Receding or resurgent? On the use of *cantase* (and *cantara*) in Galician Spanish

A Dissertation

SUBMITTED TO THE FACULTY OF THE

UNIVERSITY OF MINNESOTA

BY

Ana Maria Anderson

IN PARTIAL FULFILLMENT OF THE REQUIREMENTS

FOR THE DEGREE OF

DOCTOR OF PHILOSOPHY

Advisors: Dr. Carol A. Klee and Dr. Francisco A. Ocampo

June, 2017

Acknowledgements

Dedication

For James, who for too long has shared me with graduate school,
and little Marit Jade, who now will never have to.

Regardless of warnings, the future doesn't scare me at all.
Nothing's like before. –Utada Hikaru

Table of Contents

Table of Figures

List of Tables

Chapter 1: Introduction

The field of language contact, simply put, examines the influence of one linguistic system on another. According to Weinreich (1953), "it is immaterial whether the two systems are 'languages,' 'dialects of the same language,' or 'varieties of the same dialect'" (p. 1); all of these have the potential to influence one another. Thus, on a fundamental level, "all linguistic interaction is 'language contact', albeit between extremely similar grammars" (Bowern, 2010, p. 341).

However, and particularly when considered this broadly, the outcomes and consequences of language contact can vary greatly. For example, within my family, the euphemism "shmeckpepper"—used to express frustration when, while playing a game, an event occurs just after the moment in which it was needed—was first used by my mother, and then through interaction with her in these game-playing contexts, I introduced the word into my own repertoire. From there, it also passed into general use at both my father's house and with my husband's family. Despite the fact that the linguistic systems involved in these familiar interactions are undoubtedly instantiations of both the same language and the same dialect, an idiosyncrasy in the speech of one individual has been taken up by others, thus creating a change in their individual linguistic systems. The consequences of this particular change, however, are likely to be limited, due not only to the small number of individuals who use it (*e.g.* Thomason & Kaufman, 1988), but also to its topical restriction to the context of game-playing. On the other extreme is, for example, the case of Media Lengua, in which intense contact in most areas of life between speakers of two very different languages—Spanish and Quechua—and large numbers of highly fluent bilinguals resulted in the creation of a new linguistic system

with Quechua grammar and phonology but almost exclusively Spanish lexicon (Muysken, 1981).

One key factor in the diffusion of linguistic practices is mobility, that is, movement through social spaces on both the everyday level (*e.g.* leaving home to interact at work or at the grocery store) and on a larger level (*e.g.* moving to a new region or country) (Britain, 2010; Trudgill, 1986). This is so because, while linguistic innovations occur on an individual basis, propagation of these innovations is a necessarily social phenomenon (Croft, 2000). In the words of Linell (2009), the situation-transcending (*i.e.* societally shared) properties of language "are linked to habituality, routinization, conventionalization and institutionalization of human practices, that is, our tendencies to do things approximately in the same ways as we have done before, or seen others do, in similar situations" (p. 50). Thus, changes in societal routines, habits, conventions, and mobility practices, which all bear on the practice of language, can result in changes in linguistic systems. Such social changes have occurred in much of the Western world in the last century, related to:

> increasing urbanization and counter-urbanization; increased migration and immigration; …a shift from primary and secondary to tertiary sector employment as the backbone of the economy; …an expansion in higher levels of education (in places often well away from the local speech community);…and increasing geographical elasticity of family ties (Britain, 2010, p. 213).

And, as Britain (2010) adds, "this mobility has had sociolinguistic consequences" (p. 213).

This is particularly the case when changes like those just mentioned bring

different languages into contact, as within the Autonomous Community of Galicia, in northwest Spain. In the pre-modern era, Galician was the main language of the vast majority of those living in the region, and both the opportunity and the necessity of learning the Castilian language were notably absent outside the largest cities. In the movement toward modernity, however, mobility and social networks began to change, and along with a massive migration to urban centers came an increased presence of Castilian in daily life (Monteagudo, 2012). The Castilian language, then, became associated with the modern and urban lifestyle, while use of the Galician language increasingly became a marked choice indexing rurality, lack of education, and outdated tradition.

Despite the gains in status for the Galician language achieved under the new Spanish Constitution in 1978 and the Galician Statute of Autonomy in 1981, the patterns of language use in the region have been undeniably marked by contact between the two languages. Contact effects have been noted, for example, in the incorporation of Castilian lexicon and grammatical structures into Galician (*e.g.* Negro Romero, 2013; Silva Valdivia, 2013) and, inversely, in the transfer of aspects of the Galician sound system and intonation patterns onto Castilian (Ramallo, 2007; Faginas Souto, 2001; Perez Castillejo 2012, 2014). The exact scope of the consequences (for either language) are difficult to determine, however, as both languages emerged from closely related varieties of Late Latin. In a case such as this, where the languages involved are phylogenetically related, the boundary between what is considered native and what is not can be blurred (Epps, Huehnergard, & Pat-El, 2013). Even the classification of two genetically related linguistic systems as differing "languages" or simply as "dialects" of the same language

can be complex, as can be seen in Croft's (2000) definition of sibling languages as "two linguistic varieties that are structurally so similar that they are considered to be 'dialects of the same language', yet are perceived by the speakers—or at least by one group of speakers—as distinct languages" (p. 16). Indeed, the classification of Galician and Catalan as independent languages has historically been contested, as Spain's minority languages were at one time considered "underdeveloped, bastardized dialect[s] of Castilian, incapable of serving people's communicative needs in daily life" (Vann, 2002, p. 231).[1]

While it is not the purpose of the present thesis to enter into this debate, which was in large part resolved in the 1978 iteration of the Spanish Constitution, noting its existence serves to emphasize the great degree of similarity between the two languages. Such similarity can be problematic for proposing contact explanations for linguistic phenomena in genetically related languages. Proving that a feature exists in a language due to contact rather than to common inheritance from the parent language is difficult; contact arguments are more convincing when no typological similarity is involved (Pat-El, 2013). Linguistic drift—the generally accepted internal cause of language change— can produce essentially the same results as would be predicted from contact between closely related languages; for this reason, it is likely that some contact effects between languages such as Galician and Castilian will go unnoticed (Pat-El, 2013; Thomason, 2010). In addition, while contact between non-related languages is often noticeable due to

---

[1] More recently, a similar debate has arisen about the status of Galician with respect to European Portuguese, with some academics of the opinion that the former is a dialectal variation that ought to be reintegrated into the latter (Rei-Doval, 2013). This academic opinion appears to find some support in social reality; Beswick (2014) reports that, along the Minho River, speakers of both languages consider themselves members of a cross-border speech community and have higher in-group affect for their counterparts on the other river bank than for their own countrymen living elsewhere.

the introduction of new features in one of the languages, another possible effect of contact—more likely to occur in related languages since there is naturally more overlap in features—is a change in the rate at which a particular inherited feature is used (Silva-Corvalán, 1994; Pat-El, 2013).

The focus of this thesis is to examine one such potential effect of language contact between Galician and Castilian, a set of grammatical features in which the Galician and Castilian languages overlap in form while differing in function. Specifically, this work focuses on verbal morphemes *–ra* and *–se,* from the verb forms *cantara* and *cantase,* respectively. In Galician, the form *cantara* is normatively used to mark the simple pluperfect indicative verb form, which corresponds to the form *había cantado* in Castilian. *Cantase*, on the other hand, marks the imperfect subjunctive. In Castilian, however, both *cantara* and *cantase* are employed with the function of imperfect subjunctive. Given this asymmetrical form-function relationship, either the presence of perfective uses of *cantara* or a heightened rate of *cantase* use in Galician Spanish may indicate transfer effects from Galician (following Silva-Corvalán, 1994). The first of these two conditions has been attested in Rabanal (1967) and Pollán (2001), among others, while the second, which is the main focus of this thesis, is attested for some speakers in both Kempas (2011) and in Rojo and Vázquez Rozas (2014).

What both Kempas (2011) and Rojo and Vázquez Rozas (2014) lack, however, is a consideration of the ways in which patterns of habitual language use, whether Galician or Castilian, may be related to choice of *cantase* or *cantara* in Galician Spanish. For this reason, neither of these studies is sufficient for making a strong case either for or against contact with Galician as a motivating factor for the increased use of *cantase* in Galician

Spanish when compared with non-contact varieties of Castilian. Nor does either of these studies explore the potential effect of proximity to an urban center, which may be related to patterns of mobility and therefore language exposure, as described previously. Finally, while Kempas (2011) explores the written modality and Rojo and Vázquez Rozas (2014) examine oral interviews, neither study takes into account potential differences in production and perception of the forms in question.

