanish in Figure 31, and codeswitching in Figure 32. In fact, the text of the meme in Figure 32 points directly to the idea of seeing oneself in a meme and identifying

with it. These examples demonstrate how the characterological figure of the "self" works to further enregister U.S. Latinx millennial use of both English and Spanish and provides the possibility for followers to personally identify with internet content.

Phrase "I'm going to tell God to take you" attributed to "mother".

Figure 30. Meme depicting English used by the "self" figure reacting to "mother's" scolding in Spanish.

Phrase "I'm not even tired" attributed to "self".

Figure 31. Meme depicting use of Spanish by the "self" figure.

Phrase "It's me" attributed to "self".

Figure 32. Meme depicting a codeswitch by the "self" figure.

The bilingual status of U.S. Latinx millennials is constructed through the use of English and Spanish both separately and together, as well as examples that demonstrate the ability

to adapt one's mode of expression based on social context. Consider Figure 33, where the U.S. Latinx millennial fluidly uses English and Spanish depending on the context. The option on the left side shows the orthographic representation of laughing in English, "haha", while the option on the right side shows the orthographic representation of laughing in Spanish, "jaja".

Figure 33. Spanish and English laughing.

The characterological figures of "mother," "father," "parents," and "self" are constructed within these memes not only individually, but also in contrast to each other. Looking back at Figure 30, we see that the "parent" figures speak in Spanish, and the child, the U.S. Latinx millennial, replies in English. Another example of this is shown in Figure 34, where the "mother" speaks in Spanish, but the "self" speaks in English. Interactions like these construct the U.S. Latinx millennial in opposition to the previous Spanish-dominant generation and subsequently align with Gutiérrez's (2021) analysis that the *abuela* figure on the Pero Like channel illustrates generational differences and Latinx millennials' experiences. This example not only contrasts the difference in language use of the "mother" and the "self" figures, but also highlights generational differences in media consumption. The television program mentioned, *Primer Impacto*, is a weekday evening news program produced by the Spanish-language network Univisión that first aired in the 1990s and is known for sensationalist stories.

Phrase "that is from the devil" attributed to "mother".

Figure 34. Meme depicting use of Spanish by a mother and English by the child.

The way that the memes analyzed in the present project are instantiated is a product of several factors, including the norms of internet communication, the social and linguistic experiences of Latinx millennials engaged with the mitú platform, and the characterological

figures that they construct in the enregisterment process. These memes tend to follow similar structural patterns, such as the insertion of culturally specific Spanish words into English matrices and the use of Spanish quotatives attributed to older family members. These norms appear to represent a sort of template, which is a common characteristic of successful internet memes (Lankshear and Knobel 2007; Zappavigna 2012). Importantly, part of the success of a particular memetic template is its relevance to the experiences and identities of those who engage with it. The content and structure of the memes posted to the mitú page are not random, but are part of complex, interconnected networks that are bound to a specific context and require a particular set of referential knowledge to be wholly understood (see Miltner 2014), which also speaks to the indeterminacy of social meaning in general. More specifically, these memes tap into shared popular culture experiences (Lankshear and Knobel 2007) via their references to foods, traditions, and familial relationships that are culturally specific to Latinxs in the United States. In other words, the memes investigated in the present project not only follow widespread internet norms for the creation of memes but are also constructed around content that is specific to mitú's primary community of users, Latinx millennials in the United States.

5. Conclusions

The data presented here illustrate that the representations and use of language in English–Spanish bilingual memes posted by mitú are not static or isolated, but rather are connected to and participate in a larger sociopolitical dynamic. In particular, the content of these memes, on the one hand, is representative of the cultural and linguistic norms shared by many U.S. Latinx millennials, and, on the other hand, works to construct a particular brand of U.S. Latinx millennial identity that is based on a specific set of social signifiers. The use of culturally specific lexical insertions and the development of characterological figures through quotatives construct and reinforce a specific kind of experience of Latinx millennial identity in the U.S., including being bilingual, having Spanish-speaking parents, and having strong ties to Latinx culture. Overall, the motivations for and outcomes of the content of the memes posted to the mitú Facebook page are multi-faceted and dynamic. These memes allow the representation of Spanish–English language contact and multicultural identity on a social platform and work to construct U.S. Latinx millennial identity through the enregisterment of particular linguistic forms and cultural behaviors.

Looking forward, the findings of this project can be applied to a variety of contexts. Dovchin's (2022) work on translingual discrimination finds that the delegitimation of the practices of multilingual individuals can have wide-reaching social repercussions with regard to education, employment opportunities, and conceptions of identity. Bucholtz et al.'s (2014) framework for sociolinguistic justice presents potential recourse to these processes of delegitimation and the social consequences that result from them. Specifically, this framework calls for (1) the valorization and legitimation of the linguistic resources and practices of minoritized communities, (2) access to all varieties of language that these individuals desire to use, and (3) the affirmation of individuals' own linguistic repertoires. In particular, both Dovchin's and Bucholtz's work is applied to multilingual education contexts, stressing the importance of sociolinguistic justice in education. While the present study does not specifically address educational contexts, it may be a useful resource for linguists, educators, and communities alike by showcasing online community building and reaffirming the importance of language variation, multilingualism, and their connection to identity construction.

Notes

1 https://www.facebook.com/wearemitu (accessed on 1 September 2019).

2 I use the term "Latinx" to be inclusive to gender non-conforming and non-binary individuals, as has increasingly been the tradition in academic writing. As used in this paper, the term is meant to include Latinas, Latinos, Latines, and Latin@s, and anyone who identifies as being of Latino origin. See Guidotti-Hernández (2017) for further discussion.

3 Not only does the content posted to mitú lend itself to a multitude of areas of analysis (e.g., Morrison 2020), but the images in the memes in conjunction with the linguistic material are also worthy of future multimodal discursive analysis, which involves multiple modes of communication (such as visual and spatial) in conjunction with written text.

4 https://www.nbcnews.com/news/latino/mitus-beatriz-acevedo-wants-be-voice-millennial-generation-n396381 (accessed on 1 November 2019).

5 While it is difficult to define precisely what constitutes a millennial, a report from the Pew Research Center (Dimock 2019) categorized those born between 1981 and 1996 as millennials, which would have made them from 23 to 38 years old at the time the data were collected for this project. This means that they represent a large portion of the age range of the listed target audience for the page. According to data from the 2014 American Community Survey, approximately 14.6 million Latinxs are considered millennials, accounting for 26% of the Latinx population in the U.S. (Patten 2016). A 2019 report from the Pew Research Center (Perrin and Anderson 2019) indicates that individuals from the millennial generation fall squarely within the age range that has the highest rate of active Facebook users. For example, in 2019, 84% of 25 to 29 year olds and 79% of 30 to 49 year olds were active users, as compared to 76% of 18 to 24 year olds, 68% of 50 to 64 year olds, and just 46% of those aged 65 and older. While there is no way to verify with absolute certainty the identities of consumers of online media (Varis and Blommaert 2015), much of the content posted to mitú's page appears to be aimed, at least in part, at this expressed target audience.

6 Still, it is critical to note that "following" does not necessarily entail interaction (Zappavigna 2014, p. 156) due to the fact that users have the option to "mute" and "see less" of the content posted by Facebook pages they follow.

7 All the memes presented visually in this work are meant to be illustrative of larger trends in the data set as a whole. A full catalog of all 765 memes can be provided upon request.

8 I acknowledge my position as a white, non-Latina woman who has only ever experienced Spanish through the lens of white, English speaking privilege. I must note that despite having worked with Spanish speakers and Latinx communities throughout my life and career, I am writing about experiences that are not my own. What I am able to offer are my insights into this particular phenomenon based on my knowledge of language and society. I take full ownership of any errors, inaccuracies, and biases presented herein, and acknowledge that my interpretation of these data may be flawed and/or not represent the experiences of all U.S. Latinx individuals.

9 Future work would do well to investigate the degree to which Facebook users who interact with the meme content posted by mitú feel that it is derived from real-life connections to and experiences of U.S. Latinx millennial identity via the analysis of likes, shares, comments, etc.

10 Not only is the use of Spanish enregistered with these portrayals of "mother" and "father", the representations of the behaviors associated with each (re)enregister and reinforce stereotypical gender roles. Though beyond the scope of the present analysis, the ways in which the "mother" and "father" are associated with traditional gender roles are worthy of investigation in their own right. For review of previous work on gender roles in U.S. Latinx communities and the cultural systems that underlie them, see Raffaelli and Ontai (2004) and Miville et al. (2017). For previous work related to the portrayal of gender roles in mitú's content, see Morrison (2020) and Wallace et al. (2020). Gutiérrez (2021) notes that the characters featured by content creators on the Pero Like channel similarly mimic stereotypical behavior as a means of satire.

11 As an interesting note, the use of the word *sike* (also written *psych*) in Figure 30, a slang interjection meaning 'just kidding' that gained widespread popularity in the 1980s and 1990s (Green 2008, p. 1038), further reinforces the millenial age group as one of mitú's target audiences.

References

Agha, Asif. 2003. The Social Life of Cultural Value. *Language and Communication* 23: 231–73. [CrossRef]

Agha, Asif. 2005. Voice, Footing, Enregisterment. *Journal of Linguistic Anthropology* 15: 38–59. [CrossRef]

Agha, Asif. 2006. *Language and Social Relations*. Cambridge: Cambridge University Press. [CrossRef]

Androutsopoulos, Jannis. 2006. Introduction: Sociolinguistics and Computer-Mediated Communication. *Journal of Sociolinguistics* 10: 419–38. [CrossRef]

Androutsopoulos, Jannis. 2015. Networked Multilingualism: Some Language Practices on Facebook and their Implications. *International Journal of Bilingualism* 19: 185–205. [CrossRef]

Backus, Ad. 2001. The Role of Semantic Specificity in Insertional Codeswitching: Evidence from Dutch-Turkish. In *Codeswitching Worldwide II*. Edited by Rodolfo Jacobson. Berlin and New York: De Gruyter Mouton, pp. 125–54. [CrossRef]

Backus, Ad. 2015. A usage-based approach to code-switching: The need for reconciling structure and function. *Code-switching Between Structural and Sociolinguistic Perspectives* 43: 19–37. [CrossRef]

Baym, Nancy K. 1995. The Emergence of Community in Computer-Mediated Communication. In *CyberSociety: Computer-Mediated Communication and Community*. Edited by Stephen G. Jones. Thousand Oaks: Sage Publications, pp. 138–63.

Bentahila, Abdelâli, and Eirlys E. Davies. 1992. Code-Switching and Language Dominance. In *Advances in Psychology*. Edited by Richard Jackson Harris. Amsterdam: Elsevier, pp. 443–58. [CrossRef]

Blackmore, Susan J. 1999. *The Meme Machine*. Oxford: Oxford University Press.

Bucholtz, Mary, Audrey Lopez, Allina Mojarro, Elena Skapoulli, Chris VanderStouwe, and Shawn Warner-Garcia. 2014. Sociolinguistic Justice in the Schools: Student Researchers as Linguistic Experts. *Language and Linguistics Compass* 8: 144–57. [CrossRef]

Christiansen, M. Sidury. 2015a. 'A ondi queras': Ranchero Identity Construction by U.S. Born Mexicans on Facebook. *Journal of Sociolinguistics* 19: 688–702. [CrossRef]

Christiansen, M. Sidury. 2015b. Mexicanness and Social Order in Digital Spaces: Contention Among Members of a Multigenerational Transnational Network. *Hispanic Journal of Behavioral Sciences* 37: 3–22. [CrossRef]

Dawkins, Richard. 1976. *The Selfish Gene*. Oxford: Oxford University Press.

De Fina, Anna. 2007. Code-Switching and the Construction of Ethnic Identity in a Community of Practice. *Language in Society* 36: 371–92. [CrossRef]

Dimock, Michael. 2019. *Defining Generations: Where Millennials End and Generation Z Begins*. Washington, DC: Pew Research Center. Available online: https://www.pewresearch.org/short-reads/2019/01/17/where-millennials-end-and-generation-z-begins (accessed on 10 October 2020).

Dorleijn, Margreet, and Jacomine Nortier. 2009. Code-Switching and the Internet. In *The Cambridge Handbook of Linguistic Code-Switching*. Edited by Brenda Bullock and Almeida Jacqueline Toribio. Cambridge: Cambridge University Press, pp. 127–41. [CrossRef]

Dovchin, Sender. 2022. *Translingual Discrimination*. Cambridge: Cambridge University Press. [CrossRef]

Gal, Noam, Limor Shifman, and Zohar Kampf. 2016. "It Gets Better": Internet Memes and the Construction of Collective Identity. *New Media and Society* 18: 1698–714. [CrossRef]

Goebel, Zane. 2007. Enregisterment and Appropriation in Javanese-Indonesian Bilingual Talk. *Language in Society* 36: 511–31. [CrossRef]

Green, Jonathon. 2008. *Chambers Slang Dictionary*. Edinburgh: Chambers. [CrossRef]

Guidotti-Hernández, Nicole M. 2017. Affective Communities and Millennial Desires: Latinx, or Why My Computer Won't Recognize Latina/o. *Cultural Dynamics* 29: 141–59. [CrossRef]

Gumperz, John J. 1982. *Discourse Strategies*. Cambridge: Cambridge University Press. [CrossRef]

Gutiérrez, Arcelia. 2021. Pero Like and mitú: Latina Content Creators, Social Media Entertainment, and the Politics of Latinx Millenniality. *Feminist Media Histories* 7: 80–106. [CrossRef]

Gutiérrez, Cynthia. 2017. Exploring Latino Acculturation and Code-Switching through Self-Portraiture (Selfies). Master's thesis, Notre Dame de Namur University, Belmont, CA, USA.

Halim, Nur Syazwani, and Marlyna Maros. 2014. The Functions of Code-Switching in Facebook Interactions. *Procedia-Social and Behavioral Sciences* 118: 126–33. [CrossRef]

Hinrichs, Lars. 2006. *Codeswitching on the Web: English and Jamaican Creole in E-Mail Communication*. Amsterdam and Philadelphia: John Benjamins. [CrossRef]

Inoue, Miyako. 2003. The Listening Subject of Japanese Modernity and His Auditory Double: Citing, Sighting and Siting the Modern Japanese Woman. *Cultural Anthropology* 18: 156–93. [CrossRef]

Johnstone, Barbara. 2011. Dialect Enregisterment in Performance. *Journal of Sociolinguistics* 15: 657–79. [CrossRef]

Johnstone, Barbara. 2017. Characterological Figures and Expressive Style in the Enregisterment of Linguistic Variety. In *Language and a Sense of Place: Studies in Language and Region*. Edited by Chris Montgomery and Emma Moore. Cambridge: Cambridge University Press, pp. 283–300. [CrossRef]

Kecskes, Istvan, Olga Obdalova, Ludmila Minakova, and Aleksandra Soboleva. 2018. A Study of the Perception of Situation-Bound Utterances as Culture-Specific Pragmatic Units by Russian Learners of English. *System* 76: 219–32. [CrossRef]

Kosoff, Zoë. 2014. Code-switching in Egyptian Arabic: A sociolinguistic analysis of twitter. *Al-ʿArabiyya: Journal of the American Association of Teachers of Arabic* 47: 83–99.

Krogstad, Jens Manuel, and Ana Gonzalez-Barrera. 2015. *A Majority of English-Speaking Hispanics in the US Are Bilingual*. Washington, DC: Pew Research Center. Available online: https://www.pewresearch.org/short-reads/2015/03/24/a-majority-of-english-speaking-hispanics-in-the-u-s-are-bilingual (accessed on 10 October 2020).

Lankshear, Colin, and Michele Knobel. 2007. Researching New Literacies: Web 2.0 Practices and Insider Perspectives. *E-Learning and Digital Media* 4: 224–40. [CrossRef]

MacSwan, Jeff. 2014. *Grammatical Theory and Bilingual Codeswitching*. Cambridge: MIT Press. [CrossRef]

Mahootian, Shahrzad. 2005. Linguistic Change and Social Meaning: Codeswitching in the Media. *International Journal of Bilingualism* 9: 361–75. [CrossRef]

Miltner, Kate M. 2014. "There's No Place for Lulz on LOLCats": The Role of Genre, Gender, and Group Identity in the Interpretation and Enjoyment of an Internet Meme. *First Monday* 19: 8. [CrossRef]

Miville, Maria L., Narolyn Mendez, and Mark Louie. 2017. Latina/o Gender Roles: A Content Analysis of Empirical Research from 1982 to 2013. *Journal of Latina/o Psychology* 5: 173–94. [CrossRef]

Montes-Alcalá, Cecilia. 2007. Blogging in Two Languages: Code-Switching in Bilingual Blogs. In *Selected Proceedings of the Third Workshop on Spanish Sociolinguistics*. Somerville: Cascadilla Proceedings Project, pp. 162–70.

Morrison, Hannah Grace. 2020. Snapchat Page Mitú: Challenging the Patriarchy? Or just Obsessing over Flaming Hot Cheetos? *Prose Studies* 41: 253–71. [CrossRef]

Muysken, Pieter. 1997. Code-Switching Processes: Alternation, Insertion, Congruent Lexicalization. In *Language Choices: Conditions, Constraints, and Consequences*. Edited by Martin Pütz. Amsterdam and Philadelphia: John Benjamins, pp. 361–80. [CrossRef]

Muysken, Pieter. 2000. *Bilingual Speech: A Typology of Code-Mixing*. Cambridge: Cambridge University Press.

Myers-Scotton, Carol, and Janice L. Jake. 1995. Matching Lemmas in a Bilingual Language Competence and Production Model: Evidence from Intrasentential Code Switching. *Linguistics* 33: 981–1024. [CrossRef]

Ortman, Jennifer M., and Hyon B. Shin. 2011. *Language Projections: 2010 to 2020*; Census Working Papers. Available online: https://www.census.gov/library/working-papers/2011/demo/2011-Ortman-Shin.html (accessed on 10 October 2020).

Patten, Eileen. 2016. *The Nation's Latino population Is Defined by Its Youth*. Washington, DC: Pew Research Center. Available online: https://www.pewresearch.org/hispanic/2016/04/20/the-nations-latino-population-is-defined-by-its-youth (accessed on 10 October 2020).

Perrin, Andrew, and Monica Anderson. 2019. *Share of U.S. Adults Using Social Media, Including Facebook, Is Mostly Unchanged since 2018*. Washington, DC: Pew Research Center. Available online: https://www.pewresearch.org/short-reads/2019/04/10/share-of-u-s-adults-using-social-media-including-facebook-is-mostly-unchanged-since-2018 (accessed on 10 October 2020).

Poplack, Shana. 1980. Sometimes I'll Start a Sentence in Spanish Y TERMINO EN ESPAÑOL: Toward a Typology of Code-Switching. *Linguistics* 18: 581–618. [CrossRef]

Poplack, Shana. 1993. Variation Theory and Language Contact. In *American Dialect Research*. Edited by Dennis R. Preston. Amsterdam and Philadelphia: John Benjamins, pp. 251–86. [CrossRef]

Raffaelli, Marcela, and Lenna L. Ontai. 2004. Gender Socialization in Latino/a Families: Results from Two Retrospective Studies. *Sex Roles* 50: 287–99. [CrossRef]

Shifman, Limor. 2013. Memes in a Digital World: Reconciling with a Conceptual Troublemaker. *Journal of Computer-Mediated Communication* 18: 362–77. [CrossRef]

Squires, Lauren. 2010. Enregistering Internet language. *Language in Society* 39: 457–92. [CrossRef]

U.S. Census Bureau. 2015. *American Community Survey*. Available online: https://www.census.gov/acs/www/data/data-tables-and-tools/data-profiles/2015 (accessed on 10 October 2020).

Varis, Piia, and Jans Blommaert. 2015. Conviviality and Collectives on Social Media: Virality, Memes, and New Social Structures. *Multilingual Margins: A Journal of Multilingualism from the Periphery* 2: 31–45. [CrossRef]

Wallace, Ryan, Olga Lazitski, Kirsi Cheas, Maiju Kannisto, Noora Juvonen, Carolyn Nielsen, and Mark Poepsel. 2020. "We are the 200%": How Mitú Constructs Latino American Identity Through Discourse. *International Symposium on Online Journalism* 10: 13–33.

Wei, Li. 2011. Moment Analysis and Translanguaging Space: Discursive Construction of Identities by Multilingual Chinese Youth in Britain. *Journal of Pragmatics* 43: 1222–35. [CrossRef]

Wei, Li. 2018. Translanguaging as a Practical Theory of Language. *Applied Linguistics* 39: 9–30. [CrossRef]

Wei, Li, and Angel M. Y. Lin. 2019. Translanguaging Classroom Discourse: Pushing Limits, Breaking Boundaries. *Classroom Discourse* 10: 209–15. [CrossRef]

Woolford, Ellen. 1983. Bilingual code-switching and syntactic theory. *Linguistic Inquiry* 14: 520–36.

Zappavigna, Michele. 2012. *Discourse of Twitter and Social Media: How We Use Language to Create Affiliation on the Web*. London: Bloomsbury Publishing.

Zappavigna, Michele. 2014. Coffeetweets: Bonding Around the Bean on Twitter. In *The Language of Social Media: Identity and Community on the Internet*. Edited by Philip Seargeant and Caroline Tagg. London: Palgrave Macmillan, pp. 139–60. [CrossRef]

Article

Mediated Bricolage and the Sociolinguistic Co-Construction of *No Sabo Kids*

Salvatore Callesano

Department of Spanish and Portuguese, University of Illinois, Urbana-Champaign, Urbana, IL 61801, USA;
scalle@illinois.edu

Abstract: Sociolinguistic styles and the resultant ascribed identities are understood as the product of simultaneous variables, leading to the notion of bricolage, or the co-occurrence of variables and their collective indexical meanings. Relatively little attention has been paid to these processes as they manifest on social media platforms. The goal of the current paper is to understand which linguistic and thematic features co-occur in the online production of the *no sabo kid* style and identity, which manifests as a form of linguistic discrimination towards U.S. Latinx youth. "Hashtag communities" were used to locate posts about *no sabo kids* on TikTok (N = 95), and videos were automatically and manually coded for salient linguistic and discursive resources in the online *no sabo kid* community. The results show the co-occurrence of code-switching and phonological and lexical variation, alongside discursive themes, namely 'proficiency', 'ethnicity', and 'performative lexical gaps'. I argue that the *no sabo kid* hashtag community is a mediated manifestation of ideologies surrounding U.S. Latinx bilinguals, where a supposed lack of proficiency in Spanish and grammatical blending of Spanish and English index inauthentic ethnicity. Mediated instantiations of sociolinguistic styles shed light on how linguistic features become enregistered through multimodality and semiotic bricolage.

Keywords: indexicality; enregisterment; bricolage; Latinx language; heritage Spanish; hashtag community; bilingualism

1. Introduction

The same linguistic output can often have different social interpretations. In the context of child language acquisition, *sabo* (I know) in Spanish is an overregularization of an irregular present tense verbal inflection (saber.INF > sé.PRES.1PS). Overregularizations like these are common among children acquiring Spanish (Clahsen et al. 2002), and, when produced by young children, are often met with laughter and reflect affective indexes of 'cute', 'playful', and 'childlike'. Yet, when the same overregularization is produced by an adolescent or young adult heritage speaker of Spanish, the evaluative reactions can shift to those of 'ridicule', 'embarrassment', 'lacking Spanish proficiency', and 'inauthentic ethnic identity'. The purpose of the current paper is to critically address the latter examples, whereby U.S. Latinx social media users produce linguistic features associated with a *no sabo kid* persona. Stransky et al. (2022) note that the identity label is used in mediated spaces to make reference to Latinx youth who are dominant in English and show evidence of language mixing (i.e., borrowing, calquing), all the while functioning within a "linguistic oppressive movement that openly criticizes US Spanish and its speakers" (p. 38).

To date, extensive research has been conducted on the ideological dimensions of heritage Spanish communities and speakers (e.g., Leeman 2018; Rosa 2019), and an area of growing interest is how linguistic ideologies manifest in ethnolinguistic identity production, particularly within new media. As such, the present research directly addresses a call for greater disciplinary connection between linguistics and media studies (Sánchez-Muñoz and Retis 2022) and aims to provide a detailed expansion of Stransky et al.'s (2022), which brings awareness to the current issue and highlights the marginalizing forces of the term *no sabo kids* in social media videos.

1.1. U.S. Latinx Language and Ideology

Despite ongoing political efforts to make English the federal official language of the United States (U.S.), other languages are well represented: nearly one quarter of all school children in the U.S. speak a non-English language at home, and the number of students classified as English Language Learners is likely to increase by 50% by the year 2025 (Potowski 2021). Historically, the Spanish language has been spoken on land that today we consider the U.S. for longer than English, and currently U.S. Latinxs[1] constitute the country's second largest racio-ethnic population, with U.S.-born Latinxs outnumbering those born abroad (Funk and Lopez 2022). Between 1980 and 2019, self-rated English proficiency among U.S. born Latinxs increased by 19% and, conversely, use of Spanish in the home decreased by 10% (Funk and Lopez 2022).

This demographic information gives rise to questions about prestige attributed to the Spanish language in the U.S. context. In South Florida, where Spanish is narrativized as a language of necessity, studies show overwhelming prestige and biases in favor of English as compared to Spanish (Callesano and Carter 2022; Carter and Lynch 2014). In the U.S. Southwest, Rangel et al. (2015) show that, in terms of "status", "solidarity", and "personal appeal", Spanish and English are rated equally, while code-switching is rated lower than both monolithic understandings of the two languages. These examples reflect colonial and monolingual language ideologies (e.g., Wiley 2014), which set the stage for a psychosocial outcome—linguistic insecurity among minoritized language communities.

Linguistic insecurity among heritage language speakers is well documented, particularly in the context of heritage language education (e.g., Martinez and Petrucci 2004). However, outside of the context of education, the negative feelings that heritage speakers often experience towards their own linguistic abilities in the minoritized language—Spanish in this case—is of key importance for sociolinguistic questions of indexicality and style, as these attitudes mediate linguistic performance and use. In a study of linguistic insecurity, Tseng (2021) argues there is an "imposed deficit" attributed to second and third generation Latinx bilinguals, which leads to negative effects on their "linguistic self-identities and self-esteem" (p. 131). The everyday lived experiences of U.S. Latinx bilinguals can be described as facing an ideological double bind—racialized as low proficiency English speakers (e.g., Flores and Rosa 2015) and judged as not proficient enough in Spanish. The process of linking linguistic practices with racial/ethnic categories emphasizes the creation of difference and deficit, allowing for the subordination of multilingual, Latinx linguistic repertoires against monolingual and white language practices (García et al. 2021). The linguistic policing that happens has been described as "linguistic terrorism" in Anzaldúa (1987) and Christoffersen (2019). Bilingual speakers face hostile perceptions and actions, like unnecessary repetition requests (as the data here will show) and placement in remedial education programs, for example. The racialization processes that lead to linguistic insecurity also reflect how shame is a common outcome of linguistic socialization (Lo and Fung 2011).

Whether the perceptions that are studied among U.S. Latinx communities attend to "Spanish" and "English" as conceptual categories (e.g., Kutlu and Kircher 2021) or to specific linguistic features (e.g., Chappell 2019), language-based social biases persist within socio-historical systems that allow them to manifest in material ways. For example, Carter and Callesano (2018) note that different national dialects of Spanish (Cuban, Colombian, and Peninsular) are associated with different levels of occupational prestige and socioeconomic status. African American English (AAE) has long been studied in terms of materialized perceptions, for example in the workplace (McCluney et al. 2021), the court system (Rickford and King 2016), and automatic speech recognition programs (Martin and Wright 2023). As such, the perception of minoritized linguistic varieties—and the bodies that produce them—materialize as not eligible for higher paying employment, not credible in legal environments, and disadvantaged within technological systems, underscoring the ways in which marginalized Latinx bilingual language practices are policed.

This policing takes place within the contexts of standard language (Lippi-Green 1997), monolingual (e.g., Wiley 2014), and raciolinguistic ideologies (e.g., Rosa and Flores 2017)—the axes to which speakers and listening subjects (Inoue 2003) orient themselves when enacting, performing, or evaluating ethnolinguistic identities. To that point, we now turn to a discussion of indexicality, enregisterment, and style.

1.2. The Co-Construction of Ethnolinguistic Identity: From Indexicality to Performance

There is no linguistic variable that occurs outside of a social environment and the social meanings that are attributed to linguistic resources are fluid according to context and listener. Sociolinguists have collectively used the notions of indexicality (e.g., Silverstein 2003), enregisterment (see Agha 2003; Johnstone 2016) and bricolage (Eckert 2008; Guy and Hinskens 2016; Hebdige 1979, among others) to think through the ways in which semiotic resources (i.e., linguistic variables, music, clothing) come together in the production of social meaning. On the one hand, indexicality refers to how linguistic variants point to certain social traits, characteristics, and qualities, such as how variable realizations of syllable-final /s/ in Spanish may index—or perceptually indicate—various levels of social status (e.g., Chappell 2019). Enregisterment, on the other hand, refers to another stage in the processes of assigning social meaning to language features, where linguistic forms ultimately point to identifiable social beings. Johnstone (2011) describes how a set of phonetic and morphosyntactic variables collectively signal a Pittsburgh identity, such as "yinz" as a second-person plural pronoun, multiple negation, and monophthongization of /au/. In Pittsburghese and other language varieties, different linguistic variables are known to co-occur with one another, forming the building blocks that collaboratively index or enregister some social outcome, such as a quality, attribute, or a named social being. The co-occurrence of multiple linguistic variables and their indexical meanings is known as bricolage (Eckert 2008; Guy and Hinskens 2016; Hebdige 1979).

Furthermore, all these processes function within ideological systems about language (Schieffelin et al. 1998), meaning that those individuals and groups who exist at the privileged hegemonic center have the power to dictate, in part, the social meanings that are assigned to language. Individuals whose linguistic repertoires match those of the ostensibly socially privileged are not deemed deficient and, instead, gain capital that is tied to language (Bourdieu 1991). These processes are explored in what Eckert (2012) refers to as third-wave sociolinguistics, which seeks to elucidate how language users take advantage of the variable social meanings that are attributed to linguistic features as they perform various versions of themselves.

To exemplify the theoretical points above, I turn to Mendoza-Denton's (2008) seminal research on language and identity in two Latina youth gangs in a California high school. On the one hand, the *Norteñas* are "Americanized", and to show that, they forefront their bilingualism through translingual practices. On the other hand, the *Sureñas* orient themselves more towards Latin America and thus use Spanish more frequently. The "Norteñas 'feel' themselves Chicanas, Sureñas 'feel' themselves Mexican, and are both opposed to and in intense competition with each other in every aspect of their symbolic existence" (Mendoza-Denton 2008, pp. 250–51). One way this competition becomes clear is through semiotic resources. The *Norteñas* prioritize red and burgundy for their clothing and lipstick, have feathered hair, and listen to Oldies, which goes hand-in-hand with their prioritization of the English language. The *Sureñas* prefer blue and navy clothing with brown lipstick, wear ponytails, and listen to banda music, which is then linked with their more frequent use of Spanish. In this one example, we can see how indexicality and bricolage come together, where the social meaning of a particular linguistic repertoire is reinforced by various semiotic resources building upon each other. Furthermore, paying attention to a specific phonetic variable, Mendoza-Denton's (2008) results show that the *Norteñas* and the *Sureñas* do not differ significantly in their production of raised /ɪ/[2] and as such, the indexical values associated with raised /ɪ/ are dependent upon the

simultaneous realization of extralinguistic features—or, in other words, the co-occurrence of sociolinguistic variables.

Numerous studies have been conducted on the co-occurrence of linguistic features within ethnolinguistic varieties (e.g., Newlin-Łukowicz 2016; Oushiro 2016), but studies of a similar nature are still lacking when it comes to Spanish, particularly for U.S. Latinx bilingual speech communities. Among the few that exist, Erker and Otheguy (2016) examine the co-occurrence of pronoun use, coda /s/, subjunctive morphology, pronoun placement, and word order in New York City and argue for effects of social saliency and Spanish-language ideologies. The aforementioned studies tend to take a variationist sociolinguistic approach, quantitatively exploring how and when variables co-occur, while the data shown below, however, are understood through the notions of linguistic style and mediated performance. When working with data from performance, which are not naturalistic in the early variationist sense, it is useful to note that some videos have a clearly parodic goal and others are simply recordings of interaction. Whether in a performative context or not, linguistic variables are not only produced alongside other linguistic resources (e.g., discourse topic) but also other semiotic materials (e.g., particular clothing, food, or music choices).

In the subsequent analysis, I specifically combine Rampton's (1999) stylization, Johnstone's (2011) linguistic performance, and Zappavigna's (2015) hashtag community in order to place the social media videos in a performative and community-centered context. Rampton (1999) argues that semiotic variables are used as tools in the production of a linguistic style and that these tools then become enregistered with a person-type. Here we can think of the linguistic styles of U.S. presidents (Slatcher et al. 2007), broader styles of a European multiethnolect (Quist 2008), or "Valley Girls" versus "Ghetto Girls" (Goodwin and Alim 2010), for example. According to Johnstone (2011), linguistic performances of enregisterment hinge on paying "meta-communicative attention to [themselves], putting on display not only what the message means but how" (p. 676). Research on linguistic performance illustrates how various semiotic variables co-occur within mediated spaces (e.g., Calhoun 2019; Chun 2004; Slobe 2018). For example, parody performances on YouTube show how collections of semiotic features like uptalk, tag questions, and blond hair work together to enregister the Mock White Girl persona, but also how the performers' ideological stances towards the persona vary (Slobe 2018). Additionally, Chun (2004) demonstrates how the Mock Asian style in standup comedy performances, which features, for example, confusion of liquids and prayer bows, depends on and reproduces "racist ideologies of Asian othering and ridiculing" (p. 287). Both stylization and linguistic performance can be conceptualized in the context of mediatization, and one mediated place where we can locate examples of stylized performance abound is the hashtag community—a mediated community whose membership is determined by using and interacting with metadata, such as a hashtag (Zappavigna 2015). Thus, #nosabokids functions as a mechanism for online community organization. Social media users can choose to incorporate hashtags into their posts for several reasons, such as making the post searchable and/or providing labels or metacommentary. The myriad combinations of features, which can be found within a hashtag community, lead to embodied styles, and the data presented here aim to illustrate the sociolinguistic processes through which the *no sabo kid* identity is performed and sustained on social media.

2. Materials and Methods

The materials and tools used for the current analysis are Python, RStudio (R Core Team 2022), and publicly available videos on the social media app, TikTok.[3] The app has gained immense popularity in recent years and data from 2021 show that 21% of people in the U.S. use the app. Nearly half (48%) of all U.S. adults between the ages of 18 and 29 say they use the app (Auxier and Anderson 2021), and the data discussed here are understood within the context of mediated spaces associated with younger adults at the crossroads of the Millennial generation and Generation Z (Dimock 2019).

The data collection and coding process consisted of two stages. First, a Python script designed to access TikTok's API (Teather 2022) was modified to fit the parameters of the study (see Brown 2023). A total of 118 videos that included one of the four hashtags—#nosabo (n = 30), #nosabokids (n = 30), #yonosabokid (n = 28), #nosabokidsbelike (n = 30)—were automatically collected and coded (see Table 1 for codes). As of 12 December 2022, the hashtags listed above had the following numbers of views on TikTok: 440.9 million, 242.4 million, 937,000, and 47.2 million, respectively, and all video engagement data reported in this article are based on that date. The total runtime for the Python script was 8 min and 30 s on a 2019 Macbook Pro. After collecting this information, a post-hoc analysis was run to see whether the hashtag of interest was found alongside #mexico within the same video caption. This was motivated by the observation that numerous lexical indexes used in the videos are stereotypes of Mexican Spanish (as shown below). By querying whether #mexico is used alongside the other hashtags, results might suggest that the videos producers are taking a stance to ideologically align the *no sabo kid* identity with a particularly Mexican heritage.

Table 1. Codes for information gathered from Python script.

Python Code	Output
ID	numerical 1–118
URL	link to TikTok video
video_author	username of video producer
hashtags	list of hashtags
caption	caption text
isAd	true/false if advertisement
stickersOnItem	text of sticker(s)
privateItem	true/false if private video
commentCount	count value of comments
playCount	count value of plays
shareCount	count value of shares
searchQuery	hashtag used in search
isPairedWith	true/false if paired with #mexico

During the analysis stage, 23 videos were removed from the dataset due to irrelevance to the subject, repetition because of co-occurrence of two or more hashtags, the video's topic being on Anglo-white second language learners of Spanish, the video having been deleted from the user's profile, or the account having changed its settings from public to private. After all exclusions, the final analysis consisted of a total 95 videos produced by 65 users.

The second stage of coding was conducted manually and included the information outlined in Table 2. This stage is divided into three structural categories—phonetic, lexical, and code-switching. Phonetic codes refer to instances of bilingual sound production, lexical codes pertain to words or lexicalized phrases, and code-switching notes any discourse where Spanish and English are simultaneously present. First, each video was coded for the presence or absence of code-switching, as well as phonetic and lexical variables. After the initial round of coding, the lexical category was divided into two sub-categories: nominal (e.g., semantic extensions) and verbal (e.g., morphological blending). These categories were used based on impressionistic data from a pilot analysis conducted in 2021, where videos were qualitatively analyzed for the saliency of linguistic variables employed in the production of the *no sabo kid* identity. Additionally, the codes in Table 2 are founded in previous research on sociolinguistic bricolage and enregisterment (e.g., Johnstone 2011; Newlin-Łukowicz 2016).

Table 2. Manual coding criteria for each video.

Manual Code	Output
phonetic	presence (1), absence (0)
lexical	presence (1), absence (0)
code-switching	presence (1), absence (0)
theme	grounded qualitative coding
type	parody (p), naturalistic (n), commentary (c)

For example, let us say we found the following features in one video: production of /r/ as [ɹ] in the word *gracias* [gɹa.sjas], examples of interactional code-switching between the speaker and an interlocutor, and also an example of a lexical play on words, such as *pedo* 'fart' for *puedo* 'I can'. This example would be coded as 1 for phonetic, 1 for lexical, and 1 for code-switching.

Next, the theme of each video was determined using inductive qualitative coding, specifically following Bingham and Witowsky (2022), where emerging themes were coded by taking into account the entirety of the video—discourses, stickers, sounds, captions, and hashtags. The themes of 'performative lexical gaps', 'proficiency', and 'ethnicity', arose from within the data, but also follow previously established sociolinguistic theory regarding this particular ethnolinguistic community (e.g., Tseng 2021). One video can have multiple themes; however, for the purposes of this analysis, one theme was selected as the most prominent. Lastly, each video was coded for type; distinctions were made if the video was clearly a parody (n = 39), naturalistic (n = 15), in the sense that the recording was made of a genuine interaction with no intention to act or perform, or commentary (n = 41), where the video either consisted of a monologue style of meta-linguistic commentary about *no sabo kids* or the video was created using a meme style, where text (the stickersOnItem in Table 1) was overlaid on top of a short clip with the purpose of that text being the commentary.

When working with data from social media, and considering speakers[4] within the hashtag community, taking note of views, likes, shares, and comments helps to understand community scale, that is, how widespread the hashtag community is and how much other people engage with the material. See Table 3 for this information across the dataset. The corpus of videos shows a wide range of plays, shares, and comments, indicating that videos within the *no sabo kid* hashtag community circulate frequently throughout TikTok. It is an online community that reaches millions of views, which shows that the *no sabo kid* persona is prolific on social media.

Table 3. Corpus engagement statistics (all videos).

Count	Range	Mean	Standard Deviation
Play (number of views)	Min. 513 Max. 12,600,000	1,402,634	2,494,473
Share (number of shares)	Min. 0 Max. 125,100	9145	21,521
Comment (number of comments)	Min. 0 Max. 25,400	1834	3619

3. Results

The data presented here speak to the specific linguistic and discursive variables that are employed within the *no sabo kid* hashtag community on TikTok. In the sections that follow, I describe the indexes at a phonological level as well as the indexes considered to be lexical, which are divided into nominal and morphological categories. Then, I consider their co-occurrence alongside specific discourse themes. I note that nearly 60% of all videos demonstrate code-switching. I also argue that code-switching functions as yet another gear

in the mechanism of co-constructing the *no sabo kid* identity, following studies showing the lack of social prestige of language mixing in light of monolingual norms (e.g., Rangel et al. 2015) and how environments including both Spanish and English are overtly labeled as "ghetto" as a result of the racialization of language in southern California (Bucholtz et al. 2007). The examples provided below demonstrate the linguistic layering in the construction of a derogatory, ethnolinguistic persona.

3.1. Phonetic Indexes

Of the 95 videos, 31 included at least one instance of a phonetic production that indexes a *no sabo kid* identity. While all examples are indicative of either phonetic or phonological differences between Spanish and English, numerous cases involved different instances of rhotic variation. Other examples demonstrate consonant cluster reduction, affrication, and differences between the English and Spanish vocalic systems. Of note is that all the Spanish phonetic realizations that index *no sabo kids* do so vis-à-vis English. Whether intentional through parody or as a part of spontaneous speech, by inserting English-like phonetic realizations into Spanish discourse, the index signals proficiency in English, which has as its byproduct a perceived lack of proficiency in Spanish.

Furthermore, and to the argument for the co-occurrence of these variables within and across the social media posts, it is common to find these phonetic features couched within lexical indexes of the *no sabo kid* identity. Table 4 shows the phonetic indexes that are most present in the dataset, highlighting bilingual productions where monolingual understandings of Spanish and English phonetics are set apart from one other. Transcript 1 illustrates an example of a phonetic index being mobilized to label someone as a *no sabo kid*.

Table 4. IPA of phonetic examples.

Non-Stigmatized Production	Production Indexing *No Sabo Kids*	Example Context
[ɾ]	[ɹ]	*señora* [se.ɲo.ɹa]
[r]	[ɾ]	*perro* [pe.ɾo]
[ɾ]	[r]	*fuera* [fwe.ra]
[ɾd]	[r, l]	*perdón* [pe.ron] or [pe.lon]
[t]	[tʰ]	*taco* [tʰ.koʊ]
[dɾ]	[d͡ʒ]	*madre* [ma.d͡ʒ ɹe]
[we]	[e]	*puedo* [pe.ðo]
[o]	[oʊ]	*sabo* [sa.boʊ]

Transcript 1. #nosabokidsbelike video with 150 comments, 137,300 plays, and 49 shares

1	A: *A ver, Jimena, ¿cómo se preparan los tacos?*
2	'Let's see, Jimena, how are tacos prepared?'
3	A: *Enséñanos.*
4	'Show us'.
5	A: *¿Cómo se preparan los tacos?*
6	'How are tacos prepared?'
7	B: '*Esto es carnita*'.
8	This is meat.
9	A: You put the *carnitas* [kaɾ.ni.tas] in the *tortilla* [toɾ.ti.ʎa]
10	'You put the meat in the tortilla'
11	B: uh huh
12	A: mmk. And then?

13	B: Hold on.
14	B: *walks over to another table with toppings*
15	B: cilantro! [sə.lan.t͡ʃɹo]
16	A: huh!?
17	B: cilantro! [sɪ.lan.t͡ʃɹo]
18	A: cilantro [sɪ.lan.t͡ʃɹo]
19	A: No ss.
20	You a no sabo kid?

To contextualize, speaker A is recording the video and the viewer is limited to only hearing their voice. Speaker B is a young adult, wearing a University of California, Los Angeles sweatshirt. She is inside of a home where a party or gathering is taking place and person A is recording her as she prepares her plate of food. Based on the hashtags used in the video caption, we assume this interaction took place in Staten Island, New York (#statenisland). Speaker A uses phonological code-switching in line 9, producing the food items with Spanish phonology. Then, in line 15, speaker B realizes 'cilantro' with an English-like phonology (i.e., reduced and centralized vowel, affrication of <tr>). This is met with a jarring "huh!?" by speaker A (line 16), which prompts a repetition by speaker B. Speaker A then reiterates the word, as if to mock speaker B in line 18. Both speakers appear to be proficient bilinguals, yet it is speaker B who is being video recorded, indirectly asked to validate her Latinx identity via the process of putting together a taco, and whose linguistic production of 'cilantro' is mocked and used as the trigger for the *no sabo kid* identity. Crucially, as will be discussed further, the co-construction of the *no sabo kid* identity is not just the work of the speaker, but that of the speaker and the listener (D'Onofrio 2015; Rosa and Flores 2017), which here includes a multitude of ambient listeners on social media. Additionally, the phonological alternation across speakers A and B reflects an issue of bilingual familiarity, discussed by López (2020) and Delgado (2020), where words with "common currency" between the two languages can be realized using either phonological system, but the pronunciation of words like *carnitas*, *tortilla*, and *cilantro* with particular phonological patterns does socio-political identity work (Hill 2009). It seems to be the case, then, that there is a co-occurrence of at least two types of linguistic indexes in transcript 1, phonetic and lexical.

3.2. Lexical Indexes

The lexical indexes are divided into two categories: nominal and morphological. Approximately 25% of the videos include a nominal example of a lexical index, and lexical entries coded as nominal include three sub-types: (1) plays on words via phonetic similarity, (2) lexical stereotypes of Mexican Spanish, and (3) phonological adaptation and calquing.

On the one hand, nominal plays on words occur within Spanish and two examples are: *noches/nalgas* 'nights'/'butt cheeks' and *puedo/pedo* 'I can'/'fart'. On the other hand, some are cross-linguistic, such as false cognates (*sopa* 'soup' referencing 'soap' and *embarazado* 'pregnant' referencing 'embarrassed'). There are other examples of more creative false cognates, such as *veterinario* 'veterinarian' used for 'vegetarian', and also inverted Spanglish (Rosa 2016)—*addresso* 'address'—showing the addition of morphological -o to an English nominal root. These plays on words categorically occurred in the parodic type of TikTok video, indicating the hyper-performative nature of these indexes. To that point, it is likely that the comedic effect of phrases such as *buenas nalgas* instead of *buenas noches* or *no pedo* for *no puedo* comes from the rupture of the pragmatic frame, where listeners (viewers on TikTok) have an expectation of a linguistic exchange, and when that expectation is not met, the frame is said to be broken and a possible result is humor (Goffman 1974).

Across the set of videos analyzed here, examples of Mexican Spanish lexical stereotypes are: *güey* 'man'/'dude', *no mames* 'no way', *a la verga* 'fuck it', *que hubo* 'what's up', and *compa* 'friend/mate'. These lexical items are thought of as stereotypes in the Labovian sense, where they are not only used in stylistic variation but are also discussed in

meta-linguistic terms (Labov 1971). 55% of the videos in the dataset included #mexico, suggesting an explicit connection between the *no sabo kid* persona and Mexican heritage.

Finally, the category of phonological adaptation and calques consists of examples where either a lexical form in English is phonologically adapted to Spanish or a structure from English (e.g., skyscraper) is produced as a literal translation in Spanish (e.g., *rascacielos*). Examples of the former are the productions of *gualmar* [wal.maɾ] for English 'Walmart', and an example of the later is *chi-chi de pollo* for 'chicken breast', which uses a colloquial term for women's breasts in place of *pechuga*, the Spanish word for 'breast' in the context of food. The data also show evidence of *lonjas* [lon.xas] for English 'lungs', a semantic calque with some cross-linguistic phonological similarity. It is important to keep in mind the performative nature of these examples; phrases like *chi-chi de pollo* are unlikely to become part of the linguistic norm of the speech community, contrasting with phonologically adapted forms like *gualmar*.

The data show that the lexical indexes, whether nominal or morphological, are (a) not very frequent and (b) are found primarily in the parody and commentary video types, with just one example falling in the naturalistic category (see Figure 1). From the morphological category, examples of morphological blends overwhelmingly show English verbal roots with Spanish inflection, which aligns with Toribio and Bullock's (2016) findings from Spanish heritage speakers in Texas. The examples in Table 5 present the orthography that was used by the TikTok producers.
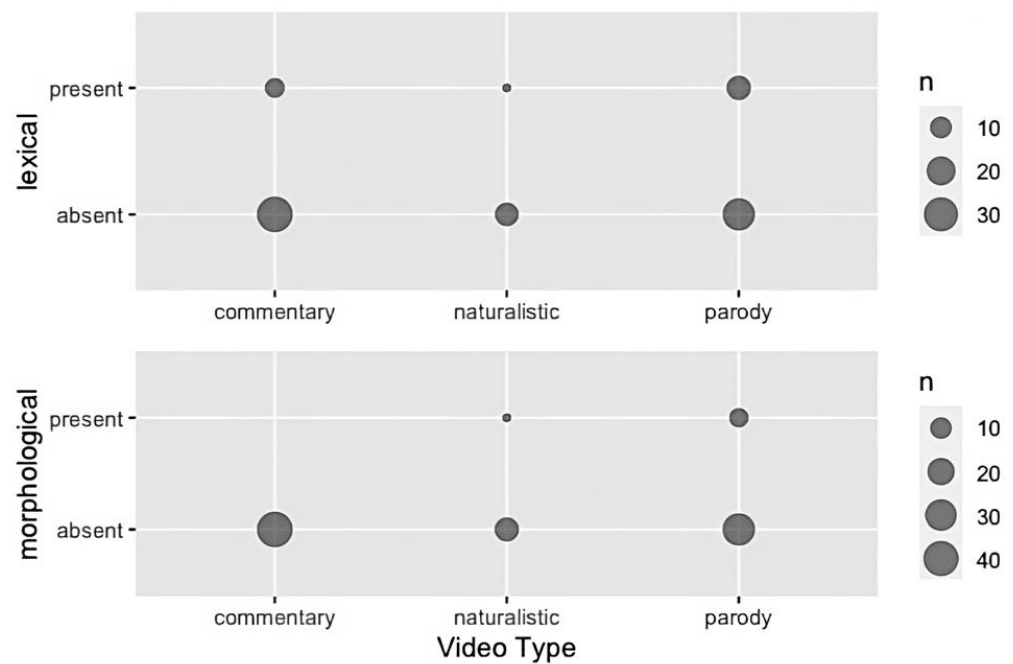

Figure 1. Lexical and morphological indexes by video type.

Table 5. Examples of morphological blends.

Verbal Category	Morphological Blend	Interpretation
Progressive	*barkiando*	barking
	jumpiando	jumping
Infinitive	*danciar*	to dance
	parkiar	to park
	catchar	to catch
	askiar	to ask
	helpiar	to help
	workiar	to work
	blamiar	to blame
	popiar	to pop
	cutiar	to cut
Finite	*blamias*	blame [+present, 2nd person singular]
	tripié	trip [+past, 1st person singular]

In the current dataset, there is exactly one instance of a morphologically blended index of the *no sabo kid* identity within the naturalistic video type category (see Figure 1). In Transcript 2, a child of approximately 4–6 years old is being recorded by an adult (assumed to be her mother or caregiver). While the young child is aware of the recording that is happening, the child is giving the adult directions about the recording placement and method, indicating a non-performative moment. The video producer (assumed to be the adult the viewers do not see and the one who posted the video to TikTok) turns the purpose of the video into a meta-linguistic commentary about the child's use of 'grabbing'. The bilingual child likely retrieved the English verbal root "grab" from the Spanish infinitive *grabar* 'to record' and attached to it the English progressive marker -ING. The caption of the video is, "She tries! She's so cute!" with two heart-type emojis and a laughing-with-tears emoji, and includes the following hashtags: #mexican, #latina, #cute, #nosabo, #nosabokidsbelike, and #texas. The video producer, speaker A (adult), placed a sticker (S) on top of the video (Line 1), which is spoken by an app-internal voice and heard by the viewer. Speaker C represents the child.

Transcript 2. #nosabokidsbelike video with 374 comments, 472,900 plays, and 974 shares

1	S: No sabo kids be like
2	C: Stay grabbing here *eye gaze towards the floor and open hand gesture to suggest to the
3	interlocutor they should not move* ==
4	A: == What's grabbing?
5	C: *Steps closer to camera and phonates something like [ps:]
6	C: Just keep (..)
7	Wait for me! ==
8	== *raises hand in open palm gesture to interlocutor*
9	But stay the camera there.
10	A: Ok

3.3. Thematic Indexes

As a part of the manual coding process, each video was coded for theme via an inductive approach. While many themes became apparent throughout these videos, I focus here on the three most common. The third most frequent theme was that of 'proficiency', (13%) with 'ethnicity' as the second most common (16%), and finally, 'performative lexical gaps' as the most common (20%).

In one proficiency-related video (2 comments, 513 views, and 0 shares), which includes the caption "insecure speaking spanish as if it wasnt my first language" and the hashtags #latina, #yonosabokid, #spanish, and #bilingual, a young adult relates that she watched a comedy skit online, and many written comments claimed the comedian had a "yo no sabo accent". The speaker of the TikTok video analyzed here makes a rhetorical move to say that she has no hate towards *no sabo kids* and that one of her friends is one, but if someone were to refer to her as a *no sabo kid*, she would cry. She moves to say that she is already the "whitewashed" one among her cousins and that is because she is "not like Cuban, y'know", with phrasal stress placed on "Cuban" to suggest that she is Cuban, but not quite as Cuban as her cousins. Next, she says it would be her figurative "last straw" if someone called her a *no sabo kid* and then the video cuts to the same speaker reading very brief segments of a book aloud, where she performs a lack of reading ability in Spanish.

In another video, one that takes the form of a commentary type, the factor of ethnicity is much more direct. In just a seven second video, with 36 comments, 13,200 plays, and 73 shares, the overall argument is that if you do not speak Spanish, you cannot be Dominican. The video sticker reads, "*Cuando me dicen que son dominicanos pero no saben hablar español*" 'When they tell me they are Dominican but they don't know how to speak Spanish', and in the video a person is sitting down within a facial gesture indicating confusion. The caption reads, "*SI USTED ES DOMINICANO Y NO HABLA ESPAÑOL UTE NO E DOMINICANO NA !!*" 'IF YOU ARE DOMINICAN AND DON'T SPEAK SPANISH, YOU ARE NOT DOMINICAN AT ALL !!'. The message here is that to be Dominican, speaking Spanish is required, according to this TikToker, and the link between language and ethnicity is reinforced in this video in two ways: first, in the interaction between the video's sticker and the facial gesture of the user and second, in the caption. Of note is how the producer of the video actively switches into a more Caribbean, or Dominican, way of speaking through the written form of the caption. With "UTE NO E DOMINICANO NA" the producer is arguing that to be Dominican you need not just speak Spanish, but your Spanish needs to sound a certain way (i.e., omit /s/ and /d/).

Finally, regarding the most common theme of 'performative lexical gaps', I include a transcript of a parody TikTok, which uses a greenscreen of a nightclub and a sticker that reads "No Sabo Kids be like:" In this interaction, while there are three distinct characters, all are acted out by the same person. S is the main character in the skit (i.e., the *no sabo kid*), *tía* is the main character's aunt, which makes a reference to women figures as sources of validation in Latinx families (Vasquez et al. 2022), and DJ is a background character in the context of the nightclub.

Transcript 3. #nosabo video with 2129 comments, 1,900,000 plays, and 6166 shares

1	Tía: *Ven, mijo, ven ven ven!*
2	'Come here, mijo[5], come come come!'
3	S: *No no tía es que*
4	'No no aunt it's that'
5	*Yo no* (.) ugh!
6	'I don't' (.) ugh*!*
7	How do you say dance in Spanish?
8	*No sabo* (.) *no sabo danciar*
9	'I don't know (.) I don't know how to dance'
10	Tía: *gasps* ~ sticker reads *shocked in Spanish* ~
11	*¡¿qué dijistes?!*
12	'What did you say!?'
13	DJ: *No mames, un pinche no sabo*
14	'No way, a fucking no sabo'
15	S: *¿Por qué todos me miran? Es que yo no sabo danciar*

16	'Why is everyone looking at me? I don't know how to dance'.
17	*No sabo danciar!*
18	'I don't know how to dance!'
19	Tía: *Mijito no se dice "danciar"*
20	'Mijito we don't say "danciar" (to dance)'
21	*Se dice bailar*
22	'We say "bailar" (to dance)'
23	S: Ah ok
24	*p p p p perrón*
25	'p p p p perrón (a play between *perrón* 'cool' and perdón 'sorry/excuse me')
26	Tía: *No te preocupes mijito yo te enseño cómo bailar*
27	'Don't worry mijito, I'll teach you to dance'
28	DJ*: ¿Alguien dijo perrón?*
29	'Did someone say "perrón"?'
30	*music starts playing with the lyrics *algo que está perrón que toda la gente brinca"*
31	*'something that is so cool that everyone jumps up'**

At the core of this parody is the supposed missing lexical entry for "to dance" in Spanish. To fill this ostensible gap, a bilingual can use their linguistic awareness to creatively come up with what might be a plausible outcome by adding Spanish verbal morphology to an English verbal root. Each and every instance of *danciar* is immediately preceded and followed by the emoji 🕺, indicating a mocking tone. In the clip we also see evidence of shaming from a familiar adult, a topic to be discussed further below, as well as other examples of sociolinguistic indexes of the *no sabo kid* identity, such as *perrón* for *perdón*. The video producer is aware of the play on words between *perrón* and *perdón*, as the song lyrics that are heard at the end of the clip use the term in a more appropriate context, as an adjective describing something as really 'cool'. While this video is certainly parodic, evidence in lines 11 and 26 indicate experience with stigmatized sociolinguistic variants in Spanish—second person singular preterit /s/ (Barnes 2012) and heritage Spanish systems, where for example English "how" is mapped onto Spanish infinitives. This illustrates an interaction of stigmatized variables, some parodic and some not, in the mediated production of the *no sabo kid*.

4. Discussion

The evidence from this mediated approach to sociolinguistic style, enregisterment, and identity shows that phonetic and lexical variables, as well as code-switching and narrative themes collectively play key roles in the production (and performance) of the *no sabo kid* persona. While the language features noted here may index various attributes (e.g., friendly, trustworthy), the collection of variables enregisters the *no sabo kid* identity, or in other words, the social being to which the features point (Johnstone 2016). In her work on enregisterment and performance, Johnstone (2011) notes that audiences are asked to consider what it could mean that a specific character (e.g., a mother) has a specific accent (e.g., Pittsburgh) while also sounding working class (p. 676). Additionally, she notes that not all audience members/listeners will perceive the indexical relationships in the same way. In the data presented here, we consider what it could mean that a particular community (e.g., Latinx young adults) speaks a certain way (e.g., mixing Spanish and English phonetics, morphology, and discourse), while also facing pressures of racialization and linguistic subordination. Across naturalistic, commentary, and parody video types, the data show, for example, how rhotic variation is couched within lexical plays on words, how verbal roots in one language are combined with verbal suffixes in another, and how the bricolage of bilingual features are encompassed by themes of proficiency, ethnicity, and performative lexical gaps.

Whether through parody or not, enregisterment of the *no sabo kid* is accomplished through a clear mechanism—the hashtag community (Zappavigna 2014, 2015). Each video

included in the analysis is overtly marked with the name of the social being to which the semiotic features point, namely #nosabokids (or any variation of the hashtag therein). While the phonetic, lexical, and discursive categories within this hashtag community were treated separately, all are in constant interaction. The hashtags themselves also function in conjunction with other hashtags, such as #mexico. Just over half of the videos in the dataset showed co-occurrence with #mexico, making a clear national origin identity link with the *no sabo kid* persona, though there is evidence of non-Mexican associations as well. Furthermore, the variables that index *no sabo kids* carry various sociopolitical weights, one of which is intergenerational shame.

Research on language socialization (Lo and Fung 2011) shows that shaming (often between parents and their children) is the result of local ideologies and sociocultural context and a "belief in the power of language" (p. 186). Furthermore, not all shaming is necessarily negative, and it can be paired with feelings of play. Yet, the interactional moves of parents, caregivers, teachers (i.e., adults in positions of power) to ask for repetition or imitate young Latinx bilinguals' linguistic productions has the potential of internalizing shame that is directly linked to language. Following Rosa (2019) and Lawrence and Clemons (2022), when that shame is linked to language, it is simultaneously linked to race and ethnicity, raising issues of not just linguistic insecurity, but also authenticity.

Models of bilingualism, especially in the context of U.S. schools, are largely based on linguistic deficits (e.g., García 2002)—or what bilinguals supposedly lack. The video productions analyzed here, which come from an app used primarily by people between the ages of 18 and 29 in the U.S., can be thought of as manifestations of imposed sociolinguistic deficits (Tseng 2021)—the result of colonial approaches to bilingualism. As schools in the U.S. attend to students of minoritized language backgrounds in ways that further deficit models of bilingualism, parents and caregivers are then likely to continue the ideological process of shaming, and when that is coupled with social media technologies like TikTok, popular videos with millions of views exemplify more of what bilinguals do not do. One linguistic skill that is evident in the TikTok videos is inverted Spanglish (Rosa 2016), where, at times, Latinx Spanish-English bilinguals produce linguistic realizations that indicate intimate knowledge of both Spanish and English, such as hyperanglicizing Spanish pronunciation (e.g., *sabo* as [sa.boʊ] or *madre* as [ma.d͡ʒɹe]). In the data here, from the deficit perspective, pronunciations like these, supposed lexical gaps, and jarring repetition requests from adults may create indexes of inauthenticity. What is more, these repetition requests can take various forms, such as the example of "huh!?" in Transcript 1 and the *gasp* in Transcript 3, and both reflect linguistic hostility (Christoffersen 2019). At the same time, the linguistic productions of *no sabo kids* give us insight into the detailed knowledge that bilinguals have in their linguistic repertoire, and additionally, the cultures and ethnicities that are inextricable from language.

The expectations of parents, caregivers, and other listening subjects are founded on an understanding of Spanish as separate from English, and moreover, the language-related purpose of the parody videos, for example, depends on positioning one language against the other. #nosabokid commentaries and performances hinge upon and uphold colonial ideologies of named languages. This work has also drawn from colonial narratives about language separateness, as discussions of code-switching inevitably reflect the grammatical and ideological marking of a boundary between two languages (Urciuoli 2008), whereas translanguaging intentionally erases the distinction (García et al. 2021). However, if bilingual speakers are viewed as drawing from one unified network of language, as is the case within a translanguaging framework (García et al. 2021), they cannot be "deficient" in any named language. Some videos in this dataset, albeit just two of 95, demonstrate a critical call to decolonize the discussions we have around heritage Spanish speakers in the U.S. One TikToker stated in a video—"I did not choose the no sabo kid life, the no sabo kid life chose me"—and the videos that call out the negative effects of the term make explicit mention of how intra-Latinx shaming furthers linguistic assimilation to English in the U.S.

Mediated bricolage—the collective construction of a style, identity, or persona via various semiotic resources within new media—is a mechanism through which the *no sabo kid* persona is produced and reproduced. Regarding how sociolinguistic styles within a mediated context play critical roles in damaging and marginalizing ethnolinguistic identity, I raise two points. First, mediated instantiations of sociolinguistic styles tell us not just about the linguistic features used within a mediascape (Appadurai 1996), but also how linguistic features become enregistered through semiotic bricolage and how mediated tools (e.g., emojis) can shed light on language and its indexes of ethnicity and authenticity. Second, I highlight the negation—the "no"—in the label *no sabo kids*. Here, the focus is on the negative, on what the speakers do not know. In other words, the ideological stances produced by these videos create meta-communicative commentary about what U.S. Latinx bilinguals do not do, rather than the myriad of grammatical and communicative skills they dominate, in both languages.

The *no sabo kid* hashtag community is a mediated manifestation of ideologies surrounding U.S. Latinx bilinguals, where a supposed lack of proficiency in Spanish and grammatical blending of Spanish and English not only index inauthentic ethnicity but also point to an idea of community, namely *no sabo kids*. In Transcript 1, the adult listener directly asks the teenager if she is a "no sabo kid" simply due to vocalic productions that sound more like English during a cultural activity linked to *latinidad*. Thus, the nexus between Latinx identity and the ability to speak Spanish in the U.S. means different things according to who is speaking and who is listening, which is a result found in both mediated and non-mediated contexts (Lawrence and Clemons 2022; Rosa and Flores 2017). Furthermore, we must note how linguistic productions by young Latinx bilinguals are grammatically and cognitively valid. To productively blended verbal morphology, for example, requires high level competency in both languages (see Toribio and Bullock's (2016) work on code-switching and proficiency) and the presence of English in discourse where Spanish is expected does not presuppose a lack of ability in Spanish.

These sociolinguistic expectations from the perspective of listening subjects (Inoue 2003) are conditioned by ideologies related to standardization (Lippi-Green 1997) and race/ethnicity (Rosa 2019). Negative attitudes towards bilingualism likely exacerbate ongoing declines of Spanish use in the home (Funk and Lopez 2022). When we consider the interconnectedness of linguistic policing and language production in light of mediatization (Androutsopoulos 2016) and how connections across digital communities are more readily available (Zappavigna 2011)—and more public—we come to understand just how prolific (i.e., millions of views) the *no sabo kid* persona, along with its ascribed attributes and qualities, can be. The speakers noted here are bilinguals not by choice but by necessity, and the social meanings of Spanish, at least in the context of U.S. Latinx communities on social media, are inseparable from the sociopolitical forces of English.

5. Conclusions

Latinx linguistic practices in the U.S. are policed under the colonial histories of both Spanish and English, where Latinxs are racialized to lack competency and proficiency in each language (e.g., Rosa 2019). Bilingual language use is framed as different and deficient under the guises of white monolingual norms and expectations, which are traceable back to the formation of the nation-state (Andresen and Carter 2016). Bilingualism and language contact are certainly not new phenomena, though in postmodern societies the circulation of language ideologies is more easily dispersed across boundaries and communities (Lynch 2022). One contemporary result of this has been the *no sabo kid* persona on social media (Stransky et al. 2022), and this article highlights the sociolinguistic mechanisms through which this persona is enregistered. The bricolage of bilingual phonology, morphology, lexical items, discursive themes, and other semiotic material, all in the context of new media, allows sociolinguists to directly consider multimodality in the construction of personae (Ilbury 2023; Mondada 2016). Additionally, this work intentionally emphasizes the interface of production and perception within third-wave sociolinguistics; the combination of how

Spanish is heard (e.g., tense /i/ heard as lax /ɪ/), which body produces that variation, and the person or institution that listens, plays a significant role in creation of ethnolinguistic identities. Finally, the linguistic productions described here exist at a blurred crosslinguistic boundary between Spanish and English, yet the reactions, perceptions, ideologies are couched within monolithic understandings of speech and speech communities. It is at this juxtaposition where the devaluation of bilingualism occurs.

Future research on *no sabo kids* will benefit from an explicit comparison of parody versus non-parody video types to understand more completely how proficiency and competency can be understood within parody. Additionally, a quantitative approach to mediated bricolage, perhaps involving acoustic analyses, to understand how variables build upon each other in statistical terms might be enlightening, but researchers should keep in mind potential acoustic limitations of social media data. Finally, if future lines of this work are to be carried out qualitatively, it will be useful to incorporate multiple coders and test for inter-rater reliability, to avoid potential single-researcher bias.

Notes

1.   The Pew Research Center uses the term "U.S. Hispanic", but for the purpose of consistency through this manuscript, the term Latinx is used throughout, following a pattern in the field of referring to Latinx language repertoires or language in U.S. Latinx communities, for example.

2.   As discussed in Mendoza-Denton (2008), raised /ɪ/ is exemplified by the pronunciation of the word 'thing' as [θiŋ] or [t̪iŋ]

3.   All TikTok usernames have been removed to anonymize the producers of the videos and recorded speakers, following IRB protocol approved by the University of Illinois, Urbana-Champaign.

4.   In this paper, the term "speaker" may refer to the individual who posted the video to TikTok or the speakers included in the video recording.

5.   "Mijo" or "mijito" in Spanish is a morphological blend of the possessive adjective *mi* (my) and the noun *hijo* (son) and it can be used to refer to someone younger and denote care and respect. *Mijo/mija* can be used to refer to someone who is not the speaker's son or daughter, as in Transcript 3.

References

Agha, Asif. 2003. The Social Life of Cultural Value. *Language & Communication* 23: 231–73. [CrossRef]

Andresen, Julie Tetel, and Phillip M. Carter. 2016. *Languages in the World: How History, Culture, and Politics Shape Language*. West Sussex and Malden: WILEY Blackwell.

Androutsopoulos, Jannis. 2016. Theorizing Media, Mediation and Mediatization. In *Sociolinguistics: Theoretical Debates*. Edited by Nikolas Coupland. Cambridge: Cambridge University Press, pp. 282–302.

Anzaldúa, Gloria. 1987. *Borderlands/La Frontera: The New Mestiza*, 2nd ed. San Francisco: Aunt Lute Books.

Appadurai, Arjun. 1996. *Modernity At Large: Cultural Dimensions of Globalization*. Minneapolis: University of Minnesota Press.

Auxier, Brooke, and Monica Anderson. 2021. Social Media Use in 2021. Pew Research Center: Internet, Science & Tech (blog). April 7. Available online: https://www.pewresearch.org/internet/2021/04/07/social-media-use-in-2021/ (accessed on 15 February 2023).

Barnes, Sonia. 2012. ¿Qué Dijistes?: A Variationist Reanalysis of Non-Standard-s on Second Person Singular Preterit Verb Forms in Spanish. In *Selected Proceedings of the 14th Hispanic Linguistics Symposium*. Edited by Kimberly Geeslin and Manuel Díaz-Campos. Somerville: Cascadilla Proceeding Project, pp. 38–47.

Bingham, Andrea J., and Patricia Witowsky. 2022. Deductive and Inductive Approaches to Qualitative Data Analysis. In *Analyzing and Interpeting Qualitative Research: After the Interview*. Edited by Charles Vanover, Paul Mihas and Johnny Saldaña. Newcastle upon Tyne: Sage, pp. 133–48.

Bourdieu, Pierre. 1991. *Language and Symbolic Power*. Reprinted. Cambridge: Harvard University Press.

Brown, Josh. 2023. TikTok Hashtag Communities. Available online: https://github.com/joshb283/TikTokHashtagCommunities (accessed on 3 April 2023).

Bucholtz, Mary, Nancy Bermudez, Victor Fung, Lisa Edwards, and Rosalva Vargas. 2007. Hella Nor Cal or Totally So Cal?: The Perceptual Dialectology of California. *Journal of English Linguistics* 35: 325–52. [CrossRef]

Calhoun, Kendra. 2019. Vine Racial Comedy as Anti-Hegemonic Humor: Linguistic Performance and Generic Innovation. *Journal of Linguistic Anthropology* 29: 27–49. [CrossRef]

Callesano, Salvatore, and Phillip M. Carter. 2022. Unidirectional Language Bias: The Implicit Association Test with Spanish and English in Miami. *International Journal of Bilingualism*. [CrossRef]

Carter, Phillip M., and Andrew Lynch. 2014. Ideological Dimensions of Speech Perception in Bilingual Miami. Paper presented at 7th Workshop of Spanish Sociolinguistics, Madison, WI, USA, April 3–5.

Carter, Phillip M., and Salvatore Callesano. 2018. The Social Meaning of Spanish in Miami: Dialect Perceptions and Implications for Socioeconomic Class, Income, and Employment. *Latino Studies* 16: 65–90. [CrossRef]

Chappell, Whitney. 2019. Caribeño or Mexicano, Profesionista or Albañil? Mexican Listeners' Evaluations of /s/ Aspiration and Maintenance in Mexican and Puerto Rican Voices. *Sociolinguistic Studies* 12: 367–93. [CrossRef]

Christoffersen, Katherine. 2019. Linguistic Terrorism in the Borderlands: Language Ideologies in the Narratives of Young Adults in the Rio Grande Valley. *International Multilingual Research Journal* 13: 137–51. [CrossRef]

Chun, Elaine W. 2004. Ideologies of Legitimate Mockery: Margaret Cho's Revoicings of Mock Asian. *Pragmatics* 14: 263–89. [CrossRef]

Clahsen, Harald, Fraibet Aveledo, and Iggy Roca. 2002. The Development of Regular and Irregular Verb Inflection in Spanish Child Language. *Journal of Child Language* 29: 591–622. [CrossRef]

D'Onofrio, Annette. 2015. Persona-Based Information Shapes Linguistic Perception: Valley Girls and California Vowels. *Journal of Sociolinguistics* 19: 241–56. [CrossRef]

Delgado, Rodrigo. 2020. Gender and Phonetics in the Spanish/English Code-Switched Determiner Phrase. Ph.D. thesis, University of Illinois Chicago, Chicago, IL, USA. [CrossRef]

Dimock, Michael. 2019. Defining Generations: Where Millennials End and Generation Z Begins. Pew Research Center (blog). Available online: https://www.pewresearch.org/short-reads/2019/01/17/where-millennials-end-and-generation-z-begins/ (accessed on 10 June 2023).

Eckert, Penelope. 2008. Variation and the Indexical Field1. *Journal of Sociolinguistics* 12: 453–76. [CrossRef]

Eckert, Penelope. 2012. Three Waves of Variation Study: The Emergence of Meaning in the Study of Sociolinguistic Variation. *Annual Review of Anthropology* 41: 87–100. [CrossRef]

Erker, Daniel, and Ricardo Otheguy. 2016. Contact and Coherence: Dialectal Leveling and Structural Convergence in NYC Spanish. *Lingua* 172–173: 131–46. [CrossRef]

Flores, Nelson, and Jonathan Rosa. 2015. Undoing Appropriateness: Raciolinguistic Ideologies and Language Diversity in Education. *Harvard Educational Review* 85: 149–71. [CrossRef]

Funk, Cary, and Mark Hugo Lopez. 2022. A Brief Statistical Portrait of U.S. Hispanics. Pew Research Center Science & Society (blog). June 14. Available online: https://www.pewresearch.org/science/2022/06/14/a-brief-statistical-portrait-of-u-s-hispanics/ (accessed on 21 September 2022).

García, Ofelia. 2002. Teaching Language Minorities in the United States: From Bilingualism as a Deficit to Bilingualism as a Liability. *International Journal of the Sociology of Language* 2002: 125–30. [CrossRef]

García, Ofelia, Nelson Flores, Kate Seltzer, Li Wei, Ricardo Otheguy, and Jonathan Rosa. 2021. Rejecting Abyssal Thinking in the Language and Education of Racialized Bilinguals: A Manifesto. *Critical Inquiry in Language Studies* 18: 203–28. [CrossRef]

Goffman, Erving. 1974. *Frame Analysis: An Essay on the Organization of Experience*. Frame Analysis: An Essay on the Organization of Experience. Cambridge: Harvard University Press.

Goodwin, Marjorie Harness, and H. Samy Alim. 2010. 'Whatever (Neck Roll, Eye Roll, Teeth Suck)': The Situated Coproduction of Social Categories and Identities through Stancetaking and Transmodal Stylization. *Journal of Linguistic Anthropology* 20: 179–94. [CrossRef]

Guy, Gregory R., and Frans Hinskens. 2016. Linguistic Coherence: Systems, Repertoires and Speech Communities. *Lingua* 172–173: 1–9. [CrossRef]

Hebdige, Dick. 1979. *Subculture, the Meaning of Style*. New Accents. London: Methuen.

Hill, Jane H. 2009. *The Everyday Language of White Racism*. Blackwell Studies in Discourse and Culture 3. Chichester: Malden: Wiley-Blackwell.

Ilbury, Christian. 2023. The Recontextualisation of Multicultural London English: Stylising the 'Roadman'. *Language in Society*, 1–25. [CrossRef]

Inoue, Miyako. 2003. The Listening Subject of Japanese Modernity and His Auditory Double: Citing, Sighting, and Siting the Modern Japanese Woman. *Cultural Anthropology* 18: 156–93. [CrossRef]

Johnstone, Barbara. 2011. Dialect Enregisterment in Performance. *Journal of Sociolinguistics* 15: 657–79. [CrossRef]

Johnstone, Barbara. 2016. Enregisterment: How Linguistic Items Become Linked with Ways of Speaking. *Language and Linguistics Compass* 10: 632–43. [CrossRef]

Kutlu, Ethan, and Ruth Kircher. 2021. A Corpus-Assisted Discourse Study of Attitudes toward Spanish as a Heritage Language in Florida. *Languages* 6: 38. [CrossRef]

Labov, William. 1971. The Study of Language in Its Social Context. In *Advances in the Sociology of Language*. Edited by J. A. Fishman. The Hague: Mouton, Volume 1, pp. 152–216.

Lawrence, Anna, and Aris Clemons. 2022. (Mis)Languaging and (Mis)Translating Identity: Racialization of Latinidad in the US Mediascape. *Latino Studies* 21: 42–63. [CrossRef]

Leeman, Jennifer. 2018. Critical Language Awareness and Spanish as a Heritage Language: Challenging the Linguistic Subordination of US Latinxs. In *The Routledge Handbook of Spanish as a Heritage Language*. Edited by Kim Potowski. London: Routledge.

Lippi-Green, Rosina. 1997. *English with an Accent: Language, Ideology, and Discrimination in the United States*. London: Routledge.

Lo, Adrienne, and Heidi Fung. 2011. Language Socialization and Shaming. In *The Handbook of Language Socialization*. Hoboken: John Wiley & Sons, Ltd., pp. 169–89. [CrossRef]

López, Luis. 2020. *Bilingual Grammar: Toward an Integrated Model*, 1st ed. Cambridge: Cambridge University Press. [CrossRef]

Lynch, Andrew. 2022. *Spanish in Miami: Sociolinguistic Dimensions of Postmodernity*. London: Routledge.

Martin, Joshua L., and Kelly Elizabeth Wright. 2023. Bias in Automatic Speech Recognition: The Case of African American Language. *Applied Linguistics* 44: 613–30. [CrossRef]

Martinez, Glenn A., and Peter R. Petrucci. 2004. Institutional Dimensions of Cultural Bias on the Texas-Mexico Border: Linguistic Insecurity Among Heritage Language Learners. *Critical Inquiry in Language Studies* 1: 89–104. [CrossRef]

McCluney, Courtney L., Myles I. Durkee, Richard E. Smith, Kathrina J. Robotham, and Serenity Sai-Lai Lee. 2021. To Be, or Not to Be…Black: The Effects of Racial Codeswitching on Perceived Professionalism in the Workplace. *Journal of Experimental Social Psychology* 97: 104199. [CrossRef]

Mendoza-Denton, Norma. 2008. *Homegirls: Language and Cultural Practice Among Latina Youth Gangs*. Hoboken: John Wiley & Sons.

Mondada, L. 2016. Challenges of Multimodality: Language and the Body in Social Interaction. *Journal of Sociolinguistics* 20: 336–66. [CrossRef]

Newlin-Łukowicz, Luiza. 2016. Co-Occurrence of Sociolinguistic Variables and the Construction of Ethnic Identities. *Lingua* 172–173: 100–15. [CrossRef]

Oushiro, Livia. 2016. Social and Structural Constraints in Lectal Cohesion. *Lingua* 172–173: 116–30. [CrossRef]

Potowski, Kim. 2021. No Child Left Monolingual: How and Why to Promote Multilingualism in US Schools. In *International Education at the Crossroads*. Edited by Deborah N. Cohn and Hilary E. Kahn. Bloomington: Indiana University Press, pp. 156–74.

Quist, Pia. 2008. Sociolinguistic Approaches to Multiethnolect: Language Variety and Stylistic Practice. *International Journal of Bilingualism* 12: 43–61. [CrossRef]

R Core Team. 2022. *R: A Language and Environment for Statistical Computing*. Vienna: R Foundation for Statistical Computing. Available online: https://www.R-project.org/ (accessed on 22 August 2023).

Rampton, Ben. 1999. Styling the Other: Introduction. *Journal of Sociolinguistics* 3: 421–27. [CrossRef]

Rangel, Natalie, Verónica Loureiro-Rodríguez, and María Irene Moyna. 2015. 'Is That What I Sound like When I Speak?': Attitudes towards Spanish, English, and Code-Switching in Two Texas Border Towns. *Spanish in Context* 12: 177–98. [CrossRef]

Rickford, John R., and Sharese King. 2016. Language and Linguistics on Trial: Hearing Rachel Jeantel (and Other Vernacular Speakers) in the Courtroom and Beyond. *Language* 92: 948–88. [CrossRef]

Rosa, Jonathan. 2016. From Mock Spanish to Inverted Spanglish: Language Ideologies and the Racialization of Mexican and Puerto Rican Youth in the United States. In *Raciolinguistics: How Language Shapes Our Ideas About Race*. Edited by H. Samy Alim, John R. Rickford and Arnetha F. Ball. Oxford: Oxford University Press, pp. 65–80.

Rosa, Jonathan. 2019. *Looking Like a Language, Sounding Like a Race*. Oxford: Oxford University Press.

Rosa, Jonathan, and Nelson Flores. 2017. Unsettling Race and Language: Toward a Raciolinguistic Perspective. *Language in Society* 46: 621–47. [CrossRef]

Sánchez-Muñoz, Ana, and Jessica Retis. 2022. Spanish in the US: A Much-Needed Dialog between Linguistics and Communication/Media Studies. In *Communicative Spaces in Bilingual Contexts: Discourses, Synergies and Counterflows in Spanish and English*. Edited by Ana Sáchez-Muñoz and Jessica Retis. Boca Raton: Taylor & Francis.

Schieffelin, Bambi B., Kathryn A. Woolard, and Paul V. Kroskrity. 1998. *Language Ideologies: Practice and Theory*. Oxford: Oxford University Press.

Silverstein, Michael. 2003. Indexical Order and the Dialectics of Sociolinguistic Life. *Language & Communication* 23: 193–229. [CrossRef]

Slatcher, Richard B., Cindy K. Chung, James W. Pennebaker, and Lori D. Stone. 2007. Winning Words: Individual Differences in Linguistic Style among U.S. Presidential and Vice Presidential Candidates. *Journal of Research in Personality* 41: 63–75. [CrossRef]

Slobe, Tyanna. 2018. Style, Stance, and Social Meaning in Mock White Girl. *Language in Society* 47: 541–67. [CrossRef]

Stransky, Daniela, Emma Donnelly, Cheyenne Stonick, María Dominguez, and Diego Pascual y Cabo. 2022. New Media Representations of Spanish Heritage Speaker: The Case of #nosabokids. In *Communicative Spaces and Media in Bilingual Contexts: Discourses, Synergies and Counterflows in Spanish and English*, 1st ed. Edited by Ana Sánchez-Muñoz and Jessica Retis. Routledge Research in Language and Communication. Abingdon and New York: Routledge, pp. 32–44.

Teather, D. 2022. TikTokAPI. Available online: https://github.com/davidteather/tiktok-api (accessed on 15 October 2022).

Toribio, Jacqueline Almeida, and Barbara Bullock. 2016. A New Look at Heritage Spanish and Its Speakers. In *Advances in Spanish as a Heritage Language*. Edited by Diego Pascual y Cabo. Amsterdam: John Benjamins Publishing, pp. 27–50.

Tseng, Amelia. 2021. 'Qué Barbaridad, Son Latinos y Deberían Saber Español Primero': Language Ideology, Agency, and Heritage Language Insecurity across Immigrant Generations. *Applied Linguistics* 42: 113–35. [CrossRef]

Urciuoli, Bonnie. 2008. Boundaries, Language, and the Self: Issues Faced by Puerto Ricans and Other Latina/o College Students. *Journal of Latin American Anthropology* 8: 152–72. [CrossRef]

Vasquez, Marissa C., Ángel De Jesus Gonzalez, Yolanda Cataño, and Fernando Garcia. 2022. Exploring the Role of Women as Validating Agents for Latino Men in Their Transfer Success. *Community College Journal of Research and Practice* 46: 488–503. [CrossRef]

Wiley, Terrence G. 2014. Diversity, Super-Diversity, and Monolingual Language Ideology in the United States: Tolerance or Intolerance? *Review of Research in Education* 38: 1–32. [CrossRef]

Zappavigna, Michele. 2011. Ambient Affiliation: A Linguistic Perspective on Twitter. *New Media & Society* 13: 788–806. [CrossRef]

Zappavigna, Michele. 2014. CoffeeTweets: Bonding around the Bean on Twitter. In *The Language of Social Media: Identity and Community on the Internet*. Edited by Philip Seargeant and Caroline Tagg. London: Palgrave Macmillan UK, pp. 139–60. [CrossRef]

Zappavigna, Michele. 2015. Searchable Talk: The Linguistic Functions of Hashtags. *Social Semiotics* 25: 274–91. [CrossRef]

(Mis)pronunciations of Hispanic Given Names in the U.S.: Positionalities and Discursive Strategies at Play

Paola Enríquez Duque

Department of Spanish and Portuguese, The Ohio State University, Columbus, OH 43210, USA; enriquezduque.1@osu.edu

Abstract: This qualitative study examines the indexical nature of given names and their role in self-positioning within diverse social contexts. The study centers on the pronunciation of Hispanic given names in the United States. The analysis is grounded in interviews with six young adults who recognize that their names have Spanish and English variants, and it demonstrates that bearers' phonological awareness plays a critical role in distinguishing name variants and mispronunciations, as evidenced through metalinguistic comments. These distinctions are additionally shaped by personal criteria. By examining the participants' narratives and one participant's discursive strategies in particular, I show that the pronunciation of given names constitutes a significant linguistic resource intentionally mobilized and managed to negotiate social positionings. Moreover, this research highlights that conferring Hispanic given names in the U.S. constitutes a sociocultural strategy that extends beyond an indexical ethnocultural naming practice across generations. This practice is found to be a means of fostering and maintaining intergenerational relationships.

Keywords: names; mispronunciation; social meaning; discursive strategies

1. Introduction

1.1. The Power of Names

Names and their corresponding phonetic and orthographic realizations are not just referential forms but also social indexes (Bucholtz 2016; Rymes 1996). Given names are a symbolic representation of "images, significations, and emotional reverberations for the giver, the bearer, and the community at large" (Zelinsky 1970, p. 748) because they carry great cultural significance (Parada 2019; Rymes 1996; Thompson 2006) and are indexes of ethnic identity and group and community membership (Ainiala and Östman 2017; Lipski 1976; Parada 2019; Rymes 1996; Sue and Telles 2007), especially in pluri- ethnic and racial contexts, which also tend to be multilingual spaces. In the context of immigrant and heritage populations, names become key indexes in the multiple communities of practice in which immigrants participate (Pavlenko 2001) and are part of the construction and negotiation of bicultural, bilingual, and *binominal*[1] identities (Thompson 2006). In this way, a given name has "the power to not only convey but to help form a sense of ethnic identity and even commitment to the language it represents" (Parada 2013, p. 304). It is common for these names to become distinctive above the mainstream language names because of their phonological patterns, pronunciation, and/or orthographic representation, and, as such, ethnic names, or even single phonemes of these, function as complex social indexes (Parada 2020).

In the context of marginalized or ethnically minoritized populations, names can function as a source of discrimination and racial aggressions materialized in offensive naming practices (Bertrand and Mullainathan 2004; Goldstein and Stecklov 2016; Kohli and Solórzano 2012; Lipski 1976). In the U.S., studies have shown that people with distinctive African American names (e.g., Lakisha or Jamal) face differential treatment in contexts such as searches for jobs (Bertrand and Mullainathan 2004), rentals (Carpusor and Loges 2006),

or temporary lodging (Edelman et al. 2017). In all three studies, African Americans received fewer callbacks or acceptances than people with White-sounding names (e.g., Emily or Greg). Similarly, proper names can prime stereotypes. Paredes (1993) focuses on the Mexican-American border and discusses how Anglo-Americans make particular use of the names *Pedro* and *Pancho* as distinctive derogatory ethnic labels to invoke a stereotypic view of Mexicans or Mexican Americans: *Pedro*, "usually pronounced with a long *a* as 'Paydro'—evokes the fat, stupid but basically harmless peon"; and *Pancho* suggests "the bandit stereotype, the Mexican with the long mustaches and the cartridge belts crossed over his chest" (33). In addition to this renaming practice (i.e., arbitrarily assigning a different name), denaming (i.e., erasing someone's name), misnaming (i.e., naming with someone else's name) (Bucholtz 2016), and mispronouncing (i.e., phonetically rendering a name into the dominant language phonetic system) (Lipski 1976) are all instances of (linguistically) violent actions. As Lipski (1976) argues, it is essential to recognize that, although mispronouncing could occur unintentionally due to "lack of phonological awareness or dexterity and laziness," it also happens deliberately with the "desire to degrade, belittle, or ridicule" (113). This phonological alteration into English could reach the point of hyperanglicization, which in some instances becomes an exaggerated imitation of a Spanish accent underlying racist jokes (Zentella 2003), as is the case with Mock Spanish (Hill 2008). Consequently, it has been argued that in immigrant-background families, the choice of a given name in the language of the host country, as opposed to a name in the parents' language, signals an early stage of acculturation in contrast to ethnic maintenance (Goldstein and Stecklov 2016; Parada 2016; Sue and Telles 2007). Additionally, in some cases, the deliberate selection of a mainstream name is understood as a strategy by parents to protect their children from racially discriminatory experiences (Souto-Manning 2007).

Research framed within third-wave sociolinguistic approaches and linguistic anthropology has shown how language use functions as a complex resource for social indexicality (Eckert 2019; Hall-Lew et al. 2021). Studies analyzing this complexity range from specific linguistic variables (e.g., Babel et al. 2021; Babel 2014; Barnes 2015; Bucholtz 2009; Chappell 2016; Walker et al. 2014) to languaging[2] practices (e.g., Babel 2018; Mendoza-Denton 2008; Rosa 2019; Zentella 1998). However, scholarship on the pronunciation of personal names as a sociolinguistic variable is still scarce. Likewise, despite the extensive breadth of the field of onomastics (the study of names) (Hough 2016; Kostanski and Puzey 2016; *Names: A Journal of Onomastics*), the subfield of socio-onomastics is still in its infancy.

Ainiala and Östman (2017, p. 1) define socio-onomastics as "the sociolinguistic study of names as linguistic elements that are not only employed as identificatory or reference devices but that are also used to accomplish a variety of culturally, socially and interactionally relevant tasks." Some studies in the field have revealed how the phonetic realization of foreign place-names indexes speakers' social categories. For instance, Hall-Lew et al. (2010) and Silva et al. (2011) examine how English speakers in the U.S. use different pronunciations of *Iraq* and *Iran* to index their political stance and signal their status as bi/multilingual speakers or as individuals who have served in the military. Hill (2008, pp. 143–44) discusses how the distinct pronunciations of *Tucson* serve as social indexes of cultural identity (e.g., Chicana) and speakerhood (e.g., speaker of Spanish). In Austin, Texas, Regan (2022) studies the pronunciation of two popular streets with Hispanic names: *Guadalupe* and *Manchaca*. The author finds that, while the English-phonology guises were perceived in terms of localness, the Spanish-phonology guises indexed lower socioeconomic status and higher social affect by participants, but perceptions varied based on the listeners' demographics, including time lived in Austin and ethnicity (i.e., Hispanic or non-Hispanic). This body of work focusing on place-names, although different in nature from personal names, shows that a sociolinguistic perspective on the study of names and name pronunciation is fertile ground for exploring indexicalization and identity construction and negotiation.

In a particularly noteworthy investigation in the understudied arena of personal names, Wolf et al. (1996) explore the pronunciation of French surnames in New Orleans as a device by which speakers manage their interactions. Based on interview analyses, the

authors contend that name bearers are not necessarily committed to any pronunciation of their names and therefore have varying attitudes toward particular mispronunciations. In this way, understanding pronunciation as (in)correct comes from a personal stance and is mediated by the context of the interaction itself. Likewise, correcting a mispronunciation can include varying strategies depending on the case, such as informing the interlocutors about the pronunciation or expecting the interlocutor to repeat and produce the corrected form of the name.

Hispanic[3] given names in the U.S. usually present two variants: a Spanish-phonology variant and an English-phonology variant. The current study focuses on the pronunciation of given names as a social variable and a linguistic resource to deploy and construct social identities.

1.2. Hispanic Names in the U.S.

Parada (2016, p. 21) contends that Hispanic naming patterns in the U.S. vary by communities' sociohistorical backgrounds, demographics, language patterns, and political and sociocultural stances. An early study of Hispanic given names in five different U.S. cities (Lavender 1988) provides an analysis of the twenty most frequent male and female names, classified as Hispanic names (Spanish spellings, e.g., Carlos), bicultural names (names spelled identically in English and Spanish and frequently used in Anglophone and Spanish-speaking cultures, e.g., Andrea), and English names (English spellings or variants, e.g., Michael). The results reflect the size, historical background, and language patterns of the five communities. Notably, Miami was the city with the highest proportion of Hispanic names among its adult population, primarily attributed to its large foreign-born demographic. Conversely, the remaining four cities, namely, San Antonio, Tampa, Denver, and Albuquerque, favor bicultural names rather than total assimilation to English names. Among these four cities, Denver and Albuquerque stood out with the largest proportion of English names.

Using data from 1995, Sue and Telles (2007) examine the most popular 500 names of children in Los Angeles county, most of whom were of Mexican origin. They argue that, when looking at naming patterns, previous studies have often oversimplified the categorization of names as *ethnic* or *nonethnic*.[4] Therefore, they introduce a five-point scale that allows a more gradient classification and understanding of the language choices of a name. The scale is in the form of a "Spanishness" continuum, which has at the left-most extreme the most English names with no Spanish equivalent, e.g., *Ashley*, and at the right-most extreme, Spanish names with no English equivalent, e.g., *Guadalupe*. The categories in between represent names that have corresponding counterparts in English and Spanish, e.g., *Michael* and *Miguel*, and the centermost category contains names native to both languages, e.g., *Andrea*. The authors found that U.S.-born Hispanics tended to give their sons more English-translatable names than first-generation immigrants, but for daughters, more exclusively English names were given. They argue that the selection of both English- and Spanish-translatable names over exclusively Spanish names allows people of Hispanic heritage to participate in both heritage and mainstream communities because the name, familiar in both languages and cultures, works as a bridge between the two.

Parada (2016), in a study of Hispanic naming patterns in Chicago, found similar results to those of Sue and Telles (2007) in Los Angeles: Hispanic names are predominant in the first generation (59%), and their proportion decreases gradually for the second (54%) and third (38%) generations. Additionally, English name proportions increase across generations: 16% to 31% to 45%. Parada suggests that this growth in English names may signal a progressive cultural and linguistic shift. Nonetheless, the data reveal that the third generation continues to bear Hispanic names, showing that "it is likely that the implicit understanding of the connections between names, ethnic affiliation, and language is at the heart of Latino immigrant parental naming practices" (32). However, as the results differ between the two communities, with LA presenting a higher adoption of English names

than Chicago, the author calls for more nuanced approaches to analyzing naming patterns and practices in each context.

Focusing on a Mexican American family in El Paso, Doran (2001) analyzes three generations (grandparents, parents, and children) and shows that grandparents tend to use Spanish phonology in personal names more consistently and about twice as often as parents and children. The author claims that, although there seems to be a certain degree of conventionalization for the choice of Spanish phonology in some instances, for most words no clear pattern as to which ones are pronounced with Spanish or English phonology is apparent. However, as one participant argues, it is necessary to "know the code" of the community, which means recognizing which pronunciation a particular name calls for (Doran 2001, p. 167). For instance, elders' names call for Spanish pronunciation more strictly than do the names of younger community members.

The trends shown in the previous studies give an idea of the complexity and significance of naming practices and what these entail at an individual, family, community, and cultural level. Furthermore, research over the past decades confirms the continuous presence of Hispanic names in the personal-names landscape of these communities. Most of these studies have provided large-scale regional descriptions and analyses, setting up the context for qualitative work centered on ideologies and attitudes toward names.

There are very few studies that examine the social meanings of Hispanic names from a qualitative perspective, with Parada (2020) being one of the most representative. Based on data from two focus groups sessions, the study addresses the nexus among names, language, and identity, and specifically, the role that name-based attitudes, ascriptions, and stylizations play in the dynamics of the negotiation of identities. Focusing on Spanish-receptive bilinguals in a Spanish L2 college classroom, the author shows how names can serve as indexes of language competence and ethnolinguistic affiliation. The study explores the students' attitudes toward their names as ethnolinguistic markers and how names and their pronunciation serve as a sociolinguistic resource in negotiating identities.

Parada identifies two positions among the students: those who feel misrepresented by their ethnic names and those who feel more neutral or positive about them. On the one hand, pride in their family and ethnic heritage resulted in a positive appreciation of their names. On the other hand, the main reasons for a more negative stance were name length, pronunciation difficulty, negative ethnic stereotypes tied to the name, and being directly racially profiled, which included the expectation of knowing Spanish. The author discusses that the pronunciation of names is a dynamic social variable whose value is determined situationally. For the receptive bilingual participants, the Spanish variant of their names seemed to function as one of the few linguistic means they use to overtly index their ethnolinguistic background to convey a sense of honoring their ethnolinguistic roots and even assert certain Spanish competence when in conversation with monolingual or Spanish-dominant speakers, such as their grandparents. Alternatively, they felt that if they were to use the Spanish variant in an English-dominant context, such as school, it would invoke an unwanted identity stereotype (i.e., positioning them as an "expert" in Hispanic heritage and the Spanish language). Ultimately, Parada's study shows that names are important indexes of ethnolinguistic identities and that their phonetic realizations are sociolinguistic tools for initiating style choices and navigating social spheres.

Parada's work aligns with Baird et al.'s (2018) study on the indexicality of lexically specific phonology switches (LSPS). LSPS occur when a Spanish loanword, that is, any lexical element such as object names (e.g., taco, piñata), personal names (e.g., Carlos, Garcia), or place-names (e.g., San Antonio, Colombia) is produced with Spanish phonology in a full English utterance. Baird et al. posit that the use of Spanish phonology in Spanish words indexes stereotypes of Hispanic communities in the U.S. and that this use makes even native English-speaking bilinguals targets of racial and cultural profiling. Kohli and Solórzano (2012, p. 47) discuss that, in the U.S. context, the mispronunciation or changing of given names of racial minorities and people of color constitute racial microaggressions, defined as "covert or everyday forms of systemic racism used to keep those at the racial margins

in their place." Their study focuses on K-12 classrooms when teachers, intentionally or not, mispronounced their students' names. Retrospective narratives of student participants from different ethnic backgrounds show that, in some cases, these microaggressions become internalized and have lasting consequences. For some Hispanic participants, what may seem like casual mispronunciations, such as placing the stress on the wrong syllable or not rolling the <r>, made them feel like outsiders or as if they were not Americans. During their childhood, these mispronunciations caused them feelings of embarrassment and contributed to the formation of negative perceptions of: (1) themselves because of how different their names appeared to be; (2) their parents, who chose those names for them; and (3) their ethnicity and culture in general.

Overall, scholarship on Hispanic names and naming practices in the U.S. demonstrates that given names are significant and complex indexes of individual and cultural identities. Hispanic communities and immigrant generations in the U.S. exhibit varying naming practices resulting from the different historical and sociocultural contexts, and individual choices of names are the result of situated factors that account for the caregivers' expectations for their children. The cross-generational presence of Hispanic given names in a politically English-dominant country is a testament of the significance of the choice of Hispanic names as an enduring naming practice in Hispanic communities in the U.S. Nevertheless, despite culturally shared understandings and experiences across communities, attitudes towards the pronunciation of names are as individual as each life story, and these result from individual interpretations of (mis)pronunciation experiences.

2. Methods

This study is based on qualitative data from six individual interviews conducted in early 2022 over Zoom, lasting on average 35 minutes each. Only the audio recordings were stored for further analysis. Three participants came from personal connections, and the other three were contacts from people in my personal social network. Following Rubin and Rubin (2011)'s qualitative interviewing framework, the conversations consisted of a semi-structured and responsive interviewing style, characterized by flexibility of design and active listening. The guiding questions were not addressed in a fixed order or posed literally in the same manner to each participant but adjusted to how the conversation developed in each case. Most lead-in questions were open-ended, and follow-up questions were posed when necessary to encourage the participants to elaborate further on their commentaries. No specific language was strictly set. Therefore, in all cases, Spanish and English were used, following the naturalness of the interaction between both parties. The following are some of the guiding questions posed in the interviews:

- How do you usually pronounce your given name?
- Are there any circumstances under which you alter this pronunciation?
- How important do you think an accurate pronunciation of a name is?
- How do you feel about other people pronouncing your name differently?
- Could you describe any incident involving the mispronunciation, changing, or disrespect of your name, if any?
- Do you know the story behind your name?
- What does your name mean to you?
- How would you name your children in the future?

The participants were three men and three women, all young adults between the ages of 19 and 28 (mean = 23.33; standard deviation = 3.28). All of them self-identified as Latino/Hispanic. Five of them were of Mexican heritage and were born in the U.S., and one was born in Peru. Following Silva-Corvalán's (1994) and Del Carpio's (2022) immigration generation groupings, three participants were second generation (i.e., parents were born abroad, and they were born in the U.S.), two were third generation (i.e., they were born in the U.S. along with at least one parent also born in the U.S.), and one, having arrived in the U.S. at the age of 8, belonged to the 1.5 generation (i.e., born in their country of origin but completed their formative periods of education in the U.S. [Rumbaut and Ima (1988,

p. 22)]). Four grew up in the Midwest, one in the Southwest, and one on the West Coast. Two participants were high school graduates, one was an undergraduate student, two were graduate students, and one was a graduate professional. All of them were multilingual and reported speaking Spanish at home, mainly with their parents. Following Sue and Telles's (2007) scale for measuring the "Spanishness" of names, four of the names fall in the fourth category, i.e., Spanish names with an English translation; one falls in the third category, i.e., a name native in both languages; and one falls in the first category, i.e., an English name with no Spanish translation. However, regardless of the category, all participants recognized that their names had English and Spanish pronunciations with which they were familiar.

Participants received the informed consent form via email before the pre-arranged appointment for the Zoom interview, and it was discussed before beginning the interview and the recording. Considering that personal names constitute the object of study here, participants agreed to make and discuss explicit references to their given names. Participants were guaranteed that any other potential personal identifier (e.g., last name, relatives' names, name of an institution, etc.) that would come up during the conversation would not be transcribed to text. In addition, this identifier would be bleeped out from the audio recording with white noise in an effort to maintain anonymity and confidentiality in disseminating the results of this study.

Following Guenther's (2009) critical discussion of the methodological decision of using real names or pseudonyms in social science studies, and Parada's (2020) reference to it, I opted for pseudonyms in the 24 excerpts chosen below due to the small sample size of the study. Efforts have been made to choose a pseudonym alternative that belongs to the same categories of "Spanishness" (Sue and Telles 2007) the original names fall into and to reflect relevant phonological features that distinguish each name variant from the bearers' perspectives, as will be discussed in the following section. Consequently, the four pseudonyms in the 'Spanish name with English translation' category are *Alberto*, *Estela*, *Leonardo*, and *Elena*; the pseudonym in the 'Native in both languages' category is *Leyla*, and the pseudonym for 'English name with no Spanish translation' is *Tristan*.

The data analysis is framed within a qualitative approach and grounded theory (Bernard 2006; Charmaz 2006; Glaser and Strauss 2006), which is a methodology in social sciences that aims to develop theory from empirical data collected systematically. As such, grounded theory methodology contrasts with logical deductions from a priori assumptions, and data analysis under this approach consists of coding for themes and synthesizing the data in an iterative process (Bernard 2006; Charmaz 2006). After the first step of data collection (in this case, conducting interviews), emergent themes were identified during the transcription process, and when this process was completed, the themes were refined, categorized, and related. The overarching category of this analysis is the meanings of (mis)pronunciations of names. It is informed by three main themes selected: (1) metalinguistic (phonological) awareness of the name variants; (2) the role of name variants in social relationships and contexts; (3) the pronunciation of names as a negotiation strategy for social positionalities.

In addition, a case study is presented, which consists of a discourse analysis of one of the participants' narratives of (mis)pronunciation experiences, selected due to the participant's particularly open and detailed retelling during the interview. The analysis focuses on the interaction of three main discursive strategies used by Elena, which support the selected themes: the usage of metalinguistic comments, the use of reported speech, and the use of the discourse particle *just*. (In Section 3.2.1, I give a brief overview of *just* and its multiple discursive functions.)

The discourse analysis is framed within the Interactional Sociolinguistics approach (Gumperz 1982, 2015), which focuses on the meaning-making and interpretation processes, or how meanings are constructed and understood in specific interactions occurring in a situated context (Bailey 2008). Within Interactional Sociolinguistics, the multiple social

categories that comprise social identities are considered dynamic and are communicatively constructed in the discourse (Gumperz and Cook-Gumperz 1982).

3. Analysis and Discussion

All participants recognized that their names are phonetically realized in multiple forms, usually referred to as the Spanish and English pronunciations. These variants involve disparities between certain phonemes contained in the names, which bearers distinguish and demonstrate by metalinguistic commentary about vowel and consonant realizations.

(1)   *Tristan*:[5] I guess it (the difference) would come down to like just the alphabet, right? Like the way, the way the language difference in language is like and in Spanish, dices [a] so [ˈtris.tan], whereas in English it would- it's [eɪ] so it's [ˈtɹɪs.tən]

(2)   *Leonardo*: in English like drop the <r> I guess, and yeah, and just slow down and make that subtle change of like the <l> sound

Participants resort to the alphabet to explain the difference between the variants. Wolf et al. (1996, p. 420) discuss how the alphabet is a significant metalinguistic device that individuals use to address (mis)pronunciation of names, as it facilitates the connection between speech and writing and thus affords speakers a tool to manage the phonology of their names. In excerpt (1), Tristan explicitly mentions that the difference between variants depends on the respective alphabets and then expands upon his explanation. Although his name contains a letter <a>, his description alludes to the fact that the orthographic vowels have different phonetic realizations in each language. In excerpt (2), Leonardo mentions what he perceives as the difference in the phonetic realization of two consonants in his name, a drop of an <r> and a change in the <l>. The metalinguistic comments reveal, from each individual's viewpoint, the meaningful differences that determine if their names are being produced in Spanish or English, which are personal rather than strictly phonetic. The individual take on the phonemes that contrast and shape the variants' distinctions demonstrates that an exhaustive phonetic account, as a linguist would provide, may be unnecessary for determining when a name is pronounced in one language or the other, for individuals seem to be attuned to a limited set of identifiable differences between the two forms. This highly personal interpretation also suggests that the consideration of the pronunciation of a name as a mispronunciation is subject to individual deliberations and is not a fixed category.

In line with Wolf et al. (1996), the participants' attitudes towards the (mis)pronunciations of their names are not homogeneous. For some participants, only the Spanish form is considered correct, while for others, both pronunciations are taken as valid, illustrated below in excerpts (3)–(6).