To address these gaps in current research, the present work examines the production and perception of 39 speakers of Galician Spanish in both written and oral modalities. Oral production is obtained through the use of a semi-guided group conversation between friends and family, and oral perception is assessed through the use of a pseudo-matched-guise task (following Silva-Corvalán, 1984). Written perception is approached through the use of an acceptability judgment task, while written production is examined through both the written corrections on the acceptability judgment task and through the completion of a fill-in-the-blank exercise. In addition to these tasks, each participant completed a background questionnaire to obtain both basic sociolinguistic information (*i.e.* age, education level, gender, and place of residence) and detailed information about use of Galician and Castilian both currently and throughout their lives (questionnaire adapted from Birdsong, Gertken, & Amengual, 2012). This data will be used to address the following research questions:

1. How is use of *cantara* and *cantase* perceived by speakers of Galician Spanish?

2. Do written and oral production of *cantara* and *cantase* correlate with perception/acceptance of these forms?

3. What social factors, if any, correlate with the use or perception of *cantara* or *cantase* in Galician Spanish?

4. What linguistic factors, if any, condition the choice of *cantara* or *cantase* in Galician Spanish?

5. What relationship, if any, exists between exposure to the Galician language and use of *cantara* and *cantase* in Galician Spanish?

The implications of this study are threefold. First, as previously mentioned, contact phenomena in the Spanish of Galicia are relatively understudied, both in comparison to studies on Galician and studies on other varieties of Spanish in contact; the present thesis begins to address this gap. Second, most theories of language contact focus on the introduction of new linguistic material into a contact language, but relatively few predictions are made about the conditions under which existing patterns will be reinforced or undermined or about which groups are likely to be the agents of these processes. By examining forms that are parallel between Galician and Spanish, but in differing distributions, I not only begin addressing this question for the Galician community, but also contribute to language contact theory more broadly. Finally, by taking into account the complex competencies that different speakers have in each language at various points in their lives, I begin to untangle the issues of language dominance and of different directions of shift that complicate the application of current theoretical frameworks to many modern contexts.

The following chapter contains an overview of the history of language contact in Galicia, the development of *cantara* and *cantase* from Latin to the present, studies of related phenomena throughout the Spanish-speaking world, a discussion of linguistic and

social variables that have previously been found to condition form choice, and the

theoretical principles that inform this thesis. Chapter 3 presents in greater detail the

methodology used to obtain and analyze the data under consideration. In Chapter 4 the

qualitative and quantitative results are presented, and the implications of these results are

discussed in Chapter 5. Chapter 6 summarizes the conclusions and contributions of this

thesis. Copies of the tasks employed and additional charts not included in Chapter 4 are

included in the Appendix.

Chapter 2: Literature Review

This chapter provides background information to situate the present study in its historical, social, linguistic, and theoretical contexts. For this reason, the following section will present a brief overview of the history of language contact in Galicia. Subsequent sections will delve into the sociolinguistic situation of the two languages since the passage of the Galician Statute of Autonomy in 1981, the historical evolution of *cantase* and *cantara*, current trends in the variable use of these two forms in Latin America, in Spain, and in Galicia more specifically, the factors associated with this variation, and the theoretical considerations which bear upon the study of these issues in a language contact situation.

I.   History of language contact in Galicia

When the Iberian Peninsula came under complete Roman control in 218BCE, Latin spread to all regions, where it coexisted with other pre-Romance languages such as Basque, Celtiberian, and Lusitanian, among others (Beltrán Lloris, 2004). Gradually, spoken vernacular Latin came to dominate and replace all other languages save Basque, and to develop a unique set of phonological, lexical, and syntactic features that distinguished it from other points on the Late Latin dialectal continuum ranging from what is now Romania, in the east, to the Iberian Peninsula in the west (Vincent, 1982). These developments did not happen uniformly within the Peninsula, however. Lloyd (1987) describes a language continuum stretching from what is now Catalonia on one extreme to modern Galicia and Portugal on the other. Castile, in the center, tended to favor innovative developments to the language, while the variety spoken in Galicia was

much more conservative and remained truer in form to its Latin origins. This trend continued throughout the reign of the Visigoths from the 5th century CE until the Muslim invasion in 711 (Beltrán Lloris, 2004).

After 711, the Muslims and Berbers quickly took control of all but the northernmost regions of the peninsula. Those few regions that were not under direct Muslim dominance, due in large part to the mountainous terrain and the lack of economic and agrarian benefit to be gained by their conquest, established themselves into three distinct kingdoms. The westernmost of the three, situated in the Cantabrian region with its capital in Oviedo, began to re-extend its territory and eventually came to control what is now Galicia as well (García de Cortázar y Ruiz de Aguirre, 2004). This territory would eventually divide itself into two distinct political regions comprised of three separate social groups. In the west, the Kingdom of León was formed of the Galician-Portuguese social group and the Asturian-Leonese group, while the eastern division, made up of Cantabria and Castile, became the *Condado de Castilla*.

Shortly after 1085CE, Alfonso VI, king of León, married his daughters to Raimundo and Enrique de Borgoña. As dowry for his daughters, he gave Galicia to Raimundo and what is now northern Portugal to Enrique (Beswick, 2002; García de Cortázar y Ruiz de Aguirre, 2004). Upon his death in 1128, Enrique's son Alfonso Henriques was named King of Portugal, and in 1143 Portugal officially declared its independence. Galicia, on the other hand, remained politically tied to the Kingdom of León. When Alfonso IX, king of León, died in 1230, Fernando III of Castile reunited the two kingdoms and irreversibly tied Galicia to the Castilian State (Beswick, 2002; Ramallo, 2007).

Throughout this time of expansion and reconquest, the linguistic varieties of each individual region had continued to develop to the point of becoming distinct languages in their own right. In Galicia and Portugal the Galician language, spoken at that time in both regions, enjoyed great prestige in all areas of life, including in the government and ruling class (García, 1986; Vladu, 2011). Indeed this prestige was not limited solely to within these territories, as according to Beswick (2002), the "twelfth to the fifteenth centuries were the 'golden age' of Galician literature, and Galician became the language *par excellence* for the lyrical troubadour poetry throughout the majority of the Iberian Peninsula" (p. 258).

However, after Pedro de Borgoña, whose reign Galician nobles had supported, was defeated by Enrique II in 1369, Galician nobles were dispossessed of their land and Castilian nobility were granted power in the region (García, 1986). These nobles brought with them their language, which gradually replaced Galician as the language of power in the region as the dispossessed Galician nobility shifted to Castilian in an attempt to recover the prestige they had lost (Beswick, 2002; Ramallo, 2007; Vladu, 2011). The dominance of the Castilian language was fully realized with the marriage of Fernando II of Aragón to Isabel of Castile in 1469 and with their subsequent unification strategy which, among other things, involved the declaration of Castilian as the official language of the new Spanish state (Beswick, 2002).

Throughout the 16th to 19th centuries, a period referred to in Galician as *Os Séculos Escuros*, Galician continued to dwindle in prestige, lost its status as the Hispanic lyrical language, and even began to differ from the language variety spoken in Portugal (Ramallo, 2007). The result was that, by the end of the 18th century, language use in

Galicia was diglossic, with Castilian as the high variety and Galician relegated to home, village, and unofficial uses (Vladu, 2011). The attitude toward Galician over the following century can be seen reflected in the statement by the Scholarly Assembly and Exposition of Pontevedra in 1893 that "*el dialecto es causa grave que se opone grandemente al aprendizaje del armonioso, rico, inimitable y melodioso idioma de nuestra España*" (González González, 1985, p. 102).

Interest in regional linguistic and cultural identity briefly surged in the mid-19th century with an armed uprising by Galician nationalists and with the *Rexurdimento* (Galician 'Renaissance'), during which time local intellectuals attempted to reestablish Galician as a literary language and as the official regional language (González González, 1985). This period saw the emergence of Galician language societies known as *Irmandades da Fala* and of the *Xogos Florais*, a Galician poetry competition begun by Rosalía de Castro (Carbolová, 2009; Even-Zohar, 2008; González González, 1985). Their efforts, though stymied by the public's lack of written literacy and by the absence of an orthography system differing from that of Castilian, culminated in 1905 in the creation of the *Real Academia Galega*. In 1931, the *Partido Galeguista*, a political party whose platform was "regional autonomy and linguistic unification", came to power, and in 1936 they ratified the region's first Statute of Autonomy (Beswick, 2002, p. 259).

The provisions of the Statute, however, were never to be realized. When Francisco Franco came to power in 1939, one characteristic of his dictatorship was linguistic persecution. Though the use of Galician was never expressly prohibited (Carbolová, 2009; Ramallo, 2007), punishments for using a linguistic system other than Castilian were common and occasionally severe (Beswick, 2002; Dominguez-Seco,

2002; González González, 1985; Vladu, 2011).