(3)   *Alberto*: I think it sounds kind of awesome because it's like, "hey, I can pronounce myself in English and Spanish, en inglés y en español". (...) And like I recently learned how to say my name in German (...). So it's like there's three ways of saying it. It sounds like... It makes you sound educated. It makes you sound nice. It's just- it sounds awesome

(4)   *Leyla*: para mí no (no hay una versión correcta y una incorrecta) porque como crecí con los dos lenguas estoy como que me acostumbré a los dos y pues los dos no no cambian para mí, es como lo mismo

        'for me, no (there's no correct or incorrect version of the name) because since I grew up with the two languages, I am kind of like used to both and, well, neither changes for me, both are like the same'

(5)   *Leonardo*: I would say, obviously, that the way that I grew up saying it is the proper way to say it, I would grow up to say, you know, [le.o.ˈnaɾ.ðo],

like, you know, but it's, but yeah I would say that that's the correct way, yeah that's the correct way

(6)　*Estela*: the correct way that my parents would pronounce it, like they'd be like "[es.ˈte.la] *ven para acá*" ('come here') or something like that

Alberto and Leyla accept both name variants but express different emotional stances toward them. For Alberto, the multiplicity of variants is positively assessed as an index of multilingualism, as indicated by the adjectives *awesome*, *educated*, and *nice*. Alberto embraces the three different pronunciations of his name as a tool to display his multiple linguistic resources, which he recognizes as valuable in his social environment. Leyla, on the other hand, expresses a less emotive stance. She accepts the correctness of both pronunciations of her name based on her familiarity with the name variants, as she grew up accustomed to hearing both realizations.

Excerpts (5) and (6) demonstrate how, for both Leonardo and Estela, the way the closest people in their most intimate circle pronounce their names has a direct impact on their definition of the 'correct' variant. For Leonardo, mentioning how he pronounced his name growing up reflects how his family would say it, which establishes the correct model. For Estela, the command phrase ('come here') following her name as a vocative gives the sense of a typical childhood interaction between her and her parents. These participants' responses illustrate how the value of correctness for the name variants usually develops throughout their childhoods.

The tight connection between the Spanish pronunciation of their names and their families is also apparent in the rationale behind the selection of their names. Some of the participants note that their Hispanic name was chosen out of consideration for those in their closest social spheres, like non-English speaking relatives, who may not feel comfortable pronouncing an English-only name.

(7)　*Leonardo*: there's not much of a story behind it (the name) other than they (the parents) keep insisting that it is a it's a name that would sound good in English and Spanish and it was a name that they just liked

(8)　*Elena*: (mi madre) quería un nombre es que dijo que tenías que tener en este tiempo un nombre que se podía pronunciar en español (...) dijo "mis tíos, mis tías, todos no podían pronunciar un nombre en inglés y aunque te quería poner... si te quería poner algo como Ashley o algo así, no se podía hacer, era algo que no aceptaban", entonces decidió este nombre porque pensó que se podía pronunciar bien en los two idiomas y que no iba a ser como difícil para los Spanish speakers and the English speakers

'(my mom) wanted a name she said that, at that time, you had to have a name that was pronounceable in Spanish (...) she said "my uncles, my aunts, everyone couldn't pronounce a name in English, and even though I wanted to call you... even if I wanted to call you Ashley or something like that, it was not possible, it was something that was not accepted," so she decided this name because she thought it was easy to pronounce in the two languages and that it was not going to be as difficult for the Spanish speakers and the English speakers'

In both cases, the parents' aspiration for their children is that they would partake in their ethnolinguistic heritage by establishing relationships with close relatives such as uncles and aunts. A Spanish form is an index of this heritage, and the availability of this variant benefits the Spanish speakers of their circle, who, in some cases, might find pronouncing an English-only name complicated. This consideration is still present in the second generation, showing that the practice of giving names pronounceable in Spanish is robust across generations, as seen in excerpt (9).

(9)　*Leyla*: quisiera que... podría que si tuviera una hija podría relacionarse con con sus grandparents, you know? con mi parte de la familia que tiene

> ese español, pero también al mismo tiempo con la parte de hablar inglés que es la que yo traigo. Y por para mí sí sería importante escoger nombres que tienen una pronunciación en español y una en inglés (...) una parte de respeto a mis padres para que ellos puedan decir los nombres de de mis hijos en una versión correctamente, you know? sin sentirse como que están diciéndolo mal, porque creo que es importante esto también

> 'I would like that... she could, if I had a kid, she could connect and interact with her grandparents, you know? with that part of my family that has that Spanish, but at the same time with the English speaking part that is the one I bring in. For me it would be important to choose names with a Spanish and an English pronunciation (...) a form of respect to my parents so they can pronounce my kids' names correctly, you know? without feeling like they are saying it wrong, because I think that's important too'

As previous studies have discussed (Baird et al. 2018; Hall-Lew et al. 2010; Regan 2022), articulating a proper name, such as a place-name, in the phonology of the source language can be perceived as a demonstration of higher social affect and respect towards the culture represented by the name. In the case of personal names, the respect is extended to individuals. In line with Parada's (2020) participants' comments, Leyla makes it explicit that a Spanish variant constitutes a way of honoring and showing respect to her closest relatives, and therefore to her family and ethnolinguistic heritage. Since Spanish might be the dominant language of previous generations (as is the case for all six participants), the Spanish pronunciation of names becomes a meaningful resource for establishing a connection among generations. An English-only name could be difficult for her parents to produce, causing them to feel linguistically unskilled when addressing their future grandchildren.

Elena's (8) and Leyla's (9) comments underscore that a Spanish name facilitates intergenerational relationships where linguistic practices may differ. Furthermore, an intentional name with both pronunciations depicts the cultural negotiation that multicultural and multilingual people engage in daily. The younger generations' regard for parents, grandparents, and people in their community, upon choosing Spanish names for their children, shows that a given name carries value beyond its referential function and that its choice is emotionally and socially motivated. In tandem with these motivations, individual social meanings for names, and particularly for the name variants, are constructed throughout a person's life (hi)story, as Leyla explains in excerpt (10).

(10)   *Leyla*: creo que es como aceptar que tengo dos i-, bueno no dos identidades, pero dos cosas que pueden definirme como a mi identidad. Como tengo mi nombre como se diría más culturalmente por parte de de mis padres, de, you know, de ser mexicana y luego por parte de haber crecido aquí. Es como, no sé, tener esas dos versiones es como explicar todo mi pasado en una una manera

'I think it is like accepting that I have two i-, well not two identities, but two things that can define me like define my identity. Like I have my name as it would be said more culturally from my parents' side, from, you know, from being Mexican, and then, on the other side, from having being raised here. It is like, I don't know, having those two versions is like explaining my entire past in a way'

Hill (2008, p. 143) posits that the social indexicality of language choice is complex in multicultural contexts where languages are in conflict, as is the case of English and Spanish in the U.S. This complexity is reflected in the way name variants index different components of an individual's positionality. The junction of family, ethnic, and national-origin heritage, the speakerhood of multiple languages, and the participation in different social networks and sociocultural contexts intervene in the formation of meaning and the constructed value of a name with multiple pronunciations. This process begins when the bearer becomes aware of the existence of the two variants, which usually happens when they compare the

linguistic practices they undertake in an educational context with the Spanish linguistic practices at home, as shown in excerpts (11)–(13).

(11)   ***Tristan*:** I guess when I started going to school because when I, when I- I was born in Peru, so Spanish is my first language, so when I came over to the U.S. and I went to school, I guess that's when and I talked with people that spoke English. That's when I realized, so I guess like elementary school (...) when I first encounter that, like the difference, I would think about it, but it's so it's so common now that it doesn't even I don't even think about it anymore

(12)   ***Elena*:** Ahh la primera vez no recuerdo muy bien, pero sé que estaba chiquita, it was elementary school, tenía medio 8 o 9 años y me estaba pensando como por qué en la escuela todos me dicen [ə.ˈleɪ.nə], like no me gusta que se suena así y todos en mi casa me dicen [e.ˈle.na]. Entonces así me- ahí me di cuenta y luego entré a high school en mi clase de español, pronunciaban mi nombre en español, y ahí estaba como "¿por qué lo podemos hacer aquí pero no a mis otras clases?"

'Uhm, the first time I don't remember clearly, but I know I was little, it was elementary school, I was like 8 or 9, and I was thinking like at school why does everybody call me [ə.ˈleɪ.nə] like I don't like that it sounds like that, and at home, they call me [e.ˈle.na]. So that's when I realized it, I noticed it, and then I started high school, and in my Spanish class, they would pronounce my name in Spanish, and I was like, "why can we do it here but not in my other classes?" '

(13)   ***Leyla*:** se me hizo tan natural que en la casa me decían [ˈlej.la] y en la escuela me decían [ˈleɪ.lə] y era como que me acostumbré rápido y llegué al a la universidad y ahí es cuando me di cuenta de que de esa diferencia porque estaba más en la universidad que en la casa, pero cuando mi mamá me llamaba por teléfono me decía [ˈlej.la] y decía "¿y por qué nadie lo dice así aquí?" es cuando me di cuenta, pero antes de eso no, no pensé mucho de eso

'it was so natural to me that at home, they called me [ˈlej.la], and at school, they called me [ˈleɪ.lə], and it was like I got used to it fast, and then I got to college, and that's when I noticed that difference because I spent more time at school than at home. But when my mom used to call me on the phone, she would call me [ˈlej.la], and I used to say, "why does nobody say it like that here?" That is when I realized it, but before then, no, I didn't think much about it'

There is a clear relationship between the Spanish pronunciation and the way names are pronounced at home. In (11), for Tristan, the difference became evident when he started learning English at school after moving to the U.S. from Peru. The then-novel realization of the English pronunciation of his name became an ordinary vocative that he grew accustomed to, to the point that he would not think about the distinction later on. Leyla reports something similar. Both pronunciations of her name always felt natural, depending on the social context. However, contrary to Tristan, Leyla would start reflecting on the difference later in college. Spending less time at home and more time at school made her notice that the Spanish variant of her name was mostly restricted to an intimate circle of people. For Elena, awareness of the two pronunciations started early in elementary school and she noticed a personal preference for the Spanish variant. However, in high school, specifically in Spanish class, she would question why she had to tolerate two different pronunciations. These experiences show that, starting at an early age, the pronunciation of participant's names is one of the identity and linguistic resources they mobilize when navigating different sociocultural contexts and relationship spheres. Not only do they manage when to allow other people to pronounce their name in one way or another, but

they themselves also decide when to alternate between the variants to conform to a given context, as highlighted in (14)–(16).

(14)   ***Alberto***: If I meet somebody, you know, an English speaker, like let's just say he's not Latino, he's an American, and he is, "Oh, hey what's your name?," "[ˈæl.bɚt]." That's it. Like, oh in Spanish, it's like "[al.ˈβeɾ.to]"

(15)   ***Elena***: Cuando hay muchas personas que solomente solamente hablan inglés dejo que lo pronuncien como en inglés que luego dicen [ə.ˈleɪ.nə]. Ahh en la escuela también dicían [ə.ˈleɪ.nə]. A- así es donde lo dejaba pasar. También muchas veces cuando voy como al doctor, al dentista, ahí no importa, lo pronuncian como pueden

'when there are a lot of people that only speak English, I let them pronounce it in English, and they say [ə.ˈleɪ.nə]. Also at school, they used to say [ə.ˈleɪ.nə]. That is when I let it happen. Also, usually when I go to the doctor, the dentist, there it doesn't matter, they pronounce it however they can'

(16)   ***Estela***: If it's someone that I just met and I know that we're not going to like have any type of contact, it's fine [ˈstɛ.lə], it's, I mean, it'll go by. But if it's someone that I meet, and we're becoming friends and they can't pronounce my name wrong [*sic*], it starts bugging me. So yes, eventually, I'd be like, 'you gotta say my name right'

Choosing which variant to introduce themselves with and which to accept in a given context is, therefore, an important component of their sociolinguistic styles and repertoires. The addressee or audience (in Bell's (1984) terminology), their corresponding social positionings, and the social relationships between the participants in the interaction are all considered to determine the name variant to be used. In some cases, the choice of the name variant could work as convergent accommodation (Bell 1984; Giles et al. 1973) to the addressee's language and social context. Zentella (1997), in her ethnographic study of a lower-working-class Puerto Rican community in East Harlem, *El Bloque*, argues that the linguistic practices of multilingual, multidialectal, and multicultural speakers are constantly accommodating to and resisting the pressures of their community's social context. These processes are flexible and dynamic and are actively part of the construction of individuals' positionings. Furthermore, as Bucholtz (2016) puts it, these strategies should not be taken as "either simple linguistic accommodation or coerced cultural assimilation. Rather, all such strategies are acts of ethnoracial agency that claim the right to name oneself as one sees fit in a given context" (278).

As comments (14)–(16) show, speakers determine a languaging strategy in the situated context, exerting their agency as multilingual speakers and individuals positioned within their communities. They are the ones who decide when it is acceptable for someone else to pronounce their name in the English form, which is primarily conditioned by intimacy. The pronunciation becomes trivial in a transitory social context or interaction, like at a doctor's office or when no significant relationship exists. However, when considering a closer relationship, the expectation is that the other person will pronounce their name according to the bearer's preferences. Nonetheless, accepting an alternative realization of their names is conditioned by individual phonetic boundaries that define a mispronunciation, triggering a correction, as in (17)–(18).

(17)   ***Leyla***: yo sí claramente cuando lo escucho que es diferente, uhmm, les digo "ese no es mi nombre", like "mi nombre es [ˈlej.la], pero es con una <e> [e], not an <a> [ej]" (with 'not an <a>' [ej], she refers to the more common alternate form Layla)

'when I clearly notice that it is different, uhmm, I tell them "that's not my name", like "my name is [ˈlej.la], but with an <e> [e], not an <a> [ej]" '

(18)   **Tristan**: every time they mispronounce it I correct them, so I guess it would be important, but whether the only reason they would quote un-quote mispronounce it is because sometimes, like in English specially, there's a [ˈtɹɪs.tən] with like an <en>, or like sometimes it's spelled with an <i> or something. So that's when I would be like "no, it's [ˈtɾis.tan], with like uh, with <a>." That's the only time I would really correct them.

Wolf et al. (1996) note that mispronunciation is an individual concept that changes from speaker to speaker or for the same speaker depending on the circumstances. What counts as a mispronunciation for one name bearer is not necessarily taken as such for another, or some may view mispronunciation as the intentional frivolous distortions of their names and not failed attempts to pronounce it correctly based on lack of knowledge (419–420). Therefore, a combination of factors determines whether an individual takes the realization of their name as a mispronunciation that they need to address: the specific context where it occurs; the type of relationship they have with the person pronouncing their name; the phonetic deviation of the realization from any of the variants they accept, such as a different vowel or consonant phoneme; or the potential misnaming due to another name with a close phonetic realization. As such, the acceptability of the pronunciation of a name is not a static category but a dynamic negotiation influenced by the context of the interaction and personal stances. Internal and external factors, such as the individual's subjectivity, the social environment, and power dynamics, mediate the reaction to a mispronunciation. Consequently, responses to overt or covert aggressions based on someone's name may display conflicting sentiments.

(19)   **Alberto**: I used to work in this moving, it's like storage and I had another friend, he was [ro.ˈlan.do], [ro.ˈlan.do], they would called him [ɹoʊ.ˈlæn.doʊ], he wouldn't mind. Then one time he (their American boss) was like, "hey, Juan and Pablo, get off the truck," he was calling me and my friend Juan y Pablo. And I was like, "okay" (laughs). I got offended, but at the same time, I thought it was funny. He was just calling me Juan and Pablo. It was not my name (...) Maybe that was racist, racial, but like I found that funny, so I don't- I didn't really get offended

This short but forthright experience depicts how the racist practice discussed in Paredes (1993) (regarding the very name of Pablo) has persisted. Alberto's boss's action reproduces the benefits that power relationships (social and structural) bring to those on top. The intentional misnaming aggression that the boss exerted on Alberto and Rolando and the choice of two stereotypical Hispanic names depict the property of names as racialized signs (Smalls 2020). From a raciolinguistic (Rosa and Flores 2017) and raciosemiotic (Smalls 2020) perspective, Juan and Pablo are enregistered linguistic forms (Agha 2003), specifically racialized signs indexical of male individuals of ethnoracialized Hispanic communities in the U.S. In his retelling of the experience, Alberto shows a conflicting attitude towards his boss's intentional renaming. First, he expresses his discomfort and bafflement with the unexpected utterance (by saying 'okay' and laughing), subsequently stating that he felt offended. But next, he admits that it was funny and argues that, although it could have been a racist joke, it felt amusing at the time and, therefore, did not offend him. Without the right to condemn Alberto's reaction to the ethnoracial and demeaning aggression when it occurred, I argue that this casual encounter and its impromptu retelling prove that names are proxies for committing racial aggressions, which in day-to-day spheres are still unidentified as such. Despite the frequent justifications for these actions as being not racist but playful jokes, and as Hill (2008, 1993) describes in her discussion of Mock Spanish, these practices are part of strategies reproducing racial hegemonies (Bucholtz 2016, p. 278) and sustaining and propagating noxious language ideologies.

3.1. Summary

The accounts of these six Hispanic young adults in the U.S. whose names have at least two variants show that positionalities towards their given names are individual, complex, and dynamic.

As personal as every name is, the meaning of its pronunciation is subject to individual interpretation. The experiences here are a sample of the multiplicity of stances regarding name pronunciation preferences, which are motivated by personal stories and experiences. Some prefer the Spanish form and consider it the correct version of their name; others prefer the Spanish but do not attribute incorrectness to the English variant; others have no preference, and therefore consider neither variant incorrect.[6] Accordingly, what constitutes a name mispronunciation is more a subjective deliberation than an objective and fixed categorization, and it is best not to overgeneralize the conditions that determine when someone would judge the realization of their name as a mispronunciation. Likewise, a Hispanic name does not necessarily imply that the bearer has a strict preference for the Spanish variant or that the anglicized version would be considered a mistake. Accordingly, out of respect, and as Bucholtz (2016, pp. 286–87) recommends, the best course of action is to make the effort to correct your ignorance by asking people how they prefer their names to be pronounced and then addressing them as such.

As all participants shared, the Spanish pronunciation of their names is tightly linked to their family heritage, in which Spanish is a significant language within their intimate circle of relationships. For this reason, the selection of a name pronounceable in Spanish indicates an enduring practice that displays more than just the well-known cultural significance of names. It also reflects a thoughtful familial concern with potentially disrespecting Spanish-dominant family members, like first-generation parents or grandparents, by giving English-only names to the next generations, as their pronunciation may pose difficulty for them. Thus, a name pronounceable in Spanish helps cultivate and preserve intergenerational relationships.

Among the outer social circles, educational contexts usually constitute one of the most significant environments where the Spanish and English variants of a name first find themselves in conflict. As mentioned by the participants, their attention was drawn to the two distinct pronunciations of their names at different stages of their educational journeys, and they reflected on and questioned them in elementary school, middle, or high school, or even later in college. Spanish classes, in particular, become a meaningful space because they facilitate an awareness of the pronunciation of their names parallel to home practices and in contrast to how teachers in other courses pronounce them. The next section centers on Elena's life story, showing how the school context is a crucial social sphere in which (mis)pronunciations of a name contribute to the negotiation and construction of positionalities.

3.2. Case Study: A Discourse Analysis of a Narrative of (Mis)pronunciation Experiences

In the following analysis of Estela's retelling of her experience with the (mis)pronunciations of her name, three main discursive strategies are considered: (1) metalinguistic commentaries on pronunciation, (2) reported speech throughout her narration, and (3) the usage of *just*. The interaction among these strategies constructs a narrative of the power of a name and shows the self-positioning dynamics toward its pronunciations. Before the analysis, a brief overview of *just* and its multifaceted values in the discourse is given to justify its selection for this case study.

3.2.1. A brief Overview of *Just*

Just is considered a polysemous form frequently used by young people (Tagliamonte 2005) that can syntactically function as an adverb as well as a discourse particle. As the adverbial form, it modifies a constituent corresponding to its primary semantic function of restrictive meaning (Lee 1987) and, accordingly, to its propositional meaning. In this case, it can be alternated with other adverbs such as 'exactly,' 'only,' or 'simply.' As a discourse particle, its function is to modify the illocutionary force of the whole utterance and, therefore, it reflects the speaker's involvement in the discourse event (Aijmer 2002). From a discourse-pragmatic perspective, *just* "has an indexical relation to the speaker's attitudes or emotions towards the discourse event," differentiating it from other particles with parallel semantic meaning. In this way, its usage implicates an instruction to the interlocutor to "interpret the utterance as the expression of an attitude" (158).

Molina and Romano (2012) proposed the pragmatic meanings of *just* as branching out in opposite directions: as a downtoning and as an uptoning marker. The former refers to readings of restriction, minimization—also described as the depreciatory meaning (Lee 1987)— and possibility, whereas the latter corresponds to readings of exactitude, emphasis, and agreement. Nonetheless, these meanings are not necessarily fixed and exclusive, and they may overlap. As such, the meaning of *just* depends on the interaction between the linguistic form and the context in which it occurs (Aijmer 2002).

In this study, the use of *just* and its downtoning and uptoning functions are analyzed as a means for marking two of Estela's stances with respect to the pronunciation of the variants of her name. Stance is a widely studied concept in sociolinguistics (Barnes 2018; Bucholtz 2009; Bucholtz and Hall 2005; Du Bois 2007; Jaffe 2009; Kiesling 2005, 2009; Ochs 1993, 1996), and for the purpose of this analysis I follow Barnes (2018, p. 2) and define stance as the way a speaker conveys their attitude or position towards the content and context of their speech. The ensuing analysis shows how the multifaceted *just* serves to deploy (1) a stance of resignation towards the English variant [ˈstɛ.lə]—corresponding to *just* as a downtoner—and (2) a stance of annoyance towards the pronunciation of the Spanish variant [es.ˈte.la]—corresponding to *just* as an uptoner.

3.2.2. Estela's (Hi)story and Discursive Strategies at Play

The following excerpt, expanding on (6), provides a clear picture of Estela's positionality, which is necessary to understand her personal story. Each variant of her name indexes complementary sides of her bilingual, bicultural, and binominal context. In (20), the interaction among phonetic realizations, languages, and ethnoracial identities is made explicit:

(20)   it's like it was two sides of me because it was either [ˈstɛ.lə], or the- the correct way that my parents would pronounce it, like they'd be like "[es.ˈte.la], ven para acá" or something like that (...) It's like at school when they called me [ˈstɛ.lə], it would be fine, because it sounds, it sounds American or whatever, and being at home, and then my parents speaking Spanish to me but calling me [es.ˈte.la] and then it's- I felt like I was like a whole new other person. Like this person that doesn't speak English at all. And I'm just like completely like Latino, but then I go to school, and it's like, "Okay, I'm not Latino anymore, I'm American," or in some ways, but then, but then people messed up my name and then I was just like, "oh no, yeah, I am Latina."

This excerpt encapsulates the complexity of the social meanings each variant of Estela's name holds. Estela is explicit and specific about the correct pronunciation of her name, corresponding to the way her parents pronounce it, [es.ˈte.la]. This directly relates to her family heritage and stresses the significance that her parents' pronunciation and linguistic practices have for her. Thus the Spanish variant of her name is tightly connected to the presence of the Spanish language in the home as a crucial component of the construction of a Latino family and, consequently, a Latino identity. In contrast, the English variant relates to a more public sphere in Estela's life where English is used: the academic context, which for Estela invokes a more "American" identity. Furthermore, she clarifies that in this social sphere, probably where she encounters Spanish speakers less frequently, her name is not pronounced in the way she considers correct.

For Estela, the variants of her name index distinct but intertwined social categories that construct her positionality. Social contexts and their linguistic practices mediate Estela's self-adscriptions to and navigation between American and Latino identities. Despite seemingly describing a compartmentalized presence of English (at school) and Spanish (at home) and an apparent corresponding pronunciation of her name in each context, Estela notes that when English speakers struggle with ("messed up") the pronunciation of her name at school, it is a signal of her Latino ethnicity. The social meanings of the (mis)pronunciations of her name are not fixed, but are rather in constant flux depending on the interactional context.

As previously discussed, the interpretation of differences between name variants is personal rather than strictly phonetic. In Estela's case, the difference resides mainly in the first phoneme of her name:

(21)   my family and I pronounce it [es.ˈte.la], but like making emphasis on the
            <e> like [es.ˈte.la], you know? (...) When I meet people and they think
            that it's complicated just saying [e]. So they just- I'll just be like:
            – "It's not hard"
            but they see they just forget so that normally they just call me [ˈstɛ.lə]

Estela metalinguistically refers to an "emphasis on the <e>." This phonological reference does not correspond to the word's stress but the actual realization of the phoneme /e/ at the beginning of her name. When contrasted to the English version, in Spanish, Estela's name includes what is commonly referred to as an *epenthetic <e>*. Spanish disallows syllable-onset consonant clusters such as /st-/, /sk-/, /sp-/, and thus, various word pairs exist in English and Spanish where the presence of this /e/ is the salient distinguishing characteristic, e.g., *stress* [ˈstɹɛs] vs. *estrés* [es.ˈtɾes] or *Stella* [ˈstɛ.lə] vs. *Estela* [es.ˈte.la] (Hualde 2014, p. 64). The awareness of this phonemic rule gives Estela the tools to point out the /e/'s role and mark it as the differentiatior between variants, hence the correct pronunciation vs. a mispronunciation. Using an indefinite 'they' or 'people,' Estela generalizes the assignation of the mispronunciation to anyone or any social group of non-Spanish speakers, given the common tendency among them to use the English variant. Still, Estela expresses her bewilderment at the use of English by suggesting that the realization of the initial phoneme /e/ should not be complicated yet seems infrequently achieved.

Moreover, excerpt (21) contains various instances of *just*. In the first instance, "they think that it's complicated *just* saying [e]," *just* serves to emphasize the requirement for pronouncing the Spanish variant. She uses reported speech to convey her reproach and dissent that adding the /e/ is supposedly difficult: "I'll *just* be like: 'it's not hard.'" This second *just* in the quotative complements the uptoning function of the previous one by emphasizing that the expectation of adding an onset /e/ is usually not met. These *justs* convey a stance of annoyance with others' inability to pronounce her name in Spanish. Immediately after, Estela justifies the mispronouncers, saying that "they *just* forget" and use the English variant ("they *just* call me [ˈstɛ.lə]"). These two downtoning *justs* convey a sense of Estela's minimization of her discontent about others not being able to pronounce her name the way she prefers, which shows Estela taking a stance of resignation towards the English variant [ˈstɛ.lə]. Excerpt (21) shows that stance marking is complex and continously constructed throughout the discourse, which aligns with previous work (Barnes 2018; Bucholtz 2009; Eckert 2019; Jaffe 2009; Kiesling 2009) that highlights stancetaking as a multidimensional and dynamic process that takes part in the construction of social identities. In this excerpt, a tension between the two main stances is apparent: the stance of annoyance about the supposed difficulty that including an /e/ poses to non-Spanish speakers and the stance of being resigned to accept the English variant reside in almost parallel discourses. Shifting from one stance to another can occur as fast as moving from one sentence to the next, and the same linguistic device (e.g., *just*) can be mapped onto different stances. The following examples will continue to demonstrate how the interaction among these discursive strategies reflects the dynamicity of the stancetaking process.

In excerpt (22), Estela's conflicting attitudes are apparent when directly addressing how she negotiates the pronunciation of her name with others, usually failing to achieve an acceptable Spanish realization. Furthermore, this excerpt reveals the indexical meaning that has been ascribed, throughout Estela's lived experiences, to the "problematic" phoneme /e/.

(22)   I know a l- a lot of people that are just... like when they hear my name,
            they know that I'm I'm Latino, just because of the [e], like they'll be like:
            – "Oh, [eːs.ˈte.la]"
            and then I'll be like:
            – "Yeah"
            and they try and say it kind of like in a... I don't know, like in a Latino

way, like they they j̲u̲s̲t̲ make so much emphasis on the [e] now and I'm
j̲u̲s̲t̲ like:
– "No, it's not so much emphasis, it's j̲u̲s̲t̲ [e] like [es.ˈte.la]"
and they'll be like:
– "[eːːs.ˈte.la]"
and I'm like:
– "No, no."

Estela explicitly states that the Spanish /e/ phoneme acts as a sociolinguistic index, differentiating the two variants and marking her Latino ethnic identity. Her metalinguistic commentary demonstrates her phonetic awareness throughout the excerpts included here. For example, in (22), she displays both her linguistic knowledge and sociocultural sensitivity as she confirms that the pronunciation of her name plays a defining role in constructing and negotiating her ethnoracial identity and positionality. This excerpt portrays a simulated interaction between Estela and non-Spanish speakers with whom she is negotiating the pronunciation of her name. It contains a chain of uptoning *justs* with which Estela takes a stance of annoyance towards the failed attempts of others to pronounce [es.ˈte.la]. First, in the statement "*j̲u̲s̲t̲* because of the /e/," *just* specifies the indexicalization of /e/ to her Latino ethnoracial identity. Then, after other people presume she is Latino because of the onset /e/ in her name, they attempt to pronounce it but do so in an exaggerated fashion. The *just* in "they j̲u̲s̲t̲ make so much emphasis on the /e/" conveys Estela's unease caused by the unnecessary vowel elongation, which she perceives as a performative attempt to pronounce her name in a "Latino way." Although Estela may not frame her commentary as an affirmation that this practice carries an injurious intention, her remark shows that the Spanish variant is easily subjected to playful linguistic practices, as is the case of Mock Spanish (Hill 2008). The final two instances of *just*, in the quotative "I'm j̲u̲s̲t̲ like" and within the quote "it's j̲u̲s̲t̲ /e/ like [es.ˈte.la]," together with metalinguistic notes, indicate specifically how she expects her name to be pronounced and, at the same time, help construct a stance of annoyance towards others mispronouncing her name.

In Estela's personal story, the need to constantly negotiate the correct pronunciation made her disdain her name. This contempt reached a stage where she gave up on trying to educate others about the Spanish pronunciation, accepted the anglicized form that was easier and more common for non-Spanish speakers, and decided to use it herself when saying her name.

(23)   There was a while back, like I said, like, oh, when I was younger I hated
my name. So when they would ask me how to pronounce it, I did j̲u̲s̲t̲ say
[ˈstɛ.lə], like I wouldn't, I didn't bother on saying it right myself like I'm
j̲u̲s̲t̲ like:
– "No, it's just [ˈstɛ.lə]"
So I've grown up with like my best friends. Uhm. They j̲u̲s̲t̲ called me
[ˈstɛ.lə] until the point we were, we came to high school.

All four *justs* in (23) serve as downtoners, marking the stance of resignation toward the English variant and minimizing her discontent, which she would mask by limiting her explanations of the pronunciation of her name to "it's *just* [ˈstɛ.lə]." In this way, in spite of her preference for [es.ˈte.la], she opted for allowing people close to her, and even herself, to refer to her as [ˈstɛ.lə] until she got to high school. As mentioned in the previous section, the different stages of scholarly contexts are significant to an individual's reflection on the pronunciation of their name (Bucholtz 2016).

Estela describes a particular experience with a substitute teacher, who was the first person to call her attention to the significance of the pronunciation of her name and make her think critically about how she had been approaching the (mis)pronunciations. After this encounter, Estela decided to introduce herself as [es.ˈte.la] and to start correcting her friends' usual mispronunciation.

(24)   he (the substitute teacher) asked me how to pronounce my name and I
was j̲u̲s̲t̲ like:

– "oh, it's just [ˈstɛ.lə],"
and he noticed that my answer was kind of just like the same old- like it
looked like I had been repeating that the- like my entire life, which I had.
And I cou- I couldn't believe he noticed because he came on to me and he
was like:
– "What do you mean it's just [ˈstɛ.lə]? like if- is it not correct?"
And then I said:
–"Well, yeah, it's correct. It's just... they, I mean, I just changed it a little"
and he was like:
– "Well, how do you actually say it?"
and I was like:
– "Well, it's actually it's, um, it is [ˈstɛ.lə] it is just I make an emphasis on
the <e> uhm and so it's [eːs.ˈte.la]."
And then he was like:
– "Oh, ok, I see. Well, you shouldn't try and uhm change the way you
pronounce your name just because other people can't say, if I can say it
I'm pretty sure everyone else can."
And so then I just, I just kind of took that in mind

While recalling and retelling this experience, Estela maintained the conversation style she
had had throughout the interview, but she made use of more complex reported speech,
with longer utterances and more elaborated content than in previous excerpts. The contrast
between those familiar moments of mispronunciation in which others did not show interest
in Estela's name preference and this encounter with the teacher, which presented Estela
with a new perspective, is reflected in how Estela constructs her narrative. As noted in
previous excerpts, Estela usually does not specify whom she refers to when talking about
people who do not use the Spanish variant of her name. In these cases, Estela employs a
generic 'they' that, based on her situated sociocultural context, likely signals those who
do not speak Spanish. In addition, when Estela attributes reported speech to interactions
between her and the undifferentiated 'they,' it tends to be a simpler utterance simulating
their regular attempt to pronounce her name, the way other people mispronounce it,
Estela's demonstration of how to pronounce it, or her discontent with the result. On
the contrary, in the dialogue in (24), when Estela retells this significant experience with
someone who expressed a view that differed from what she had anticipated, she makes the
intervention detailed with elaborate explanations. This strategy allows for the significance
of this moment to be highlighted in her personal story.

When the substitute teacher asks Estela how to pronounce her name, she immediately
replies by giving her usual answer, the English variant, "oh, it's *just* [ˈstɛ.lə]." The *just*
here plays a crucial role, as the teacher distinguishes its downtoning force and shows it by
replicating Estela's response and asking her elaborate on it: "What do you mean it's *just*
[ˈstɛ.lə]? like if- is it not correct?" The teacher recognizes that Estela is not content with her
answer because of the presence of *just*. Comparing the same utterance without *just* marks
the difference:

(a)   oh, it's *just* [ˈstɛ.lə]
(b)   oh, it's [ˈstɛ.lə]

In (a), Estela's usage of *just* functions as a mitigating device for the conflicting attitudes
towards the pronunciation of her name. By employing *just*, she adopts a resigned stance
and provides the English variant as her answer, despite her preference for the Spanish form.
This contextual and discursive meaning is not present in (b) without *just*. From evaluating
this utterance in isolation, had Estela not employed *just*, it is possible that the teacher
would not have noticed her conflicting attitudes toward the pronunciation of her name.
Accordingly, it is the presence of *just* that conveys the emotional load that the variants of
her name hold for Estela. Throughout excerpt (24), instances of *just* demonstrate the way
Estela deemphasizes the strategies she had employed to make the pronunciation of her
name easier for non-Spanish speakers ("I *just* changed it a little"). Downplaying "changing
a little" as an insignificant move aligns with the resigned stance that Estela had taken until

that moment. However, this is recognized by the teacher, and his commentary encouraged Estela to reconsider the way she had been approaching the (mis)pronunciation of her name. In the interview, she mentioned that she then began correcting people's pronunciation of her name, starting with her closest friends, who had always called her by the English variant. Although she still accepts the English variant, especially in trivial contexts, such as when she is called to pick up an order at a restaurant, she now mainly goes by the Spanish variant of her name, which is the pronunciation she uses to introduce herself.

4. Conclusions

This study has shown that a name is not *just* a referential form, but a complex indexical form that belongs to the linguistic repertoire of each individual, which is mobilized accordingly to accomplish different social and communicative goals in the situated contexts that people navigate.

The accounts of these six individuals whose names are produced with Spanish and English phonology demonstrate that metalinguistic (phonological) awareness is a crucial tool in defining the various forms of a name and that such definitions are based on personal criteria. The differentiation of name variants, such as Spanish and English forms, is highly significant in the ongoing process of self-positioning across multiple social contexts. Although the participants in this study share experiences of using different pronunciations of their names and encountering mispronunciations, their responses to such situations vary and are shaped by negotiation strategies influenced by multiple concurrent social factors, including power dynamics present in the respective contexts. Furthermore, these responses are informed by the different social meanings associated with names and their variants.

Research framed within the third wave of sociolinguistics contends that the social meanings of language are indeterminate, and they necessitate an understanding of the social and discursive contexts in which they are employed (Eckert 2008; Hall-Lew et al. 2021; Moore 2021; Podesva 2007). This current study posits that names, particularly the variants of given names with multiple pronunciations, such as Hispanic names in the U.S., are indexical forms that encompass a range of social meanings. For some individuals, multiple variants of a name are positively evaluated as indexes of multilingualism and multiculturalism, while in other cases, a Spanish variant, or even a particular phoneme in the name, serves as a direct index of a Hispanic ethnoracial identity, which may be construed positively or negatively depending on one's experiences, as discussed in Parada (2020). From a cultural perspective, a Hispanic given name transcends a mere cross-generational ethnocultural naming practice, serving as a sociocultural strategy for maintaining and facilitating intergenerational and familial relationships. The decision to pronounce a given name in a particular manner represents an act of agency that constitutes a meaningful linguistic practice, through which individuals constantly negotiate and present their positionalities.

The case study in this work has offered an initial discursive analysis employing a qualitative approach that contributes to the scarce scholarship on Hispanic given names. This analysis sheds light on how speakers utilize different discursive strategies, which are in constant interplay, to display fluctuating stances with respect to the variants of their names. Further analyses of this kind will contribute to a more comprehensive understanding of how names, the bearers' pronunciation of a name, and the pronunciations of others constitute sociolinguistic variables that are key in the construction of social identities and positionalities.

> *"When I see the people who have had the experience of having been given a name from their family, which is one of the greatest gifts that a family can give you, it is the first gift that a child, usually, when they enter the Earth, receives from their family, it is usually informed by tradition and love and the hope and aspiration the family has for that child. It is something precious and sacred, and it is a part of their identity. And when I see people fighting for the right for that to be respected and treated in a dignified way, I applaud and salute that."*
>
> Kamala Harris in Noah (2020)

Notes

[1] Binominal here refers to having different pronunciations of a name, such as one corresponding to a minoritized language and the other to the hegemonic language in a given society.

[2] The concept of *languaging* refers to the language practices exerted by language users, and it recognizes the speakers' agency in the interactive process of meaning-making, or the "simultaneous process of continuously becoming ourselves and of our language practices, as we interact and make meaning of the world" (García and Wei 2014, pp. 8–9).

[3] Following Parada (2016) and Sue and Telles (2007), in this study, the label *Hispanic* is used to denote hispanized variants of names, which refers to the Spanish phonological realization of the name, rather than a Spanish/Latin etymology of the name. Because the Spanish language is central to this study, a clarification regarding the term *Hispanic* is in order. In this paper, *Hispanic* is employed as an overarching ethnonym referring to "being of Spanish-speaking background and trace their origin or descent from Mexico, Puerto Rico, Cuba, Central and South America, and other Spanish-speaking countries" (Lopez et al. 2022). Notably, the term *Latino* is solely employed when it is explicitly cited from a reference or when a participant used the term.

[4] As one of the reviewers pointed out, it is necessary to recognize that considering a name as ethnic or nonethnic is problematic as all names are ethnic. This distinction is determined by who is labeling them as ethnic or nonethnic. Throughout this article, the usage of ethnic refers to the Hispanic ethnonym unless specified otherwise. Nonethnic is only used when reporting that other studies have used this specific label to establish the difference between names.

[5] Excerpts transcription legend:
(text) - Clarification text
[IPA] - Phonetic transcription following IPA
<letter> - Speaker mentions the specific letter. If in a Spanish phrase/utterance, it is produced with Spanish phonology, and if in an English phrase, it is produced with English phonology.

[6] In this case, none of the participants indicated having a preference for the English variant of their names. Likewise, none of the participants considered the Spanish variant as the incorrect form.

References

Agha, Asif. 2003. The social life of cultural value. *Language & Communication* 23: 231–73. [CrossRef]

Aijmer, Karin. 2002. *English Discourse Particles: Evidence from a Corpus*. Amsterdam and Philadelphia: John Benjamins Publishing, vol. 10.

Ainiala, Terhi, and Jan-Ola Östman. 2017. Socio-onomastics and Pragmatics. In *Socio-Onomastics: The Pragmatics of Names*. Edited by Terhi Ainiala and Jan-Ola Östman. Amsterdam: John Benjamins Publishing Company, pp. 1–20. [CrossRef]

Babel, Anna, Kevin McGowan, and Paola Enríquez Duque. 2021. Niveles de percepción de las vocales en contacto: El caso de una variedad de español andino en Bolivia. In *Traspasando lo lingüístico: Factores esenciales en el contacto de lenguas*. Edited by Sara Gómez Seibane, María Sánchez Paraíso and Azucena Palacios. Volume 48 of *Lengua y Sociedad en el Mundo Hispánico*. Madrid and Frankfurt: Iberoamericana/Vervuert, pp. 119–36. [CrossRef]

Babel, Anna M. 2014. Stereotypes versus experience: Indexing regional identity in Bolivian Valley Spanish. *Journal of Sociolinguistics* 18: 604–33. [CrossRef]

Babel, Anna M. 2018. *Between the Andes and the Amazon: Language and Social Meaning in Bolivia*. Tucson: University of Arizona Press.

Bailey, Benjamin. 2008. Interactional Sociolinguistics. In *The International Encyclopedia of Communication*. Malden, MA: Wiley-Blackwell. [CrossRef]

Baird, Brandon O., Marcos Rohena-Madrazo, and Caroline Cating. 2018. Perceptions of lexically specific phonology switches on spanish-origin loanwords in american english. *American Speech* 93: 79–107. [CrossRef]

Barnes, Sonia. 2015. Perceptual salience and social categorization of contact features in asturian spanish. *Studies in Hispanic and Lusophone Linguistics* 8: 213–41. [CrossRef]

Barnes, Sonia. 2018. Copula variation in Asturian Spanish and the multidimensionality of stancetaking in interaction. *Journal of Sociolinguistics* 22: 29–54. [CrossRef]

Bell, Allan. 1984. Language style as audience design. *Language in Society* 13: 145–204. [CrossRef]

Bernard, H. Russell. 2006. *Research Methods in Anthropology: Qualitative and Quantitative Approaches*, 4th ed. New York: Rowman & Littlefield Publishers.

Bertrand, Marianne, and Sendhil Mullainathan. 2004. Are emily and greg more employable than lakisha and jamal? A field experiment on labor market discrimination. *American Economic Review* 94: 991–1013. [CrossRef]

Bucholtz, Mary. 2009. From Stance to Style: Gender, Interaction, and Indexicality in Mexican Immigrant Youth Slang. In *Stance: Sociolinguistic Perspectives*. Edited by Alexandra Jaffe. Oxford: Oxford University Press, pp. 146–70. [CrossRef]

Bucholtz, Mary. 2016. On being called out of one's name: Indexical bleaching as a technique of deracialization. In *Raciolinguistics: How Language Shapes Our Ideas about Race*. Edited by H. Samy Alim, John R. Rickford and Arnetha F. Ball. Oxford: Oxford University Press, pp. 273–90. [CrossRef]

Bucholtz, Mary, and Kira Hall. 2005. Identity and interaction: A sociocultural linguistic approach. *Discourse Studies* 7: 585–614. [CrossRef]

Carpusor, Adrian G., and William E. Loges. 2006. Rental discrimination and ethnicity in names1. *Journal of Applied Psychology* 36: 934–52. [CrossRef]

Chappell, Whitney. 2016. On the Social Perception of Intervocalic /s/ Voicing in Costa Rican Spanish. *Language Variation and Change* 28: 357–78. [CrossRef]

Charmaz, Kathy. 2006. *Constructing Grounded Theory: A Practical Guide Through Qualitative Analysis*. London: SAGE Publications.

Del Carpio, Leslie. 2022. Aspectual Differences in the Preterite and Imperfect in US Spanish: An Analysis of Their Use by Three Generation English Spanish Bilinguals. Ph.D. thesis, Arizona State University, Tempe, AZ, USA.

Doran, Amanda. 2001. Language Use and Identity in a Bilingual Community: Re-Examining the English of Mexican Americans. Ph.D. thesis, University of Texas at Austin, Austin, TX, USA.

Du Bois, John W. 2007. The stance triangle. In *Stancetaking in Discourse: Subjectivity, Evaluation, Interaction*. Edited by Robert Englebretson. Amsterdam: John Benjamins Publishing Company, pp. 139–82. [CrossRef]

Eckert, Penelope. 2008. Variation and the Indexical Field. *Journal of Sociolinguistics* 12: 453–76. [CrossRef]

Eckert, Penelope. 2019. The limits of meaning: Social indexicality, variation, and the cline of interiority. *Language* 95: 751–76. [CrossRef]

Edelman, Benjamin, Michael Luca, and Dan Svirsky. 2017. Racial discrimination in the sharing economy: Evidence from a field experiment. *American Economic Journal: Applied Economics* 9: 9–22. [CrossRef]

García, Ofelia, and Li Wei. 2014. *Translanguaging: Language, Bilingualism and Education*. New York: Palgrave Macmillan UK. [CrossRef]

Giles, Howard, Donald M. Taylor, and Richard Bourhis. 1973. Towards a theory of interpersonal accommodation through language: some Canadian data. *Language in Society* 2: 177–92. [CrossRef]

Glaser, Barney G., and Anselm L. Strauss. 2006. *The Discovery of Grounded Theory. Strategies for Qualitative Research*. New Brunswick and London: AldineTransaction.

Goldstein, Joshua R., and Guy Stecklov. 2016. From Patrick to John F.: Ethnic Names and Occupational Success in the Last Era of Mass Migration. *American Sociological Review* 81: 85–106. [CrossRef] [PubMed]

Guenther, Katja M. 2009. The politics of names: Rethinking the methodological and ethical significance of naming people, organizations, and places. *Qualitative Research* 9: 411–21. [CrossRef]

Gumperz, John J. 1982. *Discourse Strategies*. Number 1 in Studies in Interactional Sociolinguistics. Cambridge: Cambridge University Press.

Gumperz, John J. 2015. Interactional sociolinguistics: A personal perspective. In *The Handbook of Discourse Analysis*. Edited by Deborah Tannen, Heidi E. Hamilton and Deborah Shiffrin. Malden, MA: John Wiley & Sons, pp. 309–23. [CrossRef]

Gumperz, John J., and Jenny Cook-Gumperz. 1982. Language and the Communication of Social Identity. In *Language and Social Identity*. Edited by John J. Gumperz. Volume 2 of *Studies in Interactional Sociolinguistics*. Cambridge: Cambridge University Press, pp. 1–21. [CrossRef]

Hall-Lew, Lauren, Elizabeth Coppock, and Rebecca L. Starr. 2010. Indexing political persuasion: Variation in the Iraq vowels. *American Speech* 85: 91–102. [CrossRef]

Hall-Lew, Lauren, Emma Moore, and Robert J. Podesva. 2021. Social Meaning and Linguistic Variation: Theoretical Foundations. In *Social Meaning and Linguistic Variation: Theorizing the Third Wave*. Edited by Lauren Hall-Lew, Emma Moore and Robert J. Podesva. Cambridge: Cambridge University Press, pp. 1–4. [CrossRef]

Hill, Jane H. 1993. Hasta la vista, baby: Anglo Spanish in the American Southwest. *Critique of Anthropology* 13: 145–76. [CrossRef]

Hill, Jane H. 2008. *The Everyday Language of White Racism*. Chichester: Wiley-Blackwell.

Hough, Carole, ed. 2016. *The Oxford Handbook of Names and Naming*. New York: Oxford University Press.

Hualde, José Ignacio. 2014. *Los Sonidos Del Español*. New York: Cambridge University Press.

Jaffe, Alexandra. 2009. The Sociolinguistics of Stance. In *Stance: Sociolinguistic Perspectives*. Edited by Alexandra Jaffe. Oxford: Oxford University Press, pp. 3–28.

Kiesling, Scott F. 2005. Variation, stance and style. *English World-Wide* 26: 1–42. [CrossRef]

Kiesling, Scott F. 2009. Style as stance: Stance as the explanation for patterns of sociolinguistic variation. In *Stance: Sociolinguistic Perspectives*. Edited by Alexandra Jaffe. Oxford: Oxford University Press, pp. 171–94.

Kohli, Rita, and Daniel G. Solórzano. 2012. Teachers, please learn our names!: Racial microaggressions and the K-12 classroom. *Race Ethnicity and Education* 15: 441–62. [CrossRef]

Kostanski, Laura, and Guy Puzey. 2016. Trends in Onomastics: An Introduction. In *Names and Naming: People, Places, Perceptions and Power*. Edited by Guy Puzey and Laura Kostanski. Bristol, Buffalo and Toronto: Multilingual Matters, pp. xiii–xxiv.

Lavender, Abraham D. 1988. Hispanic Given Names In Five United States Cities: Onomastics As A Research Tool In Ethnic Identity. *Hispanic Journal of Behavioral Sciences* 10: 105–25. [CrossRef]

Lee, David. 1987. The semantics of just. *Journal of Pragmatics* 11: 377–98. [CrossRef]

Lipski, John M. 1976. Prejudice and Pronunciation. *American Speech* 51: 109–18. [CrossRef]

Lopez, Mark Hugo, Jens Manuel Krogstad, and Jeffrey S. Passel. 2022. Who is Hispanic. Pew Research Center. Available online: https://www.pewresearch.org/fact-tank/2022/09/15/who-is-hispanic/ (accessed on 15 September 2022).

Mendoza-Denton, Norma. 2008. *Homegirls: Language and Cultural Practice among Latina Youth Gangs*. Malden and Oxford: Blackwell Publishing.

Molina, Clara, and Manuela Romano. 2012. JUST revisited: Panchronic and contrastive insights. *International Journal of English Studies* 12: 17–36. [CrossRef]

Moore, Emma. 2021. The Social Meaning of Syntax. In *Social Meaning and Linguistic Variation: Theorizing the Third Wave*. Edited by Lauren Hall-Lew, Emma Moore and Robert J. Podesva. Cambridge: Cambridge University Press, pp. 54–79. [CrossRef]

Noah, Trevor. 2020. Kamala Harris—Coronavirus, Police Reform & Unifying America | The Daily Social Distancing Show. Youtube. Available online: https://www.youtube.com/watch?v=zl0S7v0BGYM (accessed on 15 April 2021).

Ochs, Elinor. 1993. Constructing Social Identity: A Language Socialization Perspective. *Research on Language and Social Interaction* 26: 287–306. [CrossRef]

Ochs, Elinor. 1996. Linguistic resources for socializing humanity. In *Rethinking Linguistic Relativity*. Edited by John J. Gumperz and Stephen C. Levinson. Cambridge: Cambridge University Press, pp. 407–37.

Parada, Maryann. 2013. Sibling Variation and Family Language Policy: The Role of Birth Order in the Spanish Proficiency and First Names of Second-Generation Latinos. *Journal of Language, Identity & Education* 12: 299–320. [CrossRef]

Parada, Maryann. 2016. Ethnolinguistic and Gender Aspects of Latino Naming in Chicago: Exploring Regional Variation. *Names: A Journal of Onomastics* 64: 19–35. [CrossRef]

Parada, Maryann. 2019. Linguistic and onomastic variation across the sibling spectrum in Latina/o Chicago. In *Family Language Policies and the Process of Intergenerational Transmission in the Context of Migration: Methodological and Notional Approaches*. Edited by Shahzaman Haque. Paris: Editions des Archives Contemporaines, pp. 101–21.

Parada, Maryann. 2020. A socio-onomastic study of Spanish receptive bilinguals. In *Spanish in the US: Variation and Attitudes*. Edited by Scott M. Alvord and Gregory L. Thompson. London: Routledge, pp. 56–73. [CrossRef]

Paredes, Américo. 1993. The Problem of Identity in a Changing Culture: Popular Expressions of Culture Conflict Along the Lower Rio Grande Border. In *Folklore and Culture on the Texas-Mexican Border*. Edited by Richard Bauman. Austin: University of Texas Press, pp. 19–48.

Pavlenko, Aneta. 2001. "In the world of the tradition, I was unimagined": Negotiation of identities in cross-cultural autobiographies. *International Journal of Bilingualism* 5: 317–44. [CrossRef]

Podesva, Robert J. 2007. Phonation type as a stylistic variable: The use of falsetto in constructing a persona. *Journal of Sociolinguistics* 11: 478–504. [CrossRef]

Regan, Brendan. 2022. *Guadalupe* or *Guadaloop*?: Place-Name Variation and Place Identity in Austin, Texas. *American Speech* 97: 441–82. [CrossRef]

Rosa, Jonathan. 2019. *Looking like a Language, Sounding like a Race: Raciolinguistic Ideologies and the Learning of Latinidad*. Oxf Studies in Anthropology of Language Series. Oxford: Oxford University Press.

Rosa, Jonathan, and Nelson Flores. 2017. Unsettling race and language: Toward a raciolinguistic perspective. *Language in Society* 46: 621–47. [CrossRef]

Rubin, Herber J., and Irene S. Rubin. 2011. *Qualitative Interview. The Art of Hearing Data*, 3rd ed. Thousand Oaks: Sage.

Rumbaut, Rubén G., and Kenji Ima. 1988. *The Adaptation of Southeast Asian Refugee Youth: A Comparative Study*. Washington, DC: U.S. Office of Refugee Resettlement. Available online: https://papers.ssrn.com/sol3/papers.cfm?abstract_id=1881903 (accessed on 15 April 2021).

Rymes, Betsy. 1996. Naming as social practice: The case of little creeper from diamond street. *Language in Society* 25: 237–60. [CrossRef]

Silva, David J., Sharon A. Peters, Fahad Ben Duhaish, Sok-Hun Kim, Yilmin Koo, Lana Marji, and Junsuk Park. 2011. Variation in the Iraq Vowels outside the Public Forum: The Indexing of Political Persuasion Reconsidered. *American Speech* 86: 179–91. [CrossRef]

Silva-Corvalán, Carmen. 1994. *Language Contact and Change: Spanish in Los Angeles*. Oxford: Clarendon Press.