The eight years following Franco's death in 1975 were important ones for language policy and planning initiatives in Spain. After a rapid transition to a democracy, the State passed its new Constitution in 1978, of which Article 3 was key in establishing the rights of linguistic minorities (Carbolová, 2009). It reads:

1. *El castellano es la lengua española oficial del Estado. Todos los españoles tienen el deber de conocerla y el derecho a usarla.*

2. *Las demás lenguas españolas serán también oficiales en las respectivas Comunidades Autónomas de acuerdo con sus Estatutos.*

3. *La riqueza de las distintas modalidades lingüísticas de España es un patrimonio cultural que será objeto de especial respeto y protección.*

Thus Galician, for example, was able to become co-official with Castilian in Galicia, provided the Galician Statute of Autonomy declared it to be so. Accordingly, in 1981, Galicia passed its Statute of Autonomy, in which Article 5 declared Galician as Galicia's "own language" and therefore co-official with Castilian in the Autonomous Community.

II.   Trends in the use of Galician and Castilian since 1981

By the time Galician gained the status of an official language in the region, the linguistic hegemony of Castilian was already well established. This was exacerbated by a mass migration to urban centers in the 1970s in search of employment, since Castilian was already a language firmly associated with social and economic advancement (Bouzada Fernández, 2003). Additionally, Castilian is now the "unmarked" linguistic choice in Galicia outside of the smaller villages. Woolard (2008) describes *anonymity* as

the quality of being "from nowhere", with use of the linguistic system accepted as normal speech—in these terms, Castilian, rather than Galician, is the anonymous language in Galicia. Indeed, *neofalantes*—speakers who switch from speaking predominantly Castilian to predominantly Galician in public—are always noticed, and not always in a positive sense, which "den[ies their use of Galician] the invisibility and anonymity of 'just talk'" (O'Rourke & Ramallo, 2013b, p. 299).[2]

While most of the adverse reaction to *neofalantes'* adoption of Galician reportedly comes from non-speakers, indicating the continuing existence of prejudice against the Galician language in any form, some criticism is additionally leveled at these speakers by native Galician speakers themselves. This occurs mainly because *neofalantes* in the majority speak Urban Standard Galician, the standard linguistic variety developed for formal settings and taught in schools. While the creation of a standard language variety can be seen as an attempt to gain anonymity for a minority language (O'Rourke & Ramallo, 2013b; Ortega, Amorrortu, Goirigolzarri, Urla, & Uranga, 2014), these varieties are often considered inauthentic by traditional speakers. For example, Hornberger and King (1996) report that elder Quichua speakers in the community of Lagunas in Ecuador dislike and claim not to understand Unified Quichua, the standard variety read and spoken by young Quichua-speaking members of the community. Similarly, "book learned" language is frowned upon in Ojibwe and Irish-speaking communities, particularly because of the negative association between schooling and domination in the former case (King & Hermes, 2014; O'Rourke & Ramallo, 2011). This trend holds in

---

[2] O'Rourke & Ramallo (2015) report that *neofalantes* are often viewed suspiciously as having extreme nationalist views, regardless of their true political stance. Additionally, some of these speakers report having been confronted with strong and occasionally hostile or "vicious" reactions upon choosing to adopt Galician as primary language (p. 159).

Galicia: as a non-native, often book-learned language variety, the variety of Galician spoken by *neofalantes* is often seen as less "authentic" than traditional varieties of the language (Vidal Figueiroa, 1997; Kovacová-Moman, 2007; O'Rourke & Ramallo, 2011; 2013b).

Because of their determination to maintain Galician despite this stigma, *neofalantes* constitute what O'Rourke and Ramallo (2015), following Moscovici (1976), call an active minority in that they actively resist the attitudes and pressures of the majority group. However, it is important to note here that those new speakers being described as an active minority share a common set of background characteristics: they are overwhelmingly young, urban, and middle-class (Costa, 2015; O'Rourke & Ramallo, 2011). These speakers, then, enjoy some level of social prestige, even if their language variety is stigmatized as inauthentic. Perhaps because of the social profile of this group, it should be noted that, inauthentic or not, the standard variety is still often seen as more prestigious and more "correct" than the traditional varieties, particularly among older and less economically well-off speakers (O'Rourke & Ramallo, 2011, 2013a). The result is a system of double-diglossia in which traditional Galician, already subordinated to Castilian, comes to be subordinated to the prestigious Urban Standard variety of Galician as well.

Though there has been a comparative wealth of studies on the attitudes and experiences of *neofalantes*, *i.e.* those who have shifted from speaking primarily Spanish to using primarily or exclusively Galician, into which category many of the speakers of Urban Standard Galician would fall to some extent, the vast majority of studies on the Galician linguistic system have focused on the traditional varieties. Aside from reports of

being similar phonologically to Castilian (*e.g.* Vidal Figueiroa, 1997; Beswick, 2010), little is known about the characteristics of the urban variety.

In recent years, due in part to diffusion of the minority language through the education system as described above, there has been a drastic increase in bilingualism in Galicia. García (1986) reports that 85% of residents of the community were bilingual when his article was written, and the most recent *Mapa Sociolingüístico de Galicia* (MSG 2004) indicates that, as a whole, the population's abilities in using Galician in the four modes (listening, speaking, reading, writing) only lag slightly behind those in Castilian (González González, Rodríguez Neira, Fernández Salgado, Loredo Gutiérrez, & Suárez Fernández, 2007). Along with this increased formal competence has come increased prestige, especially among the younger generations of speakers (Beswick, 2002; Bouzada Fernández, 2003). While this high level of ability to use both languages and the increased prestige of Galician is encouraging, it obscures several important facts. First, this bilingualism is occurring primarily at the expense of monolingual Galician; in 1992, nearly 1/3 of the population reported being monolingual in Galician in practice, while in 2004, only 16% of the population fell into this category. In that same time span, rates of practical monolingualism in Castilian rose from 13% to nearly 26% (González González et al, 2007, p. 41)

Additionally, del Valle (2000) argues that "patterns of intergenerational transmission are unfavorable for the maintenance of Galician" (p. 115). The percentage of the population speaking Galician as their first language drops from a high of 81.8% among speakers over 65 to a low of 38.9% among speakers aged 16-25. The MSG 2004 shows similar trends in terms of habitual language across generations; aggregate data for

different age groups and for different municipality sizes are shown in Table 1. As can be

seen, habitual use of Galician is decreasing in apparent time, though the amount of

Galician one can expect to encounter in public still varies greatly depending on the

population size of a community.

Table 1: Habitual language trends in Galicia by Age and Municipality size according to the MSG 2004 (González González et al, 2008, p. 253-264)

		Habitual Language			
		Only Galician	More Galician	More Castilian	Only Castilian
Age	15-24	12.44%	16.99%	40.72%	26.92%
	25-34	11.79%	19.23%	40.22%	28.27%
	35-44	16.54%	24.86%	33.02%	24.98%
	45-54	23.42%	29.79%	26.36%	19.83%
Municipality Size	< 5k	40.55%	29.51%	20.66%	8.74%
	5k-10k	22.97%	27.20%	29.32%	20.29%
	10k-20k	11.30%	23.16%	40.56%	24.52%
	20k-50k	13.88%	22.55%	37.47%	25.50%
	> 50k	6.21%	17.57%	40.06%	35.62%
Galicia-wide Average		15.81%	22.55%	35.33%	25.83%

As the above indicates, the general trend in the population of Galicia as a whole is

a gradual shift in apparent time from Galician to Spanish-dominant bilingualism, and

eventually to Spanish monolingualism. At the same time, however, data from the MSG

2004 indicate that *neofalantes* form a small but important sector of the population: of

those whose initial language was solely or primarily Castilian, over 10% indicate

habitually using primarily or solely Galician in their daily lives (González González et al,

2007, p. 43).

It is this complex sociolinguistic environment in which multiple linguistic

varieties, historically related and with high degrees of similarity still, intersect and interact that forms the social backdrop to the present study. Before considering what language contact theory can tell us about the possibilities in such a situation, however, I turn once more to Latin, this time to trace the trajectory of the development of the grammatical forms which are the primary focus of the present work.

III.   Evolution of *cantase* and *cantara* from Latin to modern Spanish

The Latin morphological system had a rich set of verb forms corresponding to degrees of possibility/reality and differing temporal relationships in conditional sentences (Marín, 1979). Indeed, Porcar Miralles (1993) lists nine different possible combinations, including one way of describing real events in the present/future (present indicative *cantō*), one way of describing real events in the future (future indicative *cantābō*), and two ways of describing real events in the past (imperfect indicative *cantābam,* perfect indicative *cantāvī*). In addition, there were three forms, all using subjunctive, that carried a potential meaning: two in the present/future (present subjunctive *cantem*, perfect subjunctive *cantāverim*), and one in the past (imperfect subjunctive *cantarem*). Finally, there were two ways of communicating unreal modality: one in the present (imperfect subjunctive *cantarem*) and one in the past (pluperfect subjunctive *cantāvissem*).

As is evident from the above, Latin contained four subjunctive tenses: present, perfect, imperfect, and pluperfect, with the imperfect subjunctive pulling double-duty as a marker of unreality in the present and possibility in the past. (Lloyd, 1987; Porcar Miralles, 1993). In this system, the form *cantāveram* was part of the indicative verb scheme, marking the pluperfect. This system was already altered in Late Latin, however.

In part this had to do with a tendency in Latin to regularize and simplify irregular morphology (Porcar Miralles, 1993). This tendency led to reduced forms, first in the perfect indicative, and then by analogy in the other perfect tenses, with the syllable *–ve-/-vi-* tending toward elision (Lloyd, 1987). In Late Latin, then, the pluperfect indicative and subjunctive forms were *cantāram* and *cantāssem*, respectively.

This simplification led to further changes in the Latin verbal system. Already in classical Latin, the future indicative and the perfect subjunctive conjugations differed only in the first person (*cantāverō* and *cantāverim*, respectively). After simplification, these forms became *cantāro* and *cantārem*, with the latter overlapping with the imperfect subjunctive *cantārem* (Lloyd, 1987). Thus, these three forms fused and, passing to refer to desires in the future, formed what was the future subjunctive in medieval Spanish (Marín, 1979; Porcar Miralles, 1993; Villalobos, 1997; Lloyd, 1987). As this new form was oriented toward the future, the pluperfect subjunctive *cantāssem* became the only past-referent subjunctive, and as such was used not only to express unreality in the past, but also possibility in the past and unreality in the present, thus becoming disassociated with pluperfect reference and filling the void left by the former imperfect subjunctive *cantārem* (Villalobos, 1997; Lloyd, 1987).

There seems to be little doubt that the original point of entry of *cantara* into conditionals took place in the apodosis, where it was occasionally used to indicate an event that almost happened, and would have happened if not for an unexpected intervening factor (Porcar Miralles, 1993). The commonly cited example is from Cicero: "*Praeclare **viceramus** nisi spoliatum, inermen, fugientem Lepidus recepisset Antonium*", translated as "We **had practically won** a splendid victory, had Lepidus not given shelter

to Antonius who was attempting to flee and left without troops and arms" (Becker, 2008, p. 152). In this construction, the victory is all but assumed to be reality, but a blocking factor is introduced in the protasis. In these contexts, *cantara* came to be associated with counterfactuals, which occurred alongside a reanalysis of *si*-clauses from being a blocking factor to creating a hypothetical/unreal world (Becker, 2008; Porcar Miralles, 1993). From here, according to Becker (2008), *cantara* expanded into other contexts that were more or less similar to (implicit) conditional phrases, such as the following example from *Cárcel de amor* by Diego de San Pedro: "*Por cierto con mejor voluntad **caminara** para la otra vida que para esta tierra*", translated as "Of course, I **would prefer** to pass to the other life, rather than stay in this world" (Becker, 2008, p. 156). The idea here is that this sentence is an implicit conditional statement, with the condition "*si tuviese la opción*"/"If I had the choice" understood in context but left unsaid. Veiga (1996) cautions that, while clearly unreal, the form *cantara* in these contexts should not yet be considered subjunctive (as evidenced by the clear use of indicative forms such as *cantaría* in these contexts in modern Spanish).

This much of the evolutionary trajectory of the verb forms was shared by both Spanish and Portuguese, and thus presumably by Galician (Becker, 2008). However, how *cantara* passed from the apodosis of conditionals to the protasis in Spanish is debated. It has been postulated that this extension occurred because of parallelism between the protasis and apodosis (*e.g*. Marín, 1979). García de Diego claims "*el agrado acústico de la simetría verbal*", "the acoustic pleasure of verbal symmetry" as a motivating factor (1952, p. 96). Veiga, however, sharply criticizes this position, as it remains unclear just why such a "pleasing" form would stick around long enough to get *cantara* into

conditional protases before giving way to differentialized constructions, nor why *cantara* would make the transition but not *cantaría*, when the latter was far more common than the former in conditional apodoses (Rojo & Montero Cartelle, 1983; but *c.f.* Luquet, 1988; Buzelin Haro, 2010).

Veiga's (1996) explanation for the spread of *cantara* to the protasis of conditional sentences is that there was a reanalysis of the grammatical requirements of this form, and that the necessary criterion for holding this position was no longer subjunctive over indicative, but rather the ability to express unreal modality. Thus he disagrees with those such as Luquet (1988) who argue that *cantara* must have already been considered subjunctive prior to its incorporation in the protasis, citing as the basis of his disagreement comparable indicative usage in French (*si j'avais,* "If I had (imperfect indicative)"). Incidentally, similar uses of indicative in protases have been found in dialects of modern Spanish (*e.g.* Silva-Corvalán, 1982, in Covarrubias; Lavandera, 1975, in Buenos Aires).

Finally, many have cited the development of compound *había cantado* as the impetus for the transition of *cantara* from pluperfect indicative to imperfect subjunctive (*e.g.* Porcar Miralles, 1993; Becker, 2008; Marín, 1979; Lloyd, 1987). Veiga (1996) argues that, while this does explain the loss of association of *cantara* with pluperfect use, pressure from a competing form cannot on its own explain the shift from indicative to subjunctive rather than simply the disappearance of the form. Additionally, alternation between *cantase* and *cantara* began before the new form *había cantado* was well established, and some non-normative evidence of this alternation exists even in Galician, which never developed perfect forms.

Veiga's (1996) explanation of the transition from pluperfect indicative to imperfect subjunctive, the most detailed and well-developed in the literature, relies on multiple factors rather than any one presented above. As indicated above, he argues that its unreal association allowed *cantara* to be used in the protasis of unreal conditionals; thus, the first stage of development involved competing use of indicative *cantara* and subjunctive *cantase* in the protasis of unreal conditionals. This stage was undergone by both the developing Spanish and Galician-Portuguese systems (Becker, 2008). In the second stage, *cantara* spread to other, non-conditional unreal contexts previously occupied solely by *cantase*; in this stage Veiga considers that *cantara* had indeed acquired some subjunctive uses as it stood in opposition to unreal indicative forms. Finally, in the third stage, *cantara* came to identify completely with *cantase* in all subjunctive functions, and a push from the developing perfect forms *hubiese/hubiera cantado* resulted in the loss of pluperfect association of both of these forms.

Finally, it is essential to note that the evolution of verb forms in conditional sentences in Spanish clearly continues. The following section describes the current state of *cantara/cantase* variation throughout the Spanish-speaking world, and specifically in Galicia. It also contains a description of the current uses of *cantara* and *cantase* in Galician.

IV.   *Cantara* and *cantase* variation in modern Spanish

Though the verb forms *cantara* and *cantase* in modern Spanish both correspond to the imperfect subjunctive, the previous section demonstrated that this is not the original distribution of these forms. In Latin, *cantara* marked the simple pluperfect

indicative, while *cantase* corresponded to the pluperfect subjunctive. Over time the two forms gradually lost their perfective association and *cantara* came to be associated with subjunctive values, though it did not completely lose all of its indicative associations (Rojo & Vázquez Rozas, 2014). Despite this, *cantase* and *cantara* are commonly considered equivalent and interchangeable in normative Spanish in almost all contexts (Hermerén & Lindvall, 1989; Carbonero Cano, 1990). Two commonly cited exceptions to this equivalence are the modal verbs (*e.g. quisiera hablar con el dueño*), in which *quisiese* is not permissible (though the conditional *querría* would be accepted), and the few remaining indicative uses of *cantara*, most commonly found in journalistic language (Carbonero Cano, 1990; Rojo & Vázquez Rozas, 2014). Even these exceptional contexts show evidence of overlap between the two forms; Day (2011, in Rosemeyer & Schwenter, 2017) and Lunn (1995) report some *cantase* use in modal and politeness verbs, and non-subjunctive uses of the perfect form *hubiese cantado* have also been reported (Bejarano, 1962). This nearly complete convergence of *cantara* and *cantase* has led to a steady restructuring of the verb form used to mark imperfect subjunctive, with *cantara* replacing *cantase* in most areas of the Spanish-speaking world. This shift, however, has not been uniform in all regions.

While *cantase* predominated in Spain in the early 20th century, by 1973 Gili Gaya reported that *cantara* had become the dominant form in literary and in educated spoken language within the Peninsula. Even so, average rates of use of *cantase* in the Peninsula still topped 20% (Williams, 1982; Pérez Torres, 2014). In 1975 Buenos Aires, however, Lavandera found that only approximately 12% of verbs in her oral sample used the *cantase* form, and in a compared written corpus this rate only reached 8%. Similarly, by

1978, Moreno de Alba reported less than 4% use of *cantase* in Mexico, and Navarro in 1990 noted that *cantase* was used in only 9.2% of possible *cantase/cantara* alternation contexts in Valencia, Venezuela. In a study of the Mexican-American community in Houston published in 1996, no mention of *cantase* forms is made, and all references to the imperfect subjunctive are presented as equivalent to the *cantara* form (Gutiérrez, 1996). In essence, as verified by Rojo (1996), *cantase* has almost completely disappeared as a variant in the Americas.