Smalls, Krystal A. 2020. Race, SIGNS, and the body: Towards a theory of racial semiotics. In *The Oxford Handbook of Language and Race*. Edited by H. Samy Alim, Angela Reyes and Paul V. Kroskrity. New York: Oxford University Press, pp. 232–60. [CrossRef]

Souto-Manning, Mariana. 2007. Immigrant families and children (re)develop identities in a new context. *Early Childhood Education Journal* 34: 399–405. [CrossRef]

Sue, Cristina A., and Edward E. Telles. 2007. Assimilation and Gender in Naming. *American Journal of Sociology* 112: 1282–415. [CrossRef]

Tagliamonte, Sali. 2005. So who? Like how? Just what?: Discourse markers in the conversations of young Canadians. *Journal of Pragmatics* 37: 1896–915. [CrossRef]

Thompson, Riki. 2006. Bilingual, bicultural, and binominal identities: Personal name investment and the imagination in the lives of Korean Americans. *Journal of Language, Identity, and Education* 5: 179–208. [CrossRef]

Walker, Abby, Christina García, Yomi Cortés, and Kathryn Campbell-Kibler. 2014. Comparing social meanings across listener and speaker groups: The indexical field of spanish /s/. *Language Variation and Change* 26: 169–89. [CrossRef]

Wolf, George, Michele Bocquillon, Debbie de la Houssaye, Phyllis Kryzek, Clifton Meynard, and Lisbeth Philip. 1996. Pronouncing French Names in New Orleans. *Language in Society* 25: 407–26. [CrossRef]

Zelinsky, Wilbur. 1970. Cultural variation in personal name patterns in the Eastern United States. *Annals of the Association of American Geographers* 60: 743–69. [CrossRef]

Zentella, Ana Celia. 1997. *Growing up Bilingual: Puerto Rican Children in New York.* Oxford and Cambridge, MA: Wiley.

Zentella, Ana Celia. 1998. Multiple Codes, Multiple Identities - Puerto Rican Children in New York City. In *Kids Talk: Strategic Language Use in Later Childhood*. Edited by Susan M. Hoyle and Carolyn Temple Adger. New York and Oxford: Oxford University Press, pp. 95–112.

Zentella, Ana Celia. 2003. "José, can you see?" Latin@ Responses to Racist Discourse. In *Bilingual Gmes: Some Literary Investigations*. Edited by Doris Sommer. New Directions in Latino American Cultures. New York: Palgrave Macmillan, pp. 51–66. [CrossRef]

Article

Contextual Variables as Predictors of Verb Form: An Analysis of Gender and Stance in Peninsular Spanish Requests

Lori Czerwionka [1,*]**, Bruno Staszkiewicz** [1] **and Farzin Shamloo** [2]

[1] School of Languages and Cultures, Purdue University, West Lafayette, IN 47907, USA; bstaszki@purdue.edu
[2] Department of Psychological Sciences, Purdue University, West Lafayette, IN 47907, USA; farzin.shamloo@gmail.com
* Correspondence: czerwionka@purdue.edu

Abstract: The current experiment employs a variational pragmatics perspective to explore how the contextual variables of power, distance, and imposition jointly affect social groups' and individuals' choice of verb forms in requests in Madrid, Spain. Using a mixed-method approach to explore the requests of 111 Spanish speakers from Madrid, quantitative analyses determined the level of significance and hierarchical order of the predictor variables of power, distance, and imposition on verb form and also the distribution of verb forms by gender, with male and female participants exhibiting significant differences. Additionally, certain participants demonstrated decreased sensitivity to contextual factors, adopting more categorically indirect or direct request strategies. The examination of both gendered request patterns and the stances that single participants adopt through their verb-form selections contributes to our understanding of the social moves that are made by all speakers, not just those who fall within the gendered norms. The results highlight the different frames and social meanings attached to these forms at the micro- and macro-social levels, providing new insight into the complex relationship among linguistic variables, contextual factors, and social groups and individuals.

Keywords: Spanish; gender; stance; verb form; politeness; power; distance; imposition; requests

1. Introduction

Linguistic forms used in social interaction reflect and reconstruct identities, relationships, situations, and cultures. Whereas sociolinguistics has highlighted the social indexing of linguistic forms (Agha 2007; California Style Collective 1993; Eckert 2016), the field of variational pragmatics provides a complementary approach that addresses the relationships among social factors and language use in context (Acton 2021; Barron 2014). The field examines how particular uses of linguistic forms, in particular contexts and conversations, reflect and construct the social world.[1] When defining interactional contexts, previous research in pragmatics and politeness has often considered the role of the key contextual factors of power (e.g., peer vs. superior), distance (e.g., known vs. stranger), and imposition (e.g., small vs. big favor) (e.g., Brown and Levinson 1987; Félix-Brasdefer 2005, 2009; Hübscher et al. 2017), but analyses have seldom systematically examined the joint impact of contextual factors or the interrelationship among them. Therefore, questions remain about how the contextual variables of power, distance, and imposition jointly affect language use, how they shape the language choices of different speakers, and what those findings indicate about social meaning.

This investigation examines how contextual factors (i.e., power, distance, and imposition) condition verb forms used by men and women in Madrid, Spain; the variation in verb-form use by men and women; and how Peninsular Spanish verb forms contribute to individual speakers' stances when making requests in contexts that vary in power, distance, and imposition. The theoretical frame employed recognizes the connection between contextual variables and linguistic production as well as the relationship between individuals'

stances and macro-level sociolinguistic trends, aligning with variational pragmatics (Acton 2021; Barron 2014), third-wave sociolinguistic research (e.g., Bell 2016; Eckert 2012, 2016; Eckert and McConnell-Ginet 1999), and Terkourafi's (2005, 2015) frame-based approach to politeness.

The contribution of this paper is threefold. First, as one of the few studies to quantitatively investigate the effect of the interaction of power, distance, and imposition on linguistic forms, it demonstrates that contextual variables form a specific hierarchical relationship that can be conditioned by social group. Thus, the topic of variation is addressed, but unlike most prior linguistics research, this investigation highlights the variable perception of context and its effect on language. Second, this investigation offers new insight into language and gender in the Peninsular Spanish context by examining differences in the effect of contextual variables on men's and women's language, as well as differences in verb-form selection by men and women from Madrid. Third, the analyses address individual-level data that demonstrate the agency of individuals in selecting verb forms that reflect their preferred stances in particular interactional contexts. The results identify nuanced social meanings at group and individual levels that emerge from the study of contextual variables, linguistic forms, and individual stances as dynamic aspects of interaction.

2. Literature Review

2.1. Pragmatic Variation and the Social Indexing of Language

Pragmatics research has often addressed linguistic variation across situations but paid little attention to how situational variation intersects with social variation. With the distinction of the field of variational pragmatics, which initially intended to integrate pragmatics and dialectology (Barron 2005), more researchers have begun to consider the relationship between pragmatic variables and social factors such as social class, ethnicity, age, education, religion, and gender (Barron and Schneider 2009). Demonstrating the need for a pragmatic approach to social variation, D'Arcy's (2017) research highlighted the multiple functions of *like* in interaction, which do not all behave similarly in their social uses and distributions. Her analyses countered the notion that women employ *like* more than men by showing that different pragmatic uses are favored by men and women. In general, the variational pragmatics approach examines how particular uses of linguistic forms in particular contexts and conversations reflect and construct the social world. This highlights the context-dependency of language that influences language variation and reflects the individual identities of language users (Acton 2021). Additionally, variational pragmatics foregrounds the questions addressed in this article about the varying effects of situations on different speakers, as evident in their linguistic production.

The theoretical frame of this investigation blends concepts and understandings from pragmatics and variationism. As Acton (2021) argues, pragmatics and third-wave variationism share many underlying principles and provide complementary approaches to understanding social meaning. In addition to adopting a pragmatics perspective of language and context, the current research recognizes the indexical relationship between linguistic variables (e.g., verb forms) and social factors (e.g., gender), a relationship that serves to (re)create social structures. Indexicality (Agha 2003) and Silverstein's (2003) concept of indexical order expose the relationship between local uses of linguistic features in interaction and their link to styles, performances, and macro-social identities (e.g., California Style Collective 1993; Jaffe 2016; Silverstein 2003).

Concerning specifically the relationship between linguistic forms and macro-social identity categories, stance is an important mediator (Ochs 1996). It has been shown that 'sociolinguistic variants are initially associated with interactional stances and these stances become in turn associated with a social group meaning' (Kiesling 2009, p. 172) (see also Davies and Harré 1990; Du Bois 2007; Eckert 2016; Jaffe 2016). As defined by Du Bois (2007), '[s]tance is a public act by a social actor, achieved dialogically through overt communicative means, of simultaneously evaluating objects, positioning subjects (self and others), and

aligning with other subjects, with respect to any salient dimension of the sociocultural field' (p. 163). As an example of stance being a mediator between linguistic variables and social group identity, Kiesling (2004) argues that the use of the address term *dude* is first associated with a stance of 'cool solidarity,' which then is associated with masculinity (i.e., a gendered social identity) (p. 282). He states that 'if a linguistic item co-occurs frequently in the speech of a particular person or kind of person, that linguistic item will be taken to index that group' (Kiesling 2009, p. 177).[2] In this paper, the analysis of stance is limited to how verb forms position speakers and hearers with respect to each other and, also, with respect to the request being made in the interaction.

2.2. Politeness: Linguistic Expressions, Speaker, and Context

One area of study within pragmatics that has focused on social meaning is that of politeness. When being polite, the purpose is to maintain relationships and facilitate interactions by behaving in ways that are expected by other interlocutors (e.g., Kádár 2019; Spencer-Oatey 2000; Watts 2003; Terkourafi 2015). To be polite, speakers use conventionalized language that is processed as polite via generalized conversational implicatures (Terkourafi 2005, 2015).[3] Conventionalized linguistic forms are socially constituted via the three-way relationship among 'an expression, a context, and a speaker' (Terkourafi 2015, p. 15). Whereas sociolinguistic theories have explained these connections via stance, indexicality, and situations (e.g., Ochs 1996), Terkourafi (2005, 2015) relies on the concept of a frame, which is a cognitive construct based on prior experience that includes linguistic and non-linguistic information and that can be recalled from memory to aid in interaction (e.g., Ensink and Sauer 2003; Fillmore 1975; Schank and Abelson 1975). For Terkourafi (2015), frames are combinations of an expression with a minimal context that includes extra-linguistic variables such as gender, age, relationship, setting, etc. Speakers' frames provide predictive capabilities about language use in particular contexts, and they reflect individual and social group experiences, meaning that frames may vary across speakers and time (Terkourafi 2005, p. 15). Terkourafi's (2005, 2015) approach to politeness aligns with the social constructivist theories in Section 2.1 in that any aspect of the frame can be understood as socially constituted. It also offers the theoretical explanation of a cognitive frame as a resource in interaction.

Section 2.2.1, Section 2.2.2 ans Section 2.2.3 introduce aspects of the three-way relationship among the linguistic expression, speaker, and context that are relevant to the current investigation. Section 2.2.1 addresses verb forms in requests, Section 2.2.2 provides prior research related to the extra-linguistic variable of gender, and Section 2.2.3 introduces the contextual variables that are explored in this study.

2.2.1. Linguistic Expression: Requests and Verb Forms

Requests are speech acts that attempt to make a hearer do something (Austin [1962] 1975; Searle 1976). Requests can be realized with direct, conventionally indirect, or non-conventionally indirect linguistic strategies (Blum-Kulka 1987; Brown and Levinson 1987). The requests that are most direct in languages like Spanish are accomplished with imperative verbs (Blum-Kulka et al. 1989), and the imperative is more commonly found in Spanish when the speaker has more social power compared to the interlocutor (Hernández-Flores 2004). While the imperative is not uncommon in Peninsular Spanish (Ballesteros 2001; Lorenzo-Dus and Bou-Franch 2003), conventionalized indirect requests (e.g., *¿Puedes pasarme la sal?* 'Can you pass me the salt?; *¿Podrías pasarme la sal?* 'Could you pass me the salt?') seem to be the most frequent type (Márquez-Reiter 2002, 2003; Márquez-Reiter et al. 2005). In fact, the frequent use of conventionalized indirect requests is common across the Spanish-speaking world, with evidence from Uruguayan Spanish (Márquez-Reiter 2000, 2002, 2003), Mexican Spanish (Félix-Brasdefer 2005, 2009), and Costa Rican Spanish (Félix-Brasdefer 2009). While the directness or indirectness of a request is important, the verb form is a main indicator of the illocutionary force (Félix-Brasdefer 2005), and it communicates detailed meaning beyond the direct/indirect strategy distinction.

The verbal system in Spanish, with a variety of tenses and moods, provides rich semantic options for making requests and also nuanced pragmatic meanings. In general, the deictic center of the verb is commonly used to convey different degrees of politeness in requests (Koike 1989). For example, the present tense communicates a more-direct illocutionary force, whereas future tense or conditional verb forms are less imposing (Koike 1989). Conditional, past, and future forms are temporally displaced from the current moment and thus present a temporal disorder that is linked to the expression of politeness (Fleischman 1989; Koike 1989). Furthermore, verb forms such as imperfect and future forms are representative of an irrealis mood, which has implications for politeness and can 'modulate the perceived assertiveness' (Fleischman 1989, p. 8). Not only are irrealis or non-factual forms associated with more politeness due to their temporal distance from the action as compared to realis or factual forms, but politeness becomes encoded in the verb forms (Chodorowska-Pilch 1998, 2004).

Research on verb forms in varying contexts has highlighted the connections among verb forms, mitigation, and politeness. In a study on Mexican Spanish, conditional and imperfect verb forms were used as syntactic mitigators in situations that displayed a distant relationship between interlocutors (Félix-Brasdefer 2009). Modal verbs such as *poder* 'to be able to' (e.g., Can you pass me the salt? vs. Could you pass me the salt?) were also found as mitigators when addressing a distant person, expressing a higher level of deference when used in the conditional or imperfect forms compared to present tense (Félix-Brasdefer 2005). Modal verbs mitigate by suspending to some degree the assumption that the hearer is able and willing to perform the requested action (Brown and Levinson 1987; Briz Gómez 2004, pp. 72–76)—a concept that refers to the preparatory condition of directives (Searle 1976) and relates to deictic distancing.

Based on the prior literature, and following primarily Chodorowska-Pilch (1998, 2004), the conventional politeness of verb forms in Spanish can be understood as a continuum: subjunctive, modal conditional, conditional, imperfect/future, modal present, present, imperative (Figure 1).[4] The high end of the continuum represents verb forms with an irrealis mood and a deictic distance from the action. The low end of the continuum represents verb forms with a realis mood that maintain the present deictic reference. The forms closest to the high end have been described as less direct and less assertive; those near the low end have been described as more direct and more assertive (Koike 1989; Fleischman 1989). All verb forms in Figure 1 are commonly used by speakers of Spanish. Excerpts (1)–(8) from the current data set demonstrate the use of each verb form in a request.[5] As can be noted in the excerpts, a multitude of other linguistic resources are used in conjunction with verbs to make requests. While the co-occurrence of linguistic resources can be important in communicating social meaning (e.g., California Style Collective 1993; Eckert 2012), the current focus is on the verb because the verb is a main indicator of the illocutionary force (Félix-Brasdefer 2005) and the directness or assertiveness of a given utterance (Koike 1989; Fleischman 1989).
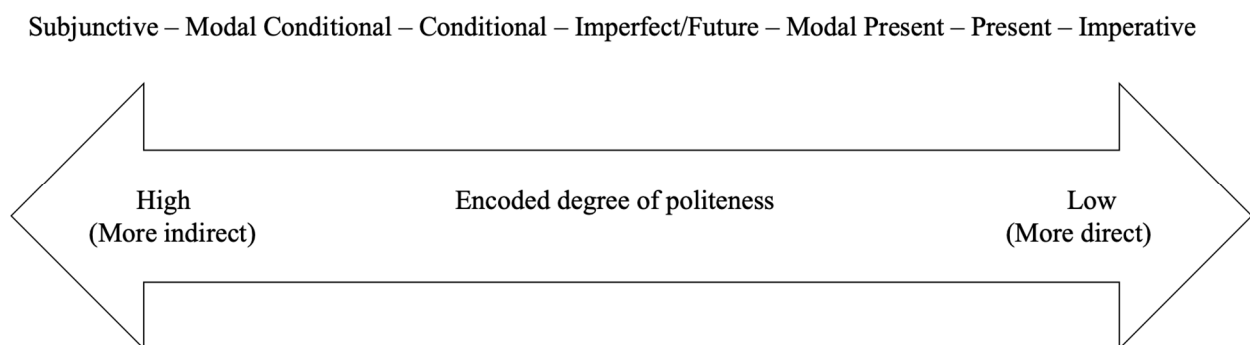
Subjunctive – Modal Conditional – Conditional – Imperfect/Future – Modal Present – Present – Imperative

High
(More indirect)   Encoded degree of politeness   Low
(More direct)

Figure 1. Verbal continuum for Spanish verbs in requests.

(1) Subjunctive.
*Perdone, necesito un manual de referencia para realizar mi trabajo, si **pudiese** dejármelo únicamente durante este fin de semana se lo agradecería.*
Excuse me, I need a reference manual to do my project, if you could lend it to me just for this weekend I would appreciate it.

(2) Modal Conditional.
*Mirian ¿me **podrías dejar** una grapadora para el trabajo, por favor?*
Mirian, could you lend me a stapler for the project, please?

(3) Conditional.
*Perdone, ¿**tendría** todavía alguna hoja de las que entregó el último día?*
Excuse me, would you still have any of the sheets you handed out the last day?

(4) Imperfect.
*Hola Rosali, oye, ¿cuál **era** el autor del artículo del otro día?*
Hi Rosali, listen, which was the author of the article the other day?

(5) Future.
*Rosa, ¿no **tendrás** un boli de sobra? que no encuentro el mío.*
Rosa, do you happen to have a spare pen? It's that I can't find mine.

(6) Modal Present.
*Luis, ¿me **puedes ayudar** con el examen?*
Luis, can you help me with the exam?

(7) Present.
*Sergio, ¿me **das** un folio?*
Sergio, will you give me a sheet of paper?

(8) Imperative.
***Déjame** un folio, tío.*
Give me a sheet of paper, man.

The verbal continuum aligns with the notion that through continual societal sanctioning of the use of certain forms in certain situations, a form becomes conventionalized (Chodorowska-Pilch 1998; Terkourafi 2005, 2015). Thus, certain forms come to index a degree of politeness (Agha 2007; Pizziconi 2011). Whereas the described pragmatic meanings of verb forms within a spectrum of directness emerge from a traditional pragmatics perspective in which these systems of meaning are available for use by all speakers, the current investigation aims to provide nuance to the question of who uses these forms and under what circumstances.

2.2.2. Speaker: The Variable of Gender

Gender, like other social identities, affects and reflects one's social experience and one's use of linguistic resources. From a performance-based perspective, identities are malleable, and they are socially constructed via language and interaction through 'the social positioning of the self and other' (e.g., Bucholtz and Hall 2005, p. 586; Gee 2014). In practice, one can demonstrate or create group membership through linguistic resources that are associated with that group; this relationship to group membership is not fixed but rather it varies across time, context, conversational topic, etc. (e.g., California Style Collective 1993; Eckert and McConnell-Ginet 1992).

The language of different gendered groups has been of interest for at least the last 50 years in linguistics research. Research on language and gender has alluded to and demonstrated in some cases correlations between gender and certain linguistic characteristics.[6] Lakoff (1975) proposed that women rely more on tentative language (e.g., hedges, qualifiers, disclaimers, tag questions, intensifiers), which 'softens an assertion' (Leaper and Robnett 2011, p. 137). Confirming these claims, Leaper and Robnett (2011) conducted a metastudy that included 29 prior investigations of tentative language. They found a small but significant difference between men and women in the use of tag questions, markers of uncertainty, and intensifiers. The effect size was larger for longer interactions (vs. shorter), undergraduates (vs. other adults), group conversations (vs. dyads), and research

settings (vs. others). The different effect sizes indicate that the construction of gendered identities may be more important during certain times of life (e.g., university) or in certain types of conversations (e.g., group conversations), demonstrating the malleability of one's linguistically marked belonging to a gendered group.

The prior findings have been interpreted as functioning to facilitate interaction, express interpersonal sensitivity, and create a supportive and affective identity (Grenoble 1999; Hancock and Rubin 2015; Holmes 1995; Leaper and Robnett 2011; Levey 2003; Wright and Hosman 1983). However, Eckert (1989) and Uchida (1992) argued that power is the underlying sociological construct associated with gender, resulting from historical and current norms of limited participation in society for certain gendered identities, including those of women. This argument aligns with Lakoff's (1975) explanation that women's tentative language may be associated with their subordinate position in society. We draw attention to the previous findings about tentative language and gender because, similar to the linguistic resources analyzed as tentative language, verb forms with a deictic distancing from the action also serve a softening function (see Section 2.2.1) and, thus, may be differentially associated with gendered performances. At the same time, the current analyses address macro-level trends and individual tendencies considering that 'gender does not have a uniform effect on linguistic behavior for the community as a whole, across variables, or for that matter for any individual' (Eckert 1989, p. 253).

In the Spanish-speaking world, the relationship between gender and language has been understudied, and few investigations of speech acts have considered gender. In an analysis of request strategies and gender in Argentine Spanish, no differences in the use of direct questions, assertions, imperative use, or conventional indirect requests were found between men and women (Yates 2015). In another study, Lorenzo-Dus and Bou-Franch (2003) found that across various speech acts, Spanish women used alerters[7] and acts of thanking more often than men, while using fewer supportive moves (utterances that provide additional explanation, justification, or support for a given speech act). In requests specifically, they found a preference for direct requests by both Spanish women and men, indicating that 'cultural behavior may be a stronger factor than gender in this particular aspect of the formulation of requests' (p. 9). They also found that women used more-direct strategies (i.e., want statements) than men in situations with asymmetrical power and distance (+power, +distance), where 30.76% of the women's responses and 15.78% of the men's were direct request strategies. These results provide mixed findings on the type of linguistic resources preferred by women versus men when performing speech acts in Spanish, indicating a need to further explore the relationship between language and gender in speech acts in the Peninsular Spanish context and in general.

2.2.3. Contextual Variables

The meaning of language is context-dependent. While various aspects of context are relevant to understanding language and social meaning, politeness research has confirmed the importance of three widely examined contextual variables: power, distance, and imposition (e.g., Brown and Gilman [1960] 1972; Brown and Levinson 1987; Brown et al. 2014; Czerwionka 2012, 2014; Félix-Brasdefer 2005, 2008, 2009; Márquez-Reiter 2002, 2003; Márquez-Reiter et al. 2005). According to Brown and Levinson (1987), power represents asymmetrical hierarchies that can be found in a society, distance is a symmetrical relationship that ranges from close to distant, and the situational variable of imposition is the degree to which an act is considered to interfere with the hearer's wants or desires to be self-determined (Brown and Levinson 1987). Each of these contextual variables should be understood to encompass a range of interpretations. For example, Spencer-Oatey (1996) reported that distance has been described using terms such as solidarity, familiarity, closeness, and relational intimacy. Thus, the meaning of power, distance, imposition, or any other contextual variable must be scrutinized.

From a theoretical perspective, it is also important to note that the interpretations of power, distance, and imposition are socially constructed and may vary among different

cultures or groups of speakers (Brown and Levinson 1987). For example, with respect to power, one can imagine distinct power differentials in an employee–boss relationship depending on the cultural, societal, or individual interpretations. With respect to imposition, particular cultures or individuals may understand the weight of a specific request or favor differently, with the same request being perceived as high imposition by some and low imposition by others. Referring to the variability in the interpretation of contextual factors and justifying the current research agenda, Spencer-Oatey and Žegarac (2017) indicated that 'little (im)politeness research has grappled with such issues' (p. 134).

While the impact or understanding of these contextual factors is variable, so too are the relationships among them. That is, the relative importance of these variables may change at different moments or from one community to another—a finding that has recently begun to emerge in empirical pragmatic research. Tamaoka et al. (2010), for example, examined the hierarchical organization of power, distance, and gender on perceived politeness in Japanese and Korean speakers. Using a decision tree analysis, the results demonstrated that Japanese participants' responses were predicted by the distance between the interlocutors, followed by the power relationship. In contrast, the Korean participants' responses were influenced by power differentials, while distance was not a significant predictor. The difference across groups suggests that the way in which Japanese and Korean speakers order the importance of these contextual variables in interactions is distinct, with Japanese speakers being most influenced by distance and Korean speakers most influenced by power.

The investigation by Tamaoka et al. (2010) supports the notion that contextual variables do not necessarily have the same social meaning across societies. The study also demonstrates that predicted hierarchical models can provide insight into how contextual factors shape behaviors and how they differentially constrain the language use of members of different social groups. Further investigation of this underexplored topic can provide insight into the socially determined understanding of the relationships among context, linguistic behaviors, and the social meaning of language.

2.3. Research Questions

This study explores the effect of power, distance, and imposition on language in Spanish, focusing on verb forms and considering the variable of gender. The primary goal is to examine how contextual factors condition the linguistic forms used by women and men in Madrid, Spain. We also examine how verb selection varies by gender and how individual-level linguistic choices portray different stances. The following research questions guide the analyses:

1. Do the effect and hierarchical order of the contextual variables of power, distance, and imposition as predictor variables of verb-form selection in requests differ between women and men?
2. Does the overall distribution of verb forms used in requests differ between women and men?
3. How do individuals' stances in requests vary in contexts that involve differing degrees of power, distance, and imposition?

Based on prior research, we hypothesize that power and distance will have a greater effect than imposition on verb selection and a higher placement in terms of the hierarchical structure (e.g., Ballesteros 2002; Félix-Brasdefer 2005, 2009; Le Pair 1996). No specific hypothesis was made about the hierarchical order across genders because of the lack of related prior research. For the second research question, there are two competing hypotheses. The first is that women will use less-direct verb forms, aligning with the tendency for women to use more tentative linguistic forms (e.g., Leaper and Robnett 2011). The second is that women will use more-direct verb forms, aligning with Lorenzo-Dus and Bou-Franch's (2003) finding that women in Spain were more likely to use direct speech acts, at least in certain contexts. Considering the third research question, we expect that variation exists, confirming that speakers have the agency to select linguistic forms that

serve to position them in ways that align with their ideologies, expectations, and identities (e.g., Du Bois 2007; Eckert 2016; Jaffe 2016).

3. Materials and Methods

This investigation relied on data collected using a discourse completion task (DCT) to prompt the production of requests in contexts with two levels of power, distance, and imposition.[8] Quantitative and qualitative analyses were conducted to understand the effect of the contextual variables of power, distance, and imposition on verb selection in requests produced by men and women in Madrid, Spain; verb selection across genders; and individuals' stances communicated with verb forms.

3.1. Participants

A total of 111 native speakers of Peninsular Spanish from the Madrid region of Spain (i.e., Community of Madrid) participated in the study after providing their consent (70 women and 41 men; Age: $M = 26.6$, $SD = 6.60$). Gender was elicited using a multiple-choice question with the options of male, female, or other, where the option of other included a fill-in-the-blank box. All participants had graduated from or were enrolled in a Spanish university degree program. The participants likely spoke the same or very similar varieties of Spanish, given the homogeneity of their geographic location and educational background. Additionally, the university-level educational experience of all participants was ideal for the current research because it ensured familiarity with the situations employed in the experiment. All reported participant names are pseudonyms.

3.2. Materials and Procedure

A DCT is an elicitation procedure where participants read a contextualized prompt and provide a response to it. DCTs have been widely used for pragmatics research because they allow for the collection of large amounts of data in controlled situations. They have also been criticized as they provide metapragmatic data rather than naturally occurring data (Golato 2003). While DCT data do not represent all aspects of a naturally occurring interaction (Félix-Brasdefer 2007; Golato 2003), DCT and naturally occurring data share more similarities when considering head acts (i.e., the part that communicates the speech act) than other aspects of the interactions (see Márquez-Reiter and Placencia's (2005) discussion, p. 226). For requests, Bataller and Shively (2011) found similarities across data types in request openings and a similar variability in request type. Given the current focus on request head acts and the need for a large amount of controlled data for statistical analysis, DCT data were ideal.

The DCT used in the current investigation included 16 situations with controlled levels of power, distance, and imposition (Situation 11, Table 1, Appendix A). All were set in an academic environment in which a student interacted with another student (−power) who was a well-known friend (−distance) or a not-very-well-known classmate (+distance) and in which a student interacted with a professor (+power) who was either well known (−distance) or not very well known to the student (+distance). The situations required that a student request something of the other interlocutor; half of the situations prompted a low-imposition request (−imposition) and half prompted a high-imposition request (+imposition). The 16 situations were balanced considering two levels of each variable: power (asymmetrical [+power], symmetrical [−power]), distance (distant [+distance], closeness [−distance]), and imposition (high [+imposition], low [−imposition]). Half of the situations were about requesting a good and half were about requesting a service (i.e., type of request (Brown and Levinson 1987)), although the effect of the type of request was not of interest in the current study. The order of the 16 situations was computer-randomized.

Situation 11. −power, +distance, and +imposition.

Tienes un examen la próxima semana y has faltado a algunas clases. [Nombre de un estudiante al que no conoces mucho] se encuentra a tu lado. Quieres que te preste sus apuntes. ¿Qué le dices?

You have an exam next week and you have missed some classes. [Name of a student that you don't know very well] is next to you. You want the student to lend you her/his notes for the exam. What do you say?

Table 1. Distribution of the situations.

Situation	Power	Distance	Imposition	Type of Request
1	Asymmetrical	Distant	High	Service
2	Asymmetrical	Distant	Low	Service
3	Asymmetrical	Distant	High	Good
4	Asymmetrical	Distant	Low	Good
5	Asymmetrical	Close	High	Service
6	Asymmetrical	Close	Low	Service
7	Asymmetrical	Close	High	Good
8	Asymmetrical	Close	Low	Good
9	Symmetrical	Distant	High	Service
10	Symmetrical	Distant	Low	Service
11	Symmetrical	Distant	High	Good
12	Symmetrical	Distant	Low	Good
13	Symmetrical	Close	High	Service
14	Symmetrical	Close	Low	Service
15	Symmetrical	Close	High	Good
16	Symmetrical	Close	Low	Good

During the task, participants were first asked to provide the names of two professors and two classmates, one who was well-known and another who was not. These names were automatically embedded in the preambles seen by the individual participant to signal the social power and distance between interlocutors. Referring to the names of classmates and professors whom participants knew resulted in greater authenticity. Then, participants read each situation and wrote what they would say in response. Written responses facilitated the data analysis, and they were sufficient for the purposes of the current study. Following the DCT, participants also completed a background questionnaire that requested information about gender, age, place of origin (city), native language, and other spoken languages. All tasks were conducted online using Qualtrics (2018).

In the creation phase of the DCT, a norming procedure was conducted to assess the perception of the contextual variables. Twelve native speakers of Spanish, who were university students in Madrid and who were not participants in the main study, provided the names of two professors and two classmates to be embedded in the preambles, as in the main experiment. They evaluated the imposition, power, and distance in each situation using a 9-point Likert scale. For imposition, the endpoints of the scale were *pedir poco* 'undemanding' or *pedir mucho* 'over-demanding' (Figure 2). For power, the scale ranged from *iguales* 'equals' to *desiguales* 'unequals.' They rated distance in terms of closeness, from *relación cercana* 'close relationship' to *relación distante* 'distant relationship'.

The perception of each variable was analyzed separately, using a hierarchical regression to control for scoring variability among raters (modeled as varying intercepts), other factors (e.g., different power levels when scoring distance), and the interactions. The norming analyses showed a difference of 5.07 ($p < 0.001$) between low- and high-imposition situations; a difference of 4.35 ($p < 0.001$) between low- and high-power situations; and a difference of 5.44 ($p < 0.001$) between low- and high-distance situations. The results suggest that the Spanish native speakers perceived the situations to describe the intended distinctions of power, distance, and imposition.
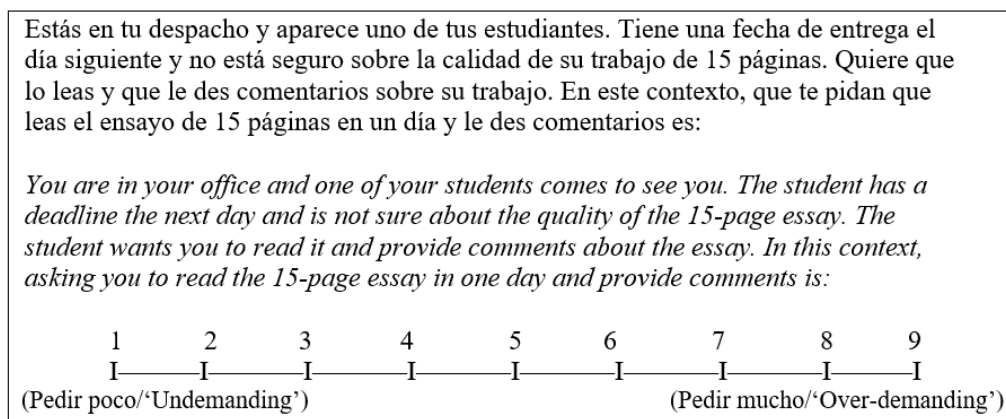
Estás en tu despacho y aparece uno de tus estudiantes. Tiene una fecha de entrega el día siguiente y no está seguro sobre la calidad de su trabajo de 15 páginas. Quiere que lo leas y que le des comentarios sobre su trabajo. En este contexto, que te pidan que leas el ensayo de 15 páginas en un día y le des comentarios es:

You are in your office and one of your students comes to see you. The student has a deadline the next day and is not sure about the quality of the 15-page essay. The student wants you to read it and provide comments about the essay. In this context, asking you to read the 15-page essay in one day and provide comments is:

| 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 |

(Pedir poco/'Undemanding')   (Pedir mucho/'Over-demanding')

Figure 2. Imposition norming-task example.

3.3. Data Coding and Analyses

All responses were coded according to the verb form in the head act of the request (i.e., the part that communicates the speech act). Considering the prior literature on verbs and politeness (Chodorowska-Pilch 1998; Félix-Brasdefer 2005; Koike 1989) and the current data, the following verb forms were coded: subjunctive; modal conditional with the verb *poder* 'to be able to'; conditional, past, future, modal present with the verb *poder* 'to be able to'; present; and imperative. Of 1776 responses, a total of 239 responses were eliminated. In 124 responses, no request was performed, and no verb was used in 4 situations. Additionally, the 111 responses from situation 10 (Appendix A) were eliminated because the situation did not prompt participants to make a request in the present moment as the others did. The gender of each participant and the contextual variables present in each situation were also coded for all responses.

To respond to the research questions, a mixed-methods approach was used to examine the data quantitatively and qualitatively. In the first part of the analysis, two separate conditional inference trees were fitted to female and male participants' data with verb form as the dependent variable (eight levels) and the predictor variables of power, distance, and imposition (two levels each) as independent variables. Conditional-inference-tree analyses determined the level of significance and the hierarchical order of each of the predictor variables on verb form by gender. Conditional inference trees are a type of random forest analysis that implements tree-structure regression models into the framework of conditional inference procedures (Hothorn et al. 2006). Within each conditional inference tree, statistical tests are performed to determine whether each split of the tree is significant or not. In addition to utilizing and providing tests of significance, the output, based on machine learning algorithms, provides a visualization of how predictors operate. The model was constructed by using the *party* package and the function *ctree()* (Hothorn et al. 2006) in R (R Development Core Team 2019). Similar statistical approaches have been used in previous sociolinguistics and pragmatics research (e.g., Sainzmaza-Lecanda and Schwenter 2017; Tagliamonte and Baayen 2012; Tamaoka et al. 2010; Rosemeyer and Schwenter 2017).

Following the conditional-inference-tree analyses, the distribution of verb forms by women and men was analyzed using a chi-square test for independence. Then, individuals' verb-form-selection tendencies were visually examined using a ridgeline plot. The unique stances adopted by each participant were also analyzed, relying on pragmatic interpretations of verb forms as communicating more or less directness with respect to the request. The benefit of adopting the notion of stance, even when the underlying concept of directness has been addressed in the prior pragmatics literature, is that it draws attention to the creation of social meaning in given interactions rather than relying on a set system of pragmatic meaning applied uniformly by all speakers. Excerpts are unaltered to reflect the original wording of the participants.

4. Results

4.1. Effects of Power, Distance, and Imposition on Verb Selection by Gender

Providing the best fit hierarchical structure, the results of two conditional inference trees show how the predictor variables (i.e., power, distance, imposition) affect verb selection in Peninsular Spanish requests for women and men in Madrid, Spain. The inner nodes represent the effect of the different predictor levels and demonstrate their significance, and the terminal nodes at the bottom of the output show the distribution of the verb forms used.

The conditional inference tree for the women's data indicated that all predictor variables examined had a significant impact on verb selection (see *p*-values in Figure 3). Power was the highest node ($p < 0.001$) (Node [1]). When a power differential was present (+power), distance was found to be the next-highest-ranked predictor ($p < 0.001$) (Node [2]), and imposition only had an effect when requests involved a person with whom the participant was familiar (−distance) ($p < 0.001$) (Node [4]). On the other side of the tree, for the situations with equal power (−power), the predictor of imposition was ranked as the next-highest-level predictor ($p < 0.001$) (Node [7]), followed by distance ($p = 0.004$ and $p < 0.001$) (Nodes [8] and [11]).

The conditional inference tree for the men's data also showed that all predictor variables had a significant impact on verb selection (see *p*-values in Figure 4). In Figure 4, power ($p < 0.001$) was found to be the main predictor of verb selection (Node [1]), as in the model of the women's data. When there was a power differential (+power), imposition ($p < 0.001$) was found to be the next-most-important predictor (Node [2]), with distance being a significant predictor only when the requests involved low imposition (−imposition) ($p < 0.001$) (Node [4]). On the other side of the tree (−power), imposition ($p < 0.001$) was the next-highest-level predictor (Node [7]), followed by distance ($p = 0.006$ and $p < 0.001$) (Nodes [8] and [11]).

Examining the inner nodes in Figure 3 (women's data) and Figure 4 (men's data), the rather high placement of imposition across both models is notable, considering that prior research has focused more on power and distance than imposition. Furthermore, the comparison across the women's and men's data highlights the inverted hierarchical order of distance and imposition in contexts of +power, with women relying on distance and men relying on imposition as a higher-level predictor. The difference in the hierarchical structure provides evidence that the way in which women and men in Madrid rely on the contextual variables to make linguistic choices is different, at least when making requests, as represented by the current data.
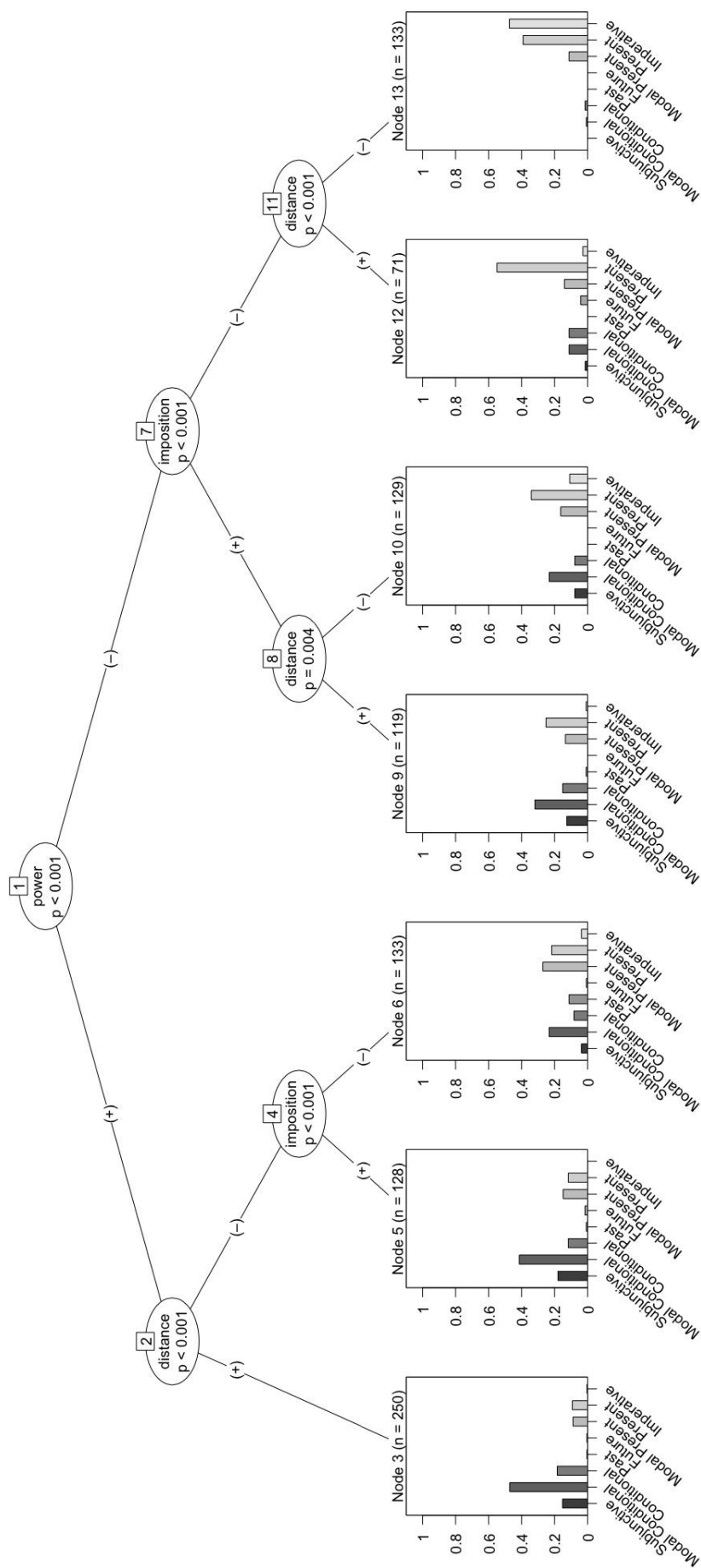
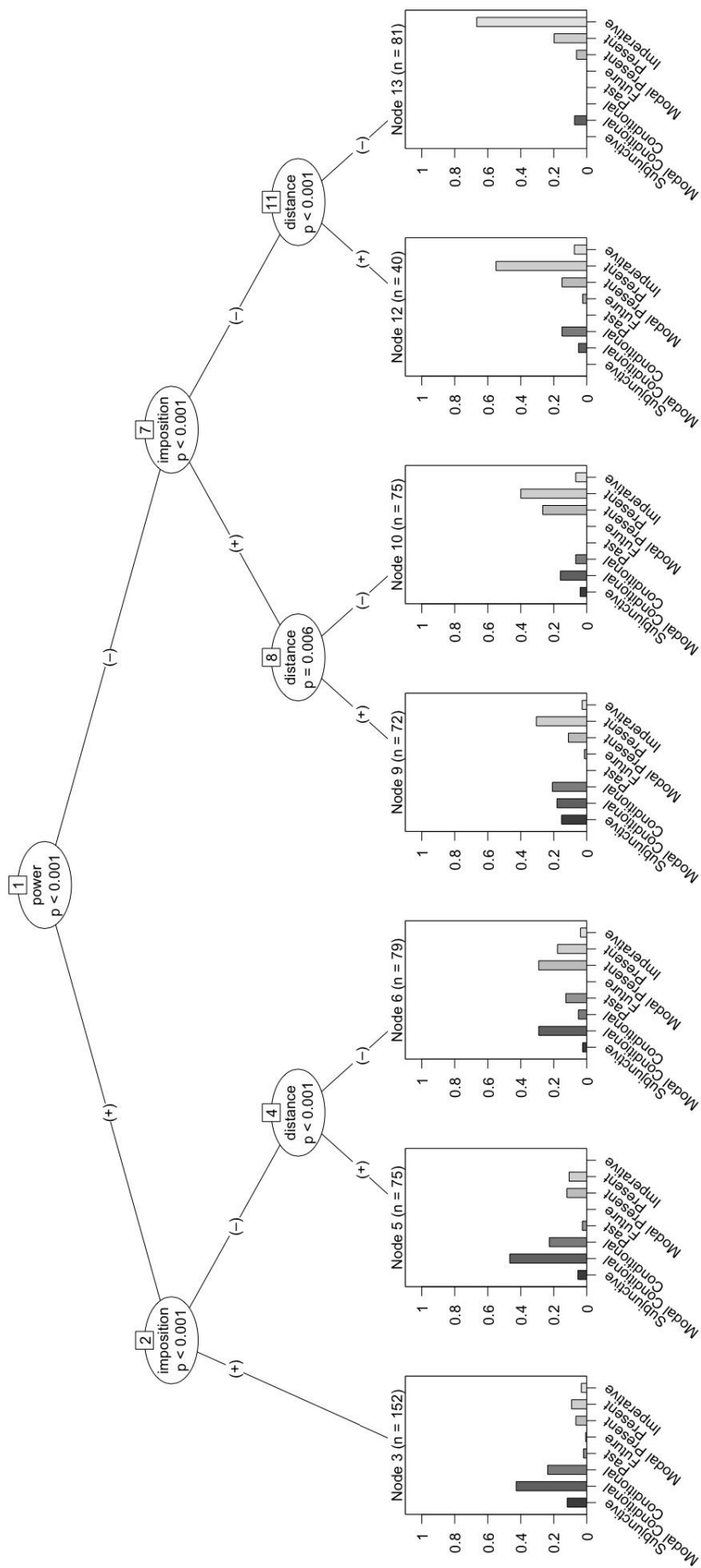**Figure 3.** Conditional-inference-tree analysis of verb-form selection for women.

Figure 4. Conditional-inference-tree analysis of verb-form selection for men.

4.2. Distribution of Verb Forms

To understand the overall distribution of verb forms by gender, the counts of verb forms used by women and men were calculated (Table 2) and a chi-square test of independence was performed.[9] The results indicated independence between gender and verb form, $X^2(6, 1536) = 12.55$, $p = 0.05$. While both groups used a wide range of verbal forms and relied most on modal conditional and present tense verb forms in requests, they exhibited slightly different verb form distributions. Women used subjunctive, modal conditional, and present slightly more often than men, while men used the conditional and imperative forms slightly more than women. Considering the verbal continuum, women used slightly more verb forms on the higher end of the continuum (e.g., less direct), and men used slightly more forms on the lower end (i.e., more direct).

Table 2. Count and percentage of verb forms by women and men.

Verb Form	Women		Men	
	n	%	*n*	%
Subjunctive	92	9.6%	38	6.6%
Modal Conditional	279	29.0%	156	27.2%
Conditional	110	11.4%	83	14.5%
Past/Future	24	2.5%	18	3.1%
Modal Present	139	14.5%	81	14.1%
Present	232	24.1%	126	22.0%
Imperative	86	8.9%	72	12.5%
Total	962	100%	574	100%

An examination of the verb forms that were most distinct by group in the terminal nodes of the conditional inference trees revealed that women employed the subjunctive more frequently than men across situations. Men used imperative forms more than women, particularly in −power, −distance, and −imposition situations (66.7% for men and 47.3% for women). In this specific situation, women were more likely to use a present tense verb than men—another indicator of the women's slight tendency to select forms towards the higher end of the continuum compared to men.

4.3. Individual-Level Analysis

To examine individuals' verb choices and related stances, each participant's set of requests was analyzed with respect to their position on the verbal continuum. Quantitatively, a mean score was calculated for each participant by assigning a 1–7 scale to the verbs on the verbal continuum (1—subjunctive, 2—modal conditional, 3—conditional, 4—imperfect/future, 5—modal present, 6—present, 7—imperative).[10]

The ridgeline plot in Figure 5 shows the distribution of the means by gender. While the group means for women ($M = 3.93$, $SD = 0.67$) and men ($M = 4.06$, $SD = 0.75$) are not notably different, the plot demonstrates various trends related to individuals' verb-selection tendencies. First, there is a similar distribution for women and men, with most data points appearing near the center of the continuum. This finding and the distribution data reported in Table 2 indicate that most participants rely on some verb forms on the high and low ends of the continuum, given that the midpoint past and future forms are seldom used. This trend is supported by the conditional-inference-tree analyses, which demonstrate that the use of verb forms on the high and low ends of the continuum aligns with specific contexts. Second, the ridgeline plot also highlights the wide range of individuals' means represented in the data, which ranged from 2.46 to 5.78 for women and 2.36 to 6.07 for men, indicating individual tendencies in verb selection for certain women and men that deviate from each group's norm. Third, whereas the women's data show a more normal distribution, the men's distribution tends towards verb forms on the low end of the spectrum, using verbs that are more direct or assertive. This finding is seen in the small increase in data points

between 6 and 7 on the scale and in the larger set of data points between 4 and 5. There also seems to be a subgroup of men that prefer to use verb forms that are higher on the continuum than the mean, indicated by the increase in data points between 3 and 3.5. Overall, these data confirm that there is a slight tendency for men to use more-direct verb forms than women, but it also highlights the great range in variability among individuals.
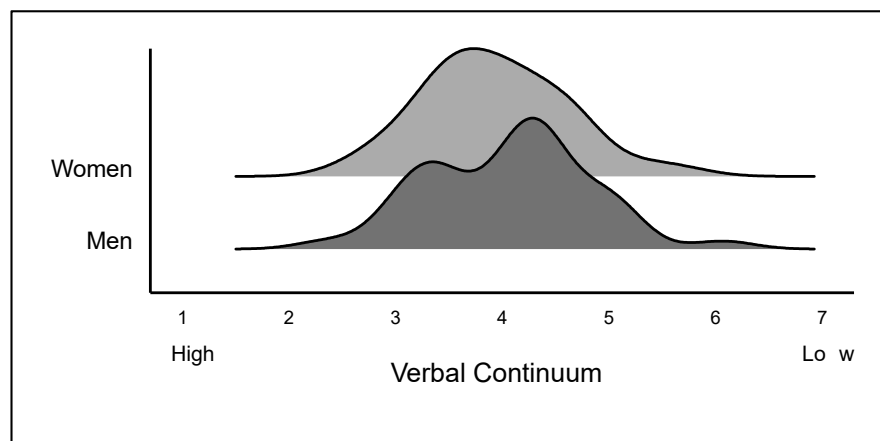

Figure 5. Individual means on the verbal continuum by gender.

To explore the variability in verb form use, data from individual participants were selected for analysis and presented in Sections 4.3.1–4.3.3. These data show that some men and women fall outside of the general and gendered norms in terms of their use of verb forms.

4.3.1. Midpoint on the Verbal Continuum: Stance Varies by Situation

The participants with individual means that aligned with the group mean tended to use verb forms across the continuum. With the infrequent use of past and future verb forms overall (Table 2), a midpoint average was achieved by using a mix of verb forms from both ends of the continuum. Demonstrating this trend, Sofía (Participant 104) had a mean of 4.0, aligning with the women's mean. In the 15 situations, she used three modal conditional verbs, four conditional verbs, one past verb, four modal present verbs, and three present tense verbs. Jorge (Participant 77) had a mean of 4.2, approximately aligning with the mean for the men's data. In his requests, he used all verb options except the future: one subjunctive form, two modal conditional forms, three conditional verbs, two past forms, two modal present forms, four present tense verbs, and one imperative.

For both participants, there was an alignment between verb form and situation. For example, they used verb forms on the high end of the continuum in the +power, +distance, and +imposition situations and verb forms on the low end of the continuum in the −power, −distance, and −imposition situations. In +power, +distance, and +imposition situations, Sofía used modal conditional verbs (Excerpts 9 and 10), whereas in −power, −distance, and −imposition situations, she used a modal present and a present tense verb (Excerpts 11 and 12). Similarly, Jorge relied on verb forms towards the high end of the continuum in the +power, +distance, and +imposition situations, where he used a subjunctive and a past form (Excerpts 13 and 14), while in the −power, −distance, and −imposition situations, he used a present tense and imperative verb to perform the requests (Excerpts 15 and 16). These participants and others whose data represent the average for their group used a wide range of verb forms, more frequently employing forms on the high end of the verbal continuum in +power, +distance, and +imposition situations and forms on the low end in the −power, −distance, and −imposition situations.[11]

(9) Sofía. Modal Conditional (+power, +distance, +imposition).
*Me gustaría optar a una beca, crees que **podrías hacerme** una carta de recomendación o sabes de alguien que pueda hacermela?*
I would like to apply for a scholarship, do you think you could give me a letter of recommendation or do you know of someone who could do it for me?

(10) Sofía. Modal Conditional (+power, +distance, +imposition).
*Cómo **podría conseguir** el manual?*
How could I get the manual?

(11) Sofía. Modal Present (−power, −distance, −imposition).
*Te **puedo coger** un folio?*
Can I take a sheet of paper from you?

(12) Sofía. Present (−power, −distance, −imposition).
Lo tiras plis?
Throw it away please?

(13) Jorge. Subjunctive (+power, +distance, +imposition).
*Buenas tardes profesor. El motivo de venir es pedirle un favor de motivo académico. Estoy interesado en solicitar una beca... Como considero que usted es uno de los especialistas de mayor renombre en nuestro país en ese campo, **quisiera** pedirle por favor que me redactara una carta de recomendación para poder solicitar dicha beca...*
Good afternoon professor. The reason for coming is to ask you for an academic favor. I am interested in applying for a scholarship... Considering that you are one of the most renowned specialists in our country in that field, I would like to ask you to please write me a letter of recommendation to be able to apply for the aforementioned scholarship...

(14) Jorge. Past (+power, +distance, +imposition).
*Disculpe profesor. **Quería** pedirle los datos de referencia de su manual para poder estudiarlo para su asignatura. Muchas gracias.*
Excuse me professor. I wanted to ask you for the reference of your manual so that I can study it for your class. Thanks very much.

(15) Jorge. Present (−power, −distance, −imposition).
*Miguel, ¿me **dejas** un par de folios?*
Miguel, will you give me a few sheets of paper?

(16) Jorge. Imperative (−power, −distance, −imposition).
*Tío, Miguel, **tírame** esto a la papelera por favor.*
Dude, Miguel, throw this in the waste basket for me please.

The verb forms that have been described as more or less direct in pragmatic terms index how individual interlocutors are positioned with respect to each other and the request (i.e., stance (Du Bois 2007)). They also communicate pragmatic meaning about the deictic point of occurrence, placing the action in a more- or less-distanced or irrealis moment. This pragmatic meaning has social implications, as noted in prior research on politeness (Chodorowska-Pilch 1998, 2004; Koike 1989). Selecting a more-direct verb form contributes to the stance of a 'direct requester'. 'Direct requesters' position themselves socially towards the hearer, as being in a relationship that requires minimal worry about impeding on the hearer, there is a greater assumption that the hearer is willing and able to perform the requested action (Searle 1976). Additionally, they position the speaker and hearer around a request that is close in terms of time or reality. The stance of a 'direct requester' can be seen in Excerpts (12), (15), and (16). On the other hand, the stance of an 'indirect requester' highlights caution around the degree to which the speaker assumes that the hearer is able or willing to perform the request, thus expanding the imagined, socially positioned space between the speaker and hearer. Indirect verb forms also present the request itself as more distanced or removed from the reality of the speaker and hearer.

In addition to viewing directness and indirectness as a stance-taking mechanism, it is important to note that the stances displayed by participants at the midpoint of the range are broadly sensitive to the contextual variables. This set of participants shows that they

sometimes behave as 'direct requesters', especially when they are in −power situations, and they behave as 'indirect requesters' at other times, especially in +power situations. The sensitivity to contextual variables aligns with hypotheses based on pragmatic theories and the prior literature about verb forms and contextual variables.

4.3.2. High End of the Verbal Continuum: Indirect Requesters

Some individuals used verb forms on one end or the other of the verbal continuum. Elena (Participant 49) had an average of 2.71, aligning her verb forms across situations with the high end of the continuum, which expresses more-indirect requests. She used subjunctive forms two times, modal conditional forms in six requests, conditional verbs four times, and present tense verbs in two situations. Similarly, Rafael (Participant 25) had an average of 2.36. In his requests, he used modal conditional verbs in 12 situations, and he used a conditional verb and a present tense verb one time each. These participants used verbs to make requests in a way that is different from those who relied on the full spectrum of forms; they more consistently employed forms that theoretically have been proposed to be less direct and communicated with less force. Excerpts (17)–(19) show Rafael's application of a modal conditional verb in a range of different situations with different combinations of contextual variables. A connection between situation type and verb form was not clearly observable. Thus, it seems that some individuals prefer verb forms in requests that more consistently present them as 'indirect requesters,' conveying tentativeness about the assumption that the hearer is willing and able to perform the request and by distancing it from the interlocutors via deictic displacement.

(17) Rafael. Modal Conditional (+power, +distance, +imposition).
Gómez, necesito solicitar la beca para el año que viene. **Podrías hacerme** *una carta de recomendación, por favor? Gracias*
Gómez, I need to apply for the scholarship for next year. Could you give me a recommendation letter, please? Thanks.

(18) Rafael. Modal Conditional (+power, −distance, −imposition).
Elena, ¿Recuerdas los artículos que nombraste en clase? ¿Me **podrías decir** *cuáles eran, por favor?*
Elena, do you remember the articles that you named in class? Could you tell me which ones they were, please?

(19) Rafael. Modal Conditional (−power, −distance, −imposition).
Pablo, me podrías **dejar** *unos folios, por favor?*
Pablo, could you give me some sheets of paper, please?

4.3.3. Low End of the Verbal Continuum: Direct Requesters

The data also indicated that some people prefer to use verb forms in requests that align with the low end of the continuum—verb forms that are more direct, such as imperative and present tense verbs. María Carmen (Participant 85) had an average of 5.47. She used one modal conditional form, three conditional forms, six present tense verbs, and five imperative verbs in her requests. Alberto (Participant 97) had an average of 6.07 on the verbal continuum, and he used one modal conditional, two modal present, four present, and seven imperative verb forms.

María Carmen used imperative forms in various situations, from those with a low-imposition request to her professor, with whom she was not very familiar (Excerpt 20), to those that involved a low-imposition request to her close peer (Excerpt 21), but her use of the imperative was not limited to situations of low imposition. She relied on present and imperative requests in many more situations than most participants, and she showed minimal sensitivity to the contextual variables. Similarly, Alberto's data did not show any observable distinction across situations. Both María Carmen and Alberto relied quite heavily on the most-direct verb forms, using the present or imperative in 11 of the 15 situations, indicating a quite consistent 'direct requester' stance. This stance was communicated by maintaining the request within the referential time period of the

present/realis moment and by not indicating that they questioned the assumption that the hearer is willing and able to carry out the requested action.

(20) María Carmen. Imperative (+power, +distance, −imposition).
      Qué se ha escrito sobre este tema? **Recomiéndame** *algo bueno.*
      What's been written about this topic? Recommend something good for me.
(21) María Carmen. Imperative (−power, −distance, −imposition).
      Tíralo, *por favor.*
      Throw it out, please.

The qualitative samples of data in Sections 4.3.1–4.3.3 confirmed that not all participants behaved in the same way. While most of the participants were similar to Sofía and Jorge, who used verb forms on the high and low ends of the verbal continuum in accordance with the contextual variables, as predicted by prior research, some participants preferred a less-direct stance, like Elena and Rafael, and some preferred a more-direct stance, like María Carmen and Alberto. Taken collectively, these data suggest that many participants adopt direct and indirect stances in response to contextual variables, as expected (e.g., more direct with −power, less direct with +power), but others take on direct or indirect stances more consistently and with less dependence on contextual variables. The participants who are more responsive to contextual variables take different stances depending on the context in systematic ways, while those who are more consistent in their stance-taking maintain more uniformity in how they position themselves when making requests. These findings give rise to new questions about why some individuals are more affected by contextual variables than others and how the outliers may shape the future norms of a given population.

5. Summary of Findings and Discussion

This investigation examined how contextual factors of power, distance, and imposition collectively condition verb forms used by women and men in Madrid, Spain, and how verb selection varies by gender and by individual. In this section, a response to each research question is provided, followed by a discussion of the main contributions.

5.1. Effect of Contextual Variables

In response to the first research question about whether the effect and hierarchical order of the contextual factors of power, distance, and imposition differed for women and men when predicting verb forms in requests, the results indicated that there was a difference in the effect and hierarchical order across groups. The women's data showed that distance was a more important predictor than imposition in situations of asymmetrical power (+power). For the men's verb selection in the same asymmetrical power situations, imposition was a more important predictor than distance.

The difference in the hierarchical structure of predictor variables demonstrates the variability in how certain social groups rely on contextual factors to make linguistic choices. Whereas Tamaoka et al. (2010) showed that particular cultural groups differentially value power and distance when interpreting politeness, the current data show that, beyond the cultural level, individual social groups can interpret context differently. The women and men who participated in the current experiment were all exposed to the same situations and yet the results point to different ways of relying on the contextual variables to make linguistic choices.

The conditional inference trees for women and men provide evidence of different frames (Terkourafi 2005, 2015). Within these unique frames, women's verb-form selections in situations with an asymmetrical power relationship (+power) are more tightly connected to whether the hearer is a more- or less-distant person; men's verb forms in those same situations are more tightly connected to the degree of imposition of the request. Thus, even when two individuals select the same verb form in the same situation involving a power differential (+power), the results demonstrated that their frame-based meanings may be distinct. In other words, when the predictive hierarchies are different, the verb forms utilized by the women and men whose data contributed to the overarching results

theoretically index diverse positionings around notions of distance or imposition. These associations between contextual variables and linguistic resources, which are based on the interlocutors' predictive abilities about frames (Terkourafi 2015), provide nuanced shifts in the social meaning of the linguistic resources used. Whereas variationist approaches have been common in linguistics research for decades, mostly focusing on the variation in linguistic resources (e.g., subjunctive vs. indicative; [s] vs. [h]), the current findings draw attention to the possibility that the perception of the context is also a variable factor that impacts linguistic production and meaning.

The results of the conditional-inference-tree analyses also shed light on a more basic question about the importance of power, distance, and imposition for understanding language variation. Of the three, power has perhaps been the most studied in pragmatics research (e.g., Félix-Brasdefer 2005, 2009; Hernández-Flores 2004), and theoretical discussions have also pointed to power being a primary influence on language and society (e.g., Eckert 1989; Uchida 1992). While distance has also been considered frequently, less systematic attention has been paid to imposition (e.g., Ballesteros 2001; Blum-Kulka et al. 1989; Félix-Brasdefer 2005, 2009; Le Pair 1996), under the assumption that it is a less important contextual variable. The results of the conditional-inference-tree analysis confirmed the importance of power, and they demonstrated that imposition was often a more influential variable than distance when predicting verb forms in requests. In an applied sense, this finding suggests that the Spaniards who participated are more sensitive to imposition than may have been expected, and future research should consider imposition when examining the effect of contextual variables.

5.2. Verb Form and Gender

The second research question addressed the distribution of verb forms used by women and men. Chi-square results indicated a significant difference in the distribution of verb forms across women's and men's requests. The results provided some support for the notion that women may prefer to use softened linguistic forms more than men (e.g., Leaper and Robnett 2011). In the current data, women used subjunctive and modal conditional forms slightly more than men, and men used imperative verbs—the most-direct verb form examined—more than women and in more situations. These observations did not align with Lorenzo-Dus and Bou-Franch's (2003) finding that Spanish women tended to use more-direct requests than men in situations with asymmetrical power and distance; rather, the current results indicate subtle differences in the selection of verb forms at the extremes of the directness continuum associated with politeness (i.e., subjunctive, modal conditional, imperative).

5.3. Individual Stances with Verbs

The final research question addressed the variability in individuals' verb forms and the stances related to directness that they portray. Most men and women used varied verb forms that reflected more- and less-direct stances in different situations (e.g., more direct with −power, less direct with +power). Most participants' range of verb forms was fairly equally distributed across the verbal continuum, although some participants maintained a more- or less-direct stance overall. These participants acted more as 'direct requesters' or 'indirect requesters' independent of the situation. The analysis at the individual level aligned with the understanding that macro-social trends are representative of large groups of individuals who share common linguistic practices (e.g., Bell 2016), and it also confirmed the agency of individuals. As clearly indicated in third-wave sociolinguistic research (Eckert 2012), the analysis of individuals' language use is fundamental for understanding the social moves that are made by all speakers, not just those that fall within the norm. In other words, the 'direct requesters', 'indirect requesters', and the rest of the participant population who relied on contextual variables to a greater degree to select verb forms are all positioned vis-à-vis each other; the social meaning of these performances is best interpreted by considering the relationship to the other performances. Furthermore, the

social moves performed via language, including those made by the outliers in these data, have the potential to affect the request norms and the continued (re)construction of gender in Peninsular Spanish.

While the individual agency was apparent, the current data were not sufficient to explain what ideologies, stances, identities, or frames explain the individual-level variation. For example, it is possible that participants who represent the outliers have different underlying hierarchies when processing contextual variables or different responses to them. It is also possible that these explanations, related to the perception of context, may be intertwined with speakers' ideologies and identities.

Reflecting on the results as a whole, one may question the degree to which interlocutors are aware of their linguistic choices that rely on interrelated contextual variables or the nuanced meanings that emerge from the coupling of a linguistic form with a specific interpretation of a context. Eckert (2019) proposes that 'most of what we do . . ., we do unconsciously' (p. 758). She suggests that speakers may be more aware of the persona that they are trying to communicate rather than the individual features of a style. Following this line of argumentation, it seems unlikely that speakers or hearers would be conscious of the meaning contributed by the hierarchical structure of predictive contextual variables, but this does not mean that it is not unconsciously calculated. In other words, they may consider themselves to be 'direct communicators', 'indirect communicators', or even communicators who do not like to impose or value close relationships, but they may not be aware of how they use linguistic resources or how they rely on contextual variables when constructing these positionalities.

6. Conclusions

This investigation examined the intertwined impact of power, distance, and imposition on verb selection in Spanish for women and men in Madrid, Spain, and explored individual strategies in the formulation of requests. The results of this study showed differences across the women's and men's data in the hierarchical structure of the three contextual variables as predictors of verb form, which highlights contextual variability as a factor that should be considered in the study of the social meaning of language. Other findings demonstrated differences in verb use not only between Spanish women and men but also among individuals. Taken as a whole, the results of this investigation are a snapshot of the participants' complex calculations involving language, context, and interlocutors, and they support the notion that frames, politeness, and gender are domain-specific, dynamic constructs in communities (Mills 2003; Terkourafi 2015).

The experimental paradigm and mixed-methods approach in this study offered the benefits of isolating specific contextual variables, collecting a large amount of data in controlled situations, and providing new insight into the complex relationship among linguistic variables, contextual variables, and social groups and individuals. Despite these strengths, the experimental data were also limited in that they do not represent naturally occurring speech. Another limitation was the scope of the investigation, which was focused on verb forms. Future research should consider other types of data to corroborate and further the current findings. For example, while prior research has shown that the verb is a fundamental indicator of illocutionary force and politeness (Chodorowska-Pilch 1998, 2004; Félix-Brasdefer 2005), future research may seek to confirm the hierarchical structure of contextual variables through the analysis of other linguistic forms, as they may be predicted by the same or a distinct organization of contextual factors. This line of research will continue to demystify the understanding of language and context. Additionally, while the current project determined that some individuals adopt more-direct or -indirect strategies regardless of context, future research should explore the social explanations for these results. Analyses of metalinguistic data or ethnographic data related to individual participants could provide necessary insight to explain the local and social meanings of the linguistic resources in use.

Appendix A

Table A1. Classification of the Situations and the Contextual Variables.

Situation	P	D	I	Situation—Action	Situación—Acciones
1	+	+	High	You want to apply for a scholarship. You want a letter of recommendation from [name of a professor you don't know very well], who was your professor in one class last year. At that point, she/he has office hours. What do you say?	Quieres solicitar una beca. Quieres una carta de recomendación de [nombre de un profesor al que no conoces mucho], que te dio una clase el año pasado. En ese momento, tiene horas de tutoría. ¿Qué le dices?
2	+	+	Low	You are interested in a topic in which [name of a professor you don't know very well] is an expert. You want to know more about the topic, so you go to her/his office hours and you ask that he/she recommend a book to you. What do you say?	Estás interesado en un tema del que [nombre de un profesor al que no conoces mucho] es un experto. Quieres saber más sobre el tema, así que vas a sus horas de tutoría y le pides que te recomiende un libro. ¿Qué le dices?
3	+	+	High	You have a due date for next Monday, but you don't have the textbook that's used in the class. You see [name of a professor you don't know very well]. You want to use his/her textbook over the weekend. What do you say?	Tienes una fecha de entrega para el próximo lunes, pero no tienes el libro de texto que se usa en la clase. Ves a [nombre de un profesor al que no conoces mucho]. Quieres usar su libro de texto durante el fin de semana. ¿Qué le dices?
4	+	+	Low	You missed [name of a professor you don't know very well]'s class, in which the professor handed out a worksheet. You want the worksheet. What do you say?	Faltaste a una clase de [nombre de un profesor al que no conoces mucho] en la que entregó una hoja de actividades. Quieres la hoja de actividades. ¿Qué le dices?
5	+	−	High	You are not sure about the quality of a 15-page essay you have written. Tomorrow is the due date. You go to [name of a professor you know]'s office to ask her/him to read the entire essay and give you comments. What do you say?	No estás seguro sobre la calidad de un ensayo de 15 páginas que has escrito. Mañana es la fecha de entrega. Vas al despacho de [nombre de un profesor al que conoces] para pedirle que lea el ensayo entero y te dé comentarios. ¿Qué le dices?
6	+	−	Low	You want to find an article that was mentioned in class, but you can't remember the author's name. [Name of a professor you know] talked about that article. You see that she/he is in the office during office hours, and you want to know the reference. What do you say?	Quieres encontrar un artículo que fue mencionado en clase, pero no recuerdas el nombre del autor. [Nombre de un profesor al que conoces] habló sobre ese artículo. Ves que está en el despacho en horas de tutoría, y quieres saber la referencia. ¿Qué le dices?

Table A1. *Cont.*

Situation	P	D	I	Situation—Action	Situación—Acciones
7	+	−	High	You want the PowerPoints that [name of a professor you know] has used in class throughout the course. You go to his office during his tutoring hours. What do you say?	Quieres los PowerPoint que ha utilizado [nombre de un profesor al que conoces] en la clase a lo largo del curso. Vas a su despacho en sus horas de tutoría. ¿Qué le dices?
8	+	−	Low	You need a stapler to staple a final paper that you have to submit to [name of a professor you know]. You know that [name of a professor you know] has a stapler in the office. What do you say?	Necesitas una grapadora para grapar un trabajo final que tienes que entregar a [nombre de un profesor al que conoces]. Sabes que [nombre de un profesor al que conoces] tiene una grapadora en la oficina. ¿Qué le dices?
9	−	+	High	You are in the library. You have an exam, and you are not very good at the topic. You see [name of a classmate that you don't know very well] studying for the same test. You want to prepare for the exam with [name of a classmate that you don't know very well]. What do you say?	Estás en la biblioteca. Tienes un examen y el tema no se te da muy bien. Ves a [nombre de un estudiante al que no conoces mucho] estudiando para el mismo examen. Quieres preparar el examen con [nombre de un estudiante al que no conoces mucho]. ¿Qué le dices?
10	−	+	Low	In a class, you did not hear when the deadline to turn in the final project is. Next to you is [name of a student you don't know very well]. You want to know when the due date is. What do you say?	En una clase, no escuchaste bien cuándo es la fecha límite para la entrega del trabajo final. A tu lado está [nombre de un estudiante al que no conoces mucho]. Quieres saber cuándo es la fecha de entrega. ¿Qué le dices?
11	−	+	High	You have an exam next week and you have missed some classes. You are in class, and you have [Name of a student you don't know very well] next to you. You want her/him to lend you her/his notes. What do you say?	Tienes un examen la próxima semana y has faltado a algunas clases. Estás en clase, y tienes a [Nombre de un estudiante al que no conoces mucho] a tu lado. Quieres que te preste sus apuntes. ¿Qué le dices?
12	−	+	Low	You're in class and you can't find a pen. [Name of a student you don't know very well] is next to you. You want her/him to lend you a pen. What do you say?	Estás en clase y no encuentras ningún bolígrafo. [Nombre de un estudiante al que no conoces mucho] está a tu lado. Quieres que te preste un bolígrafo. ¿Qué le dices?
13	−	−	High	You have a final exam in two days and a presentation with [name of student you know well]. You want [name of a student you know well] to be in charge of preparing the presentation for both of you because you don't have time. What do you say?	Tienes un examen final en dos días y una presentación con [nombre de un estudiante al que conoces mucho]. Quieres que [nombre de un estudiante al que conoces] se encargue de preparar la presentación por los dos porque no tienes tiempo. ¿Qué le dices?
14	−	−	Low	You are in class, and you want to throw a paper in the trash. [Name of a student that you know well] is next to you. You want her/him to do it for you. What do you say?	Estás en clase y quieres tirar un trozo de papel a la papelera. [Nombre de un estudiante al que conoces] está sentado a tu lado. Quieres que lo haga por ti. ¿Qué le dices?
15	−	−	High	You have an exam after the weekend, and you lost your manual. You see that [name of a student you know well] has that manual. You want her/him to lend it to you over the weekend, even though [name of student you know well] has to study too. What do you say?	Tienes un examen después del fin de semana, y perdiste tu manual. Ves que [nombre de un estudiante al que conoces bien] tiene ese manual. Quieres que te lo preste durante el fin de semana, aunque [nombre de un estudiante al que conoces] también tiene que estudiar. ¿Qué le dices?
16	−	−	Low	You are in class, and you run out of paper. [Name of a classmate that you know well] is seated next to you. You want her/him to lend you a sheet of paper. What do you say?	Estás en clase y se te acaban los folios. [Nombre de un estudiante al que conoces] está sentado a tu lado. Quieres que te preste un folio. ¿Qué le dices?

Notes

1   The field of variational pragmatics represents a shift from traditional pragmatics research, which aims to explore non-truth conditional meanings that are common across speakers. See Eckert (2019) for discussion.

2   See Kiesling (2009) for a review of the related concepts of stance, style, persona, and identity.

3   See Terkourafi (2005, 2015) for an explanation of how particularized conversational implicatures may also be used to be polite or impolite when conventionalized forms are not used.

4   Imperfect/future forms are placed together because of the lack of prior research available to distinguish their comparative placement on the continuum.

5   Spanish excerpts represent the wording of the participants. Translations in English do not always reflect the same verb form as in Spanish.

6   While gender is not a binary category, this article addresses the language of women and men because these categories represent the participants in this investigation, as indicated by them.

7   Alerters are terms of address, such as the addressee's name, or other linguistic resources, such as *perdona* 'excuse me', to draw attention to an upcoming speech act (Blum-Kulka et al. 1989).

8   Two levels of the contextual variables were used for practical, experimental purposes, despite each variable being scalar in nature.

9   Imperfect/future forms were calculated together because they are considered to be in the middle of the politeness continuum. Additionally, this prevented any cell from having fewer than five counts.

10   While the distances between verb forms may not be exactly the same, the assignation of a numerical value to the verbal continuum suffices for the current purpose of understanding the participants' stances in terms of directness.

11   The excerpts highlight the fact that not all requests that use the same verb form have the same request structure. For example, Excerpts (9) and (10) both contain a modal conditional verb, but the request strategies that were used are different (see Blum-Kulka et al. (1989) for a review of request strategies). Given the systematicity of the current results, the analysis of verb forms is relevant, and it also provides additional detail that is not captured by a request strategy analysis (e.g., conventionalized indirect requests may use present modals or conditional modals). The excerpts also draw attention to the use of other linguistic variables that may systematically relate to situation type, verb form, or identity. Jorge's excerpts demonstrate a clear differentiation between the *usted* form in the +power, +distance, and +imposition situations and the *tú* form in the −power, −distance, and −imposition situations. Furthermore, preliminary analyses indicated that *por favor* 'please' was most commonly used in conjunction with imperative verb forms, as in Excerpt (16). While related to the topics discussed in this article and special issue, these questions lie outside the scope of the current paper.

References

Acton, Eric K. 2021. Pragmatics and the third wave: The Social Meaning of Definites. In *Social Meaning and Linguistic Variation: Theorizing the Third Wave*. Edited by Lauren Hall-Lew, Emma Moore and Robert J. Podesva. Cambridge: Cambridge University Press, pp. 105–26. [CrossRef]

Agha, Asif. 2003. The social life of cultural value. *Language & Communication* 23: 231–73. [CrossRef]

Agha, Asif. 2007. *Language and Social Relations*. Cambridge: Cambridge University Press. [CrossRef]

Austin, John Langshaw. 1975. *How to Do Things with Words*. Edited by J. O. Urmson and Marina Sbisà. Oxford: Oxford University. First published 1962.

Ballesteros, Francisco J. 2001. La cortesía española frente a la cortesía inglesa. Estudio pragmalingüístico de las exhortaciones impositivas. *Estudios Ingleses de la Universidad Complutense* 9: 171–207.

Ballesteros, Francisco J. 2002. Mecanismos de atenuación en español e inglés. Implicaciones pragmáticas en la cortesía. In *Círculo de Lingüística Aplicada a la Comunicación*. Madrid: Universidad Complutense de Madrid.

Barron, Anne. 2005. Variational pragmatics in the foreign language classroom. *System* 33: 519–36. [CrossRef]

Barron, Anne. 2014. Variational Pragmatics. In *The Encyclopedia of Applied Linguistics*. Edited by Carol A. Chapelle. Oxford: Willey-Blackwell Publishing, pp. 187–99. [CrossRef]

Barron, Anne, and Klaus P. Schneider. 2009. Variational pragmatics: Studying the impact of social factors on language use in interaction. *Intercultural Pragmatics* 6: 425–42. [CrossRef]

Bataller, Rebeca, and Rachel Shively. 2011. Role plays and naturalistic data in pragmatics research: Service encounters during study abroad. *Journal of Linguistics and Language Teaching* 2: 15–50.

Bell, Allan. 2016. Succeeding Waves: Seeking Sociolinguistic Theory for the Twenty-first Century. In *Sociolinguistics: Theoretical Debates*. Edited by Nicolas Coupland. Cambridge: Cambridge University Press, pp. 391–416. [CrossRef]

Blum-Kulka, Shoshana. 1987. Indirectness and politeness in requests: Same or different? *Journal of Pragmatics* 11: 131–46. [CrossRef]

Blum-Kulka, Shoshana, Juliane House, and Gabriele Kasper. 1989. *Cross-Cultural Pragmatics: Requests and Apologies*. Norwood: Ablex.

Briz Gómez, Antonio. 2004. Cortesía Verbal Codificada y Cortesía Verbal Interpretada en la Conversación. In *Pragmática Sociocultural: Estudios Sobre el Discurso de Cortesía en Español*. Edited by Diana Bravo Moreno and Antonio Briz Gómez. Madrid: Ariel, pp. 67–94.

Brown, Lucien, Bodo Winter, Kaori Idemaru, and Sven Grawunder. 2014. Phonetics and politeness: Perceiving Korean honorific and non-honorific speech through phonetic cues. *Journal of Pragmatics* 66: 45–60. [CrossRef]

Brown, Penelope, and Stephen C. Levinson. 1987. *Politeness: Some Universals in Language Usage*. Cambridge: Cambridge University Press.

Brown, Roger, and Albert Gilman. 1972. The Pronouns of Power and Solidarity. In *Language and Social Context*. Edited by Pier Paolo Giglioli. New York: Penguin Books, pp. 252–81. First published 1960. [CrossRef]

Bucholtz, Mary, and Kira Hall. 2005. Identity and interaction: A sociocultural linguistic approach. *Discourse Studies* 7: 585–614. [CrossRef]

California Style Collective (Jennifer Arnold, Renée Blake, Penelope Eckert, Mako Iwai, Norma Mendoza-Denton, Cecilia Morgan, Livia Polanyi, Joan Solomon, and Thomas Veatch). 1993. Variation and Personal/Group Style. Paper presented at New Ways of Analyzing Variation NWAV-22 Conference, Ottawa, ON, Canada, October 14–17.

Chodorowska-Pilch, Marianna. 1998. Encoding of Politeness in Spanish and Polish: A Cross-Linguistic Study. Ph.D. dissertation, University of Southern California, Los Angeles, CA, USA.

Chodorowska-Pilch, Marianna. 2004. The Conditional: A Grammaticalised Marker of Politeness in Spanish. In *Current Trends in the Pragmatics of Spanish*. Edited by Rosina Márquez-Reiter and María Elena Placencia. Amsterdam: John Benjamins, pp. 57–75. [CrossRef]

Czerwionka, Lori. 2012. Mitigation: The combined effects of imposition and certitude. *Journal of Pragmatics* 44: 1163–82. [CrossRef]

Czerwionka, Lori. 2014. Participant perspectives on mitigated interactions: The impact of imposition and uncertainty. *Journal of Pragmatics* 67: 112–30. [CrossRef]

D'Arcy, Alexandra. 2017. *Discourse-Pragmatic Variation in Context: Eight Hundred Years of LIKE*. Philadelphia: John Benjamins. [CrossRef]

Davies, Bronwyn, and Rom Harré. 1990. Positioning: The discursive production of selves. *Journal for the Theory of Social Behaviour* 20: 43–63. [CrossRef]

Du Bois, John W. 2007. The Stance Triangle. In *Stancetaking in Discourse: Subjectivity, Evaluation, Interaction*. Edited by Robert Englebretson. Philadelphia: John Benjamins, pp. 139–82. [CrossRef]

Eckert, Penelope. 1989. The whole woman: Sex and gender differences in variation. *Language Variation and Change* 1: 245–67. [CrossRef]

Eckert, Penelope. 2012. Three waves of variation study: The emergence of meaning in the study of sociolinguistic variation. *Annual Review of Anthropology* 41: 87–100. [CrossRef]

Eckert, Penelope. 2016. Variation, Meaning and Social Change. In *Sociolinguistics: Theoretical Debates*. Edited by Nikolas Coupland. Cambridge: Cambridge University Press, pp. 68–85. [CrossRef]

Eckert, Penelope. 2019. The limits of meaning: Social indexicality, variation, and the cline of interiority. *Language* 95: 751–76. [CrossRef]

Eckert, Penelope, and Sally McConnell-Ginet. 1992. Think practically and look locally: Language and gender as community-based practice. New generalizations and explanations in language and gender research. *Annual Review of Anthropology* 21: 461–90. [CrossRef]

Eckert, Penelope, and Sally McConnell-Ginet. 1999. New generalizations and explanations in language and gender research. *Language in Society* 28: 185–201. [CrossRef]

Ensink, Titus, and Christoph Sauer. 2003. Social-functional and Cognitive Approaches to Discourse Interpretation: The Role of Frame and Perspective. In *Framing and Perspectivising in Discourse*. Edited by Titus Ensink and Christoph Sauer. Philadelphia: John Benjamins, pp. 1–21. [CrossRef]

Félix-Brasdefer, J. César. 2005. Indirectness and Politeness in Mexican Requests. In *Selected Proceedings of the Hispanic Linguistics Symposium*, 7th ed. Edited by David Eddigton. Somerville: Cascadilla Proceedings Project, pp. 66–78.

Félix-Brasdefer, J. César. 2007. Natural speech vs. elicited data: A comparison of natural and role play requests in Mexican Spanish. *Spanish in Context* 4: 159–85. [CrossRef]

Félix-Brasdefer, J. César. 2008. Sociopragmatic variation: Dispreferred responses in Mexican and Dominican Spanish. *Journal of Politeness Research* 4: 81–100. [CrossRef]

Félix-Brasdefer, J. César. 2009. Pragmatic variation across Spanish(es): Requesting in Mexican, Costa Rican, and Dominican Spanish. *Intercultural Pragmatics* 6: 473–515. [CrossRef]

Fillmore, Charles J. 1975. An alternative to checklist theories of meaning. *Proceedings of the First Annual Meeting of the Berkeley Linguistics Society* 1: 123–31. [CrossRef]

Fleischman, Suzanne. 1989. Temporal distance: A basic linguistic metaphor. *Studies in Language* 13: 1–50. [CrossRef]

Gee, James Paul. 2014. *An Introduction to Discourse Analysis: Theory and Method*, 4th ed. London: Routledge. [CrossRef]

Golato, Andrea. 2003. Studying compliment responses: A comparison of DCTs and recordings of naturally occurring talk. *Applied Linguistics* 24: 90–121. [CrossRef]

Grenoble, Lenore A. 1999. Gender and Conversational Management in Russian. In *Slavic Gender Linguistics*. Edited by Mary Mills. Amsterdam: John Benjamins, pp. 113–30. [CrossRef]

Hancock, Adrienne, and Benjamin Rubin. 2015. Influence of communication partner's gender on language. *Journal of Language and Social Psychology* 34: 46–64. [CrossRef]

Hernández-Flores, Nieves. 2004. Politeness as 'Face' Enhancement. In *Current Trends in the Pragmatics of Spanish*. Edited by Rosina Márquez-Reiter and María Elena Placencia. Amsterdam: John Benjamins, pp. 265–84. [CrossRef]

Holmes, Janet. 1995. *Women, Men, and Politeness*. London: Routledge. [CrossRef]

Hothorn, Torsten, Kurt Hornik, and Achim Zeileis. 2006. Unbiased recursive partitioning: A conditional inference framework. *Journal of Computational and Graphical Statistics* 15: 651–74. [CrossRef]

Hübscher, Iris, Joan Borràs-Comes, and Pilar Prieto. 2017. Prosodic mitigation characterizes Catalan formal speech: The frequency code reassessed. *Journal of Phonetics* 65: 145–59. [CrossRef]

Jaffe, Alexandra. 2016. Indexicality, Stance and Fields in Sociolinguistics. In *Sociolinguistics: Theoretical Debates*. Edited by Nikolas Coupland. Cambridge: Cambridge University Press, pp. 86–112. [CrossRef]

Kádár, Dániel Z. 2019. Introduction: Advancing linguistic politeness theory by using Chinese data. *Acta Linguistica Academica* 66: 149–64. [CrossRef]

Kiesling, Scott F. 2004. Dude. *American Speech* 79: 281–305. [CrossRef]

Kiesling, Scott F. 2009. Style as Stance: Stance as the Explanation for Patterns of Sociolinguistic Variation. In *Stance: Sociolinguistic Perspectives*. Edited by Alexandra Jaffe. Oxford: Oxford University Press, pp. 171–94. [CrossRef]

Koike, Dale A. 1989. Requests and the role of deixis in politeness. *Journal of Pragmatics* 13: 187–202. [CrossRef]

Lakoff, Robin T. 1975. *Language and Women's Place*. New York: Harper and Row.

Le Pair, Rob. 1996. Spanish request strategies: A cross-cultural analysis from an intercultural perspective. *Language Sciences* 18: 651–70. [CrossRef]

Leaper, Campbell, and Rachel D. Robnett. 2011. Women are more likely than men to use tentative language, aren't they? A meta-analysis testing for gender differences and moderators. *Psychology of Women Quarterly* 35: 129–42. [CrossRef]

Levey, Stephen. 2003. He's like 'do it now!' and I'm like 'no!'. *English Today* 73: 24–32. [CrossRef]

Lorenzo-Dus, Nuria, and Patricia Bou-Franch. 2003. Gender and Politeness: Spanish and British Undergraduates' Perceptions of Appropriate Requests. In *Género, Lenguaje y Traducción*. Edited by José Santaemilia. Valencia: Universitat de Valencia/Dirección General de la Mujer, pp. 187–99.

Márquez-Reiter, Rosina. 2000. *Linguistic Politeness in Britain and Uruguay: A Contrastive Study of Requests and Apologies*. Amsterdam: John Benjamins. [CrossRef]

Márquez-Reiter, Rosina. 2002. A contrastive study of conventional indirectness in Spanish: Evidence from Peninsular and Uruguayan Spanish. *Pragmatics* 12: 135–51. [CrossRef]

Márquez-Reiter, Rosina. 2003. Pragmatic Variation in Spanish: External Request Modifications in Peninsular and Uruguayan Spanish. In *A Romance Perspective on Language Knowledge and Use*. Edited by Rafael Núñez-Cedeño, Luis López and Richard Cameron. Amsterdam: John Benjamins, pp. 167–80. [CrossRef]

Márquez-Reiter, Rosina, and María Elena Placencia. 2005. *Spanish Pragmatics*. New York: Palgrave Macmillan. [CrossRef]

Márquez-Reiter, Rosina, Isobel Rainey, and Glenn Fulcher. 2005. A comparative study of certainty and conventional indirectness: Evidence from British English and Peninsular Spanish. *Applied Linguistics* 26: 1–31. [CrossRef]

Mills, Sara. 2003. *Gender and Politeness*. Cambridge: Cambridge University Press. [CrossRef]

Ochs, Elinor. 1996. Linguistic Resources for Socializing Humanity. In *Rethinking Linguistic Relativity*. Edited by John J. Gumperz and Stephen C. Levinson. Cambridge: Cambridge University Press, pp. 407–37.

Pizziconi, Barbara. 2011. Honorifics: The Cultural Specificity of a Universal Mechanism in Japanese. In *Politeness in East Asia*. Edited by Dániel Z. Kádár and Sara Mills. Cambridge: Cambridge University Press, pp. 45–70. [CrossRef]

Qualtrics. 2018. Qualtrics Software. Available online: https://www.qualtrics.com (accessed on 15 January 2018).

R Development Core Team. 2019. *R: A Language and Environment for Statistical Computing*. Vienna: R Foundation for Statistical Computing. Available online: http://www.R-project.org/ (accessed on 1 February 2020).

Rosemeyer, Malte, and Scott A. Schwenter. 2017. Entrenchment and persistence in language change: The Spanish past subjunctive. *Corpus Linguistics and Linguistic Theory* 15: 167–204. [CrossRef]

Sainzmaza-Lecanda, Lorena, and Scott A. Schwenter. 2017. Null Objects with and without Bilingualism in the Portuguese- and Spanish-speaking World. In *Multidisciplinary Approaches to Bilingualism in the Hispanic and Lusophone World*. Edited by Kate Bellamy, Michael W. Child, Paz González, Antje Muntendam and María del Carmen Parafita Couto. Amsterdam: John Benjamins, pp. 95–119. [CrossRef]

Schank, Roger C., and Robert P. Abelson. 1975. Scripts, plans, and knowledge. In *Proceedings of the 4th International Joint Conference on Artificial Intelligence*. Cambridge: Morgan Kaufmann Publishers, vol. 1, pp. 151–57.

Searle, John. 1976. A classification of illocutionary acts. *Language in Society* 5: 1–23. [CrossRef]

Silverstein, Michael. 2003. Indexical order and the dialectics of sociolinguistic life. *Language & Communication* 23: 193–229. [CrossRef]

Spencer-Oatey, Helen. 1996. Reconsidering power and distance. *Journal of Pragmatics* 26: 1–24. [CrossRef]

Spencer-Oatey, Helen. 2000. Rapport Management: A Framework for Analysis. In *Culturally Speaking: Managing Rapport through Talk across Cultures*. Edited by Helen Spencer-Oatey. New York: Continuum, pp. 11–46.

Spencer-Oatey, Helen, and Vedrana Žegarac. 2017. Power, Solidarity and (Im)politeness. In *The Palgrave Handbook of Linguistic (Im)politeness*. Edited by Jonathan Culpeper, Michael Haugh and Dániel Z. Kádár. London: Palgrave McMillan, pp. 119–41. [CrossRef]

Tagliamonte, Sali A., and R. Harald Baayen. 2012. Models, forests, and trees of York English: Was/were variation as a case study for statistical practice. *Language Variation and Change* 24: 135–78. [CrossRef]

Tamaoka, Katsuo, Hyunjung L. Yamaguchi, Yayoi Miyaoka, and Sachiyo Kiyama. 2010. Effects of gender-identity and gender-congruence on levels of politeness among young Japanese and Koreans. *Journal of Asian Pacific Communication* 20: 23–45. [CrossRef]

Terkourafi, Marina. 2005. Beyond the micro-level in politeness research. *Journal of Politeness Research. Language, Behaviour, Culture* 1: 237–62. [CrossRef]

Terkourafi, Marina. 2015. Conventionalization: A new agenda for im/politeness research. *Journal of Pragmatics* 86: 11–18. [CrossRef]

Uchida, Aki. 1992. When "difference" is "dominance": A critique of the "anti−power-based" cultural approach to sex differences. *Language in Society* 21: 547–68. [CrossRef]

Watts, Richard J. 2003. *Politeness*. Cambridge: Cambridge University Press. [CrossRef]

Wright, John W., and Lawrence A. Hosman. 1983. Language style and sex bias in the courtroom: The effects of male and female use of hedges and intensifiers on impression information. *Southern Speech Communication Journal* 48: 137–52. [CrossRef]

Yates, Allison B. 2015. Pragmatic variation in service encounters in Buenos Aires, Argentina. *IULC Working Papers* 15: 128–58.

Indexing Deficiency: Connecting Language Learning and Teaching to Evaluations of US Spanish

Gabriella Licata

Latino and Latin American Research and Studies Center, University of California, Riverside, CA 92521, USA; glicata@ucr.edu

Abstract: The examination of language attitudes towards US Spanish variables unearths indexical meanings rooted in deficit perspectives, particularly in educational contexts. Standard language ideologies undergird pedagogical practice and learning experiences in second language (L2) and heritage language (HL) Spanish classes. The present study utilizes dual research paradigms of social cognition (matched guise technique (MGT); implicit association test (IAT)) to determine if varying experiences with (Spanish) standard language ideologies in academic settings condition bias towards standardized Spanish (SS) and US Spanish (USS) repertoires. L2 and HL students as well as teachers of Spanish ($n = 81$) have more positive associations of SS in both the MGT and IAT, demonstrating that standard language ideologies influence perceptions of language acquisition and academic language learning. No correlations between the bias measures were reported yet attitudes did not differ, suggesting that attitudes are stable and reflected in both early learnings of social information and lived experiences throughout formative education. These results contribute to a growing body of research that examines how monoglossic ideologies reinforce and reproduce the stigma associated with features of US Spanish(es).

Keywords: heritage languagers; deficit perspectives; raciolinguistics; social cognition; language education

1. Introduction

The so-called inherent relationship between standardized language (i.e., an idealized, hypothetical construct of accepted language use; see (Crowley 2003); Lippi-Green 2012) and academic registers has undergirded how language education in the United States is organized (Flores 2020a; Rosa and Flores 2017), despite the fact that a universal acceptance of what 'academic language' is does not exist (R. A. Martínez and Mejía 2020). With this in mind, the indexical field (see Eckert 2008; Silverstein 2018) of US Spanish (commonly referred to as Spanglish; see Otheguy and Stern 2011) in educational contexts is conditioned by what is institutionally considered to be "appropriate" or "academic" (Rosa and Flores 2017). These registers may well be defined by what they are *not*; that is, teachers and curriculum designers may decide what language is suitable in a given context, and thus certain sounds, lexical items, and phrasing considered 'inappropriate' may be viewed unfavorably in academic contexts where they are deemed unsuitable by educators and students alike. A rejection of what Bakhtin (1981) termed *heteroglossia*—the simultaneous use of different kinds of forms or signs (Bailey 2012)—privileges monoglossia and legitimizes the monolingual languager[1] as well as linguistic variables indexing standardized forms (García 2011, p. 115). The standardizations of language production and learning are presented as monoglossic objects that can be impartially measured through standardized testing (Siordia and Kim 2022); however, the over-representation of racialized students in language remediation programs demonstrates how implicit biases condition the perception of their home varieties, as the assessment itself is "reflective of larger settler colonial frameworks embedded in linguistic standards that continue to drive education and language ideologies/practices globally... especially in U.S. schools" (Cioè-Peña 2022, p. 25).

These ideological assumptions generally condition the formative experiences of multilingual youth in US schools whose first or home language is deemed less prestigious or less useful by hegemonic standards. The institutional negation of their dynamic language abilities results in a process that co-constructs the pathologization of their communication and the racialization of their identities (Rosa and Flores 2017). These damaging co-constructions of a student and their linguistic repertoire may be mediated by other social factors, including gender and economic class (Solorzano and Yosso 2001). Resultantly, raciolinguistic ideologies are hegemonically sustained, as white-perceiving subjects (Rosa and Flores 2017) view the indexical meanings of Latinx students' language repertoires through deficit lenses.

The results of these ideologies manifest in the homogenization of language teaching and learning, paving the way for the commodification of language through a standardized scope and sequence of grammar and vocabulary (VanPatten 2015, p. 10) that conceptualizes language variation as exceptional and is marked in comparison to the standardized repertoire, which is more fixed (Train 2020). As such, Spanish as a language of instruction in the United States has undergone a hyperstandardization (Train 2003) that ignores empirically observed community use of language, particularly varieties in the United States. For example, in Northern California (and beyond), the term *rentar* and *alquilar* ("to rent") are synonymous, but the latter is the term often found in textbooks as the standardized or academic option with perhaps *rentar* denoted as a 'colloquial' or 'informal' option, though the latter is in widespread use in many domains. As a result, students who grow up using Spanish in the household are expected to 'master' an academic or prestige variety that stands in contrast to their home and community language practices, which in many regards depicts the latter varieties as 'incomplete' or 'insufficient' (see Rosa and Flores 2017).

The separation of home from school varieties upholds standard language ideologies and also relies on the notion that students can and should keep these repertoires separate. The distinctiveness of named languages—varieties that undergo codification and are labeled (often related to nation building; see García 2019)—is used to create language boundaries in the classroom, which ignores a "usage-based account of language instead which does not take linguistic systems for granted but construes them based on the single linguistic unit" (Wasserscheidt 2021). This damaging discourse has undergirded research concerning US Spanish languagers while ignoring the dynamic and lived reality of *translanguaging*, or when one utilizes their full linguistic repertoire in communication (García and Otheguy 2020; García et al. 2021). While codeswitching examines how multiple named languages may be used in an utterance, translanguaging is "a perspective that prioritizes the meaning-making of bi/multilinguals in ways that challenge normative language ideologies" (Flores 2020b).

As a result, multilingual students' abilities to dynamically language are ignored and undervalued as demonstrative of communicative ability because they do not communicate within the bounds of named languages. Thus, heritage language[2] programs are often geared towards preparing students to achieve double monolingualism (Heller 2006; Rosa and Flores 2017). Second language programs alike aim to have students develop a hyperstandardized "language-elsewhere"(Mena 2022), or rather a variety of Spanish disembodied from the surrounding communities. In understanding how these damaging ideologies are epistemologically situated in the classroom and in language learning, experimental sociolinguistics rooted in anthropological frameworks that investigate language ideologies can inform us of the explicit and implicit biases present in education. In viewing schools as one of the principal sites where damaging language ideologies are born and perpetuated, I investigate how early learned bias and Spanish language training may influence perceptions of US Spanish features that are historically stigmatized to index 'deficits.' The present study utilizes dual models of social cognition to determine if the differentiated experiences of second language learners, heritage learners, and Spanish language teachers have differing bias towards standardized Spanish and US Spanish repertoires. This experimental line combines the matched guise technique (Lambert et al. 1966; MGT) and implicit association test (Greenwald et al. 1998; IAT) to determine if any groups are

exhibiting attitude shifts via a divergence of more malleable attitudes (elicited in the MGT) from deeply rooted implicit bias (elicited in the IAT). Findings reveal that all groups' bias measures favor standardized Spanish over US Spanish, assigning, to the former, positive qualities and associations related to complete acquisition and academic quality. However, bias measures did not correlate, despite their similar trajectories, which indicates that in this study, explicit attitudes elicited from the MGT are stable and similar to those elicited from the IAT, despite being processed differently.

Results offer new insights towards the deficit perspectives that plague US Spanish languagers and the lack of legitimization of US Spanish(es) in academic settings. Accordingly, I urge educators to re-examine how traditional models of language teaching maintain raciolinguistic ideologies and deficit perspectives in and outside of the classroom.

2. Colonial Underpinnings of Language Separateness

A brief examination of colonial language epistemologies aids us in understanding how multilingual languagers are framed through deficit perspectives. Initial linguistic research legitimized US Spanish languagers and their repertoires by developing a typology of codeswitching (Poplack 1980). This research was groundbreaking at the time, as it affirmed that US Spanish languagers (and other multilinguals) could crosslinguistically navigate Spanish and English with ease, demonstrating that grammatical constraints of more than one language could be learned and employed simultaneously. However, the countless sociolinguistic and anthropological studies that followed and sought to strengthen positive positions of multilingualism have not thwarted the deficit perspectives that plague racialized students in the US (Rosa and Flores 2017).

These views rely on an idealization of codified language, or standard language ideology, defined as "bias toward an abstracted, idealized, homogeneous spoken (and written) language which is imposed and maintained by dominant bloc institutions" (Lippi-Green 2012, p. 64), which in the US is undergirded by the politicization and economization of language. A prominent tool of unabashed colonialism, standard language ideology is a "social construct of the nation-border" (Martínez 2003), reinforcing language boundaries and rejecting the naturalness of translanguaging through the intentional subordination of nonhegemonic and/or nonstandardized languages. Resultantly, these constructs erase linguistic and cultural heterogeneity, facilitating the convenient placement of languages and people into specific demographic categories, and serve as remnants of colonial organization (also known as coloniality of power; see Quijano 2000) that sought and continue to maintain the compartmentalization and homogenization of languages to elevate the idealized speaker.[3] This invalidates those whose forms of communication are weaponized, mocked, and potentially erased, resulting in monoglossic ideologies that are crystallized and interwoven into hegemonic institutions (Rosa and Flores 2017).

The limitations to understanding multilingual languagers' dynamic process result in models of separateness that attempt to codify a languager's production by how frequently they use one variety or the other, such as the matrix language frame (see Myers-Scotton 1997). Socio- and psycholinguistic researchers who describe the alleged 'deficits' of bi- or multilinguals position monolingualism as the norm and Spanish and English as distinct structural and cognitive entities, a process that reproduces and reiterates the identification of people and places with bordered territories (García 2019; Otheguy et al. 2015). The epistemological positioning of languages as separate cognitive systems paves the way for the theorization of multilinguals as *semilingual*, meaning that they have not fully acquired one or more of the named languages under examination (Martin-Jones and Romaine 1986), and many language programs seek to 'remedy' this supposed deficit.

Spanish Language Education in the United States

A deep examination of Spanish language learning models in US contexts reveals colonial epistemologies that situated Spanish language teaching and learning programs in ideologies of appropriateness in both heritage and L2 classes. Burns (2018) examines Span-

ish 'foreign' language pedagogy (textbooks and direct instruction) in a research university, finding that the reproduction of standard language ideologies and variation erasure served to, at times, explicitly delegitimize US Spanish. The very label of 'foreign' ignores and erases the long history of the Spanish language in present-day US territories (Lozano 2018; Train 2007). Similarly, Spanish language education is designed and modeled around an idealized white speaker (Flores and Rosa 2022), as language at the intersection with *latinidad* often exceptionalizes Blackness and indigeneity and erases colonial histories (Chávez-Moreno 2021), which confounds representation in course materials (Anya 2020, 2021; Austin 2022; de los Heros 2009; Padilla and Vana 2022). The consequences of these epistemologies have resulted in the homogenization of Spanish language teaching that privileges the "language-elsewhere" (Mena 2022), and US Spanish languagers who deviate from this academic hyperstandard (Train 2007) are often unjustly classified as having less social and economic capital (Alonso and Villa 2020), suffering from "word gaps" (Avineri et al. 2015), being designated 'long-term English learners' (Flores et al. 2015), and not 'belonging' to any culture or place (Anzaldúa 1987), reiterating white supremacy on every front.

As such, the damage of monoglossia and deficit perspectives affects those who are racialized by hegemonic systems that feign impartiality, for example, standardized testing (Siordia and Kim 2022). Teachers also reiterate and reproduce language ideologies, much in line with the demands of various institutions (e.g., school administrations and federal and state requirements), which trickle down to students and are reiterated in new contexts. However, scholars, language activists, and educators have emphasized the need to move away from these harmful ideologies by dissolving the artificial boundaries between languages and between identities to instead value the natural occurrence of translanguaging and multiculturalism (García et al. 2021). Researchers report on the benefits for students of counter-hegemonic language ideology in heritage language planning (Loza 2017), sociolinguistic awareness and anti-racist training in teacher preparation (Seltzer 2022), and translanguaging in both cognitive and affective domains (Carstens 2016) to reduce deficit perspectives imposed on US Latinx students. Nonetheless, many others maintain that despite decades of teacher training and education, deficit perspectives continue to dominate language learning narratives, maintaining systemic barriers to language access and expression and placing the onus on individuals to change (Flores et al. 2018). The examination of explicit to implicit biases can shed light on these inequities.

3. Evaluating Implicit Social Cognition and Linguistic Bias

Quantitative perception studies can elucidate covert language biases that might not otherwise be expressed when participants are explicitly questioned. Studies in sociolinguistic perception have demonstrated that perceiving subjects do not judge speech utterances in isolation, but rather, they develop ideas about language in conjunction with ideas about who is using the language (Rosa and Flores 2017) and the perceived social information associated with the languagers that precedes language judgment (Kang and Rubin 2009). Thus, individual perceiving subjects may reproduce and reiterate broad social stereotypes in their evaluations with little information on the person (see Inoue 2003). Experimental sociolinguistics has sought to reveal biases ranging from those that are explicitly expressed to those that are more automatic or non-verbalizable. Implicit attitudes are considered to be more stable and unchanging, as they are acquired and learned slowly in earlier socialization, while explicit attitudes are deemed more malleable, susceptible to external influence and fast learning, thus regulated, purposeful, and effortful (Petty et al. 2009; Evans 2008; Karpen et al. 2012). To elicit these biases as they relate to language, a variety of attitude assessments have been utilized in sociolinguistic study. Direct methods involve asking participants to answer questions from which explicit bias is displayed. A common and important measure for gauging attitudes, the direct approach is often employed in questionnaire or interview form and is a long-held research paradigm (Garrett 2001, p. 159), and information gathered from direct questioning has been useful in language planning (p. 159).

A common indirect methodology used for gauging attitudes in linguistic research is the *matched guise technique* (MGT; Lambert et al. 1966; Kircher 2016). Participants listen to voices and evaluate them based on a series of social scales (for example, *This person sounds friendly*). The MGT is considered to be an indirect approach to collecting attitudes because participants may be aware that their attitudes are being tested but are not sure of the precise attitudinal object. Thus, they may rate the variants of a linguistic variable differently on a number of characteristics that reveal that one is stigmatized (Solís Obiols 2002). They are also unaware that the attributes they apply to each speaker will later be interpreted as ideological stances towards the linguistic variable (Garrett 2001). MGT studies in small language communities have shown that listeners will issue discrepancies in judgment even when they know that they are listening to the same speaker using different language forms (Soukup 2013) and may even issue more severe ratings when they are made conscious of the fact that they are rating speech with stigmatized language variants (Rosseel et al. 2019).

Attitudinal research assessing both explicit and implicit biases towards US Spanish has shed light on how biases shift depending on the indexical field (see Section 4). In keeping with the notion that "linguistic and social information comes packaged in a single complex signal" (Craft et al. 2020, p. 390), a linguistic variable's indexical field is also shaped by nonlinguistic information perceived during the process of meaning making (Eckert 2008). In fact, perceiving subjects may already make their linguistic judgments of languagers before experiencing their language production, demonstrating how *reverse linguistic stereotyping* (Kang and Rubin 2009) and raciolinguistic ideologies (Rosa and Flores 2017) motivate perceptual categorization. As such, linguists have sought to reveal automatic biases that are less susceptible to manipulation, incorporating the *implicit association test* (IAT; Greenwald et al. 1998; Lane et al. 2007), a research paradigm that estimates the strength of quickly accessed mental associations between concepts (e.g., Spanish/English) and attributes (e.g., Good/Bad) by measuring the differences in response latency to each concept with one of the specific attributes; that is, if a person responds more quickly to the association of English + Good than Spanish + Good, then they have a stronger positive association between the first two.