Spain has maintained the traditional *cantase* form at a higher rate than has Latin America, but this retention is still in general far from overwhelming: Rojo (1996) reported that Spain still maintained an average use of *cantase* around 18.4%. Even in Spain, however, the transition between *cantase* and *cantara* has progressed far from homogenously, as Rojo's data ranged from a low of 0% usage in urban Sevilla to highs of 20% or more in other regions such as Madrid.

It is in Galicia where the conservation of *cantase* appears to be strongest. Kempas (2011), for example, included Galicia in his study of imperfect subjunctive variation in 14 regions in Spain. Using fill-in-the-blank elicitation exercises, he examined not only the frequency of each variant in participant responses, but also considered how many participants used one or the other variant categorically. In comparison to other regions, Galicia yielded the highest percentages of *cantase* usage (44.4%, compared with 22.9% overall) and also had the lowest overall rate of categorical use (34.8%, compared with approximately 60% categorical use nationwide). Of the eight Galicians whose choice of form was categorical, five used only the *cantara* form. This predominance of categorical *cantara* fits with the general tendency Kempas notes overall: *"la tendencia de incluso*

aquellos que prefieren cantase *a utilizar también* cantara *de vez en cuando*" (p. 253). That

is, many of those who use *cantase* forms will also at times use *cantara*, while several of

those who prefer the latter form will never employ the former.

Rojo and Vázquez Rozas (2014) studied this issue in oral Galician Spanish using

a corpus of interviews of educated speakers, thus also including sociolinguistic factors in

their analysis. They found that men and youth tended to favor the *cantase* form more than

women and older speakers. However, when they looked at the data for each individual,

they discovered that examining overall response rates by group can be highly misleading.

Individuals tended to use one form or the other heavily, and each grouping of age and

gender tended to contain at least one individual whose preference was *cantase* and

another whose categorical choice was *cantara*, indicating that idiosyncratic tendencies

should be taken into account in studies of this phenomenon. The only exception to this

was the oldest group of women, all three of whom showed a marked preference for

cantara. Despite this split, however, there was still an evident predominance of *cantara*

forms, as 75% of those who used *cantara* preferentially completely excluded *cantase*

from their usage, while none of those who preferred *cantase* used this form exclusively.

This corresponds neatly to the findings in Kempas (2011).

In Galician, the situation appears to be the inverse, with *cantase* employed in 82%

of imperfect subjunctive contexts and 18% of these contexts employing *cantara* in the

Corpus de referencia do galego actual (Rojo & Vázquez Rozas, 2014). This may be

related to the perception that use of *cantara* as imperfect subjunctive is a Castilianism;

indeed, *cantase* is the only form of the imperfect subjunctive provided in the *Normas*

ortográficas e morfolóxicas do idioma galego. In a short aside, mention is made that

cantara can also be found as imperfect subjunctive; however, it is recommended "*manter ben discriminados os usos dos dous tempos (-ra coma antepretérito de indicativo e –se coma pretérito de subxuntivo*)"[3] (RAG/ILG, 2012, p. 109). However, to the best of my knowledge, no systematic sociolinguistic study on the use of these two forms in Galician exists to date.

V.   Factors associated with use of *cantase* and *cantara*

A variety of factors have been found to be associated with variation between *cantara* and *cantase* in the Spanish of different areas; the following subsections present an overview of linguistic and social factors, respectively that have reportedly been found to correlate with form choice.

A. *Linguistic correlates of form use*

Despite the aforementioned common belief that *cantara* and *cantase* are equivalent in normative Spanish in most contexts, some linguistic factors have been reported to condition the choice of form in various dialects. In their study of a translation of *Nils Holgerssons underbara resa genom Sverige* from Swedish into Spanish, for example, Hermerén and Lindvall (1989) found that high frequency verbs *tener*, *saber*, and *hacer* were never used in the –*se* form, and verbs *ver, estar, poder, haber* and *querer* were found in this form only sparingly (less than 12%). In contrast, *ser/ir*, *dar*, and other non-specified –*ar, -er,* and –*ir* verbs were found in the –*se* form from 24% to 44% of the times they occurred. In contrast, Rosemeyer and Schwenter (2017) found that the highest

---

[3] Translation: "To maintain well-separated the use of the two tenses (-ra as pluperfect indicative and –se as imperfect subjunctive".

frequency lexical items they considered, *ser, estar,* and *haber*, showed higher frequency use of *cantase* than lower-frequency verbs. Thus, it appears that the verb involved and/or its frequency may have an influence on the form chosen, though the direction of that influence is not clear from previous studies. Person and number morphology may also play a role, with both third person and singular forms having been reported to correspond to increased *cantase* usage (Rosemeyer & Schwenter, 2017).

From a semantic standpoint, Bolinger argues that "-*se* implies 'remoteness, detachment, hypothesis, lack of interest, vagueness, greater unlikelihood,' while –*ra* brings everything into relatively sharper focus" (1956, p. 346). However, his arguments were based on the intuitions of one speaker, making their extension to other linguistic contexts, or even their reliability for the dialect considered, dubious at best. Even so, his conclusion bears strong resemblance to that of both Lavandera (1975) and Silva-Corvalán (1984) in their respective studies on the extension of conditional forms into the protasis of *si*-clauses. In her dissertation study on conditional statements in Buenos Aires, Lavandera (1975) considered present indicative, imperfect subjunctive, and conditional as possible protasis variants, and examined the occurrence of each with respect to the degree of probability of the event being discussed. She found that the present tense in the protasis was heavily linked to statements viewed as facts or nearly facts, the imperfect subjunctive was favored when statements were clearly contrary-to-reality, and the conditional was favored in those in-between cases that were neither highly probable nor clearly impossible. Similarly, in her study in Covarrubias, Spain, Silva-Corvalán (1984) noted large-scale replacement of the imperfect subjunctive by the conditional in the protasis of *si*-clauses, which she attributed to the "principle of distance". Effectively, this

principle states that because the forms in question are closely related, "the form which is farthest away from the speaker, in the sense that it refers to objects or events which are the farthest from him in his objective (*e.g.* actual distance) or subjective (*e.g.* possibility of actualization) world, will be lost" (p. 596). While neither of these studies distinguished between *cantara* and *cantase*, the principle that the most "distant" form—the *irrealis* form—tends to be lost could help to explain the diminishing use of *cantase* if Bolinger's (1956) assertion that *cantase* is less real can be shown to hold. Some support of this arises in Asratián's (2007) finding that in Caracas, Venezuela negated clauses favored the use of *cantase* while affirmative clauses favored use of *cantara*; other studies such as Rosemeyer and Schwenter (2017), however, have not found a significant effect of clause polarity.

An additional factor that merits consideration for historical reasons is the type of clause in which the subjunctive form is found. Despite the fact that neither Carbonero Cano (1990), in a study of popular speech in Seville, nor Rosemeyer and Schwenter, in their 2017 study of the *Corpus del español,* found clause type to be relevant to choice of *cantara* or *cantase*, it is generally accepted, as attested in Section III of the present chapter, that *cantara* first came into competition with *cantase* in conditional phrases and only later extended to other contexts. Thus it is plausible that, if a change is in progress in the Spanish of Galicia, it will manifest itself first or more extensively in some syntactic contexts than in others.

Finally, priming—that is, the effect of a previous word or form in the replication of that same structure in a subsequent utterance—is a linguistic factor that appears to be highly significant in determining which form will be employed. Rosemeyer and

Schwenter (2017) found that a prior use of either *cantase* or *cantara* significantly conditioned the subsequent choice of that same form. This effect was far stronger for *cantase* than for *cantara*, from which they infer that obsolescing constructions tend to have stronger priming effects than frequent constructions. Their study also found that priming resulted in a reduction of paradigmatic restrictions on subsequent iterations of the form; in other words, while the first use of a *cantase* form was far more likely to occur in a third person singular form as mentioned previously, the preference for that form was reduced or disappeared entirely in subsequent tokens of the same form. It is additionally important to note that structural priming can occur without lexical repetition (*e.g.* the use of *cantase* could influence later choice of *amase* despite the different verb involved), though lexical repetitions greatly strengthen the influence of structural priming (Pickering & Branigan, 1998). Finally, the proximity of a token to its prime must be taken into consideration as greater distance between the two occurrences has been shown to correspond to a sharp decrease in priming effects (*e.g.* Rosemeyer & Schwenter, 2017), though priming effects have also been shown to persist across nine or more intervening segments for some structures in English (Bock & Griffin, 2000).

B. *Social correlates of form use*

In terms of social factors related to imperfect subjunctive variation, findings have been inconsistent. In an appendix to her dissertation research, Lavandera (1975) analyzed *cantara*/*cantase* variation in spoken speech in Buenos Aires and found that the primary users of the *cantase* form in Buenos Aires were men, the middle-aged, and non-college-educated speakers. In contrast, Navarro (1990) noted on the basis of recordings of spontaneous speech in Valencia, Venezuela that, while gender did not appear to play a

role in conditioning form choice, *cantase* was favored in general by the older, more educated, and wealthier speakers. However, Lavandera (1975), at least, presents some doubt that these results are practically relevant, calling the tokens of *cantase* in her data "a 'trace' left by an older stage of variation and change" (p. 372-373). In Spain, Kempas (2011) reported a largely non-significant effect for gender, with men favoring *cantase* and women favoring *cantara*, though she did not analyze any potential interaction between gender and region within Spain. Social factors do also play a role in Galicia, however; as previously mentioned, Rojo and Vázquez Rozas (2014) found that men and youth tend to favor the *cantase* form, despite the fact that these group trends obscure large intra-group variation due to individual differences. Such individual differences are not frequently examined in the literature, perhaps due to the difficulty in obtaining sufficient tokens of syntactic variables to permit a robust statistical analysis even without a consideration of individual variation. However, similar to Rojo and Vázquez Rozas (2014), Rosemeyer and Schwenter (2017) also found that adding this factor to an analysis of imperfect subjunctive use in the *Corpus del Español* increased the predictive power of their model by nearly fifteen percent.

Finally, an important potential factor that has not to my knowledge been studied with respect to *cantara*/*cantase* variation in Galicia is language contact. In considering the motivation behind the extension of the conditional to imperfect subjunctive contexts, Silva-Corvalán (1982) promotes a hypothesis of complex causation in which it is a combination of semantic ambiguity and language contact that causes the shift. She postulates that the shift of *cantara*—which as Espinoza (1930) noted often alternated with conditional forms in certain contexts—from indicative to subjunctive caused a lack

of clarity in the meaning difference between subjunctive *cantara* and indicative *cantaría*.

Language contact also plays a role, she claims, as in Lavandera's (1975) study it was

Spanish speakers of Italian origin, and in her own study Spanish speakers of Basque

background, who most frequently employed the conditional forms in a parallel manner to

their language-of-origin.[4]

Though Silva-Corvalán's hypothesis was specifically regarding the extension of

conditional to imperfect subjunctive contexts, her arguments can logically be extended

also to *cantara*/*cantase* variation. She argues that "given two forms, one indicative and

one subjunctive, with almost identical meaning and distribution, the universal tendency is

for the subjunctive form to fall into disuse, probably because of an overall lower

frequency of occurrence in discourse" (Silva-Corvalán, 1982, p. 92). This appears to be

the same historical trend followed by verb forms *cantase* and *cantara*: while *cantara* was

originally indicative, as its meaning drew closer to that of *cantase* and it took on

subjunctive modality, it came to predominate over *cantase* perhaps because it still

retained some indicative presence that lent it higher frequency of use. It can additionally

be hypothesized that, if influence from parallel structures in Basque and Italian could

accentuate and accelerate a shift to conditional, contact between Spanish and a language

such as Galician that retains the modality difference between *cantase* and *cantara* might

impede the ongoing shift from the former to the latter in imperfect subjunctive contexts—

or, conversely, the shift apparent in Castilian might become the norm in Galician as well.

Because one of the guiding questions behind this project relates to the potential

---

[4] Silva-Corvalán (1982) refers to N'Diaye (1970) and to Wright (1932) to support her assertion of similarities between the examined conditional use and reported phenomena in Basque and Italian, respectively.

relevance of contact with Galician to the elevated use of *cantase* in the region, the final section of this chapter briefly presents the theoretical concepts from the field of language contact that inform the present study design and analysis.

VI.   Theoretical considerations from the field of language contact

Having developed from closely related varieties of late Latin spoken in the Peninsula and therefore being typologically similar languages, it can be expected that both Galician and Castilian would be highly permeable with respect to influences from one another (Thomason & Kaufman, 1988). However, predictions as to the type, extent, and direction of influence are not always easy to make based on language contact theory. Thomason and Kaufman's framework, for example, predicts differing results of contact between two languages based on whether speakers are maintaining their original language, in which case they may borrow lexical and perhaps structural items from the other language, or are shifting to the new language, in which case their former language can be expected to "interfere" with many levels of the target language. Similarly, van Coetsem (2000) talks about contact effects being determined by the agency of speakers dominant in one of the two languages. If speakers are dominant in the receiving language, they will borrow primarily lexical elements from the source language and incorporate them in their language use. If speakers are dominant in the source language, however, they will impose structural (and sometimes lexical) elements of their dominant language onto their weaker, receptive language.

It is difficult to apply these theoretical constructs in a systematic way to the contact situation in Galicia, for two reasons. First, because many speakers in the region

grow up bilingual, determining which language is dominant, and therefore agentitive in van Coetsem's framework, is often infeasible. When bilinguals are highly competent in both languages, they enter a state called *neutralization* in van Coetsem's framework in which theoretical restrictions on the type and direction of contact effects no longer apply. Indeed, even on a practical level, there is little consensus among researchers on how best to operationalize or measure the concept of dominance in bilinguals or on whether any one such proposed measure is adequate (Unsworth, 2015). Second, although the historical choice in the region would have been between maintenance of Galician or shift to Castilian, a reverse shift is also occurring. Many young speakers who were raised predominantly in Castilian are choosing to adopt Galician as their nearly exclusive operating language, to show pride in and solidarity with regional identity (see, for example, Ramallo (2013)). Speakers' attitudes such as these play a huge role in determining the outcome of language contact; Thomason (2001) argues that "attitudes can be either barriers to change or promoters of change. In other words," she continues, "the reason contact-induced language change is unpredictable is that speakers are unpredictable" (p. 85).

Rather than attempting to predict changes, then, more promising in this context are the possible indications of transfer between languages in contact presented in Silva-Corvalán (1994). Without constraining the direction in which transfer can occur between languages or commenting on the likelihood of such transfer, she describes the following conditions (among others) that may indicate that transfer has taken place:

- When two languages X and Y share a form that is structurally similar but with different functions in X and Y, the function of the form in language X may

become associated with the parallel form in language Y, even if language Y already has a different form with the same function.

- When two forms in language Y are in competition, a contact-variety of language Y may have more frequent use of the form that is most similar to that of language X, as compared to a possibly categorical or highly preferred alternate form in non-contact varieties.

Based on these conditions, the following two phenomena, if found to hold true for the Spanish spoken in Galicia, may indicate transfer from Galician to Spanish:

1. The *cantara* form in the Spanish of Galicia may take on the perfective meaning of the Galician form (attested in Rabanal (1967), Pollán (2001)).

2. The *cantase* form of imperfect subjunctive, far less frequent than the *cantara* form in most varieties of Spanish including that of Madrid (Nowikov 1984), may be more common in Galicia, where it is analogous to the *cantase* form of the Galician imperfect subjunctive.

It is the latter of these two possibilities—that is, the potential impact of language contact on the imperfect subjunctive system—that is the focus of this thesis. Matras (2010) presents reason to believe that modality may be particularly prone to transfer effects, due to its high degree of cognitive vulnerability; however, Thomason (2010) argues that features such as the subjunctive are also more prone to language drift, since they "impose a burden on learning…and are therefore likely to be diachronically unstable" (p. 44). It is clear that theory alone, then, is insufficient to predict whether language contact is a factor contributing to the reportedly elevated rates of *cantase* in Galician Spanish. It is with an eye toward resolving this conundrum that I now turn to

describe the present study.

Chapter 3: Methodology

One seeming paradox in the study of languages is the need, on the one hand, to describe languages as relatively fixed systems—if a language were not largely systematic and conventionalized, it would be useless for communication—while at the same time, on the other, recognizing that linguistic systems are constantly evolving (Tagliamonte, 2006; Linell, 2009). These changes often occur over large time spans, and as such, what is considered a unified "language" will at any given point in time be replete with heterogeneities and variations, both between and within individual speakers (Croft, 2000; Linell, 2009). Language change is a characteristic even of linguistic systems that are not in substantial contact with others (Croft, 2000; Thomason, 2010)—a phenomenon known as linguistic drift—but many linguistic changes can be argued to be the result of influence from a contact language (Thomason & Kaufmann, 1988; van Coetsem, 1988). Thus, one of many possible queries when examining language variation is how much of that variation is attributable to contact effects.

Variationist sociolinguistics provides an appropriate approach to address this and other questions about linguistic production within a speech community. According to Tagliamonte (2006), "the variationist enterprise is essentially, and foremost, the study of the interplay between variation, social meaning and the evolution and development of the linguistic system itself" (p. 5). In other words, a variationist study explores patterns in linguistic production within a community by considering both linguistic factors—such as pragmatic context, priming, or syntactic structure of the utterances in question, to name a few—and social factors—such as gender, age, education, and linguistic background—as possible variables conditioning the choice of one variant over others (Poplack, 1993).

Because the research questions of this study deal with both linguistic and social variables, then, a variationist sociolinguistic approach will be adopted here.