Different processing models of social cognition help us to understand if there exists a correlation of implicit to explicit attitudes (Lane et al. 2007). As the likelihood of a correlation between different measures of social cognition varies greatly from one environment to another (Fazio and Olson 2003), sociolinguistic research has accordingly exploited this variability of outcomes, demonstrating how spaciotemporal factors may determine a convergence (McKenzie and Gilmore 2017) or a divergence (M. Babel 2010; Pantos and Perkins 2013; Calamai and Ardolino 2020) of evaluations. Those that demonstrate divergence shed light on possible attitude changes in progress (McKenzie and Carrie 2018), exhibiting how linguistic variables can index different meanings depending on the type of social cognitive process accessed. This lends support to the implicit–explicit attitudinal discrepancy hypothesis (IED), which indicates that a divergence of explicit from implicit attitudes can be attributed to "an attitude change in progress at a given point in time" (Karpen et al. 2012; McKenzie and Carrie 2018). The IED hypothesis accounts for the evidence that long-held implicit evaluations generally remain stable, even if explicit attitudes about a concept have changed significantly (e.g., a smoker who no longer smokes may continue to hold positive implicit bias towards the habit). Thus, more divergence of explicit from implicit attitudes indicates that the explicit attitudes under examination are changing from early learned biases, while lower levels of divergence indicate stability across explicit to implicit attitudes (McKenzie and McNeill 2022, p. 21).

Language bias studies assessing explicit and implicit biases have utilized direct and indirect attitude elicitation (respectively) to assess correlation and potential divergences; that is, direct questioning and—most commonly—the IAT, finding weak relations between explicit and implicit attitude results (McKenzie and Carrie 2018). This study aims to understand how attitudes indirectly derived from the MGT diverge from or converge with bias elicited from the IAT to determine stable or changing attitudes among heritage

and L2 learners and teachers. I seek to demonstrate if language learning experiences condition implicit bias towards a standardized Spanish and a US Spanish repertoire, and whether implicit biases are stored differently, providing new insights into the malleability of bias derived from indirect methods and the possibility of experience to condition later learned biases. The next section uses deficit and raciolinguistic perspectives to examine the indexical field of US Spanish features.

4. Indexing Deficits

The perceiving subject (Inoue 2003, 2006) does not perceive all Spanish languaging subjects equally. That is, features of US Spanish index different meanings depending on relevant nonlinguistic information that perceivers use to make their judgments, and who is considered to have a linguistic deficit is subject to scrutiny. Perceiving subjects are capable of quickly identifying socially salient linguistic variables via analyses of verbal communication (i.e., signed or spoken language) in conjunction with nonlinguistic social information about the languager (Campbell-Kibler 2009, 2010; Drager and Kirtley 2016; Hay and Drager 2010). The ideologies that bar access to positive associations to an indexical field may be mediated by social factors; that is, societal constructs associated with the languager like gender, economic class, and ethnicity can delimit access to an indexical field, and speakers may (sub)consciously choose to eschew variants in certain instances where they may face scrutiny (see Chappell 2016). For example, if US Spanish languagers are constantly corrected or criticized for their use of so-called broken Spanish, they may avoid the use of US Spanish features in environments where they do not feel safe using them (e.g., the classroom). Thus, the hegemonic pressures that underpin language use can privilege or oppress the linguistic behavior of certain groups, providing some with liberty to navigate the positive meanings of an indexical field and others not.

Deficit perspectives and monoglossic ideologies condition individual and collective attitudes about US Spanish throughout the United States. Rangel et al. (2015) used the matched guise technique to gauge the attitudes of multilingual languagers towards standardized Spanish, standardized American English, and US Spanish in southern Texas, finding that participants rated US Spanish unfavorably. Both standardized American English and standardized Mexican Spanish—when juxtaposed with a codeswitching repertoire—elicited more positive evaluations from participants in the towns of Laredo and Edinburgh (Texas) with regard to solidarity, status, and personal appeal, though participants offered more positive solidarity ratings towards standardized Mexican Spanish over English, demonstrating long-held community ties with a prestige variety of Spanish. The examination of particular linguistic variables within US Spanish(es) also sheds light on the indexical meanings available to languagers. Heritage languagers have demonstrated similar attitudes to monolingual speech with respect to specific variables, such as aspiration (Chappell 2021a). However, in Chappell's (2021b) examination of Mexican American and Mexican languagers' reactions to voices of the aforementioned groups, language experience and identity conditioned perceptions. That is, Mexican Americans had more gradient and nuanced opinions towards concepts like 'bilingualism' and 'language proficiency' than Mexicans, likely due to their experiences navigating different cultural and linguistic communities, while the latter groups' positions were rooted in epistemologies of language separateness and hierarchy that created more exclusive language groups.

Relatedly, identity has been shown to condition perceptions in the classroom. In a matched guise test, both heritage and L2 students rated monolingual Spanish speakers more positively than heritage and L2 speakers, with heritage speakers being described as using the "'least formal' and 'incorrect' variety in comparison to the L2 variety due to dominant stereotypes and ideologies, and the incorporation of lexical characteristics of US Spanish" (Vana 2020, p. ii). This demonstrates how the stigmatization of lexical features of US Spanish, when juxtaposed with standardized features, may be explicitly or implicitly taught to L2 learners. Similarly, as L2 (Spanish) instruction increases attention towards sociolinguistic competence (Van Compernolle and Williams 2012; Sun 2014), understanding

L2 learners' attitudes towards linguistic variables sheds light on how exposure to variation and both linguistic and cultural proficiency can mediate the social salience of variants and attitudes towards them (Chappell and Kanwit 2022). Quan (2020) found that despite the positive effects of integrating Critical Language Awareness (CLA) materials in a Spanish language course for L2 learners (e.g., compassion for Latinx communities), students still explicitly expressed "ongoing agreement with standard language ideologies related to bilingualism and foreign accents" (p. 915), evidencing the importance of CLA in all language courses and a restructuring of language education. Attitudinal study can also be used to critically investigate teacher bias towards the language variation their students employ. In a matched guise test examining teacher attitudes towards syntactic and lexical features of US Spanish, teachers in general expressed negative attitudes towards the utterances, describing them as 'incorrect' even when they were not sure how to 'correct' them (Román et al. 2019). Bilingual teacher candidates in South Texas similarly reproduce harmful standard language ideologies in their own teaching practices. These were revealed to be reiterations that are learned in their home from their mothers (i.e., their language socializers) due to the societal and institutional linguistic violence that they themselves faced in youth (Ek et al. 2013).

As increased attention is given to the incorporation of implicit attitude measures in the study of (US) Spanish(es), linguists have begun incorporating the implicit association test in attitude studies. Callesano and Carter (2022) found that, when Spanish and English were contrasted in the IAT, Miami-based participants more quickly associated English + Good over Spanish + Good, an effect that was mediated by increased time spent in Miami, demonstrating how binaries shift when hegemonies shift. Similarly, Ianos et al. (2020) issued an IAT and direct questionnaire to adolescents to evaluate bias towards Catalan and Spanish. IAT results showed an overall positive bias towards Catalan over Spanish, though positive bias shifted when tested against the adolescent's home language (i.e., Spanish used at home produced positive associations with Spanish). These findings demonstrate how the strength of solidarity and institutional support can combat hegemonic pressures to assimilate at the local or regional level. To date, the IAT research paradigm has not been used to examine biases towards lexical features of US Spanish (or Spanglish), specifically in comparison with explicit attitudes. This study adds to the existing literature by using an innovative research design to evaluate the effects of language training on attitudes towards a standardized Spanish and a US Spanish repertoire, measuring and comparing bias via different models of social cognition. I seek to answer the following questions:

1. Are US Spanish lexical features salient to groups that have different experiences with Spanish language learning, and what do they index in terms of perceived acquisition and academic-ness when juxtaposed with standardized Spanish lexical features?
2. To what extent is differential participation in language education programs (L2 students, heritage students, and teachers) correlated with more positive explicit attitudes toward US Spanish features?
3. Do the attitudes and associations elicited from indirect methods (i.e., the MGT) and automatic response (i.e., the IAT) indicate that explicit and implicit biases result from distinct cognitive processes?

To address these questions, I investigate explicit and implicit biases among heritage and L2 learners and teachers with the goal of determining if any group holds more favorable bias towards US Spanish, which would signal a greater acceptance of variation, or if they continue to uphold deficit perspectives.

5. The Present Study

Participants completed all sections of the experiment in Qualtrics (Qualtrics, Provo, UT), an online survey platform, with IAT integration using Iatgen (Carpenter et al. 2019). They began with a short demographic survey (see Section 5.2) and were able to complete the experiment in 20 to 30 min. Participants were also given breaks in between experiment blocks.

5.1. Participant Recruitment

Eighty-one participants that qualified into one of the following groups completed the experiment: students who had taken either heritage or L2 Spanish classes and teachers of Spanish. Heritage learners were either born in the US or immigrated to the US by the age of five and had taken at least 1 year of heritage Spanish language education. L2 learners began learning after the age of 14 and had taken at least 2 years (four semesters) of Spanish as a second language in the US. Teachers[4] based in the US of any level were invited to participate, including primary (*n* = 3), secondary (*n* = 10), and postsecondary (*n* = 10). Participants were recruited via the crowdsourcing site Prolific (www.prolific.co) and through word of mouth (particularly with the teacher group, in which a limited number of Prolific users fit the requirements). The number of participants and demographic information are found in Table 1.

Table 1. Participants' demographic information and language background.

n = 81	Gender Identification	Age Range	Birthplace	L1 Spanish	Mean Years of Spanish Language Learning in School
Heritage *n* = 28	Female: 16 Male: 9 Nonbinary: 2 Gender fluid: 1	18–25 = 16 26–35 = 9 36–45 = 2 46–55 = 1	Argentina: 1 Cuba: 1 Guatemala: 2 Mexico: 1 Puerto Rico: 1 USA: 22	Yes: 28 No: 0	7.2
L2 *n* = 30	Female: 19 Male: 8 Nonbinary: 2 Fluid feminine: 1	18–25 = 13 26–35 = 6 36–45 = 6 46–55 = 2 56–60 = 3	Romania: 1 South Korea: 1 USA: 28	Yes: 0 No: 30	6.2
Teacher *n* = 23	Female: 17 Male: 6	18–25 = 1 26–35 = 8 36–45 = 8 46–55 = 5 56–60 = 1	Italy: 1 Mexico: 1 Peru: 1 Spain: 2 USA: 18	Yes: 11 No: 12	7.6

5.2. Stimuli and Design

5.2.1. Experiment 1: Matched Guise Test

Ten speakers provided the audio, five male-identifying and five female-identifying. All were college students or recent graduates in their twenties who were born in California or came to California at an early age (before age 10). All learned Spanish in the household and English in early childhood or learned both Spanish and English in the home, and all speak a variety of US Spanish that was most influenced by Mexican or Guatemalan Spanish. Speakers read a short passage giving directions to a familiar person (see Appendix A), either in standardized Spanish (SS) typical of that which a student might learn from a United States Spanish textbook or a US Spanish repertoire (USS). The USS guise passages contained lexical and lexicalized items common to Spanish(es) in contact with American English in the United States. Passages were written by the author, edited by both Northern and Southern Californian US Spanish languagers, and judged by three Californian Spanish languagers. Two individuals who produced the designated guises (one male and one female) read the same passage in both SS and USS, providing four target guises, which were separated using the audio of the other eight speakers, who each read one passage, serving as 'fillers' to distract listeners from the similarity of voices across the designated guises. Speakers were told to read the passages as naturally [5] as possible, which resulted in all using Spanish-like phonology in their renditions (e.g., 'ticket' pronounced similarly to ˈtiket as opposed to ˈtɪkɪt). The audio sequence is visualized in Figure 1, which contains the two guises, one in each language variety (male, positions 3 and 9; female, positions 6 and 12). Audio samples were cleaned (i.e., background noise removed) in Audacity

(V.3.0.0). The audio samples were organized for a single group of judges (Stefanowitsch 2005), meaning that all participants heard and rated all audio samples (four compared guises and eight fillers) in a within-subjects design for a total of 324 observations.
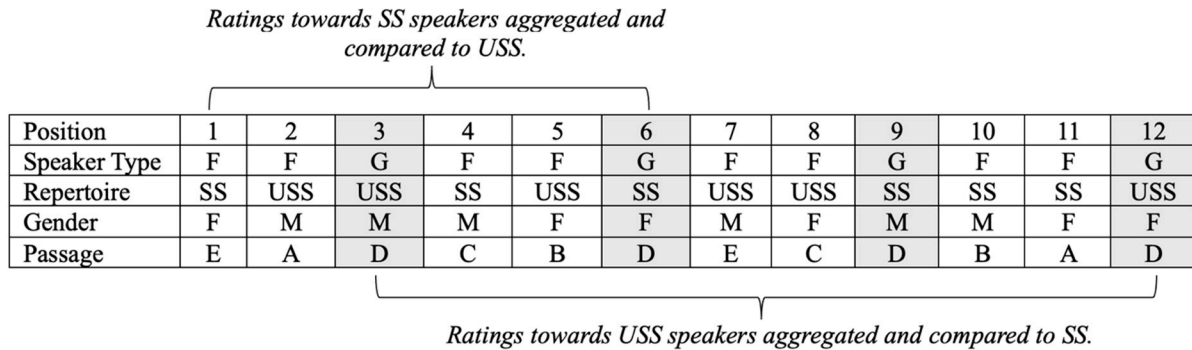
Ratings towards SS speakers aggregated and compared to USS.

Position	1	2	3	4	5	6	7	8	9	10	11	12
Speaker Type	F	F	G	F	F	G	F	F	G	F	F	G
Repertoire	SS	USS	USS	SS	USS	SS	USS	USS	SS	SS	SS	USS
Gender	F	M	M	M	F	F	M	F	M	M	F	F
Passage	E	A	D	C	B	D	E	C	D	B	A	D

Ratings towards USS speakers aggregated and compared to SS.

Figure 1. Matched guise sequence for single group of judges. (Position = Audio in ordered sequence; Speaker Type = Filler (F) and Guise (G); Repertoire = Standardized Spanish (SS) and US Spanish (USS); (Speaker) Gender = Female (F) and Male (M); Passage = Story read (A–E; see Appendix A)).

Participants were prompted to evaluate the speakers on the six social scales below. The language varieties under examination were not mentioned explicitly so as to not prime the participants, aside from prompt four, which avoided mentioning a specific country that may elicit other biases. Participants addressed how much they agreed or disagreed with each statement using a six-point Likert scale[6] (see Figure 2). The prompts are as follows:

1. This person speaks fluently.
2. This person has not fully acquired their language.
3. This person learned their language not only through speaking, but also through reading and writing.
4. This person could communicate easily in a Spanish speaking country.
5. This person is still learning their language.
6. This person would be able to use their language in a professional environment.

S4.Q4

This person learned Spanish not only through speaking, but also through reading and writing.

I agree ○ ○ ○ ○ ○ ○ I **disagree**

Figure 2. Sample matched guise test question with a six-point Likert scale.

5.2.2. Study 2: Implicit Association Test

Two IAT tests that utilized written labels were issued to participants. The concepts must be relatable to participants completing the experiment; with this in mind, SS is labeled as Spanish and USS is labeled as Spanglish. The first IAT test assessed the strength of associations between the concepts of Spanish and Spanglish with the set of attributes Good/Bad; the second contains the same concepts with the attributes of Academic/Not Academic. The concepts are subjected to named language practices for two reasons: (1) laypeople adhere to the separateness of languages in explicit understanding and are less likely to comprehend what a 'standardized' or 'nonstandardized' repertoire means in cognitive and social terms, and (2) the cognitive load in IAT tests must be reduced for participants to complete the task in a timely manner. The exemplars chosen for Spanish/Spanglish are high-frequency words and typical of a standardized Spanish repertoire taught in a US language classroom and of US Spanish varieties in California, respectively. The attributes present positive/negative binaries, with five exemplars in each category (see Table 2). The chosen exemplars are those

often associated with 'complete' forms of communication (i.e., Good/Academic) and those associated with 'incomplete' or 'broken' forms of communication (i.e., Bad/Not Academic).

Table 2. Concepts, attributes, and respective exemplars for both IAT experiments.

IAT#	Names	Exemplars
#1, #2	Spanish	*sin embargo* ('however'), *la camioneta* ('truck'), *el almuerzo* ('lunch'), *alquilar* ('to rent'), *las facturas* ('bills')
#1, #2	Spanglish	*pero like* ('however'), *la troca* ('truck'), *el lonche* ('lunch'), *rentar* ('to rent'), *los biles* ('bills')
#1	Academic	books, tests, school, formal, scholar
#1	Not Academic	slang, informal, street, uneducated, illiterate
#2	Complete	entire, full, intact, whole, perfect
#2	Not Complete	deficient, lacking, fragmented, imperfect, partial

Both IATs are composed of seven blocks. Blocks 1, 2, and 5 are practice trials that consist of sorting the exemplars with their concept (1 and 5) or attribute (2) head to get acquainted with the terms. Trials 3 and 4 pair Spanish with the 'positive' attribute (Spanish + Good/Academic) and Spanglish with the 'negative' attributes (Spanglish + Bad/Not Academic). In accordance with best practices attested in IAT tests (Lane et al. 2007), concept labels are positioned in the upper left- or right-hand corner of the frame with the attribute below, with the exemplar presented in the center of the screen (see Figure 3).

(**a**)   (**b**)

Figure 3. Screenshots from IAT #1. (**a**) Screenshot from Block 3, test trial, sorting positive and negative attributes to Spanglish + Bad and Spanish + Good. (**b**) Screenshot from Block 4, test trial, sorting language exemplars to Spanglish + Bad and Spanish + Good.

5.3. Statistical Models

The responses from the matched guise test were converted into numerical values, meaning that a positive evaluation receives a rating of '6' (e.g., *this person has fully acquired their language*) and negative evaluation '1' (e.g., *this person has not fully acquired their language*), and these ratings were centered. Raw data were then submitted to an exploratory factor analysis (EFA; Helms 2020), a model employed to reveal latent factors by examining the underlying correlations among scales. EFA helps to reduce the number of correlated measures to a set of latent social variables, which serve as the dependent variables in the regression models. A loading value of >0.4 indicates that a variable 'belongs' to the factor and is reported as such (Weatherholtz et al. 2014).[7]

The data were then submitted to mixed-effects linear regression models using the *lmerTest* package (Kuznetsova et al. 2017) in R (R Core Team 2018) with the predictors of

the participant group (heritage; L2; teacher), the guise variety (SS; USS)[8], and potential interactions among the two independent variables. An individual participant was included as a random effect.

All IAT score information was calculated using Iatgen (Carpenter et al. 2019). Participants' response latencies provided in the two IATs were converted into *D* scores (Greenwald et al. 2003; Lane et al. 2007), a measure of the within-subject difference between the compatible and incompatible block means, divided by a pooled standard deviation. Each participant's *D* score represents the subtle differences in effect size. *D* scores range from −2.0 to 2.0, whereby zero represents no difference in response latencies between conditions, a positive score indicates bias towards the expected 'compatible' pairing (i.e., Spanish + Good/Academic), and a negative score signifies bias towards the 'incompatible' pairing (i.e., Spanglish + Good/Academic).

Correlation analyses of the ratings of each grouped dependent variable from the MGT and results from each IAT were carried out in R using the Pearson correlation formula to determine whether linear relationships exist between the implicit bias measures derived from both tests.

6. Results

6.1. Eliciting Attitudes from the MGT

The raw MGT data submitted to EFA provided two salient factors as determined using Cattell's scree plot. Four evaluative scales (Q1, Q2, Q4, and Q5) grouped together into Factor 1, which pattern together under perceived *acquisition* of the language variety by the speaker. Two scales (Q3 and Q6) are grouped into Factor 2, as they patterned similarly under the perceived attribute of *academic-ness* of the language learned and produced (i.e., learned the language in school); see Table 3.

Table 3. Loadings of rating scales for guises in EFA on Factor 1, *acquisition*, and Factor 2, *academic-ness*. Loadings above an absolute value of 0.4 are bolded.

Rating Scale	Factor 1: *Acquisition*	Factor 2: *Academic-Ness*
Q1. Is fluent	**0.7**	0.49
Q2. Incomplete acquisition	**0.84**	0.26
Q3. Learned by speaking, reading, and writing	0.22	**0.55**
Q4. Speaks 'globally understood' language	**0.6**	0.51
Q5. Done learning language	**0.59**	0.24
Q6. Learned language in school	0.56	**0.77**

The regression model fit onto the factor *acquisition* demonstrates a main effect of the guise variety, and all participant groups rated SS more positively than USS ($p < 0.0001$), meaning that when speakers employed SS, they were perceived to have acquired their language more fully than when they employed a USS repertoire (see Table 4 and Figure 4).

Table 4. Summary of Mixed Effects Linear Regression Model Fitted to perceived *acquisition*.

	β Coefficient	Standard Error	*t*	*p* Value
(Intercept)	−0.592	0.111	5.314	0.0001
Participant—L2	0.322	0.155	2.082	0.03
Participant—Teacher	0.229	0.166	1.381	0.2
Guise Language—SS	0.885	0.127	6.974	0.0001

The intercept for this model is heritage participants evaluating USS speakers. Positive β values indicate that the participant evaluated the speaker as demonstrating more complete acquisition of their variety. The estimated variance of the random effect of the listener is 0.122.
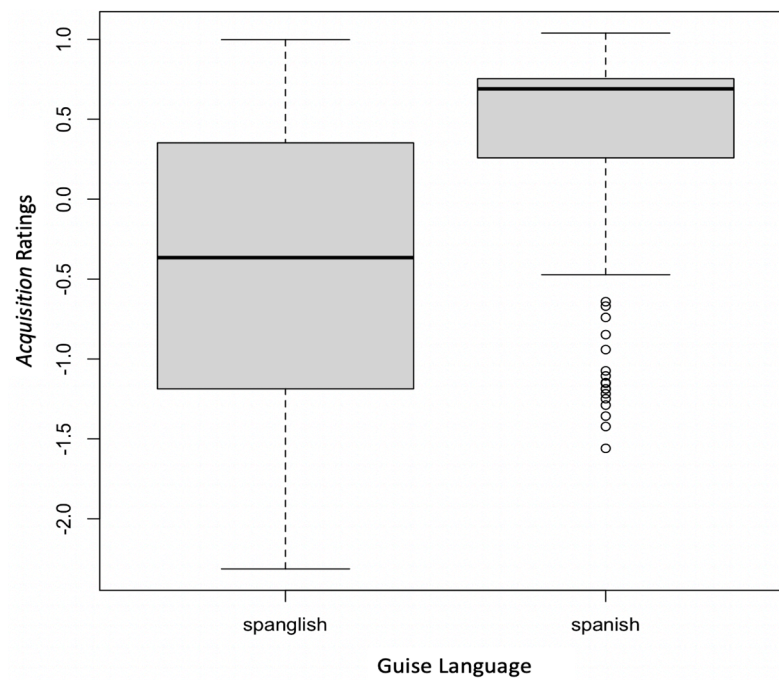
Figure 4. Boxplot showing listener evaluations of speakers' perceived *acquisition*, conditioned by the guise variety.

The regression model fit onto the *academic-ness* factor (see Table 5) once again demonstrates a significant main effect of the guise variety ($p < 0.0001$), as all participant groups rate USS less favorably than SS on the basis of perceived *academic-ness* of the speaker's learned variety (see Figure 5).
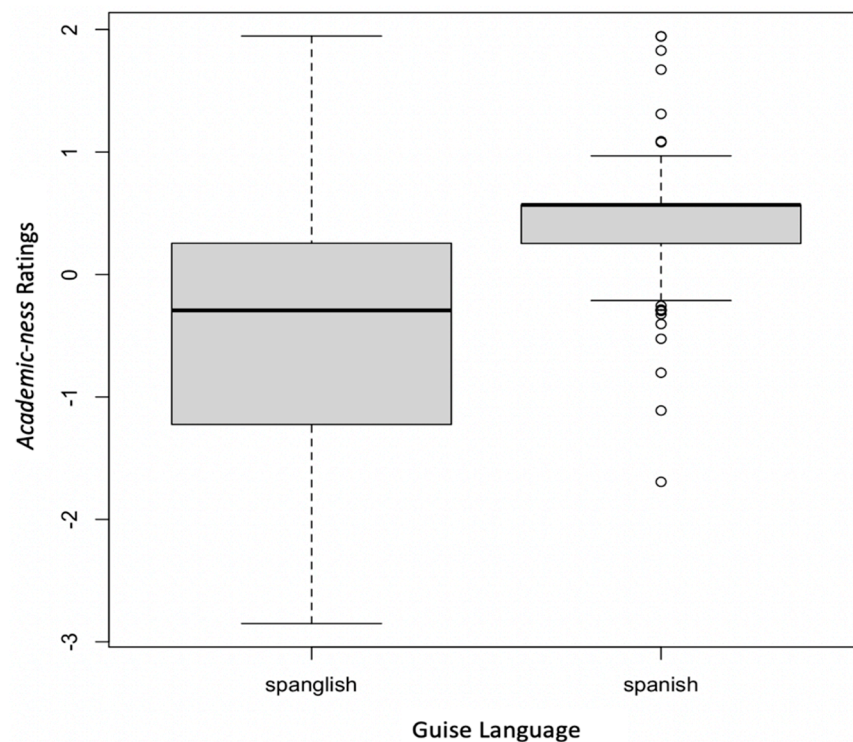

Figure 5. Boxplot showing listener evaluations of speakers' perceived *academic-ness*, conditioned by the guise variety.

Table 5. Summary of Mixed Effects Linear Regression Model Fitted to perceived *academic-ness*.

	β **Coefficient**	**Standard Error**	*t*	*p* **Value**
(Intercept)	0.591	0.1052	−5.615	0.0001
Participant—L2	0.213	0.146	1.454	0.148
Participant—Teacher	0.287	0.157	1.834	0.06
Guise Language—SS	1.0241	0.121	8.48	0.0001

The intercept for this model is heritage participants evaluating USS speakers. Positive β values indicate that the participant evaluated the speaker's variety as having been acquired in an academic setting. The estimated variance of the random effect of the listener is 0.116.

6.2. Assessing Implicit Bias

In the Spanish/Spanglish + Good/Bad IAT (see Table 6), positive *D* scores indicate that all participant groups associate Spanish more strongly with 'Good' than Spanglish due to their faster reaction times when categorizing exemplars to Spanish + Good than to Spanglish + Good (see *p* values). The timeout rate was low at <0.008%, the drop rate (number of participants dropped for overly fast responding under 300 ms) was zero for the heritage and L2 groups and one for teachers, and the error rate (percent of trials that were incorrect) was under 0.09%.[9] A one-way ANOVA modeled to these results suggests that the L2 group had significantly lower *D* score means than the heritage and teacher groups.

Table 6. Spanish/Spanglish + Good/Bad IAT data information for all participant groups.

Participants, *n*	*D* **Score Mean**	*t***-Test**	*p***-Value**
Heritage, 28	0.51	10.39	0.00001
L2, 30	0.35	4.1	0.0003
Teacher, 22	0.49	4.32	0.0003

In the Spanish/Spanglish + Academic/Not Academic IAT (see Table 7), positive scores determine that participants associate Spanish more strongly with Academic due to their faster reaction times when the two were paired. These reaction times differed significantly from the reaction times when Spanglish was paired with Academic (see *p* values). The timeout rate was low at <0.002% and the drop rate was zero for the heritage and L2 groups and one for teachers. The error rate was low at 0.08%. A one-way ANOVA once again demonstrates that the L2 group had significantly lower *D* score means than the heritage group.

Table 7. Spanish/Spanglish + Academic/Not Academic IAT data information for all participant groups.

Participants, *n*	*D***-Score Mean**	*t***-Test**	*p***-Value**
Heritage, 28	0.63	9.43	0.00001
L2, 30	0.54	7.29	0.00001
Teacher, 22	0.58	7.27	0.00001

6.3. Correlation Analyses: MGT to IAT

Correlation analyses were carried out to determine the relationship of MGT ratings to IAT results. Among the three listener groups (heritage, L2, and teacher), each MGT result set (*acquisition* and *academic-ness*) was separated by the guise variety (SS and USS), which were each paired with each IAT result set. All models demonstrated weak and/or non-significant relationships between explicit attitudes elicited indirectly from the MGT and automatic bias elicited from the IAT. Values closer to −1 indicate a strong negative correlation and values closer to +1 a strong positive correlation. Correlations below +/−0.4 are considered weak using social science standards, even in cases in which they are presented with *p* values less than 0.05 (Evans 1996).

However, there are some correlations approaching practical significance (>0.4) that should be discussed. Both the L2 and teacher groups have a moderately weak negative correlation (-0.36 and -0.34, respectively) between their Spanish + Academic IAT scores and MGT *academic-ness* scores for USS speakers. That is, as IAT scores increased (favoring Spanish, not Spanglish), *academic-ness* evaluations of USS speakers went down (favoring SS). Also, both the teacher and heritage groups have a moderately weak positive correlation (both 0.37) between their Spanish + Academic IAT scores and MGT *acquisition* evaluations for SS speakers, meaning that as IAT scores increased (favoring Spanish), *acquisition* evaluations of SS speakers increased as well (favoring SS). Interestingly, the L2 group also demonstrated a moderately weak negative correlation of Spanish + Academic (IAT) with the MGT *acquisition* evaluations for SS speakers (-0.36). This indicates that IAT scores went down as MGT scores went up, meaning that while they increasingly favored SS over USS speakers, their positive implicit bias of Spanish + Academic weakened (i.e., they had slower reaction times in their associations of the two). This is evidenced in their lower IAT scores of the three groups (see Table 8).

Table 8. Correlation analyses of MGT (separated by language) to IAT results, presented by group.

		Spanish		US Spanish	
	Results	Acquisition (MGT)	Academic-ness (MGT)	Acquisition (MGT)	Academic-ness (MGT)
Heritage learners	Spanish + Good (IAT)	$r = 0.37$ $p < 0.004$	$r = 0.16$ $p < 0.25$	$r = -0.13$ $p < 0.34$	$r = -0.02$ $p < 0.86$
	Spanish + Academic (IAT)	$r = 0.13$ $p < 0.34$	$r = 0.8$ $p < -0.03$	$r = -0.03$ $p < 0.83$	$r = -0.16$ $p < 0.25$
L2 learners	Spanish + Good (IAT)	$r = 0.09$ $p < 0.5$	$r = -0.06$ $p < 0.62$	$r = 0.18$ $p < 0.15$	$r = -0.36$ $p < 0.003$
	Spanish + Academic (IAT)	$r = 0.11$ $p < 0.41$	$r = -0.36$ $p < 0.005$	$r = -0.06$ $p < 0.66$	$r = -0.06$ $p < 0.62$
Teachers	Spanish + Good (IAT)	$r = 0.05$ $p < 0.75$	$r = -0.06$ $p < 0.69$	$r = -0.23$ $p < 0.12$	$r = -0.17$ $p < 0.27$
	Spanish + Academic (IAT)	$r = 0.37$ $p < 0.01$	$r = 0.21$ $p < 0.18$	$r = -0.25$ $p < 0.09$	$r = -0.34$ $p < 0.02$

7. Discussion

This study is the first to employ both the MGT and IAT to examine the effects of Spanish language training (i.e., heritage, L2, and teacher) and bias on perceptions of both standardized Spanish and US Spanish repertoires in the United States. The MGT was used as an indirect method to gauge bias and provide insight into how lexical items index differentiated social meanings depending on the repertoire presented. The dependent variables, perceived *acquisition* and *academic-ness* of the speaker's variety, were heavily conditioned by the guise variety, as SS was rated more positively in each attribute with no listener profile group differences. With regard to *acquisition*, these findings demonstrate that lexical items of nonstandardized or nonhegemonic language in the case of US Spanish index ideologies of incomplete acquisition that ignore the dynamic possibilities of multilingual languagers. This steadfast bias is rooted in the separateness of languages as inherent to the alleged learning, storage, and production of language. In specific relation to language variation research, this lends new insights into the indexical field of US Spanish and its salient lexical features, an investigation that should be replicated and expanded with other variables of US Spanish (see Román et al. 2019). With regard to the factor of *academic-ness*, once again all participants evaluated the SS speakers as having learned their variety in an academic setting. These results associate USS with non-academic settings for learning languages, such as the home or community, but also might indicate a conscious knowledge

that USS is not 'formally' learned in the classroom, which is in line with how heritage language programs generally promote the acquisition of an 'academic' register. Similarly, these perspectives are highly affected by monoglossic influences, whereby assimilation to SS in an academic setting is normalized and expected. Thus, SS and USS seem to be mutually exclusive codes that index different social meanings when employed by the same person. As participants demonstrated in their MGT evaluations, USS speakers are significantly less likely to have learned their repertoires in school, as 'academic' language in this case lines up neatly with the boundaries of named, codified languages. As such, the use of USS in the classroom appears incongruent with academic language use.

In the assessment of associations between the concepts of Spanish/Spanglish and the attributes of Good/Bad and Academic/Not Academic, all participant groups demonstrated faster reaction times when Spanish was paired with the positive qualities. These findings overwhelmingly indicate that participants more strongly associate Spanish with the positively charged exemplars related to Good and Academic. These subconscious biases are affirmed in the MGT, where ideologies of 'complete' *acquisition* and *academic-ness* in language variety expression are again associated with SS. As indicated in the ANOVA model, heritage participants had stronger positive bias towards Spanish + Good and Spanish + Academic (i.e., faster reaction times in associating the positive attributes to Spanish over Spanglish) than the L2 group, and teacher averages patterned similarly to the heritage group. Though all groups still demonstrated more positive bias to Spanish over Spanglish, L2s' lower averages for Spanish + Good may indicate that students do not learn these associations with Spanish repertoires as early or as directly as the heritage group, who have been exposed to Spanish and have developed metalinguistic awareness of their own Spanish(es) since an early age. "Likewise, the difference between L2 students' averages for Spanish + Good and those of teachers approached significance ($p = 0.06$), suggesting that teachers pattern similarly to the heritage group. This could be due to the teachers' exposure to sociolinguistic ideologies in school, even in the absence of the home exposure that characterizes the experience of heritage students.". These results highlight the expectations that SS, or the named language 'Spanish,' is a more 'complete' and 'whole' language affiliated with an academic realm. Heritage students hold these biases and may also learn them early in life within the community and family, L2 learners develop them further in Spanish language study, and teachers may implicitly (or explicitly) promote them in their teaching. A deeper dive into the experiences of heritage languagers as students *and* heritage teachers can shed light on how their metalinguistic awareness of the indexical meanings associated with USS is cultivated in youth (see Ek et al. 2013).

This study also attempted to reveal if data trends from each MGT variable (i.e., *acquisition* and *academic-ness*) correlated with each IAT (Spanish + Good/Spanish + Academic). Results presented generally weak correlations; however, there were some significant correlations that approached practical significance. It has been attested that the associations that participants make in the IAT are less susceptible to change (Petty et al. 2009; Evans 2008; Karpen et al. 2012), and that explicit attitudes are more apt to change with influence. Thus divergences of more malleable attitudes are possible with social and systemic change.

In light of the results discussed above, the research questions can now be addressed. The first research question asked if US Spanish lexical features were salient to L2 and heritage learners as well as teachers of Spanish. Findings revealed that US lexical items are indeed salient to all groups. This was demonstrated in the MGT evaluations, wherein Spanglish speakers were rated lower on both social attributes (*acquisition* and *academic-ness*). Similarly, the results of L2 learners also display that, even if this group had less exposure to SS ideologies in their early learning, they have learned enough about stigmatized lexical features of US Spanish to be able to identify them in the audio and rate the USS speakers less favorably. In response to the second research question, which asked to what extent differential participation in language education programs (L2 students, heritage students, and teachers) correlates with more positive explicit attitudes toward US Spanish features, the statistical analysis found no clear correlation between program type and more positive

attitudes. These findings demonstrate that L2 and heritage learners as well as teachers of Spanish exhibit similarly patterned bias across testing conditions, which suggests that heritage language education, whether it seeks to value translanguaging or not, is not associated with a change to heritage languagers' malleable attitudes (elicited from the MGT) from deeply rooted bias (from the IAT).

Finally, the third research question posed whether the attitudes and associations elicited from indirect methods (i.e., the MGT) and automatic response (i.e., the IAT) suggest that explicit and implicit biases result from distinct cognitive processes, in line with previous research examining explicit to implicit bias. Results pattern similarly to previous findings comparing explicit to implicit bias ratings, whereby correlations were weak to moderately weak, indicating that the processing of varying biases—even if they present similar findings (e.g., US Spanish or Spanglish is seen as less academic)—is complex and not monolithic, contributing insights into theories that examine implicit and explicit attitudes as structurally distinct (Greenwald and Nosek 2009). Examinations of distinct biases garnered from attitude measures like the MGT and the IAT are "able to capture distinct levels of linguistic attitudes which are potentially conflicting" (McKenzie and Carrie 2018, p. 837); however, in the case of the experimental line presented here, results are stable across measures (i.e., favor SS/Spanish over USS/Spanglish) even if they do not correlate. The biases derived from the MGT and the IAT, while they do not strongly correlate in this study, are also not diverging with regard to longstanding deficit perspectives of US Spanish. While the results of this study do not support the IED hypothesis, the demonstration of weak or nonexistent correlations between the bias measures contributes to the previous language attitude literature that has examined explicit and implicit bias as different processes (see McKenzie and Carrie 2018; Ianos et al. 2020).

As this study demonstrates, as well as the literature evaluating raciolinguistic ideologies in schools, US Spanish speakers who navigate their linguistic repertoires easily and effectively are still thought to index linguistic deficiency by their perceiving subjects, who are composed in this study of both L1 and L2 Spanish languagers, demonstrating that raciolinguistic ideologies pathologize language users. Despite the increase in heritage language and bilingual programs in the US in recent decades, standard language and monoglossic ideologies continue to dominate the ways in which language education in the United States is conceptualized and formalized through early teachings of who is a 'valid' languager in an academic setting. As mentioned previously, these ideologies rely heavily on the naming conventions of idealized languages that undergo standardization, exhibiting a perhaps subconscious inability for all groups to view speakers of languages 'in between' as fluent and competent in their communication. Relatedly, participants' perceptions delegitimize translanguagers in the US and beyond, a position that is undergirded by both folk ideologies and empirical research that maps the political boundaries of named languages onto real language production, disparaging language change and variation, which are inherent to all language communities. It has even been posited that the term 'Spanglish' somewhat iconicizes (see Gal and Irvine 2019) deficits or symbolizes such disparities (see Otheguy and Stern 2011), particularly within an academic context, as the very term visibly indicates hybridity or 'incompleteness' within the bounds of one language. These deficit perspectives contribute to the raciolinguistic ideologies that plague how racialized students are perceived, how they are educated in the classroom, and their trajectory in the US education system. This status quo leads multilingual Latinx students to experience *languagelessness*, or the incapability of producing a legitimate language according to hegemonic standards (Rosa 2016), leading them to view their own repertoires as insufficient for an academic context. As mentioned earlier, there is no consensus as to what 'academic' language is (Flores 2020a; Martínez and Mejía 2020; Valdés 2004). These biases must be discussed in conjunction with the so-called "word gap" literature that targets racialized children early on in their education, disenfranchising them from equitable learning experiences that should center their dynamic multilingual practices as natural language development and expression.

8. Conclusions

This study shows that L2 and heritage students as well as teachers demonstrate stable explicit attitudes (elicited indirectly in the MGT) and implicit attitudes (elicited from the IAT) towards indexical meanings linked with US Spanish lexical features. Likewise, results exhibit that varying implicit biases may not be stored and processed in the same place and manner, providing hope that positive changes in how languages are conceptualized and taught can shift semi-implicit attitudes and eliminate discriminatory meanings associated with the indexical fields of USS lexical items. The study, however, presented some limitations that may inspire further exploration of the topics of bias and standard language ideologies. First, the design of the study presents the two language varieties in a dichotomous fashion, which is not how language and language communities are encountered in real life. Also, the stimuli presented in the MGT were audio while those in the IAT were written. A comparison of similar stimuli or replication of this study design would shed light on these effects. Relatedly, the Spanish phonologically inflected lexical items in the USS guise passages index particular and unique indexical meanings, which can shift and change with differing stimuli; in other words, if this study had focused on morphosyntactic variation or the inclusion of phonetic features associated with Californian English, the salience of variables may have shifted as well as the indexical field of the variables under examination. The order of guises can also affect perception; that is, hearing the male USS guise before the same male SS guise may affect how the second is heard (and same for the female guise in both varieties). Lastly, a larger participant pool may strengthen the analysis and highlight the potential interactions among independent variables.

I encourage scholars engaging with this topic to accept the possibilities of language liberation and to reject discriminatory linguistic policies and practices that continue to racialize and exceptionalize multilingual languagers as in need of remediation. Many people are already doing vital work to shift the harmful associations with US Spanish(es) through the reframing of Spanish language education. Prada (2019), Flores (2020b), and Seltzer and Wassell (2022) demonstrate how translanguaging allows for oppressive structures in the classroom to be reconfigured and reimagined to center students' natural language expression. De Los De Los Ríos et al. (2021) explore the connection of cultural practices and translingualism in fostering language exploration and identity expression. Train (2020) promotes the teaching of variation as a means to achieve social justice, Holguín Holguín Mendoza (2018) privileges the home and community repertoires in the planning of a heritage language program, and Bucholtz et al. (2018) explore the empowerment of Latinx youth through community language work and participatory action research. Anya (2021) demonstrates the benefits of critical race pedagogy for more inclusive world language education, which Austin (2022) takes a step further by countering anti-Black racism in teacher training and instilling reflexivity. Relatedly, (Licata et al. Forthcoming) exhibit the benefits of a course on raciolinguistic theory and practice to increase metalinguistic awareness in both language teachers and students. There are many more examples of people laying this vital groundwork, and we must interrogate how widespread implementation of such curricula across language programs and teacher training can shift perceptions of US Spanish(es) in academic contexts via the collective rejection of deficit perspectives towards US Spanish(es) and Latinx students. Only then can we reimagine safer and more affirmative learning experiences for our heritage learners, as well as foster the development of empathy among teachers and L2 students.

Institutional Review Board Statement: The study was conducted in accordance with the Declaration of Helsinki and approved by the Institutional Review Board of the University of California, Berkeley (protocol code: 2018-04-10957, approved on 24 September 2020) for studies involving humans.

Informed Consent Statement: Informed consent was obtained from all subjects involved in the study.

Appendix A. Guise Stories (English Words Bolded Refer to SS and Those <u>Underlined</u> Refer to USS)

	Standardized Spanish (SS) Repertoire	US Spanish (USS) Repertoire	English Translation
Story A	**Bueno**, para ir al **muelle**, tienes que doblar a la derecha en la **calle** Retiro. Puedes dar vuelta en Pacheco, pero hay mucho **tráfico** ahí **todo el tiempo**. **Ya sabes**, es mejor evitar **los semáforos**. Te espero en **la camioneta verde** de mi papá. **Bueno**, nos vemos ahí.	<u>So</u>, para ir al <u>pier</u>, tienes que doblar a la derecha en <u>Retiro</u> <u>Street</u>. Puedes dar vuelta en Pacheco, pero hay mucho <u>traffic</u> ahí <u>all the time</u>. <u>You know</u>, es mejor evitar las <u>traffic lights</u>. Te espero en <u>la troca verde</u> de mi papá. <u>Ok pues</u>, nos vemos ahí.	**Well**/<u>So</u>, to go to the pier, you have to turn right on Retiro Street. You can turn on Pacheco, but there is a lot of traffic there all the time. You know, it's better to avoid the traffic lights. I'll be waiting for you in my dad's green truck. Ok, see you then.
Story B	Para llegar a la tienda, tienes que **tomar la carretera** que va al **centro**. **Sin embargo**, habrá mucha gente porque es la hora **del almuerzo**. Todos irán a los restaurantes durante **el descanso** para **comprar comida**. Yo estoy **lleno/a y** no voy a comer, **pero** si quieres, podemos **llegar** al restaurante.	Para llegar a la tienda, tienes que subirte al <u>freeway</u> que va al <u>downtown</u>. <u>Pero like</u>, habrá mucha gente porque es la hora <u>del lonche</u>. Todos irán a los restaurantes durante <u>el break</u> para <u>agarrar</u> comida. Yo estoy <u>full so</u> no voy a comer, <u>pero like</u>, si quieres, podemos <u>parar</u> al restaurante.	To arrive at the store, you need to get on the freeway and go downtown. **However**/<u>However</u> or <u>But like</u>, there will be a lot of people because it's lunch time. Everyone will be going to restaurants during break to get food. I'm full so I won't be eating, however/<u>but like</u>, if you want, we can stop at the restaurant.
Story C	**Pues** para llegar al parque, debes **dar vuelta a la izquierda** en **la calle** Olivares. Puedes dejar el carro en **la estructura de estacionamiento** de la esquina, pero no sé si **está abierta**. **De hecho**, mejor pasa por la calle San Andrés y por ahí puedes entrar. Si prefieres, podemos hacer las compras para **la fiesta** en la tienda cerca **de la parada**. También, tengo que **devolver** unas cosas que ya no me sirven.	<u>So</u>, para llegar al parque, debes <u>hacer una izquierda</u> en Olivares <u>Street</u>. Puedes dejar el carro en <u>el parking lot</u> de la esquina, pero no sé si <u>están abiertos</u>. <u>Actually</u>, mejor pasa por la calle San Andrés y por ahí puedes entrar. Si prefieres, podemos hacer las compras para <u>el party</u> en la tienda cerca <u>del bus</u>. También, tengo que <u>regresar</u> unas cosas que ya no me sirven.	So, to get to the park, you need to turn left on Olivares Street. You can leave your car in the parking lot on the corner, but I don't know if they are open. Actually, it's better to pass through San Andrés Street and you can enter through there. If you want, we can shop for the party in the store near the bus stop. Also, I have to return some things that I don't need anymore.
Story D	**Bueno**, la **película** empieza a las ocho. Si quieres comer antes, podemos ir al **restaurante** que está cerca. **Pero**, Daniel no puede entrar porque hay un bar y todavía está en **la escuela secundaria**. Nos reunimos enfrente del **supermercado**. Primero voy **de compras** con Elena, quien también quiere **platicar** sobre los planes para este **fin de semana**. Una cosa más: ¡no te olvides de los **boletos de entrada**!	<u>So</u>, la <u>muvi</u> empieza a las ocho. Si quieres comer antes, podemos ir al <u>restaurán</u> que está cerca. <u>Pero like</u>, Daniel no puede entrar porque hay un bar y todavía está en <u>la high school</u>. Nos reunimos enfrente de la <u>marketa</u>. Primero me voy <u>shopping</u> con Elena, quien también quiere <u>discutir</u> sobre los planes para este <u>weekend</u>. Una cosa más: ¡no te olvides de los <u>tickets</u>!	**Well**/<u>So</u>, the movie starts at eight. If you want to eat before, we can go to the restaurant that is nearby. **But**/<u>However</u> or <u>But like</u>, Daniel can't go in because there's a bar and he's still in high school. Let's meet in front of the market. First, I'll go shopping with Elena, who also wants to discuss plans for this weekend. One more thing: don't forget the tickets!

| Story E | **Pues**, la ruta más rápida es por la **Avenida** Paloma. Pero quizás **esté** cerrada, **así que** puedes también pasar por la calle Francisco. Aunque tal vez **llegues** tarde—**ya ves**, siempre hay mucho tráfico y poco **estacionamiento**. Javier nos va a acompañar porque **renunció** a su trabajo y ya no tiene que trabajar **por** las noches. Cuando estés **listo**, ¡**envíame un mensaje**! | So pues, la ruta más rápida es por Paloma Avenue. Pero quizás estará cerrada, **so** puedes también pasar por la calle Francisco. Aunque tal vez estés tarde—you know, siempre hay mucho tráfico y poco parkin. Javier nos va a acompañar porque cuitió su trabajo y ya no tiene que trabajar en las noches. Cuando estés ready, ¡textéame! | **So**/So like, the fastest route is down Paloma Avenue. But maybe it will already be closed, so you can also go down Francisco Street. Even if you get there late—you know, there's always a lot of traffic and little parking. Javier is going to accompany us because he quit his job and now, he doesn't have to work at night. When you are ready, text me! |

Notes

[1] In this paper, and more broadly, I utilize *languager* to generally indicate any person that communicates and perceives language. I use *speaker* when the form of communication is specifically related to oral production and hearing.

[2] In line with how Spanish language programs are often organized, heritage language learning encompasses classes for languagers who are raised in a home where a non-hegemonic language is used, and who can use or comprehend the home language to some degree, and who are, to any degree, bilingual in that language and in the hegemonic variety (Valdé s et al. 1999). Second language (L2) learning refers to courses in which languagers are learning or have learned an additional language in school. With regard to this study, L2 learners began learning Spanish in school after the age of 14.

[3] The generalized idealized communicator is a hearing subject (see Henner and Robinson 2021), thus colonial epistemologies also take phonocentric, ableist stances.

[4] L1 variability within the teacher group did not significantly mediate bias differences.

[5] See Tamminga (2017), who found no evidence that social evaluations of /ɪŋ/$\sim$/ɪn/ variation differed across frame utterance styles.

[6] A six-point scale has been shown to increase discrimination and reliability than a five-point scale (Chomeya 2010).

[7] An exploratory factor analysis is an unsupervised language model that requires human expertise in the analysis of factors. As some evaluative scales may score a loading factor above 0.4 for more than one factor, researcher judgment ultimately determines factor grouping.

[8] The only statistical comparisons of relevance to my research questions are those concerned with whether or not each group of listeners differentiated each of the target language guises (SS/USS) by the attitudinal category, namely a main effect of the guise variety or an interaction effect between the guise variety and listener profile group. Accordingly, I limit my reporting and discussion of results to these effects and potential interactions. Any potential main effect of the listener profile group would reflect non-substantive comparisons of averaged *acquisition* or *academic-ness* ratings by group (i.e., there is no meaningful interpretation of one participant group having higher ratings over other groups unless those higher ratings interact with differentiated ratings mediated by the guise variety).

[9] Errors were replaced with participant block means of correct trials plus 600 ms (or the D600 procedure; Greenwald et al. 2003).

References

Alonso, Lara, and Laura Villa. 2020. Latinxs' Bilingualism at Work in the US: Profit for Whom? *Language, Culture and Society* 2: 37–65. [CrossRef]

Anya, Uju. 2020. African Americans in World Language Study: The Forged Path and Future Directions. *Annual Review of Applied Linguistics* 40: 97–112. [CrossRef]

Anya, Uju. 2021. Critical Race Pedagogy for More Effective and Inclusive World Language Teaching. *Applied Linguistics* 42: 1055–69. [CrossRef]

Anzaldúa, Gloria. 1987. *Borderlands/La Frontera: The New Mestiza*. San Francisco: Spinsters/Aunt Lute.

Austin, Tasha. 2022. Linguistic Imperialism: Countering Anti Black Racism in World Language Teacher Preparation. *Journal for Multicultural Education* 16: 246–58. [CrossRef]

Avineri, Netta, Eric Johnson, Shirley Brice-Heath, Teresa McCarty, Elinor Ochs, Tamar Kremer-Sadlik, Susan Blum, Ana Celia Zentella, Jonathan Rosa, and Nelson Flores. 2015. Invited Forum: Bridging the 'Language Gap'. *Journal of Linguistic Anthropology* 25: 66–86. [CrossRef]

Babel, Molly. 2010. Dialect Divergence and Convergence in New Zealand English. *Language in Society* 39: 437–56. [CrossRef]

Bailey, Benjamin. 2012. Heteroglossia. In *The Routledge Handbook of Multilingualism*. Edited by M. Martin-Jones, A. Blackledge and A. Creese. London: Routledge, pp. 511–19.

Bakhtin, Mikhail Mikhaĭlovich. 1981. *The Dialogic Imagination: Four Essays*. Austin: University of Texas Press.

Bucholtz, Mary, Dolores Inés Casillas, and Jin Sook Lee. 2018. California Latinx Youth as Agents of Sociolinguistic Justice. In *Language and Social Justice in Practice*. Edited by Eric J. Johnson, Jonathan Rosa, Laura R. Graham, Netta Avineri and Robin Conley. London: Routledge, pp. 166–75.

Burns, Katharine E. 2018. Marginalization of Local Varieties in the L2 Classroom: The Case of US Spanish. *L2 Journal* 10: 20–38. [CrossRef]

Calamai, Silvia, and Fabio Ardolino. 2020. Italian with an Accent: The Case of 'Chinese Italian' in Tuscan High Schools. *Journal of Language and Social Psychology* 39: 132–47. [CrossRef]

Callesano, Salvatore, and Phillip M. Carter. 2022. Unidirectional Language Bias: The Implicit Association Test with Spanish and English in Miami. *International Journal of Bilingualism* 13670069221121160. [CrossRef]

Campbell-Kibler, Kathryn. 2009. The Nature of Sociolinguistic Perception. *Language Variation and Change* 21: 135–56. [CrossRef]

Campbell-Kibler, Kathryn. 2010. Sociolinguistics and Perception. *Language and Linguistics Compass* 4: 377–89. [CrossRef]

Carpenter, Thomas P., Ruth Pogacar, Chris Pullig, Michal Kouril, Stephen Aguilar, Jordan LaBouff, Naomi Isenberg, and Alek Chakroff. 2019. Survey-Software Implicit Association Tests: A Methodological and Empirical Analysis. *Behavior Research Methods* 51: 2194–208. [CrossRef] [PubMed]

Carstens, Adelia. 2016. Translanguaging as a Vehicle for L2 Acquisition and L1 Development: Students' Perceptions. *Language Matters* 47: 203–22. [CrossRef]

Chappell, Whitney. 2016. On the Social Perception of Intervocalic /s/ Voicing in Costa Rican Spanish. *Language Variation and Change* 28: 357. [CrossRef]

Chappell, Whitney. 2021a. Heritage Mexican Spanish Speakers' Sociophonetic Perception of /s/ Aspiration. *Spanish as a Heritage Language* 1: 167–97. [CrossRef]

Chappell, Whitney. 2021b. Mexicans' and Mexican-Americans' Perceptions of Themselves and Each Other: Attitudes toward Language and Community. In *Topics in Spanish Linguistic Perceptions*. Edited by Luis Alfredo Ortiz-López and Eva-María Suárez Büdenbender. London: Routledge, pp. 138–60.