As its name implies, the focus of a variationist study is on instances of multiple forms within a language performing the same function (Poplack, 1993; Tagliamonte, 2006). As described in previous chapters, this study specifically examines variation in the use of competing forms *cantara* and *cantase* of the Spanish imperfect subjunctive and seeks to discover which social and linguistic factors are relevant in conditioning the distribution of these forms in the variety of Spanish in contact with Galician. The rest of this chapter describes the methodology used to obtain the data on which this study is based, including the profile of participants, the tasks they were asked to complete, and the analysis to which the data was subjected. I end with a detailed exploration of the results of the Bilingual Language Profile, including both statistical and logical arguments for the way social variables were combined and coded for each task.

I.   Participants

This study incorporates data from 39 individuals living in the Galician region from a very young age.[5] Participants were obtained through the anchor group technique, following the method used in Perez Castillejo's (2014) study on Galician intonation.[6] In this technique, an original group of participants known to the researcher is identified and serves as the anchor group, and these individuals then invite the participation of their own friends, family, and acquaintances. This process is repeated, with new participants

---

[5] The majority of participants (35) were born and raised in Galicia. Four individuals were not born in Galicia, but were born elsewhere to Galician parents and returned to Galicia in childhood.
[6] This is also known as the 'friend-of-a-friend' technique (Tagliamonte, 2006).

spreading the word to others in their own social networks. All 39 participants were obtained in this way.[7]

Participants were obtained in two different *comarcas* in the province of Pontevedra: O Salnés and Pontevedra. These *comarcas* are similar in population, with roughly 108,000 and 121,000 inhabitants, respectively. Within these regions, participant solicitation was focused around the municipalities of O Grove in O Salnés and Marín in Pontevedra.[8] In addition to being the hometowns of friends of the researcher and therefore ideal locations for carrying out the anchor group technique, these two localities are both port towns whose nearest urban center is the city of Pontevedra. As such, both localities share industrial similarities. They differ, however, in size and proximity to the urban center. O Grove has a population of around 11,000, and is located roughly 20 miles from the city of Pontevedra, while Marín is only about 5 miles from the urban center and has roughly 26,000 inhabitants. The location of these municipalities is shown in Figure 1.

Linguistic practices for each municipality and for the city of Pontevedra, based on data from the *Mapa Sociolingüístico de Galicia 2004* (González González et al, 2008) are displayed in *Table 2*. Municipality-specific data was only available for the seven largest cities in Galicia, including Pontevedra, so the data for O Grove and Marín reflect the data reported for municipalities of similar sizes. As can be seen, locations the size of O Grove and Marín share similar linguistic practices, with 34.46% and 36.43% of the population, respectively, reporting habitual use of mostly or only Galician. In the city of Pontevedra,

---

[7] One statistical drawback to this method is that it is not a random sampling method, so results are not necessarily generalizable to the larger population. Additionally, if the distribution of social variables in the sample is not random, this can result in bias in the results.

[8] In a few cases, participants were actually living outside the municipal limits, but their social networks (such as friendship and church attendance) centered around the municipality in which they were included.

however, Galician is the primary language of only 22.6% of the population.

Figure 1: Map of Galicia showing the municipalities in which participants were sought

Table 2: Habitual language in the targeted municipalities and in the nearest urban center, according to the MSG 2004 (González González et al, 2008)

Municipality	Only Galician	More Galician	More Spanish	Only Spanish	Other
Between 10,001 & 20,000 inhabitants (O Grove)	11.30%	23.16%	40.56%	24.52%	0.45%
Between 20,001 & 50,000 inhabitants (Marín)	13.88%	22.55%	37.47%	25.5%	0.61%
Pontevedra	6.72%	15.88%	39.34%	36.77%	1.29%

These data reflect another important consideration in the choice of locales: while linguistic practices in O Grove and Marín themselves are likely quite similar, Marín is located so near to Pontevedra that jobs in or trips to the latter are a common part of the mobility practices (*i.e.* Britain, 2010) of those living in the former, while the same does not hold true for residents of O Grove. Thus, it is quite possible that residents of Marín,

though speaking Galician in similar rates to residents of O Grove, have less overall exposure to Galician.

One challenge in using the anchor group technique is that individuals can tend to associate with relatively homogeneous groups, thus leading to a lack of diversity in the participant base. However, effort was made to obtain as diverse a sample as possible in terms of age, gender, and education levels in each municipality. Ages ranged from 18 to 75, and education ranged from primary only to completion of a graduate degree. The breakdown of participants by locale, gender, age, and education level is shown in Table 3.

Table 3: Breakdown of participants by Location, Gender, Age, and Education

Location	O Grove															
Gender	Male								Female							
Age	Under 40				40+				Under 40				40+			
Education	P	S	U	G	P	S	U	G	P	S	U	G	P	S	U	G
			1				2	1		1	2	1			3	

Location	Marín															
Gender	Male								Female							
Age	Under 40				40+				Under 40				40+			
Education	P	S	U	G	P	S	U	G	P	S	U	G	P	S	U	G
		3	2		1	3	1			1	2	5	2	5	2	

P = primary, S = secondary, U = undergraduate, G = graduate

II. Tasks

Participants in this study completed a total of five tasks: a bilingual language profile, a semi-structured group conversation, a pseudo-matched-guise task, a fill-in-the blanks task, and an acceptability rating task. Sessions were held in groups ranging from two to five participants. The tasks, described in greater detail below, were carried out in the order listed in order to maintain participants' unawareness of the focus of the study for as long as possible. The bilingual language profile, the group conversation, and the matched-guise tasks were considered highly unlikely to reveal the focus on imperfect subjunctive, while the fill-in-the-blanks task drew participant attention to verb forms in

general. The last task, the acceptability judgment task, was designed to lead participants to focus specifically on *cantara* and *cantase* forms, and thus was carried out after all other tasks.

   A. Bilingual language profile

   This task consisted of a 5-page background questionnaire adapted from Birdsong, Gertken, & Amengual (2012) to the Galician context; this task is included in the Appendix. The first page requested the following demographic information: age, gender, education level, profession, place of birth, and current place of residence. The second page focused on the participant's linguistic background, including age of exposure to Spanish and Galician and the number of years of formal use (*e.g.* education, work) of each language. The third page asked about current levels of use of each language in different social and personal contexts, and the fourth asked participants to rate their competence in each language with respect to speaking, aural comprehension, reading, and writing. The last page contained questions about the participant's personal affect toward/affiliation with each language. This task was included to address research questions 3 and 5, restated below, by examining how social factors and language background relate to outcomes in the other tasks, and results are discussed at the end of the current chapter.

Question 3: What social factors, if any, correlate with the use or perception of *cantara* or *cantase* in Galician Spanish?

Question 5: What relationship, if any, exists between exposure to the Galician language and use of *cantara* and *cantase* in Galician Spanish?

B. *Semi-structured group conversation*

In this task, groups of two to five participants were recorded in conversation with the researcher about a variety of daily topics, including family, work, and current social issues such as emigration. An interview protocol was designed in advance in order to provide possible topics of conversation related to areas of general interest, particularly those about which hypothetical situations might be considered, such as the reaction of parents to various disciplinary problems or social issues. The complete interview protocol is included in the Appendix; however, it should be noted that the researcher used this protocol only as a loose guideline, and attempted whenever possible to pursue specific topics brought up by participants in the conversation. Additionally, participants were told that they could talk amongst themselves about whatever they wanted, and the researcher only inserted topics from the research protocol when conversation was waning. This avoidance of researcher control of the conversation was due to the desire to obtain "spontaneous unreflecting speech in its natural context", which is the main goal of sociolinguistic endeavors (Sankoff, 1982, p. 677). However, each conversation concluded with the researcher asking participants to respond to the questions "*¿Te consideras una persona feliz? ¿Bajo qué condiciones podrías llegar a ser aún más feliz?*" to create continuity between this conversation and the pseudo-matched-guise task described in the next section. This group conversation task is directed at beginning to address research

questions 2, 3, and 4, about oral production of *cantara* and *cantase* and about social and linguistic correlates of the use of both.

Group conversations were intended to last around one hour, though when participants were so inclined they often stretched longer, ranging from one hour to 2.5 hours in length. These conversations were carried out in groups rather than individually with the interviewer in order to elicit the most natural speech possible. The goal of variationist investigations of spontaneous speech is to access informal, vernacular speech styles (Tagliamonte, 2006). However, best practice for obtaining this type of speech is "that the raw data be collected by skilled interviewers who not only are, but are also perceived by informants to be, in-group members, and whose own linguistic repertoires feature the same phenomena we are attempting to elicit" (Poplack, 1993, p. 260). Because the researcher neither is nor is perceived by Galicians to be a Galician herself, the presence of multiple Galician participants in the conversation was ensured to minimize linguistic accommodation to the non-native speech of the researcher. Though it is perhaps inevitable that the context of being recorded may have caused discomfort and therefore unnatural speech patterns for some participants, the conversations obtained frequently contained laughter, risqué joking, prolonged arguments/shouting between participants, and even codeswitching between Castilian and Galician, all of which may be

taken to indicate a high degree of comfort and openness on the part of many speakers

> Question 2: Do written and oral production of *cantara* and *cantase* correlate with perception/acceptance of these forms?
>
> Question 3: What social factors, if any, correlate with the use or perception of *cantara* or *cantase* in Galician Spanish?
>
> Question 4: What linguistic factors, if any, condition the choice of *cantara* or *cantase* in Galician Spanish?