Chappell, Whitney, and Matthew Kanwit. 2022. Do Learners Connect Sociophonetic Variation with Regional and Social Characteristics?: The Case of L2 Perception of Spanish Aspiration. *Studies in Second Language Acquisition* 44: 185–209. [CrossRef]

Chávez-Moreno, Laura C. 2021. The Problem with Latinx as a Racial Construct Vis-à-Vis Language and Bilingualism: Toward Recognizing Multiple Colonialisms in the Racialization of Latinidad. In *Handbook of Latinos and Education*. Edited by Enrique G. Murillo Jr., Dolores Delgado Bernal, Socorro Morales, Luis Urrieta Jr., Eric Ruiz Bybee, Juan Sánchez Muñoz, Victor B. Saenz, Daniel Villanueva, Margarita Machado-Casas and Katherine Espinoza. London: Routledge, pp. 164–80.

Chomeya, Rungson. 2010. Quality of Psychology Test Between Likert Scale 5 and 6 Points. *Journal of Social Sciences* 6: 399–403. [CrossRef]

Cioè-Peña, María. 2022. The master's tools will never dismantle the master's school: Interrogating settler colonial logics in language education. *Annual Review of Applied Linguistics* 42: 25–33. [CrossRef]

Craft, Justin T., Kelly E. Wright, Rachel Elizabeth Weissler, and Robin M Queen. 2020. Language and Discrimination: Generating Meaning, Perceiving Identities, and Discriminating Outcomes. *Annual Review of Linguistics* 6: 389–407. [CrossRef]

Crowley, Tony. 2003. *Standard English and the Politics of Language*. Basingstoke: Palgrave Macmillan.

de los Heros, Susana. 2009. Linguistic Pluralism or Prescriptivism? A CDA of Language Ideologies in Talento, Peru's Official Textbook for the First-Year of High School. *Linguistics and Education* 20: 172–99. [CrossRef]

De Los Ríos, Cati V., Kate Seltzer, and Arturo Molina. 2021. 'Juntos Somos Fuertes': Writing Participatory Corridos of Solidarity through a Critical Translingual Approach. *Applied Linguistics* 42: 1070–82. [CrossRef]

Drager, Katie, and M. Joelle Kirtley. 2016. Awareness, Salience, and Stereotypes in Exemplar-Based Models of Speech Production and Perception. In *Awareness and Control in Sociolinguistic Research*. Edited by Anna M. Babel. Cambridge: Cambridge University Press, pp. 1–24.

Eckert, Penelope. 2008. Variation and the Indexical Field. *Journal of Sociolinguistics* 12: 453–76. [CrossRef]

Ek, Lucila D., Patricia Sánchez, and Patricia D. Quijada Cerecer. 2013. Linguistic Violence, Insecurity, and Work: Language Ideologies of Latina/o Bilingual Teacher Candidates in Texas. *International Multilingual Research Journal* 7: 197–219. [CrossRef]

Evans, James D. 1996. *Straightforward Statistics for the Behavioral Sciences*. Pacific Grove: Thomson Brooks/Cole Publishing Co.

Evans, Jonathan St. B. T. 2008. Dual-Processing Accounts of Reasoning, Judgment, and Social Cognition. *Annual Review of Psychology* 59: 255–78. [CrossRef] [PubMed]

Fazio, Russell H., and Michael A. Olson. 2003. Implicit Measures in Social Cognition Research: Their Meaning and Use. *Annual Review of Psychology* 54: 297–327. [CrossRef]

Flores, Nelson. 2020a. From Academic Language to Language Architecture: Challenging Raciolinguistic Ideologies in Research and Practice. *Theory into Practice* 59: 22–31. [CrossRef]

Flores, Nelson. 2020b. @nelsonlflores. Translanguaging Does Not Mean Using Multiple Named Languages in One Utterance. If You Define It That Way Just Stick with Code-Switching. It is, Instead, a Perspective That Prioritizes the Meaning-Making of Bi/Multilinguals in Ways That Challenge Normative Language Ideologies. Twitter. January 24. Available online: https://twitter.com/nelsonlflores/status/1220857637073494027?s=20 (accessed on 12 February 2023).

Flores, Nelson, and Jonathan Rosa. 2022. Undoing Competence: Coloniality, Homogeneity, and the Overrepresentation of Whiteness in Applied Linguistics. *Language Learning*. [CrossRef]

Flores, Nelson, Tatyana Kleyn, and Kate Menken. 2015. Looking Holistically in a Climate of Partiality: Identities of Students Labeled Long-Term English Language Learners. *Journal of Language, Identity & Education* 14: 113–32.

Flores, Nelson, Mark C. Lewis, and Jennifer Phuong. 2018. Raciolinguistic Chronotopes and the Education of Latinx Students: Resistance and Anxiety in a Bilingual School. *Language & Communication* 62: 15–25.

Gal, Susan, and Judith T Irvine. 2019. *Signs of Difference: Language and Ideology in Social Life*. Cambridge: Cambridge University Press.

García, Ofelia. 2011. *Bilingual Education in the 21st Century: A Global Perspective*. Hoboken: John Wiley & Sons.

García, Ofelia. 2019. Decolonizing Foreign, Second, Heritage, and First Languages. In *Decolonizing Foreign Language Education: The Misteaching of English and Other Colonial Languages*. Edited by Donaldo Macedo. New York: Routledge, pp. 152–68.

García, Ofelia, and Ricardo Otheguy. 2020. Plurilingualism and Translanguaging: Commonalities and Divergences. *International Journal of Bilingual Education and Bilingualism* 23: 17–35. [CrossRef]

García, Ofelia, Nelson Flores, Kate Seltzer, Li Wei, Ricardo Otheguy, and Jonathan Rosa. 2021. Rejecting Abyssal Thinking in the Language and Education of Racialized Bilinguals: A Manifesto. *Critical Inquiry in Language Studies* 18: 203–28. [CrossRef]

Garrett, Peter. 2001. Language Attitudes and Sociolinguistics. *Journal of Sociolinguistics* 5: 626–31. [CrossRef]

Greenwald, Anthony G., and Brian A. Nosek. 2009. Attitudinal Dissociation: What Does It Mean? In *Attitudes: Insights from The New Implicit Measures*. Edited by R. E. Petty, R. H. Fazio and P. Brinol. Hove: Psychology Press, pp. 65–82.

Greenwald, Anthony G., Debbie E. McGhee, and Jordan L. K. Schwartz. 1998. Measuring Individual Differences in Implicit Cognition: The Implicit Association Test. *Journal of Personality and Social Psychology* 74: 1464–80. [CrossRef]

Greenwald, Anthony G., Brian A. Nosek, and Mahzarin R. Banaji. 2003. Understanding and Using the Implicit Association Test: I. An Improved Scoring Algorithm. *Journal of Personality and Social Psychology* 85: 197. [CrossRef]

Hay, Jennifer, and Katie Drager. 2010. Stuffed Toys and Speech Perception. *Linguistics* 48: 865–92. [CrossRef]

Heller, Monica. 2006. *Linguistic Minorities and Modernity: A Sociolinguistic Ethnography*, 2nd ed. New York: Continuum.

Helms, Annie. 2020. Machine Learning for Linguistics: Exploratory Factor Analysis [Computer Software]. Available online: https://github.com/anniehelms/machine_learning_for_linguistics (accessed on 12 February 2023).

Henner, Jon, and Octavian Robinson. 2021. Unsettling Languages, Unruly Bodyminds: Imaging a Crip Linguistics. Available online: https://psyarxiv.com/7bzaw/ (accessed on 31 July 2022).

Holguín Mendoza, Claudia. 2018. Critical Language Awareness (CLA) for Spanish Heritage Language Programs: Implementing a Complete Curriculum. *International Multilingual Research Journal* 12: 65–79. [CrossRef]

Ianos, Maria Adelina, Andrei Rusu, Ángel Huguet, and Cecilio Lapresta-Rey. 2020. Implicit Language Attitudes in Catalonia (Spain): Investigating Preferences for Catalan or Spanish Using the Implicit Association Test. *Journal of Multilingual and Multicultural Development* 44: 214–29. [CrossRef]

Inoue, Miyako. 2003. The Listening Subject of Japanese Modernity and His Auditory Double: Citing, Sighting, and Siting the Modern Japanese Woman. *Cultural Anthropology* 18: 156–93. [CrossRef]

Inoue, Miyako. 2006. *Vicarious Language: Gender and Linguistic Modernity in Japan*. Berkeley: University of California Press, vol. 11.

Kang, Okim, and Donald L. Rubin. 2009. Reverse Linguistic Stereotyping: Measuring the Effect of Listener Expectations on Speech Evaluation. *Journal of Language and Social Psychology* 28: 441–56. [CrossRef]

Karpen, Samuel C., Lile Jia, and Robert J. Rydell. 2012. Discrepancies between Implicit and Explicit Attitude Measures as an Indicator of Attitude Strength. *European Journal of Social Psychology* 42: 24–29. [CrossRef]

Kircher, Ruth. 2016. The Matched-Guise Technique. In *Research Methods in Intercultural Communication*. Edited by Zhu Hua. Hoboken: Wiley-Blackwell, pp. 196–211.

Kuznetsova, Alexandra, Per B. Brockhoff, and Rune H. B. Christensen. 2017. lmerTest package: Tests in linear mixed effects models. *Journal of Statistical Software* 82: 1–26. [CrossRef]

Lambert, Wallace E., Hannah Frankle, and G. Richard Tucker. 1966. Judging Personality through Speech: A French-canadian Example. *Journal of Communication* 16: 305–21. [CrossRef] [PubMed]

Lane, Kristin A., Mahzarin R. Banaji, Brian A. Nosek, and Anthony G. Greenwald. 2007. Understanding and Using the Implicit Association Test: IV: What We Know (so Far) about the Method. In *Implicit Measures of Attitudes*. New York: The Guilford Press.

Licata, Gabriella, Tasha Austin, and Aris Moreno Clemons. Forthcoming. Raciolinguistics and Spanish language teaching. Edited by Mara Fuertes Gutiérrez, Rosina Márquez-Reiter, Ben Evans, Aris Moreno Clemons. Decolonizing Spanish Language Teaching and Social Justice: Theoretical Approaches, Teaching Competences and Practices. *Journal of Spanish Language Teaching*.

Lippi-Green, Rosina. 2012. *English with an Accent: Language, Ideology and Discrimination in the United States*. London: Routledge.

Loza, Sergio. 2017. Transgressing Standard Language Ideologies in the Spanish Heritage Language (SHL) Classroom. *Chiricù Journal: Latina/o Literature, Art, and Culture* 1: 56–77. [CrossRef]

Lozano, Rosina. 2018. *An American Language: The History of Spanish in the United States*. Berkeley: University of California Press, vol. 49.

Martínez, Glenn A. 2003. Perceptions of Dialect in a Changing Society: Folk Linguistics along the Texas-Mexico Border. *Journal of Sociolinguistics* 7: 38–49. [CrossRef]

Martínez, Ramón Antonio, and Alexander Feliciano Mejía. 2020. Looking Closely and Listening Carefully: A Sociocultural Approach to Understanding the Complexity of Latina/o/x Students' Everyday Language. *Theory into Practice* 59: 53–63. [CrossRef]

Martin-Jones, Marilyn, and Suzanne Romaine. 1986. Semilingualism: A Half-Baked Theory of Communicative Competence. *Applied Linguistics* 7: 26–38. [CrossRef]

McKenzie, Robert M., and Erin Carrie. 2018. Implicit–Explicit Attitudinal Discrepancy and the Investigation of Language Attitude Change in Progress. *Journal of Multilingual and Multicultural Development* 39: 830–44. [CrossRef]

McKenzie, Robert M., and Alexander Gilmore. 2017. 'The People Who Are out of 'Right'E Nglish': J Apanese University Students' Social Evaluations of E Nglish Language Diversity and the Internationalisation of Japanese Higher Education. *International Journal of Applied Linguistics* 27: 152–75. [CrossRef]

McKenzie, Robert M., and Andrew McNeill. 2022. *Implicit and Explicit Language Attitudes: Mapping Linguistic Prejudice and Attitude Change in England*. Abingdon: Taylor & Francis.

Mena, Mike. 2022. The Language-Elsewhere: A Friendlier Linguistic Terrorism. In *The Spanish Language in the United States*. Edited by José A. Cobas, Bonnie Urciuoli, Joe Feagin and Daniel J. Delgado. London: Routledge, pp. 80–95.

Myers-Scotton, Carol. 1997. Codeswitching. In *The Handbook of Sociolinguistics*. Edited by Florian Coulmas. Oxford: Blackwell.

Otheguy, Ricardo, and Nancy Stern. 2011. On so-called Spanglish. *International Journal of Bilingualism* 15: 85–100. [CrossRef]

Otheguy, Ricardo, Ofelia García, and Wallis Reid. 2015. Clarifying Translanguaging and Deconstructing Named Languages: A Perspective from Linguistics. *Applied Linguistics Review* 6: 281–307. [CrossRef]

Padilla, Lillie, and Rosti Vana. 2022. The 'Other' Latinx: The (Non) Existent Representation of Afro-Latinx in Spanish Language Textbooks. *Journal of Language, Identity & Education*, 1–15. [CrossRef]

Pantos, Andrew J., and Andrew W. Perkins. 2013. Measuring Implicit and Explicit Attitudes toward Foreign Accented Speech. *Journal of Language and Social Psychology* 32: 3–20. [CrossRef]

Petty, Richard E., Pablo Briñol, Chris Loersch, and Michael J. McCaslin. 2009. The Need for Cognition. In *Handbook of Individual Differences in Social Behavior*. New York: The Guilford Press.

Poplack, Shana. 1980. Sometimes i'll Start a Sentence in Spanish y Termino En Espanol: Toward a Typology of Code-Switching1. *Linguistics* 18: 581–618. [CrossRef]

Prada, Josh. 2019. Exploring the Role of Translanguaging in Linguistic Ideological and Attitudinal Reconfigurations in the Spanish Classroom for Heritage Speakers. *Classroom Discourse* 10: 306–22. [CrossRef]

Quan, Tracy. 2020. Critical Language Awareness and L2 Learners of Spanish: An Action-Research Study. *Foreign Language Annals* 53: 897–919. [CrossRef]

Quijano, Anibal. 2000. Coloniality of Power and Eurocentrism in Latin America. *International Sociology* 15: 215–32. [CrossRef]

R Core Team. 2018. *R: A Language and Environment for Statistical Computing*. Vienna: R Foundation for Statistical Computing. Available online: https://www.R-project.org/ (accessed on 12 February 2023).

Rangel, Natalie, Verónica Loureiro-Rodríguez, and María Irene Moyna. 2015. 'Is That What I Sound like When I Speak?': Attitudes towards Spanish, English, and Code-Switching in Two Texas Border Towns. *Spanish in Context* 12: 177–98. [CrossRef]

Román, Diego, Alberto Pastor, and Deni Basaraba. 2019. Internal Linguistic Discrimination: A Survey of Bilingual Teachers' Language Attitudes toward Their Heritage Students' Spanish. *Bilingual Research Journal* 42: 6–30. [CrossRef]

Rosa, Jonathan. 2016. Standardization, Racialization, Languagelessness: Raciolinguistic Ideologies across Communicative Contexts. *Journal of Linguistic Anthropology* 26: 162–83. [CrossRef]

Rosa, Jonathan, and Nelson Flores. 2017. Unsettling Race and Language: Toward a Raciolinguistic Perspective. *Language in Society* 46: 621–47. [CrossRef]

Rosseel, Laura, Dirk Speelman, and Dirk Geeraerts. 2019. The Relational Responding Task (RRT): A Novel Approach to Measuring Social Meaning of Language Variation. *Linguistics Vanguard* 5: 20180012. [CrossRef]

Seltzer, Kate. 2022. Enacting a Critical Translingual Approach in Teacher Preparation: Disrupting Oppressive Language Ideologies and Fostering the Personal, Political, and Pedagogical Stances of Preservice Teachers of English. *TESOL Journal* 13: e649. [CrossRef]

Seltzer, Kate, and Beth Wassell. 2022. Toward an Antiracist World Language Classroom: A Translanguaging Approach. In *The Antiracist World Language Classroom*. Edited by Cécile Accilien and Krishauna Hines-Gaither. Abingdon: Taylor & Francis.

Silverstein, Michael. 2018. Monoglot 'Standard' in America: Standardization and Metaphors of Linguistic Hegemony. In *The Matrix of Language*. Edited by Donald Brenneis and Ronald K. S. Macaulay. New York: Routledge, pp. 284–306.

Siordia, Celestina, and Kathy MinHye Kim. 2022. How Language Proficiency Standardized Assessments Inequitably Impact Latinx Long-term English Learners. *TESOL Journal* 13: e639. [CrossRef]

Solís Obiols, Marina. 2002. The Matched Guise Technique: A Critical Approximation to a Classic Test for Formal Measurement of Language Attitudes. *Noves SL.: Revista de Sociolingüística* 1: 1–6.

Solorzano, Daniel G., and Tara J. Yosso. 2001. Critical Race and LatCrit Theory and Method: Counter-Storytelling. *International Journal of Qualitative Studies in Education* 14: 471–95. [CrossRef]

Soukup, Barbara. 2013. On Matching Speaker (Dis) Guises–Revisiting a Methodological Tradition. In *Language (De) Standardisation in Late Modern Europe: Experimental Studies*. Edited by Tore Kristiansen and Stefan Grondelaers. Oslo: Novus, pp. 267–85.

Stefanowitsch, Anatol. 2005. The Matched Guise Technique. *Empirical Methods in Linguistics*, 1–2.

Sun, Dongyun. 2014. From Communicative Competence to Interactional Competence: A New Outlook to the Teaching of Spoken English. *Journal of Language Teaching & Research* 5: 1062–1070.

Tamminga, Meredith. 2017. Matched Guise Effects Can Be Robust to Speech Style. *The Journal of the Acoustical Society of America* 142: EL18–EL23. [CrossRef]

Train, Robert. 2003. The (Non) Native Standard Language in Foreign Language Education: A Critical Perspective. In *The Sociolinguistics of Foreign Language Classrooms: Contributions of the Native, the near-Native and the Non-Native Speaker*. Edited by Carl Blyth. Boston: Thomson Place, pp. 3–39.

Train, Robert. 2007. 'Real Spanish:' Historical Perspectives on the Ideological Construction of a (Foreign) Language. *Critical Inquiry in Language Studies* 4: 207–35. [CrossRef]

Train, Robert. 2020. Contesting Regimes of Variation: Critical Groundwork for Pedagogies of Mobile Experience and Restorative Justice. *Critical Multilingualism Studies* 8: 251–300.

Valdés, Guadalupe. 2004. Between Support and Marginalisation: The Development of Academic Language in Linguistic Minority Children. *International Journal of Bilingual Education and Bilingualism* 7: 102–32. [CrossRef]

Valdé s, Guadalupe, J. Peyton, J. Ranard, and S. McGinnis. 1999. Heritage language students: Profiles and possibilities. *Heritage Languages in America: Preserving a National Resource*, 37–80.

Van Compernolle, Remi A., and Lawrence Williams. 2012. Reconceptualizing Sociolinguistic Competence as Mediated Action: Identity, Meaning-making, Agency. *The Modern Language Journal* 96: 234–50. [CrossRef]

Vana, Rosti Frank. 2020. Learning with an Attitude?!: Heritage and L2 Students' Language Attitudes toward Spanish Language Varieties in the Advanced Mixed Class. Ph.D. dissertation, Arizona State University, Tempe, AZ, USA. Available online: https://www.proquest.com/docview/2408544341/abstract/FA3BAFAB5A5242B5PQ/1 (accessed on 12 February 2023).

VanPatten, Bill. 2015. 'Hispania' White Paper: Where Are the Experts? *Hispania* 98: 2–13. [CrossRef]

Wasserscheidt, Philipp. 2021. A Usage-Based Approach to 'Language' in Language Contact. *Applied Linguistics Review* 12: 279–98. [CrossRef]

Weatherholtz, Kodi, Kathryn Campbell-Kibler, and T. Florian Jaeger. 2014. Socially mediated syntactic alignment. *Language Variation and Change* 26: 387–420. [CrossRef]

Article

The Indeterminacy of Social Meaning Linked to 'Mexico' and 'Texas' Spanish: Examining Monoglossic Language Ideologies among Heritage and L2 Spanish Listeners

Brendan Regan * and Jazmyn L. Martinez

Department of Classical & Modern Languages & Literatures, Texas Tech University, Lubbock, TX 79409, USA; jazmyn.l.martinez@ttu.edu
* Correspondence: brendan.regan@ttu.edu

Abstract: This study examines how implied speaker nationality, which serves as a proxy for bilingual/monolingual status, influences social perception and linguistic evaluation. A modified matched-guise experiment was created with the speech of eight bilingual U.S. Spanish speakers from Texas talking about family traditions; the speech stimuli remained the same, but the social information provided about the speakers–whether they were said to be from Mexico (implied monolingual) or from Texas (implied bilingual)–varied. Based on 140 listeners' responses (77 L2 Spanish listeners, 63 heritage Spanish listeners), quantitative analyses found that overall listeners evaluated 'Mexico' voices as more able to teach Spanish than 'Texas' voices. However, only heritage listeners perceived 'Mexico' voices as being of higher socioeconomic status and of more positive social affect than 'Texas' voices. Qualitative comments similarly found that heritage listeners evaluated 'Mexico' voices more favorably in speech quality and confidence than 'Texas' voices. The implications are twofold: (i) the social information of implied monolingualism/bilingualism influences listeners' social perceptions of a speaker, reflecting monoglossic language ideologies; and (ii) there exists indeterminacy between language and social meaning that varies based on differences in lived experiences between L2 and heritage Spanish listeners. Extending on previous findings of indeterminacy between linguistic variants and meaning, the current study shows this also applies to (implied) language varieties, demonstrating the role of language ideologies in mediating social perception.

Keywords: sociolinguistics; social meaning; social perception; language ideologies; language attitudes; bilingualism; Texas Spanish; U.S. Spanish; indeterminacy

1. Introduction

Speech perception studies have been fundamental to understanding the social meaning[1] of linguistic variation and language varieties. It has been shown that linguistic information affects social perceptions of speakers (Barnes 2015; Campbell-Kibler 2007; Chappell 2016; Regan 2022c; Walker et al. 2014; Wright 2021a) and that social information affects linguistic perception as well (Barnes 2019; Hay et al. 2006a, 2006b; Hay and Drager 2010; Koops et al. 2008; Niedzielski 1999). For example, subtle differences in linguistic information,[2] such as hearing affricate [tʃ] or fricative [ʃ] for the voiceless prepalatal /tʃ/ in Andalusian Spanish (Regan 2020), can affect the social perception[3] of the speaker. Social information, such as implied speaker nationality (Hay et al. 2006a; Niedzielski 1999) or implied speaker ethnicity (Rubin 1992; Gutiérrez and Amengual 2016), has also been found to affect social perception and/or linguistic evaluation.[4] The majority of studies thus far have focused on the role of linguistic information in social perception, with fewer studies examining the role of social information in social perception and/or linguistic evaluations, and the present study seeks to build on this body of work. Following previous research (Hay et al. 2006a; Rubin 1992; among others), we seek to determine how the presence

of social information affects how listeners perceive a speaker and evaluate their speech. This approach allows sociolinguists to shed light on how language attitudes and language ideologies mediate the social meaning of language.

The current study focuses on the context of Texas, where both monolingual and bilingual varieties of Spanish are well represented. More specifically, it investigates the effect of a speaker's implied nationality (i.e., 'from Mexico' or 'from Texas'), which here serves as a proxy for bilingual or monolingual status, on listeners' social and linguistic evaluations. Doing so sheds light on the role of monoglossic ideologies (see Section 2.2) in influencing how speakers are perceived, on the one hand, and explores whether differences in bilingualism type (second language versus heritage[5] language speakers/listeners) affect one's perception of monolingual[6] and bilingual varieties of Spanish, on the other. To set the stage, Section 2 presents the background information, and the methodology is provided in Section 3. Section 4 reports the results of quantitative and qualitative analyses. Finally, Section 5 discusses the results in relation to the research questions, previous research, and theoretical implications.

2. Background

2.1. Social Information in Speech Perception

Rather than a linguistic variable having static meaning, third-wave[7] approaches to language variation maintain that a variable's social meaning is constantly renegotiated in various contexts and styles (Eckert 2005, p. 94). Eckert (2008), building on Silverstein's (2003) notion of indexical order, proposed the indexical field to theorize the social meaning of linguistic variation. The indexical field is a "constellation of meanings that are ideologically linked. As such it is inseparable from the ideological field and can be seen as an embodiment of ideology in linguistic form" (Eckert 2008, p. 464). The multitude of social meanings[8] attached to linguistic variants has been supported by several sociolinguistic perception studies, where the manipulation of a single phonetic variant (Barnes 2015; Campbell-Kibler 2007, 2008, 2011; Chappell 2016, 2018, 2019, 2020, 2021a; Regan 2020, 2022b, 2022c; Walker et al. 2014; Wright 2021a, 2021b) or of a single word (Baird et al. 2018; Regan 2022a) between guises affects listeners' evaluations. Other sociolinguistic perception studies have examined the indexical fields of language varieties[9] (Callesano and Carter 2019; Carter and Callesano 2018; Chappell and Barnes 2023; Niedzielski and Preston 1999), demonstrating that speech perception research can further our understanding of the social meaning of language varieties as well.

While most speech perception work has examined the role of linguistic information on social perception, there is a growing number of studies exploring the effect of social information on linguistic and social perception. For instance, implied age, portrayed through the visual stimuli of photos of older and younger speakers, has been used to examine the linguistic perception of sound change, such as the NEAR-SQUARE merger-in-progress in New Zealand (Hay et al. 2006b) and the split-in-progress of the PIN-PEN merger in Houston, TX (Koops et al. 2008), finding that listeners were aware of the phonetic distinction among older speakers in New Zealand and in younger speakers in Houston, respectively. Barnes (2019) used speaker photos (one urban, one rural) to examine notions of urban-ness/rural-ness on the linguistic perception of a feature from Asturian Spanish (a contact variety in Spain) and found that the Spanish variant was heard more with the urban cosmopolitan photo. Implied ethnicity is perhaps the most studied social factor influencing speech perception; these studies tend to use speaker photos to suggest different ethnicities (Babel and Russell 2015; Chappell and Barnes 2023; Gutiérrez and Amengual 2016; Kutlu 2020; Rubin 1992; Staum Casasanto 2010), resulting in different evaluations of accentedness, comprehensibility, or social qualities (like religiousness).

Especially relevant to the present paper are studies that have analyzed how implied nationality affects speech perception. For example, Niedzielski (1999) found that Detroit listeners who were presented with a 'Canadian' label heard more raised diphthong /aw/ than listeners with a 'Detroit' label, even though the raised variant is common on both

sides of the border, showing that the labels activated social stereotypes that attribute this pronunciation to Canadians. Similarly, national labels (Carter and Callesano 2018; Hay et al. 2006a), stuffed animals (Hay and Drager 2010), and negative/positive information about a nation (Walker et al. 2018) have been shown to affect linguistic and social perception. For example, Carter and Callesano (2018) found that the inclusion of national labels (Colombian, Cuban, Peninsular Spanish) affected Miami listeners' perceptions of a speaker's family wealth and salary, indicating that social stereotypes about different countries influenced socioeconomic judgements about speakers, even with a label–input mismatch.

A recurring finding from perception studies is that the relationship between linguistic form and social meaning has a "multiplicity and indeterminacy of indexical relations" (Johnstone and Kiesling 2008, p. 5). For example, in examining the social perception of monophthongal /aw/ as an index of localness in Pittsburgh, Pennsylvania, Johnstone and Kiesling (2008) found that those who produce the least amount of /aw/ monophthongization were more likely to associate audio clips with the /aw/ monophthong with Pittsburgh while those who produced more /aw/ monophthongization did not perceive /aw/ monophthong as being from Pittsburgh. They posit that indeterminacy between linguistic forms and social meaning exists because of differences in lived experiences. Specifically, they state that

> "It is people's lived experiences that create indexicality. Since every speaker has a different history of experience with pairings of context and form, speakers may have many different senses of the potential indexical meanings of particular forms. Indexical relations are forged in individuals' phenomenal experience of their particular sociolinguistic worlds." (Johnstone and Kiesling 2008, p. 29)

Numerous differences in lived experiences have been shown to create an indeterminacy between language and social meaning. For example, in examining the social perception of Spanish place names in Texas (with English phonology versus Spanish phonology), Regan (2022a) found that, while all listeners perceived Spanish place names with Spanish phonology similarly in some regards (e.g., as more respectful and friendlier), only Hispanic listeners perceived them as more educated, local to Austin, and older, while non-Hispanic listeners perceived them as non-local. In other words, the indexical meaning of place names in Austin was highly dependent on the listener's background. As another example, after observing a gender effect in which only male listeners produced more local variants when presented with a stuffed animal associated with a sports rival, Hay and Drager (2010) suggest this result could be related to sports fandom, which serves as a strong marker of nationalism in New Zealand and interacts with gender (see also Drager et al. 2010).[10] Previous perception studies of Costa Rican Spanish (Chappell 2016) and Andalusian Spanish (Regan 2022c) have found that female voices are judged more negatively than male speakers for using less institutionally prestigious features, and Chappell (2016, p. 372) suggests that awareness of the differential social payout for using local forms leads women to avoid them more than men.

One difference in life experience that warrants more research is that of bilingualism. In exploring attitudinal differences between U.S.-born bilinguals and Spanish-dominant Mexican listeners toward monolingual Mexican and bilingual heritage speech, Chappell (2021b) found that Mexican listeners "exhibited an in-group preference for the Mexican speakers' Spanish", while also taking a more "critical tone" in highlighting perceived "incorrect" aspects of the heritage Spanish speakers' Spanish (Chappell 2021b, p. 153). The U.S.-born bilinguals, on the other hand, valued both Mexican and U.S.-born voices. As Chappell (2021b) indicates, the Mexican listeners demonstrated a more hierarchical view based on language, using Spanish as a proxy for status and education, while U.S.-born bilinguals saw Spanish as serving more of a "communal, familial, and cultural role" (p. 154). Perhaps one of the only studies to examine differences in the lived experiences of L2 listeners is that of Chappell and Kanwit (2022), who examined L2 Spanish listeners' social evaluation of sociophonetic variation (coda /s/ as [s] and [h]). They found that advanced L2 Spanish listeners were capable of acquiring the indexical values of phonetic variants

in their second language, especially those that had previously taken a phonetics course and participated in study abroad in a coda-/s/ aspirating region. In addition to examining the effect of language ideologies on social perceptions, the present study extends this last line of work, exploring how the differences in lived experiences between L2 and heritage listeners impact their perception of language varieties.

2.2. Language Ideologies in Sociolinguistic and Language Attitude Studies

Linguistic anthropologists have long examined language ideologies (Kroskrity 2004; Irvine and Gal 2000; Schieffelin et al. 1998; Woolard 1998, 2008; Woolard and Schieffelin 1994), or people's "beliefs, or feelings, about languages as used in their social worlds" (Kroskrity 2004, p. 498), and Milroy (2004) and Woolard (2008) have called for language ideologies to have a more prominent role in sociolinguistics. One such ideology that has been frequently studied from a qualitative perspective is the MONOGLOSSIC LANGUAGE IDEOLOGY (Silverstein 1996; Lippi-Green 2012; Fuller and Leeman 2020; Leeman 2004), which we use here to refer to the notion that monolingual varieties are valued as "more correct" than bilingual varieties due to their lack of contact with another language. By default, such an ideology is not an additive but rather a deficit bilingual perspective.[11] It is worth noting, however, that ideologies are plural (Kroskrity 2004) and many times overlapping. The notion that a monolingual variety is "more correct" than a bilingual variety due to lack of contact also overlaps with a STANDARD LANGUAGE IDEOLOGY (Lippi-Green 2012; Milroy 2001; Milroy and Milroy 1999), which Lippi-Green (2012, p. 67) defines as "a bias towards an abstracted, idealized, homogeneous spoken language which is imposed and maintained by dominant bloc institutions, and which names as its model the written language, but which is drawn primarily from the spoken language of the upper middle class". This ideology is based heavily on the notion of correctness. As Milroy (2001, p. 535) indicates, "when there are two or more variants of some word or construction, only one of them can be right. It is taken for granted as common sense that some forms are right and others wrong". For example, the devaluing of a word such as *la troca* in bilingual U.S. Spanish in Texas (a linguistic borrowing from the English word *truck*, with phonological adaptation) as opposed to *el camión* or *la camioneta* (depending upon the dialect) is an example of the intersection of a monoglossic language ideology and a standard language ideology. While such a monoglossic language ideology[12] will be examined within the bilingual context of Texas, this ideology can operate anywhere there exists language contact.

Several studies have examined language attitudes of bilingual U.S. Spanish that reflect such a monoglossic language ideology. These studies have found that, although Mexican-Americans have positive attitudes toward remaining bilingual for both communicative and identity purposes (Mejías and Anderson 1988; Galindo 1995; Rangel et al. 2015), features of bilingual speech, such as accented English (Ryan and Carranza 1977) or code-switching (Rangel et al. 2015), have been shown to elicit less favorable evaluations than monolingual practices. Riegelhaupt and Carrasco (2000) found that the use of just a few features of bilingual Spanish were generalized by monolingual speakers to label the speaker as uneducated or of low social status. Furthermore, Goble (2016) and Tseng (2021) found that third- and second-generation U.S. Spanish speakers, respectively, tend to feel linguistic insecurity (see also Martínez and Petrucci 2004) when speaking Spanish with older generations who are viewed as having "native-like" Spanish, and this insecurity is further intensified by familial teasing. Self-perceptions of linguistic abilities may affect heritage speakers' interactions with Spanish monolinguals and/or Spanish-dominant speakers as well (Guerrero-Rodríguez 2021).

The internalization of monoglossic language ideologies that create linguistic insecurity for bilingual U.S. Spanish speakers has also been attributed in part to socialization in the education system (Leeman 2012), in which more value is given to L2 bilingualism than heritage language bilingualism[13] (Beaudrie and Loza 2023; Valdés et al. 2003) or, alternatively, to monolingual Spanish as opposed to bilingual varieties of Spanish (Achugar and Pessoa 2009; Valdés et al. 2003). Other work (Lowther Pereira 2010; Loza 2019) has

demonstrated that instructional practices disfavor U.S. Spanish. Research in this area has adopted a raciolinguistic perspective (Flores and Rosa 2015) to show that racialized speakers, such as Latinx speakers in the U.S., are often "perceived as linguistically deficient even when engaging in language practices that would likely be legitimized or even prized were they produced by white speaking subjects" (Rosa and Flores 2017, p. 628).

2.3. Research Questions

The current project examines implied nationality, which here serves as a proxy for perceived bilingual/monolingual status, to examine the role of monoglossic language ideologies in mediating the social meaning of bilingual Spanish in Texas in two different populations of listeners. The project was guided by two main research questions: (i) What is the effect of social information, namely whether a speaker is said to be from 'México' or 'Texas', on the social perception and linguistic evaluation of the speaker? and (ii) How do speaker and listener characteristics affect these social perceptions and linguistic evaluations? Most importantly, do L2 and heritage Spanish listeners differ in their evaluations?

3. Methodology

3.1. Stimuli

A modified matched-guise experiment (Lambert et al. 1960) was created with stimuli taken from informal sociolinguistic interviews, following previous studies (Campbell-Kibler 2007; Regan 2020, 2022b, 2022c). As noted by Campbell-Kibler (2007, p. 34), spontaneous speech sacrifices control of content, but also provides for more naturalistic data. To keep the content relatively similar between speakers, only clips from the sociolinguistic interviews that dealt with family traditions, holidays, and foods were included (see Appendix A for more information).

Eight Spanish-English bilingual speakers (four female, four male) produced the stimuli for the study. They were all pursuing undergraduate degrees at a large Texas public university, were between the ages of 20 and 24 (*Mean*: 21.6; *SD*: 1.2), and were all born in Texas with parents from Mexico. Thus, according to Silva-Corvalán's (1994) notion of sociolinguistic generation, they would all be considered second-generation (G2). Participants were recorded with a Marantz PMD660 solid-state digital recorder and a Shure WH20XLR head-worn dynamic microphone with a sampling rate of 4.1 kHz (16-bit digitization). The sociolinguistic interviews were conducted by the first author[14] in the sound-treated *Sociolinguistics & Bilingualism Research Lab* in the fall of 2018 and winter of 2019. The interviews ranged between 40 and 60 min and were conducted in Spanish, but participants were told that they should feel free to code-switch between languages whenever they wanted. The speakers were asked open-ended questions about their studies, professional future plans, their home city/town in Texas and what they liked about it, family traditions (holidays, birthdays, *quinceañeras*, etc.), traditional family foods, trips to visit family in Mexico, sports, and identity.

The first author selected two clips lasting between 8 and 12 seconds long from each participant. These clips did not contain any sections of code-switching into English or salient English influence to avoid confounds. Additionally, the selections avoided any repetitions or pauses. These clips were then presented to the first author's colleague, a linguist who is a Mexican Spanish speaker with extensive Spanish language teaching experience within Texas and is thus familiar with both monolingual Mexican varieties and Texas bilingual varieties. She listened to each clip and provided her input on which sounded more fluid for the purposes of the project, and the audio file she selected (one per speaker) was incorporated into the study (see Appendix B). The final clips ranged from 8.51 to 12.23 seconds long (*Mean*: 10.67, *SD*: 1.45). Individual audio files were normalized for intensity (dB) in Praat (Boersma and Weenink 2019) using the *Modify > Scale Intensity* function[15] in order to bring all sound files to an overall range of 65 to 70 dB.

3.2. Experimental Design

The eight audio files were uploaded into an online survey in Qualtrics (2005–2023). Following previous studies (Barnes 2015; Regan 2020, 2022a, 2022b, 2022c), two versions of the survey were created and branched so that each voice was only heard once by each listener. This helps to reduce the overall time of completion and voice recognition, as each speaker has a unique utterance. Each speaker (i.e., audio clip) was given a pseudonym (Natalia, María, Sofía, Rosa, José, Juan, Alejandro, Pedro). These specific names were selected because they are some of the most frequent names in Mexico (bbmundo 2017; W Radio 2017) and, by default, are also quite common among Spanish speakers in Texas. While the audio was the same for Versions A and B for each speaker, the two social guises varied based on implied nationality: "María from Texas sounds…" or "María from México sounds…". The complete experimental design can be seen in Table 1. The audio files within each block were randomized and participants were shown the blocks in random order, such that some participants were presented with block 1 and then block 2, while others started with block 2 and then heard block 1.

Table 1. Experimental design.

BLOCK	GENDER	PSEUDONYM	VERSION A	VERSION B
1	Female	Natalia	Texas	México
	Female	María	México	Texas
	Male	José	Texas	México
	Male	Juan	México	Texas
2	Female	Sofía	Texas	México
	Female	Rosa	México	Texas
	Male	Alejandro	Texas	México
	Male	Pedro	México	Texas

Upon consenting to the survey terms, participants (see Section 3.3 for listener participant recruitment) were asked to wear headphones and were told that they would hear eight short audio files ranging from 8 to 12 seconds long (see Appendix B for audio files). They were able to listen to each recording as many times as they liked and then responded to a series of questions to evaluated each speaker on a six-point Likert scale,[16] as seen in Figure 1. The first six questions elicited evaluations of perceived social class, educational level, intelligence, friendliness, confidence, and eloquence[17] of speech. Similar to Chappell (2021b, p. 143), although the recordings were in Spanish, the questions were presented in English, as the L2 listeners and most heritage listeners were English-dominant bilinguals. The final question (*Do you think they could help you with your Spanish?*) was designed to prompt reflection about whether or not listeners thought the speaker's Spanish could serve as a pedagogical model. There was also an optional open-ended question for each voice. Finally, after completing all eight evaluations, listeners answered basic demographic questions about themselves, including their gender (male, female, self-identify [write-in]), age, home city, years lived in Texas, number of trips to Mexico, current Spanish class, and whether they identified as a Spanish heritage speaker or a second-language speaker of Spanish.

3.3. Implementation and Participants

All listener participants were undergraduate students currently enrolled in Spanish courses at a large public university in Texas. Over the course of 1.5 weeks in March of 2019, instructors of each upper-level Spanish course who agreed to have their classes participate brought their students to the department's *Language Lab & Research Center*, where each participant had their own desktop computer and headphones. The survey link was placed

on the Blackboard website of each course for the duration of the class period. Upon clicking the link, participants were asked to consent to the survey terms and confirm they were 18 years or older. Those who consented and confirmed their eligibility continued with the study, while skip logic took ineligible participants to the end of the survey, which prevented their participation. Responses were necessary to continue in the study, with the exception of the optional open-ended question (see Figure 1). Only completed surveys were used in the analysis.

Figure 1. Screenshot of survey questions for one audio file.

There were 140 listeners (110 female, 30 male), ranging in age from 18 to 57 (*Mean*: 21.6, *SD*: 4.0), who participated in the experiment. They were roughly balanced by bilingualism type with 77 listeners who were L2 Spanish speakers and 63 listeners who were heritage Spanish speakers. At the time of participation, all were enrolled in third- (junior-level) and fourth- (senior-level) year Spanish classes. The majority (*n* = 132) of the participants were from Texas[18] (Amarillo: 3, Austin: 16, Dallas/Fort Worth: 38, El Paso: 9, Houston: 18, Lubbock: 26, McAllen/Brownsville: 4, Midland/Odessa: 2, Presidio: 1, San Antonio: 13, Waco: 1, Wichita Falls: 1), with a few (*n* = 8) from other states (California: 1, "East Coast": 1, Florida: 1, Mississippi: 1, New Mexico: 4).

3.4. Quantitative (Statistical) Analysis

All six-point Likert scales were centered on zero and then subject to a principal component analysis (PCA) and factor analysis in R (R Core Team 2023) to determine whether there were any correlations between the dependent measures and, if so, combine

the correlated measures. The factor analysis revealed that the seven measures could be reduced to four factors, with three accounting for the majority of the variation ($p < 0.01$). Following Weatherholtz et al.'s (2014, p. 400) cutoff of 0.4 to determine whether a variable loaded onto a factor, the first factor strongly loaded for education (0.779) and intelligence (0.795). Thus, these two factors were combined to form "perceived education". The second factor loaded for friendliness (0.635), confidence (0.805), and eloquence (0.556), which were then combined into a single factor entitled "perceived social affect". A third factor only strongly loaded for socioeconomic status (0.861), and ability to help with one's Spanish did not load strongly onto any of the three aforementioned factors. For this reason, both socioeconomic status and ability to help with Spanish were considered separately. As a result, there were a total of four continuous dependent measures: socioeconomic status, education (education and intelligence combined), social affect (friendliness, confidence, and eloquence combined), and ability to help with one's Spanish.

Each dependent variable was subject to mixed-effects linear regression modeling using the *lme4* (Bates et al. 2015) and *lmerTest* (Kuznetsova et al. 2017) packages in R with the random intercepts of speaker and listener. The independent variables tested in the modeling were (i) guise ('México' voices,[19] 'Texas' voices); (ii) speaker gender (male, female); (iii) listener gender (male, female); (iv) listener city/town (border area, non-border area); (v) study abroad experience in a country where Spanish is the majority language (yes, no); (vi) frequency of trips to Mexico; (vii) course level (third-year, fourth-year); and (viii) listener bilingualism type (L2, heritage). Frequency of trips to Mexico was treated as a categorical variable of never, 1–5 times, 5–10 times, or more than 10 times. While it would have been ideal to treat this variable as a continuous measure, some listeners wrote comments (instead of raw numbers) such as "too many times to count". Thus, we were unable to assign them an exact number of times and these listeners were coded as more than 10 times. After further analysis, we observed that both listener city/town and frequency of trips to Mexico demonstrated high collinearity with bilingualism type, as very few L2 listeners lived in border regions or had visited Mexico frequently. Thus, bilingualism type was included in the regression models while listener city/town and frequency of trips[20] to Mexico were not. Model construction began with all independent variables and each social factor in interaction with guise, and non-significant factors were gradually removed. Three-way interactions that included bilingualism type and guise with all other social factors were also tested. Non-significant interactions with guise were removed from subsequent models, and in the case of interactions with more than two categorical levels, estimated marginal means (Lenth et al. 2018) were implemented to conduct post hoc analyses.

3.5. Qualitative Analysis

As previous researchers have stated (Baird et al. 2018; Campbell-Kibler 2010), some quantitative Likert-scale questions may not be able to uncover all language attitudes. For this reason, and following previous studies (Baird et al. 2018; Kirtley 2011; Nance 2013), the qualitative comments in response to "Anything else that occurs to you about this person based on their speech?" were subjected to word clouds. Rather than listing all of the words within the word cloud, the authors coded for any underlying themes. Thus, when possible, semantic themes were used in place of longer phrases, but only comments that were truly of the same semantic theme were combined. For example, comments related to quality of speech were deemed either "speaks well" or "speaks poorly" while observations related to speech rate were classified as either "speaks fast" or "speaks slowly". To ensure objectivity in organizing these semantic themes, both authors separately coded and classified each comment. Given that the semantic category was more important than the individual word, all Spanish comments were translated into English semantic themes so that they would be represented in the same category for the word clouds. Of the possible 168 themes, the coding of both authors aligned on 120 specific descriptors, which constitutes a 71.4% agreement rate. The authors reviewed together the 48 semantic codes for which they did not have the exact same descriptors, most often due to a difference in synonyms. The

finalized semantic themes for the 'Mexico' and 'Texas' voices were subjected to word clouds using the *wordcloud()* function in R. The size of the word/phrase in each word cloud is representative of its frequency, with larger words/phrases being more frequent in number than the smaller words/phrases.

4. Results

4.1. Quantitative Results

The results for each mixed-effects model are presented in Table 2, which displays the estimate, the standard error (*SE*), *t*-value, and *p*-value. Negative estimates indicate a lower rating than the reference level, while positive estimates indicate a higher rating than the reference level. Within the table, each model also has marginal R-squared (R^2m) and conditional R-squared (R^2c) values to assess how well the model explains the variation (Nakagawa and Schielzeth 2013). There were no significant main effects of guise or significant interactions with guise for perceived education (education and intelligence combined). As such, it will not be discussed further.

Table 2. Summary of mixed-effects linear regression models for PERCEIVED SOCIOECONOMIC STATUS, PERCEIVED SOCIAL AFFECT, and PERCEIVED ABILITY TO HELP WITH ONE'S SPANISH, speaker and listener as random intercepts, *n* = 1120 for each model. Reference levels are 'Mexico' for guise, L2 for listener bilingualism type, female for speaker gender, and female for listener gender.

Predictor	Estimate	SE	*t*-Value	*p*-Value
PERCEIVED SOCIOECONOMIC STATUS (R^2m: 0.01, R^2c: 0.35)				
(Intercept)	0.04	0.14	0.29	0.777
Guise = Texas	0.06	0.07	0.99	0.324
Bilingualism Type = Heritage	−0.05	0.11	−0.50	0.617
Guise: Bilingualism Type	−0.21	0.10	−2.17	0.031 *
PERCEIVED SOCIAL AFFECT (R^2m: 0.05, R^2c: 0.46)				
(Intercept)	0.23	0.14	1.64	0.129
Guise = Texas	0.05	0.07	0.74	0.458
Bilingualism Type = Heritage	−0.12	0.11	−1.10	0.273
Speaker Gender = Male	−0.35	0.18	−1.96	0.087
Listener Gender = Male	0.16	0.12	1.34	0.18
Guise: Bilingualism Type	−0.23	0.11	−2.19	0.029 *
Guise: Speaker Gender	−0.10	0.10	−0.97	0.333
Bilingualism Type: Speaker Gender	0.003	0.11	0.031	0.976
Guise: Listener Gender	−0.20	0.09	−2.14	0.033 *
Guise: Bilingualism Type: Speaker Gender	0.34	0.15	2.26	0.024 *
PERCEIVED ABILITY TO HELP WITH ONE'S SPANISH (R^2m: 0.01, R^2c: 0.55)				
(Intercept)	0.07	0.09	0.83	0.414
Guise = Texas	−0.15	0.04	−3.65	0.0003 ***

Note: * = $p < 0.05$, ** = $p < 0.01$, *** = $p < 0.001$.

The model for perceived socioeconomic status revealed a significant interaction between guise and listener bilingualism type (see Figure 2A). Post hoc analyses revealed that there was no significant difference in guises for the L2 listeners (*p* = 0.324) but that heritage listeners perceived voices with a 'Mexico' label as being of a higher socioeconomic class than those with a 'Texas' label (*p* < 0.05). There was no significant difference in the evaluation of the 'Mexico' voices between L2 and heritage listeners (*p* = 0.617). However,

heritage listeners perceived 'Texas' voices as being of a lower socioeconomic status than L2 listeners ($p < 0.05$).
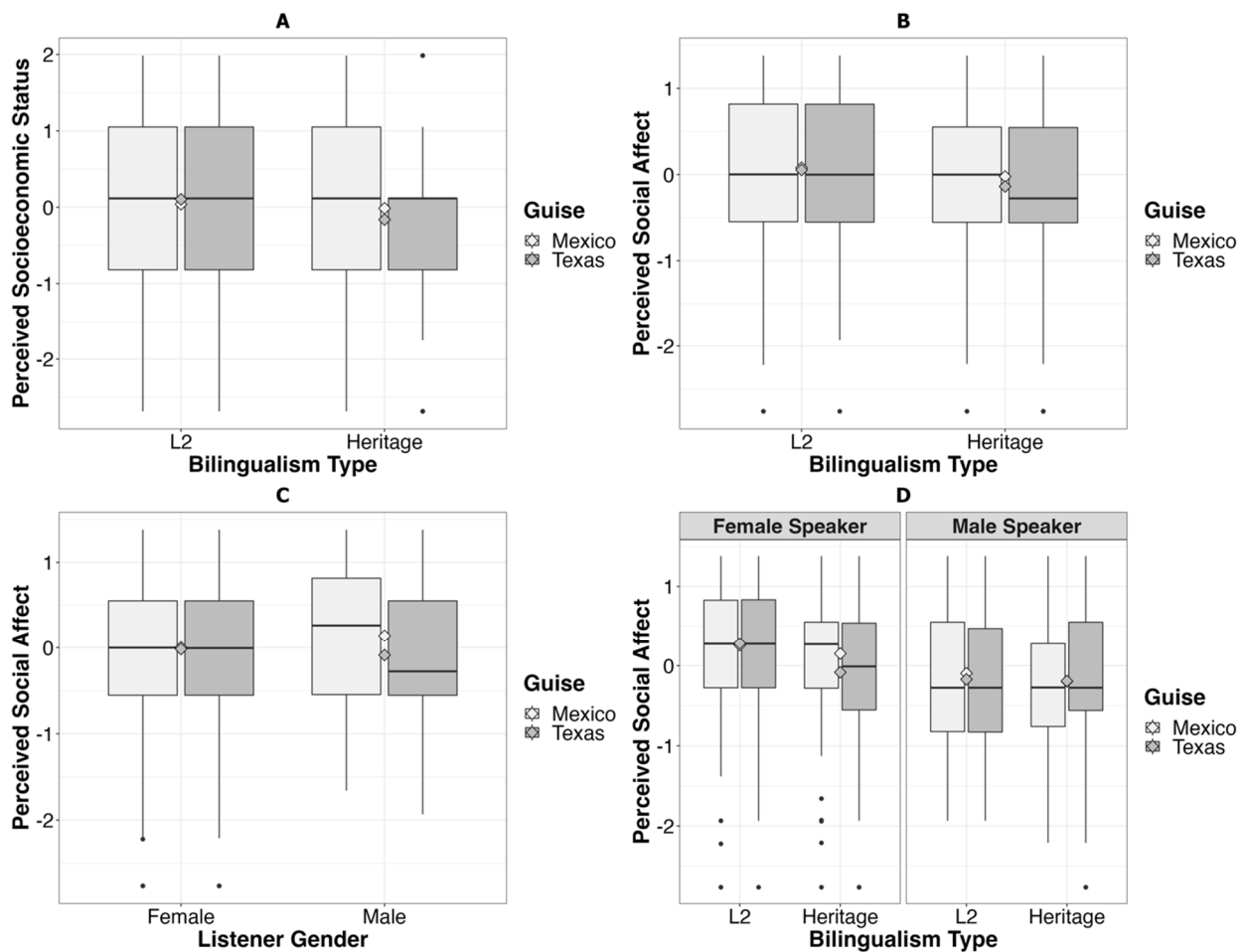

Figure 2. Guise by listener bilingualism type for perceived socioeconomic status (**A**); Guise by listener bilingualism type for perceived social affect (**B**); Guise by listener gender for perceived social affect (**C**); Guise by listener bilingualism type and speaker gender for perceived social affect (**D**).[21] Note: bold horizontal black lines denote medians and the diamonds denote means.

The model for perceived social affect revealed two significant two-way interactions between guise and bilingualism type, on the one hand, and between guise and listener gender, on the other, as well as a significant three-way interaction between guise, speaker gender, and listener bilingualism type. Post hoc analyses revealed that L2 listeners did not perceive any significant difference ($p = 0.28$) between guises while heritage listeners perceived 'Mexico' voices as having higher positive social affect than 'Texas' voices ($p < 0.01$) (see Figure 2B). There were no significant differences between L2 and heritage speakers' evaluations of 'Mexico' voices ($p = 0.22$) and 'Texas' voices ($p = 0.06$). Post hoc analyses found that female listeners ($p = 0.53$) did not evaluate speakers differently based on the guise, while male listeners perceived speakers with the 'Mexico' label as having more positive social affect than speakers with the 'Texas' label ($p < 0.01$) (see Figure 2C). The perception of social affect was not significantly different between male and female listeners for 'Mexico' ($p = 0.18$) and 'Texas' voices ($p = 0.78$). Regarding the three-way interaction, post hoc analyses indicate that L2 listeners did not perceive any significant difference in social affect between 'Mexico' or 'Texas' female voices ($p = 0.57$) nor between 'Mexico' or 'Texas' male voices ($p = 0.07$). While heritage listeners also did not perceive any significant differences in social affect between guises for male voices ($p = 0.65$), they perceived female

'Mexico' voices as having higher positive social affect than male 'Mexico' voices ($p = 0.0006$) (see Figure 2D).

The model for perceived ability to help with one's Spanish demonstrated a main effect of guise in which speakers with the 'Mexico' label were perceived as more able to help with one's Spanish than those with the 'Texas' label (see Figure 3).
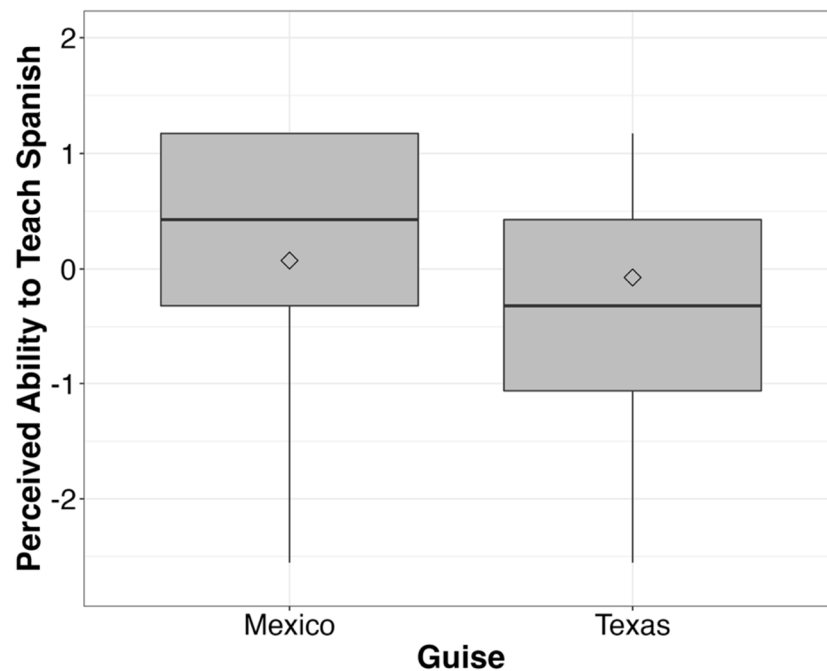

Figure 3. Main effect of guise for perceived ability to teach Spanish. Note: bold horizontal black lines denote medians and the diamonds denote means.

To briefly summarize these results, the interaction between guise and bilingualism type for perceived socioeconomic status indicated that heritage Spanish listeners perceived 'Mexico' voices as being of higher socioeconomic status than 'Texas' voices, while L2 listeners did not perceive a significant difference in socioeconomic status between guises. There were three interactions for perceived positive social affect (friendliness, confidence, and eloquence). Speakers with 'Mexico' voices were evaluated as having higher social affect than those with 'Texas' voices, but only by heritage listeners. 'Mexico' voices were given higher social affect ratings only by male listeners. Additionally, for the three-way interaction, the only difference for guises was found among heritage listeners who evaluated female speakers with the 'Mexico' label as having higher social affect than those with the 'Texas' label. Finally, for perceived ability to teach Spanish, the main effect revealed that all listeners perceived speakers with the 'Mexico' label as more able to teach Spanish than those with the 'Texas' label.

4.2. Qualitative Open-Ended Comments

Given the differences in perception based on bilingualism type in the quantitative analysis, we separated comments for L2 and heritage listeners to examine any qualitative differences in perception between groups. Regarding the 'Mexico' voices, the L2 listeners' most frequent comment was *accent*, followed by *confident*, *not confident*, *family-oriented*, *speaks well*, *slurred-pronunciation* and *low socioeconomic status (SES)* as seen in Figure 4. The word *accent* should be taken with caution as, without the full context of what listeners meant, it could imply they have a regional accent or possibly that they speak L2-accented Spanish. However, given that the listeners are L2 bilinguals, we can reasonably assume that they are referring to a Mexican-sounding accent. While *confident* and *not confident* stand in contradiction to one another, *speaks well* indicates a positive evaluation of speech quality.

Slurred-pronunciation suggests that listeners had (or perceived that they had) difficulty in distinguishing words, which may indicate the perception of a fast speech rate. Finally, while not language-related, these listeners believe the speakers are *family-oriented* and of *lower SES*. Heritage listeners' most frequent comment for the 'Mexico' voices was clearly *confident*, followed by *speaks fast* and *intimidating*, as seen in Figure 5. Of particular interest is that the descriptor *intimidating* was not used once by the L2 listeners. While there are differences between the two groups, overall, the comments are positive toward speech quality and speech rate.

Figure 4. Word cloud of comments for 'Mexico' voices by L2 listeners.

Figure 5. Word cloud of comments for 'Mexico' voices by heritage listeners.

In terms of the 'Texas' voices, the L2 listeners' most frequent comment was *speaks slowly*, followed by *speaks fast*, *speaks well*, *native*, and *difficult to understand*, as seen in Figure 6. The comments on speech rate (fast versus slow) appear to contradict one another, but then *speaks well* and *native* indicate an overall high evaluation of speech quality. For the heritage listeners, the most frequent words used to describe the 'Texas' voices were *not fluid*, *lacks proficiency*, and *not confident*, followed by *confident* and *repetition*, as seen in Figure 7. While the less frequent words demonstrate contradicting terms, the two most common descriptors are not only more negative toward speaking abilities than those used by the L2 listeners but also much more negative than the heritage listeners' comments for the 'Mexico' voices. That is, the heritage listeners' two most common descriptors for each guise stand in complete contrast: *speaks fast* and *confident* for 'Mexico' voices and *not fluid* and *not confident* for 'Texas' voices. The association of *not fluid* and *speaks slowly* with 'Texas' voices reflects the notion that monolingual speakers are more adept speakers of the language. For example, one full-length comment for "Pedro from Texas" was that he "sounds confident, but speaks rather slowly which I have noticed is a difference in speakers from Texas and Mexico. Speakers from Mexico typically are speaking faster and so that may cause them to be more intimidating" (P344,[22] 20-year-old female heritage speaker, Miami, FL).

Figure 6. Word cloud of comments for 'Texas' voices by L2 listeners.

In addition to the open-ended comment after each guise, there was also one final open-ended comment section at the end of the study for listeners to complete if they decided they wanted to add anything else. Excerpt 1 demonstrates that one listener (a second language learner) thought that all speakers were highly proficient in their speech.

(1)

They all seemed advanced in Spanish and all spoke with the same level on [*sic*] confidence and eloquence in my opinion. (P269, 21-year-old female L2 listener, Dallas, TX)

However, others indicated that they heard or perceived differences between speakers 'from Mexico' and 'Texas', as seen in Excerpt 2.

(2)

The accents and tones were different amongst Mexican speakers and non-Mexican speakers in my opinion. (P351, 20-year-old female heritage listener, Austin, TX)

Another participant reflected on the fact that she believed knowing where the participants were from may have affected how she rated each speaker, as seen in Excerpt 3.

(3)

I think that knowing where the person was from before listening to the audio clip may have influenced how I rated their Spanish, and I am not sure if that was purposeful but it is just something that came to mind. I tried not to let that influence me, but I can definitely see how it still might have affected my answers subconsciously. (P344, 20-year-old female heritage listener, Miami, FL)

One participant even acknowledged her own bias against bilingual Spanish due to her experience as a bilingual speaker, as seen in Excerpt 4.

(4)

I think adding whether the person was from Texas or from Mexico influenced my expectation about how their speech should sound. I grew up being constantly corrected on my Spanish since San Antonio has very mixed Spanish, so I feel like I've been conditioned to think that those that have more of an accent tend to be less educated especially if their parents speak Spanish at home. (P156, 20-year-old female heritage listener, San Antonio, TX)

This listener (P156) has internalized monoglossic and standard language ideologies that imply that her bilingual variety of Spanish is not as "correct" as monolingual varieties of Spanish due to San Antonio Spanish being "mixed". As mentioned previously, there were no common morphosyntactic features of bilingual U.S. Spanish in the recordings. Thus, the inclusion of a speaker's supposed nationality was enough social information to activate monoglossic and standard language ideologies, which were intensified in light of her own experience.

Figure 7. Word cloud of comments for 'Texas' voices by heritage listeners.

5. Discussion

5.1. Revisiting the Research Questions

In this section, we first discuss the results in relation to the research questions and previous literature and then the findings' theoretical implications. Regarding the first research question (*What is the effect of implied nationality on the social perception and linguistic evaluation of the speaker?*), the most notable finding was the perceived ability to help with one's Spanish: 'Mexico' voices were evaluated significantly higher than 'Texas' voices. It is important to remember that the audio guises were not digitally manipulated and

that they were all produced by bilingual U.S. Spanish speakers, indicating that the social information of supposed monolingual/bilingual speech itself affected the differences in listener perception. This is the clearest indication of the pervasiveness of the monoglossic language ideology (Silverstein 1996; Pavlenko 2002; Lippi-Green 2012; Fuller and Leeman 2020; Leeman 2004, 2018), as an implied monolingual speaker is perceived to be more adept at teaching or explaining a language than an implied bilingual speaker. Such an ideology would increase the linguistic insecurity of bilingual speakers regarding their Spanish skills as somehow not being as adequate as those of monolingual speakers. This was evidenced by some of the qualitative comments from the heritage listeners (Excerpts 2–4), as well as the heritage listener word clouds.

In terms of the second research question (*How do speaker and listener characteristics affect these social perceptions and linguistic evaluations?*), listener bilingualism type interacted with guise for two of the dependent measures. More specifically, only heritage listeners perceived differences in socioeconomic status between guises, with 'Mexico' voices ranked as being of higher socioeconomic status than 'Texas' voices. This may demonstrate that heritage listeners are more acutely aware of the fact that in the U.S. context, the public discourse associates U.S. Spanish with a lower socioeconomic status (Urciuoli 1996, p. 26; Fuller and Leeman 2020, p. 85). Whether or not it relates to their own life experience, at-large public discourse may influence how they view (bilingual) U.S. Spanish speakers versus (monolingual) Mexican Spanish speakers. If we connect the results from the perceived socioeconomic status to those from the perceived ability to teach Spanish, we are able to observe what Zentella (2007, pp. 25–26) states: "Above all, distinct ways of being Latina/o are shaped by the dominant language ideology that equates working-class Spanish speakers with poverty and academic failure, and defines their bilingual children as linguistically deficient and cognitively confused (Zentella 2002)".

The results for perceived positive social affect demonstrated additional interactions between bilingualism type and guise. Heritage listeners evaluated 'Mexico' voices as having higher positive social affect than 'Texas' voices, while L2 listeners did not perceive a significant difference between guises. This is interesting, as one may expect an in-group preference (Preston 1993) among the heritage listeners with regards to solidarity or perhaps tendencies similar to Chappell (2021b), where U.S.-born listeners demonstrated a broad conceptualization of community in which they positively evaluated both the Mexican and Mexican-American speakers. The three-way interaction would indicate that heritage listeners were more likely to perceive this difference in social affect between guises with female voices. Returning to the intersection of monoglossic and standard language ideologies, in which there is one variant or variety that is considered more correct or prestigious than the other, it may be the case that female speakers are evaluated more critically, at least by heritage listeners, for the use of any bilingual features or even—as in the current study—the mere implication of bilingual features. Previous studies (Chappell 2016; Gordon 1997; Eckert and McConnell-Ginet 1999; Regan 2022c) have found that women are judged more negatively than male speakers for using less institutionally prestigious features. Thus, because monoglossic and standard language ideologies would position bilingual varieties as less institutionally prestigious than monolingual varieties, this may explain why heritage listeners perceived 'Mexico' female voices as having more positive social affect than 'Texas' female voices. Regarding listener gender, only male listeners perceived differences in guises, evaluating 'Mexico' voices as having more social affect than 'Texas' voices. This finding should be taken with caution given there were only 30 male listeners in comparison to 110 female listeners.

5.2. Theoretical Implications

While evidence of a monoglossic language ideology was found among all listeners regarding one's ability to teach Spanish, it is notable that, in the quantitative and qualitative analyses, there were differences in the evaluations between the L2 and heritage listeners. Specifically, heritage listeners evaluated 'Mexico' voices as being of higher socioeconomic

status and having more positive social affect than 'Texas' voices, while L2 listeners generally did not perceive any other differences between guises. Although both groups are exposed to monoglossic and standard language ideologies, the findings of the current study indicate that heritage listeners may have more exposure to these ideologies than L2 listeners due to differences in lived experiences such as (i) contact with Spanish-dominant and monolingual Spanish speakers in and outside of the U.S., as well as (ii) experiences in the educational system.

While heritage speakers are not a monolith (Pascual y Cabo and Rothman 2012, pp. 451–52), it is important to consider that, in general, they have more contact with Spanish-dominant and/or monolingual Spanish speakers than L2 speakers, allowing for more internalization of monoglossic language ideologies. Previous studies have found that familial teasing for bilingual language features by more Spanish-dominant family members increases one's linguistic insecurity (Carruba-Rogel 2018; Goble 2016; Tseng 2021), as monolinguals often expect bilinguals to behave like monolinguals (Riegelhaupt and Carrasco 2000). Tseng (2021) found this particularly true among second-generation speakers who were criticized for their pronunciation. She also found that second-generation U.S. bilinguals had more linguistic insecurity in the presence of Spanish-dominant speakers and therefore would avoid speaking. Thus, "purist language beliefs imposed deficiency identities on second-generation speakers regardless of actual language use" (Tseng 2021, p. 129). It is quite possible that these ideologies are being reinforced in interactions with members of their community, as shown by the listener from San Antonio in Excerpt 4. When the listeners first came into contact with these ideologies may play a role, given that the heritage listeners were all simultaneous bilinguals or early-sequential bilinguals while the L2 listeners were late-sequential bilinguals. As a result, L2 learners would have had significantly fewer years of exposure to these Spanish-specific attitudes in childhood, furthering differences in perceptions between the two groups.

Aside from the individual's age of acquisition of Spanish, the educational system also has a role in reinforcing monoglossic language ideologies, cyclically recycling and reaffirming them. As Leeman (2012, p. 44) states, "school is a key site where young people are socialized into hegemonic value systems" such as "which kind of Spanish is 'best'". These language ideologies become "naturalized and come to be understood as common sense" (Leeman 2012, p. 46), such that "even individuals who are negatively affected by particular conceptions of language may embrace the very ideologies that subordinate them" (Leeman 2012, p. 44). Just as monolingual Spanish is granted a privileged status over bilingual varieties of Spanish (Achugar and Pessoa 2009; Valdés et al. 2003), it has been shown in multiple university[23] contexts that more value[24] is given to L2 bilingualism than to heritage language bilingualism (Beaudrie and Loza 2023; Valdés et al. 2003). For example, within a bilingual creative writing graduate program in El Paso, Texas, Achugar and Pessoa (2009) found that, while there were overall highly favorable attitudes toward bilingualism, local (bilingual) varieties of Spanish were viewed as inferior to monolingual Latin American varieties of Spanish. Similarly, in focus group interviews with professors and graduate instructors of Spanish departments across several universities, Valdés et al. (2003) observed that monolingual varieties of Spanish (Spain and Latin American) were considered the most correct varieties of Spanish and, while L2 Spanish was also viewed positively, the educators held the most negative views toward bilingual U.S. Spanish speakers. Valdés et al. (2003, p. 24) concluded,

> "[. . .] these departments are complicitous—although perhaps unconsciously—with the deep values and linguistic beliefs of American monolingualism that continue to view the United States as a profoundly English-speaking country. Both directly and indirectly, such departments transmit ideologies of nationalism (one language, one nation), standardness (a commitment to linguistic purity and correctness), and monolingualism and bilingualism (assumptions about the superiority of monolingual native speakers)."

These studies have demonstrated the role of the educational system in modeling the ideal variety of language as a monolingual one, which in turn leads to a devaluing of one's own bilingual variety. Given the "paradox of Spanish" in the U.S. (Carter 2018), in which L2 bilingualism is valued more than heritage bilingualism, it would seem from the differences in perception based on bilingualism type that heritage speakers may be more exposed to such deficit ideologies in the educational system, making them more ingrained in their evaluations of language varieties, including their own.

L2 and heritage listeners may also differ in linguistic proficiency in Spanish, which could have influenced the participants' evaluations.[25] We may assume that—at least in terms of speaking and listening—the heritage listeners here are more advanced than L2 listeners given their exposure to a wide variety of monolingual and bilingual Spanish speakers. If this is the case, the L2 listeners may have simply perceived all speakers as more advanced than they were and thus did not feel comfortable assessing any potential differences based on status and/or social affect. As mentioned previously, Chappell and Kanwit (2022) found that only more advanced L2 listeners (as opposed to less advanced L2 listeners) were able to associate coda /s/ aspiration with a geographical distribution and social status, especially among those with phonetics courses and study abroad experience, respectively. However, even the advanced L2 listeners were not able to perceive the more nuanced social meanings that L1 listeners attribute to coda /s/ aspiration. Perhaps similar to their findings, if proficiency differences existed among the two groups, this may have led to L2 listeners not being able to (or feeling qualified to) evaluate speakers on more nuanced social properties.

Finally, building on Johnstone and Kiesling's (2008, p. 25) idea of the "indeterminacy of relations between forms and meanings", the current results reveal that, even when hearing the same linguistic input, the social interpretation of speech varies due to differences in listeners' lived experiences (Johnstone 2011). While all participants in the study, L2 and heritage listeners alike, perceived the 'Mexico' and 'Texas' voices similarly for their ability to teach Spanish, the divergence in perceived socioeconomic status and social affect, as well as the qualitative comments, indicates a difference in the social meaning of bilingual varieties based on the listeners' type of bilingualism. The lived experience of heritage listeners, which includes a greater exposure to monoglossic ideologies, may lead them to evaluate bilingual varieties differently from L2 listeners. What is of particular interest is that, in this experiment, the labels of 'Mexico' and 'Texas' were enough to activate associations within monolingual (or Spanish-dominant) and bilingual speakers. These results align with Regan's (2022a, pp. 467–68) finding of two partially overlapping indexical fields for the perception of Spanish place names in Austin, TX, which varied based on listener ethnicity. While some social meaning was shared between Hispanic and non-Hispanic listeners, there were also differences between them, in which Hispanic listeners perceived Spanish phonology with Spanish place names just as local as English phonology, while non-Hispanic listeners[26] only perceived the English phonology as local to Austin. As third-wave sociolinguistic studies continue to theorize the social meaning of linguistic variation (Hall-Lew et al. 2021), more emphasis should be placed on the role of differences in the lived experiences of listeners and speakers. That is, as Johnstone and Kiesling (2008, p. 29) state, researchers should pay "attention to the multiplicity and indeterminacy of indexical relations and to the way in which such relations arise in lived experience, [which] can lead to a more nuanced account of the social meanings of variant forms in a speech community". Studies in bi/multilingual communities that only examine the social perception of linguistic variants in a broad, community-based sense may overlook this indeterminacy based on differences in lived experiences such as bilingualism type (L2, heritage), proficiency, trips to the country of family origin, cultural and emotional connection to Spanish, etc. Thus, these findings highlight the need for studies in multilingual settings to explore sociodemographic factors that may result in differential perceptions among bi/multilingual speakers and listeners. As the current study indicates, this applies not only to specific linguistic variants but entire language varieties.

It has been said that social perception is where linguistic variants and language varieties become associated with social meaning (Walker et al. 2014, p. 169). However, the current study demonstrates that commonly held language ideologies, such as monoglossic language ideologies (Silverstein 1996; Fuller and Leeman 2020; Leeman 2004) and standard language ideologies (Lippi-Green 2012; Milroy 2001; Milroy and Milroy 1999), can influence one's social perception and linguistic evaluation even without any modification to the linguistic input. That is, language ideologies may activate indexical fields of social meaning related to language varieties without the presence of the variety itself. Given that language ideologies become an entrenched "common sense" notion (Leeman 2012, p. 46), a simple social prompt of being from 'Mexico' or being from 'Texas' affects how one is socially perceived and linguistically evaluated. This has real-life implications in which bilinguals are judged based on their linguistic status as a bilingual Spanish speaker and less so on their actual linguistic practices. This finding supports a raciolinguistic perspective (Flores and Rosa 2015; Rosa and Flores 2017) in the U.S. context in which U.S. bilingual Latinxs are viewed as having deficient forms of speaking. Rosa and Flores (2017, p. 628) indicate that "language ideologies associated with social categories produce the perception of linguistic signs". That is, regardless of the actual linguistic input, ideologies associated with social categories (such as U.S. bilingual Latinxs) can shape how linguistic practices are perceived. Specifically, Rosa and Flores (2017, p. 628) indicate that these raciolinguistic ideologies produce "racialized language practices that are perceived as emanating from racialized subjects". As such, the mere suggestion that a speaker is 'from Texas', that is, a U.S. bilingual Latinx, is enough to evoke negative social and linguistic meanings for listeners. From the results of the study, this appears to be more ingrained among heritage listeners and less so among L2 listeners. Consequently, there is much work to do in K-12 and university education to continue to show the value of bilingual varieties. Following a Critical Language Awareness approach (Leeman and Serafini 2016), L2 and heritage language curricula should actively include concepts of language ideologies to examine how they mediate the social perception of language.

6. Conclusions

Using quantitative and qualitative analyses, this study has demonstrated that (i) the social information of implied monolingualism/bilingualism influences listeners' social perceptions of a speaker, reflecting monoglossic and standard language ideologies; and (ii) there exists indeterminacy between language and social meaning that varies based on differences in lived experiences between L2 and heritage listeners. Extending on Johnstone and Kiesling's (2008) finding of indeterminacy between linguistic variants and meaning, the current study shows this also applies to (implied) language varieties, highlighting the role of language ideologies in mediating social perception.

Future studies would do well to include more metalinguistic questions in the experimental design, such as the quality of speech (speaks well/speaks poorly) and the speech rate (speaks slowly/speaks quickly). Additionally, future work should attempt to disentangle exactly how participants interpret implied nationality ('Mexico', 'Texas'), as they may not truly be a proxy for a speaker's monolingualism or bilingualism. For example, it is possible that some listeners based their evaluations on national stereotypes rather than notions of monolingualism versus bilingualism. Other listeners may assume that speakers from Mexico have had more years of formal education in Spanish while speakers from Texas have received most of their formal education in English, and therefore Mexican speakers are viewed as more qualified to teach Spanish based on this factor alone. Thus, future research should continue to tease apart these factors as they explore the role of language ideologies in the evaluation of speakers and their speech.

Author Contributions: Conceptualization, B.R.; methodology, B.R.; formal analysis, B.R. and J.L.M.; writing—original draft preparation, B.R. and J.L.M.; writing—review and editing, B.R. All authors have read and agreed to the published version of the manuscript.

Appendix A. Online Instructions for the Experiment

Instructions

Social psychology studies have demonstrated that one can infer a great deal about a person by briefly hearing her/his speech. To ensure a consistent topic, each speaker was interviewed in Spanish with the same question about Mexican/Mexican-American family traditions, holidays, and foods. You will hear clips of **8 different speakers** who are from México or from Texas. Each recording is an **8–12 second clip** taken from a larger conversation. Listen to the recording as many times as you would like. Respond to the question following each recording.

Make sure to be in a quiet place, to <u>wear headphones</u>, and to complete the study individually. The study lasts **<u>10–12 minutes</u>** Don't overthink it, go with your first instinct.

Appendix B. Stimuli Phrases

1. Natalia: *Toda mi familia viene a mi casa y vamos a misa el veinticuatro y miramos como una danza que siempre hacen en la misa.* 'All my family comes to my house and we go to mass on the 24th and we see a dance that they always do at the mass'.

2. María: *Sí sí, y luego siempre hay lumbre afuera, carne asada, ponche, tamales, de to- de todo, [risa], pero así nomás.* 'Yes yes, and then there's always a fire outside, carne asada, punch, tamales of all kinds [laughter], but just like that'.

3. José: *Para el almuerzo, así como un burrito o, este: huevos rancheros, y luego para la cena, como tacos o tamales, a veces estamos en el invierno los tamales.* 'For lunch, like a burrito or um:[27], huevos rancheros and then for the dinner like tacos or tamales, sometimes when we're in the winter the tamales'.

4. Juan: *Pero por la Navidad es cuando mi mamá siempre saca el- el niño Dios, y pues nosotros ponemos a rezar, y hay tamales, menudo, de todo, y luego más comer y más risa.* 'But for Christmas is when my mother always takes out the baby Jesus, and well we begin to pray and there are tamales, menudo, and everything, and then more eating and more laughter'.

5. Sofía: *Pero tuvimos la quinceañera, la ceremonia, la fiesta, y como le dije, a mí me encanta bailar me encanta, encanta, entonces.* 'But we had the quinceañera, the ceremony, the party, and like I told you, I love dancing, I love, love dancing, so'.

6. Rosa: *Hay ponche de fruta, hay arroz con leche, chocolate [risa], de eso, sí sí, eh: hacemos mucho entonces en la Navidad hay más variedad en la comida.* 'There is fruit punch, there is arroz con leche, chocolate [laughter], yes yes, uh: we make a lot of that so at Christmas there is more variety in the food'.

7. Alejandro: *En Cuaresma, creo que, es un poco más, pescado solamente porque no puedes comer carne los viernes, entonces siempre es pescado.* 'In Lent, I believe, it's a bit more, fish only because you can't eat meat on Fridays, so it's always fish'.

8. Pedro: *Pos, mi papá también cocina afuera, hace fajitas o hace como una discada para tacos, también unas enchiladas verdes también me gustan mucho, pero eso se hace con pollo.* 'Well, my father also cooks outside, he makes fajitas or makes a meat stew/roast for tacos, also some enchiladas verdes, also I really like them, but he makes them with chicken'.

Notes

1   Here, SOCIAL MEANING is defined as "the set of interferences that can be drawn on the basis of how language is used in a specific interaction" (Hall-Lew et al. 2021, p. 3).

2   Here, LINGUISTIC INFORMATION refers to the subtle changes in the audio input. Much of the sociolinguistic literature has focused on phonetic variation at the subsegmental, segmental, and suprasegmental levels.

3   SOCIAL PERCEPTION is defined as the social characteristics that listeners attribute to speakers such as perceived educational level, friendliness, respectfulness, etc.

4   Here, we distinguish linguistic perception from LINGUISTIC EVALUATION, which is a more global or holistic evaluation (not-segmental specific) such as perceived accentedness (Rubin 1992) or speech intelligibility (Babel and Russell 2015).

5   A HERITAGE SPANISH SPEAKER/LISTENER within the U.S. context is someone who grew up speaking or hearing Spanish at home while receiving their schooling in English (Valdés 2000). While this is one term, we could also refer to heritage Spanish speakers as U.S. SPANISH SPEAKERS (see Erker and Otheguy 2021, pp. 199–200).

6   While we use the term "monolingual speaker" for the 'Mexico' guises, "Spanish-dominant speaker" could easily be employed as well. This acknowledges that in Mexico, there are L1 speakers of indigenous languages (L2 Spanish), as well as L1 Spanish speakers who are bi/multilingual.

7   The authors would like to state that a sequence of sociolinguistic waves does not indicate one wave is inherently better than the other, but rather each methodological approach depends upon one's research questions. To this point, we strongly agree with Schilling (2013, p. 343), who states, "as far as we may have sailed over the first and second waves of variation to reach the third, we would do well to remember that the three "waves" are part of the same ocean, that elements of all three 'waves' of study were present from the outset of variation study (see Eckert 2005), and that the best current studies will approach the social meaning of linguistic variation from a range of perspectives [...]".

8   It has been shown that a combination of features—or the perception of a combination of features—can also be perceived as indexing a type of persona, such as what Inoue (2006) describes as "schoolgirl speech" for Japanese women.

9   Perceptual dialectology studies (Alfaraz 2002, 2014; Alfaraz and Mason 2019; Montes-Alcalá 2011) have examined the language attitudes of different language variants using the participants' intuitions on questionnaires and/or "draw-a-map" tasks (Preston 1999).

10 This topic was explored in a production study by Drager et al. (2010) in which they found that while the condition (good information, bad information, no information of Australia) had an effect on New Zealanders' speech, this interacted with sports fandom in which sports fans in the bad condition favored the Australian variant while non-sports fans in the good condition favored the Australian variant. Here, sports fandom was a stronger predictor than gender.

11 See Erker and Otheguy (2021) for an excellent example of quantitative evidence against the deficit bilingual perspective as seen in the U.S. context.

12 Of note is that while the focus here is on the monoglossic language ideology, there is also overlap with the standard language ideology in this context.

13 This has been referred to as the "paradox of Spanish" in the U.S. context (Carter 2018).

14 The researcher's variety is Andalusian Spanish, which in theory could influence the bilingual Mexican Spanish speakers. However, given that Mexican Spanish is the dominant norm in West Texas, this is unlikely.

15 The rationale for a range instead of a fixed number is so that no audio clip was adjusted more than 5 dB (up or down) from the original recording.

16 Following previous studies (Campbell-Kibler 2007; Barnes 2015; Chappell 2016, 2018; Regan 2020, 2022a, 2022b, 2022c), an even number was selected to avoid neutral responses.

17 The first author was informed by several participants after the experiment that they were not very familiar with the term *eloquent*. Future studies should employ a synonym.

18 Of note, these cities refer to the greater metropolitan areas and surrounding towns from each city.

19 Similar to Niedzielski's (1999) use of 'Canadian'/'Detroit' voice, here we use 'Mexico'/'Texas' voice in quotations, as all the speakers are truly bilinguals from Texas, but their voices were presented to listeners with two different nationality labels.

20 In future studies these two variables should be explored further with a larger sample size of listeners to examine whether differences in communities and frequency of trips to Mexico play a role in differences in perceptions among heritage listeners.

21 Figures 2 and 3 were created using *ggplot2* (Wickham 2016).

22 This is the participant code (Excerpts 1–4 also provide participant codes).

23 This has also been found within the K-12 context in multiple school districts (see Clemons 2022).

24 Such ideological valuing of L2 Spanish over heritage Spanish is often reflected in the lack of institutional support for Spanish heritage programs (Beaudrie and Loza 2023).

25 We can only conjecture, as we did not provide any way to measure participant proficiency.

26 Of note, while this was the quantitative finding, Regan (2022a, p. 646) found that even among non-Hispanic listeners, there were individual listeners that demonstrated in the qualitative section that, due to their social contacts, they perceived Spanish phonology with Spanish place names to be just as local as English phonology, indicating the importance of lived experience in one's perception.

27 "*Este*", with vowel elongation of [e], here is a discourse marker (such as "uh" or "um" in English) to indicate to the interlocutor that the speaker is thinking of their next response and maintaining their speaking turn.

References

Achugar, Mariana, and Silvia Pessoa. 2009. Power and place: Language attitudes towards Spanish in a bilingual academic community in Southwest Texas. *Spanish in Context* 6: 199–223. [CrossRef]

Alfaraz, Gabriela G. 2002. Miami Cuban perceptions of varieties of Spanish. In *Handbook of Perceptual Dialectology*. Edited by Daniel Long and Dennis R. Preston. Amsterdam: John Benjamins, vol. 2, pp. 1–11.

Alfaraz, Gabriela G. 2014. Dialect perceptions in real time: A restudy of Miami-Cuban perceptions. *Journal of Linguistic Geography* 2: 74–86. [CrossRef]

Alfaraz, Gabriela G., and Alexander Mason. 2019. Ethnicity and perceptual dialectology: Latino awareness of U.S. regional dialects. *American Speech* 94: 352–78. [CrossRef]

Baird, Brandon, Marcos Rohena-Madrazo, and Caroline Cating. 2018. Perceptions of lexically specific phonology switches on Spanish-origin loanwords in American English. *American Speech* 93: 79–101. [CrossRef]

Babel, Molly, and Jamie Russell. 2015. Expectations and speech intelligibility. *The Journal of the Acoustic Society of America* 137: 2823–33.

Barnes, Sonia. 2015. Perceptual salience and social categorization of contact features in Asturian Spanish. *Studies in Hispanic and Lusophone Linguistics* 8: 213–41. [CrossRef]

Barnes, Sonia. 2019. The role of social cues in the perception of final vowel contrasts in Asturian Spanish. In *Recent Advances in the Study of Spanish Sociophonetic Perception*. Edited by Whitney Chappell. Amsterdam and Philadelphia: John Benjamins, pp. 15–38.

Bates, Douglas, Martin Mächler, Ben Bolker, and Steve Walker. 2015. Fitting linear mixed-effects models using lme4. *Journal of Statistical Software* 67: 1–48. [CrossRef]

bbmundo. 2017. Available online: https://www.bbmundo.com/comunidad/noticias/los-nombres-mas-populares-en-mexico/ (accessed on 23 January 2019).

Beaudrie, Sara M., and Sergio Loza. 2023. Insights into SHL program direction: Student and program advocacy challenges in the face of ideological inequity. *Language Awareness* 32: 39–57.

Boersma, Paul, and David Weenink. 2019. Praat: A System for Doing Phonetics by Computer. (Version 6.0.04) [Computer Program]. Available online: http://www.praat.org (accessed on 6 February 2019).

Callesano, Salvatore, and Phillip M. Carter. 2019. Latinx perceptions of Spanish in Miami: Dialect variation, personality attributes and language use. *Language & Communication* 67: 84–98.

Campbell-Kibler, Kathryn. 2007. Accent, (ING), and the social logic of listener perceptions. *American Speech* 82: 32–64. [CrossRef]

Campbell-Kibler, Kathryn. 2008. I'll be the judge of that: Diversity in social perceptions of (ING). *Language in Society* 37: 637–59. [CrossRef]

Campbell-Kibler, Kathryn. 2010. Sociolinguistics and Perception. *Language and Linguistics Compass* 4: 377–89. [CrossRef]

Campbell-Kibler, Kathryn. 2011. The sociolinguistic variant as a carrier of social meaning. *Language Variation and Change* 22: 423–41. [CrossRef]

Carruba-Rogel, Zuleyma. 2018. The complexities in seguir avanzando: Incongruences between the linguistic ideologies of students and their familias. In *Feeling It: Language, Race, and Affect in Latinx Youth Learning*. Edited by Mary Bucholtz, Dolores Inés Casillas and Jin Sook Lee. New York: Routledge, pp. 149–65.

Carter, Phillip M., and Salvatore Callesano. 2018. The social meaning of Spanish in Miami: Dialect perceptions and implications for socioeconomic class, income, and employment. *Latino Studies* 16: 65–90. [CrossRef]

Carter, Phillip. 2018. Spanish in U.S. language policy and politics. In *The Routledge Handbook of Spanish as a Heritage Language*. Edited by Kim Potowski. New York: Routledge, pp. 36–52.

Chappell, Whitney, and Matthew Kanwit. 2022. Do learners connect sociophonetic variation with regional and social characteristics?: The case of L2 perception of Spanish aspiration. *Studies in Second Language Acquisition* 44: 185–209. [CrossRef]

Chappell, Whitney, and Sonia Barnes. 2023. Stereotypes, language, and race: Spaniards' perception of Latin American immigrants. *Journal of Linguistic Geography* 11: 104–18. [CrossRef]

Chappell, Whitney. 2016. On the social perception of intervocalic /s/ voicing in Costa Rican Spanish. *Language Variation and Change* 28: 357–78. [CrossRef]

Chappell, Whitney. 2018. Caribeño or mexicano, profesionista o albañil? Mexican listeners' evaluations of /s/ aspiration and maintenance in Mexican and Puerto Rican voices. *Sociolinguistic Studies* 12: 367–93. [CrossRef]

Chappell, Whitney. 2019. The sociophonetic perception of heritage Spanish speakers in the United States: Reactions to labiodentalized <v> in the speech of native and heritage voices. In *Recent Advances in the Study of Spanish Sociophonetic Perception*. Edited by Whitney Chappell. Amsterdam: John Benjamins, pp. 239–64.

Chappell, Whitney. 2020. Perceptions of Spanish(es) in the United States: Mexican's sociophonetic evaluations of [v]. In *Spanish in the United States: Attitudes and Variation*. Edited by Scott Alvord and Gregory Thompson. New York: Routledge, pp. 31–55.

Chappell, Whitney. 2021a. Heritage Mexican Spanish Speaker's Sociophonetic Perception of /s/ Aspiration. *Spanish as a Heritage Language* 1: 167–97. [CrossRef]

Chappell, Whitney. 2021b. Mexican and Mexican-Americans' perceptions of themselves and each other: Attitudes towards language and community. In *Topics in Spanish Linguistic Perceptions*. Edited by Luis A. Ortiz-López and Eva-María Suárez Büdenbender. New York: Routledge, pp. 138–60.

Clemons, Aris. 2022. Benefits vs. Burden: A raciolinguistic analysis of world language mission statements and testimonios of bilingualism in the United States. *Journal of Postcolonial Linguistics* 6: 69–91.

Drager, Katie, Jennifer Hay, and Abby Walker. 2010. Pronounced rivalries: Attitudes and speech production. *Reo, Te* 53: 27–53.

Eckert, Penelope, and Sally McConnell-Ginet. 1999. New generalizations and explanations in language gender research. *Language in Society* 28: 185–201. [CrossRef]

Eckert, Penelope. 2005. Three waves of variation study: The emergence of meaning in the study of sociolinguistic variation. *Unpublished manuscript*.

Eckert, Penelope. 2008. Variation and the indexical field. *Journal of Sociolinguistics* 12: 453–76. [CrossRef]

Erker, Daniel, and Ricardo Otheguy. 2021. American myths of linguistic assimilation: A sociolinguistic rebuttal. *Language in Society* 50: 197–233. [CrossRef]

Flores, Nelson, and Jonathan Rosa. 2015. Undoing appropriateness: Raciolinguistic ideologies and language diversity in education. *Harvard Educational Review* 85: 149–71. [CrossRef]

Fuller, Janet M., and Jennifer Leeman. 2020. *Speaking Spanish in the U.S.: The Socio-Politics of Language*, 2nd ed. Bristol: Multilingual Matters.

Galindo, D. Leticia. 1995. Language attitudes toward Spanish and English varieties: A Chicano perspective. *Hispanic Journal of Behavioral Sciences* 17: 77–99. [CrossRef]

Goble, Ryan. 2016. Linguistic insecurity and lack of entitlement to Spanish among third-generation Mexican Americans in narrative accounts. *Heritage Language Journal* 13: 29–54. [CrossRef]

Gordon, Elizabeth. 1997. Sex, speech, and stereotypes: Why women use prestige speech forms more than men. *Language in Society* 26: 47–63. [CrossRef]

Guerrero-Rodríguez, Paola. 2021. Telecolaboración: Discutiendo sus efectos e implicaciones pedagógicas para hablantes de herencia de español. Doctoral dissertation, University of Florida, Gainesville, FL, USA.

Gutiérrez, Maria Elena, and Mark Amengual. 2016. Perceptions of standard and nonstandard language varieties: The influence of ethnicity and heritage language experience. *Heritage Language Journal* 13: 55–79. [CrossRef]

Hall-Lew, Lauren, Emma Moore, and Robert Podesva. 2021. Social meaning and linguistic variation: Theoretical foundations. In *Social Meaning and Linguistic Variation: Theorizing the Third Wave*. Edited by Lauren Hall-Lew, Emma Moore and Robert J. Podesva. Cambridge: Cambridge University Press, pp. 1–20.

Hay, Jennifer, and Katie Drager. 2010. Stuffed toys and speech perception. *Linguistics* 48: 865–92. [CrossRef]

Hay, Jennifer, Aaron Nolan, and Katie Drager. 2006a. From fush to feesh: Exemplar priming in speech perception. *The Linguistic Review* 23: 351–79. [CrossRef]

Hay, Jennifer, Paul Warren, and Katie Drager. 2006b. Factors influencing speech perception in the context of a merger-in-progress. *Journal of Phonetics* 34: 458–84. [CrossRef]

Inoue, Miyako. 2006. *Vicarious Language: Gender and Linguistic Modernity in Japan*. Berkeley: University of California Press.

Irvine, Judith T., and Susan Gal. 2000. Language ideology and linguistic differentiation. In *Regimes of Languages: Ideologies, Polities, and Identities*. Edited by Paul Kroskrity. Santa Fe: School of American Research Press, pp. 35–83.

Johnstone, Barbara, and Scott F. Kiesling. 2008. Indexicality and experience: Exploring the meanings of /aw/-monophthongization in Pittsburgh. *Journal of Sociolinguistics* 12: 5–33. [CrossRef]

Johnstone, Barbara. 2011. Dialect enregisterment in performance. *Journal of Sociolinguistics* 15: 657–79. [CrossRef]

Kirtley, M. Joelle. 2011. Speech in the U.S. Military: A Sociophonetic Perception Approach to Identity and Meaning. Master's thesis, University of Hawai'i at Manoa, Honolulu, HI, USA.

Koops, Christian, Elizabeth Gentry, and Andrew Pantos. 2008. The effect of perceived speaker age on the perception of PIN and PEN vowels in Houston, Texas. *University of Pennsylvania Working Papers in Linguistics* 14: 93–101.

Kroskrity, Paul V. 2004. Language ideologies. In *A companion to Linguistic Anthropology*. Edited by Alessandro Duranti. Malden: Blackwell, pp. 496–517.

Kutlu, Ethan. 2020. Now you see me, now you mishear me: Raciolinguistic accounts of speech perception in different English varieties. *Journal of Multilingual and Multicultural Development* 44: 511–25. [CrossRef]

Kuznetsova, Alexandra, Per B. Brockhoff, and Rune H. B. Christensen. 2017. lmerTest Package: Tests in Linear Mixed Effects Models. *Journal of Statistical Software* 82: 1–26. [CrossRef]

Lambert, Wallace E., Richard C. Hodgson, Robert C. Gardner, and Samuel Fillenbaum. 1960. Evaluational reactions to spoken language. *Journal of Abnormal and Social Psychology* 60: 44–51. [CrossRef]

Leeman, Jennifer. 2004. Racializing language: A history of linguistic ideologies in the US Census. *Journal of Language and Politics* 3: 507–34. [CrossRef]

Leeman, Jennifer. 2012. Investigating Language Ideologies in Spanish as a Heritage Language. In *Spanish as a Heritage Language in the United States*. Edited by Sara M. Beaudrie and Marta Fairclough. Washington, DC: Georgetown University Press, pp. 43–59.

Leeman, Jennifer. 2018. It's all about English: The interplay of monolingual ideologies, language policies and the U.S. Census Bureau's statistics on multilingualism. *International Journal of the Sociology of Language* 252: 21–43. [CrossRef]

Leeman, Jennifer, and Ellen J. Serafini. 2016. Sociolinguistics for heritage language educators and students. In *Innovative Strategies for Heritage Language Teaching: A Practical Guide for the Classroom*. Edited by Marta Fairclough and Sara M. Beaudrie. Washington, DC: Georgetown University Press, pp. 56–79.

Lenth, Russell, Henrik Singmann, Jonathan Love, Paul Buerkner, and Maxime Herve. 2018. *Emmeans: Estimated marginal means, aka least-squares means*. R package revions 1.7.5. Available online: https://cran.r-project.org/web/packages/emmeans/emmeans.pdf (accessed on 5 September 2023).

Lippi-Green, Rosina. 2012. *English with an Accent: Language, Ideology, and Discrimination in the United States*, 2nd ed. New York: Routledge.

Lowther Pereira, Kelly Anne. 2010. Identity and Language Ideology in the Intermediate Spanish Heritage Language Classroom. Ph.D. dissertation, University of Arizona, Tucson, AZ, USA.

Loza, Sergio. 2019. Exploring language ideologies in action: An analysis of Spanish heritage language oral corrective feedback in the mixed classroom setting. Doctoral dissertation, Arizona State University, Tempe, AZ, USA.

Martínez, Glenn A., and Peter R. Petrucci. 2004. Institutional dimensions of cultural bias on the Texas-Mexico border: Linguistic insecurity among heritage language learners. *Critical Inquiry in Language Studies* 1: 89–104. [CrossRef]

Mejías, Hugo A., and Pamela L. Anderson. 1988. Attitude toward use of Spanish on the South Texas border. *Hispania* 71: 401–7. [CrossRef]

Milroy, James, and Lesley Milroy. 1999. *Authority in Language: Investigating Standard English*, 3rd ed. London: Routledge.

Milroy, James. 2001. Language ideologies and the consequences of standardization. *Journal of Sociolinguistics* 5: 530–55. [CrossRef]

Milroy, Lesley. 2004. Language ideologies and linguistic change. In *Sociolinguistic Variation: Critical Reflections*. Edited by Carmen Fought. Oxford: Oxford University Press, pp. 161–77.

Montes-Alcalá, Cecilia. 2011. ¿Mejor o peor español? Actitudes lingüísticas de universitarios hispanohablantes en Estados Unidos. *Studies in Hispanic and Lusophone Linguistics* 4: 35–54. [CrossRef]

Nakagawa, Shinichi, and Holger Schielzeth. 2013. A general and simple method for obtaining R^2 from generalized linear mixed-effects models. *Methods in Ecology and Evolution British Ecological Society* 4: 133–42. [CrossRef]

Nance, Claire. 2013. Phonetic Variation, Sound Change, and Identity in Scottish Gaelic. Ph.D. dissertation, University of Glasgow, Glasgow, UK.

Niedzielski, Nancy, and Dennis R. Preston. 1999. *Folk Linguistics*. Berlin and New York: Mouton de Gruyter.

Niedzielski, Nancy. 1999. The effect of social information on the perception of sociolinguistic variables. *Journal of Language and Social Psychology* 18: 62–85. [CrossRef]

Pascual y Cabo, Diego, and Jason Rothman. 2012. The (il)logical problem of heritage speaker bilingualism and incomplete acquisition. *Applied Linguistics* 33: 450–55. [CrossRef]

Pavlenko, Aneta. 2002. 'We have room but for one language here': Language and national identity at the turn of the 20th century. *Multilingua* 21: 163–96. [CrossRef]

Preston, Dennis R. 1999. A language attitude approach to the perception of regional variety. In *Handbook of Perceptual Dialectology*. Edited by Dennis R. Preston. Amsterdam: John Benjamins, vol. 1, pp. 359–73.

Preston, Dennis R., ed. 1993. *American Dialect Research*. Philadelphia: John Benjamins.

Qualtrics. 2005–2023. Qualtrics Online Survey [Computer Program]. Available online: www.qualtrics.com (accessed on 5 December 2022).

R Core Team. 2023. *R: A Language and Environment for Statistical Computing*; Vienna: R Foundation for Statistical Computing. Available online: www.r-project.org (accessed on 20 September 2023).

Rangel, Natalie, Verónica Loureiro-Rodríguez, and María Irene Moyna. 2015. 'Is that what I sound like when I speak?' Attitudes towards Spanish, English, and code-switching in two Texas border towns. *Spanish in Context* 12: 177–91. [CrossRef]

Regan, Brendan. 2020. 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: 82–101. [CrossRef]

Regan, Brendan. 2022a. Guadaloop or Guadalupe?: Place-name variation and place identity in Austin, Texas. *American Speech* 97: 441–82. [CrossRef]

Regan, Brendan. 2022b. La percepción social del seseo sevillano y sus implicaciones para la norma sevillana. *Lingüística en la Red* XIX: 1–27.

Regan, Brendan. 2022c. The social meaning of a merger: The evaluation of an Andalusian Spanish consonant merger (*ceceo*). *Language in Society* 51: 481–510. [CrossRef]

Riegelhaupt, Florencia, and Roberto Luis Carrasco. 2000. Mexico host family reactions to a bilingual Chicana teacher in Mexico: A case study of language and culture clash. *Bilingual Research Journal* 24: 405–21. [CrossRef]

Rosa, Jonathan, and Nelson Flores. 2017. Unsettling race and language: Toward a raciolinguistic perspective. *Language in Society* 46: 621–47. [CrossRef]

Rubin, Donald L. 1992. Nonlanguage factors affecting undergraduates' judgements of nonnative English-speaking teaching assistants. *Research in Higher Education* 33: 511–31. [CrossRef]

Ryan, Ellen Bouchard, and Miguel A. Carranza. 1977. Ingroup and outgroup reactions toward Mexican American language varieties. In *Language, Ethnicity, and Intergroup Relations*. Edited by Howard Giles. London: Academic Press, pp. 59–82.

Schieffelin, Bambi B., Kathryn A. Woolard, and Paul Kroskrity, eds. 1998. *Language Ideologies: Practice and Theory*. New York: Oxford University Press.

Schilling, Natalie. 2013. Investigating Stylistic Variation. In *The Handbook of Language Variation and Change*, 2nd ed. Edited by Jack K. Chambers and Natalie Schilling. Oxford: Wiley, pp. 325–49.

Silva-Corvalán, Carmen. 1994. *Language Contact and Change: Spanish in Los Angeles*. Oxford: Clarendon.

Silverstein, Michael. 1996. Monoglot 'Standard' in America: Standardization and Metaphors of Linguistic Hegemony. In *The Matrix of Language: Contemporary Linguistic Anthropology*. Edited by Donald Brenneis and Ronald K.S. Macaulay. Boulder: Westview, pp. 284–306.

Silverstein, Michael. 2003. Indexical order and the dialectics of sociolinguistic life. *Language & Communication* 23: 193–229.

Staum Casasanto, Laura. 2010. What do listeners know about sociolinguistic variation? *University of Pennsylvania Working Papers in Linguistics* 15: 40–49.

Tseng, Amelia. 2021. 'Qué barbaridad, son latinos y deberían saber español primero': Language ideology, agency, and heritage language insecurity across immigrant generations. *Applied Linguistics* 42: 113–35. [CrossRef]

Urciuoli, Bonnie. 1996. *Exposing Prejudice: Puerto Rican Experiences of Language, Race, and Class*. Boulder: Westview.

Valdés, Guadalupe, Sonia V. González, Dania López García, and Patricio Márquez. 2003. Language Ideology: The Case of Spanish in Departments of Foreign Languages. *Anthropology & Education Quarterly* 43: 3–26.

Valdés, Guadalupe. 2000. Introduction. In *Spanish for native speakers. AATSP Professional Development Series Handbook for Teachers K–16*. New York: Harcourt College, vol. 1, pp. 1–20.

W Radio. 2017. Available online: https://wradio.com.mx/programa/2017/08/16/en_buena_onda/1502845447_093649.html (accessed on 23 January 2019).

Walker, Abby, Christina García, Yomi Cortés, and Kathryn Campbell-Kibler. 2014. Comparing social meanings across listener and speaker groups: The indexical field of Spanish /s/. *Language Variation and Change* 26: 169–98. [CrossRef]

Walker, Abby, Jennifer Hay, Katie Drager, and Kauyumari Sanchez. 2018. Divergence in speech perception. *Linguistics* 56: 257–78. [CrossRef]

Weatherholtz, Kodi, Kathryn Campbell-Kibler, and T. Florian Jaeger. 2014. Socially-mediated syntactic alignment. *Language Variation and Change* 26: 387–420. [CrossRef]

Wickham, Hadley. 2016. ggplot2: Elegant Graphics for Data Analysis. New York: Springer. Available online: https://ggplot2.tidyverse.org (accessed on 5 September 2023).

Woolard, Kathryn A. 1998. Introduction: Language ideology as a field of inquiry. In *Language Ideologies: Practice and Theory*. Edited by Bambi B. Schieffelin, Kathryn A. Woolard and Paul Kroskrity. New York: Oxford University Press, pp. 3–47.

Woolard, Kathryn A. 2008. Why *dat* now?: Linguistic-anthropological contributions to the explanation of sociolinguistic icons and change. *Journal of Sociolinguistics* 12: 432–52. [CrossRef]

Woolard, Kathryn A., and Bambi B. Schieffelin. 1994. Language ideology. *Annual Review of Anthropology* 23: 55–82. [CrossRef]

Wright, Robyn. 2021a. Madrileños on the ejque: Perceptions of velarlized /s/. *Studies in Hispanic and Lusophone Linguistics* 14: 207–33. [CrossRef]

Wright, Robyn. 2021b. Regional perceptions of the ejque: Exploring language attitudes toward velarized /s/ outside of Madrid. *Borealis: An International Journal of Hispanic Linguistics* 10: 103–29. [CrossRef]

Zentella, Ana Celia. 2002. Latina/o Languages and Identities. In *Latinos: Remaking America*. Edited by Marcelo M. Suárez-Orozco and Mariela Páez. Berkeley: University of California Press.

Zentella, Ana Celia. 2007. "Dime con quién hablas, y te diré quién eres": Linguistics (In)security and Latina/o Unity. In *A Companion to Latina/o Studies*. Edited by Juan Flores and Renato Rosaldo. Malden and Oxford: Blackwell, pp. 25–38.