(Labov, 1972).

C. Pseudo-matched-guise task

The purpose of this task is to access aural perception of the *cantara* and *cantase* forms, in line with research questions 1, 2, and 3.

> Question 1: How is use of cantara and cantase perceived by speakers of Galician Spanish?
>
> Question 2: Do written and oral production of *cantara* and *cantase* correlate with perception/acceptance of these forms?
>
> Question 3: What social factors, if any, correlate with the use or perception of *cantara* or *cantase* in Galician Spanish?

In a true matched-guise task (*i.e.* Lambert, Hodgson, Gardner & Fillenbaum, 1960), speakers who are able to use two different languages or dialects are recorded producing the same text in both linguistic varieties, and a series of these matched recordings are played for naïve listeners unaware that they are hearing the same people in different languages. Listeners then rate the speakers on Likert scale with respect to a variety of different personality traits, such as sense of humor, intelligence, and

trustworthiness. Because the same speakers have been used, differences in timbre and voice quality are presumably controlled for, meaning that differences in listener evaluations of the two recordings should be due to the language used rather than to actual speaker differences.

The present study follows Silva-Corvalán (1984) in adopting a modified version of this task suitable to studying differences in the verb form used. To create the recordings, a woman from Galicia, aware of the focus of this study and therefore instructed to use the imperfect subjunctive in her response but otherwise unscripted, was recorded responding spontaneously to the questions "*¿Te consideras una persona feliz? ¿Bajo qué condiciones podrías llegar a ser aún más feliz?*" Her response was then transcribed to produce two texts, each identical to the other and to her original production except that one text only included the imperfect subjunctive form *cantara* and the other only included *cantase*. The Galician woman was then recorded reading both texts, and these two recordings were the stimulus to which participants reacted in this task.

Each group of participants, directly following the group conversation, listened to one of the two versions of the recording and was asked to rate the speaker on a Likert scale from 1 to 6 based on their perceptions of eighteen personal characteristics. The response form containing the list of characteristics, included in the Appendix, was the matched-guise questionnaire designed by the *Seminario de Sociolingüística de la Real Academia Galega* research team, to be certain that the characteristics included were relevant to and appropriate for the Galician context. An attempt was made to ensure that participants with similar backgrounds listened to different recordings, so as to neutralize

any bias in the responses due to participant characteristics.

The purpose of a matched-guise task is to allow the researcher to access subjective reactions to language use, thereby allowing participant behaviors to be compared to their implicit judgments of the language forms under consideration. However, it should be noted that participant evaluations in this type of task are not reliable predictors of linguistic behavior (Poplack, 1993). For this reason, the present study incorporates the matched-guise as only one of several tasks related to the status of *cantara* and *cantase* in Galician Spanish.

D. Fill-in-the-blanks task

To evaluate written production of *cantara* and *cantase*, participants were asked to complete a written conversation between *Dra. Sánchez* and *Don Ismael* by filling in the blanks with a conjugated form of a verb, given in parenthesis.[9] This task was adapted in two ways from a Spanish language textbook activity from the chapter in which the imperfect subjunctive in hypothetical statements is introduced. The first adaptation involved replacing all conjugated verbs in the conversation with a blank and an infinitive verb form in parenthesis yielding a total of 22 blanks to fill; these additional blanks were meant to distract the participant from identifying the specific focus of this study. Additionally, the conversation was examined by a woman from Galicia for naturalness, and any words or expressions that were out of place for the Galician context were modified based on her recommendations. This task was included to address research questions 3, 4, and 5.

---

[9] Despite the fact that this task depicted a conversation, the fact that the task was realized in writing combined with the formal nature of the interaction between doctor and patient justifies consideration of this task as much more formal than the group conversation.

Question 3: What social factors, if any, correlate with the use or perception of *cantara* or *cantase* in Galician Spanish?

Question 4: What linguistic factors, if any, condition the choice of cantara or cantase in Galician Spanish?

Question 5: What relationship, if any, exists between exposure to the Galician language and use of *cantara* and *cantase* in Galician Spanish?

E. Acceptability rating task

In this task, participants were given a series of 33 sentences to be evaluated for grammatical acceptability. Of these statements, 24 were the conditional statements targeted by this project, and 9 were sentences requiring a pluperfect indicative form. The latter type were included to ensure that participants remained focused on the meaning of the sentences throughout the task, as *cantara* is the standard form of the pluperfect indicative in Galician, while it is not associated with this function in standard, non-journalistic Spanish. All statements were verified by a Galician woman to ensure that there were no lexical or grammatical errors outside of the choice of verb form, and a prior version of this task was piloted in a separate study.

The 24 conditional statements consisted of twelve distinct sentences, each included twice with different verb forms in each iteration. In these statements, the verb form in the protasis rotated between imperfect subjunctive *cantara*, imperfect subjunctive *cantase*, and conditional *cantaría*, while the form in the apodosis was either conditional *cantaría* or imperfect *cantaba*. Additionally, half of the conditions were clearly contrary-to-fact (*e.g. "Si yo fuera tú…"*), while the other half were not clearly contrary-to-fact (*e.g. "Si yo ganase la lotería…"*). The composition of the 24 sentences is summarized in

Table 4.

Table 4: Composition of conditional statements in acceptability task

		Protasis						
		Contrary-to-Fact			Not Contrary-to-Fact			
		-ra	-se	cond	-ra	-se	cond	Totals
Apodosis	cond	2	2	2	2	2	2	12
Apodosis	imp	2	2	2	2	2	2	12
Totals		4	4	4	4	4	4	**24**
		Total C-t-F: 12			Total Not C-t-F: 12			

Given these sentences, participants were asked to rate a series of three statements related to each on a four-point descriptive scale, with options being "*Estoy muy en desacuerdo*", "*Estoy más o menos en desacuerdo*", "*Estoy más o menos de acuerdo*", and "*Estoy muy de acuerdo*". The series of statements rated for each sentence was:

1. *Conozco a gente que lo diría así.*

2. *Yo lo diría así.*

3. *Está bien dicho.*

Participants were instructed that they were to rate the sentences based on their grammatical correctness, not on their propositional content. If participants responded that they disagreed with statements 2 or 3, they were then asked to reformulate the phrase so that it would be correct and/or something they themselves would say. The purpose of this was twofold. First, it allows for verification that rejected phrases were evaluated in this way due to their verb form rather than due to lexical or propositional issues. Second, the proposed corrections supplement the written data from the fill-in-the-blank task discussed

previously, allowing for more insight into written production of the forms in question. Thus, this task is designed to contribute to an understanding of research questions 1, 2, 3, 4, and 5 shown below, related to social and linguistic factors contributing to both written perception and production of the two possible imperfect subjunctive forms. The form used for this task is included in the Appendix.

Question 1: How is use of cantara and cantase perceived by speakers of Galician Spanish?

Question 2: Do written and oral production of *cantara* and *cantase* correlate with perception/acceptance of these forms?

Question 3: What social factors, if any, correlate with the use or perception of *cantara* or *cantase* in Galician Spanish?

Question 4: What linguistic factors, if any, condition the choice of cantara or cantase in Galician Spanish?

Question 5: What relationship, if any, exists between exposure to the Galician language and use of *cantara* and *cantase* in Galician Spanish?

III. Methods of analysis

As stated in Chapters 1 and 2, one of the main questions of the present study is the relationship between written and spoken perception and production of the forms *cantara* and *cantase*. To this end, the last four of the tasks just described were intended to address oral production, aural perception, written production, and written production/perception, respectively. For all statistical tests, the significance level chosen was $p < 0.01$, with p-

values less than 0.05 considered to be approaching significance.[10] The different methods

employed for each task, however, make direct statistical comparison of the results of

these tasks a difficult, if not impossible, endeavor. Because of this, each task was

analyzed individually, and larger connections between results on each task are presented

descriptively in Chapter 5.

A. *Oral production*

Oral production was assessed through the data obtained from the group

conversation task. Each conversation was listened to in full and instances of *cantara* and

cantase were identified and transcribed along with at least fifteen seconds of preceding

and following utterances. Uses of *cantara* or *cantase* other than the subjunctive (*e.g.*

pluperfect indicative; modal uses synonymous with the conditional) are not included in

the present analysis.[11]

The dependent variable in this analysis is the choice of verb form. Each instance

of *cantara* and *cantase* was coded for the following linguistic factors, which served as

independent variables:

- Possibility: whether or not the proposition expressed with the imperfect

  subjunctive is contrary-to-fact (following Lavandera, 1975)

- Polarity: whether or not the form is preceded by a "*no*" or other form of negation

---

[10] Although it is common practice in linguistics to use p < 0.05 as the cutoff for significance, the sheer number of statistical tests run in this study made it prudent to adopt a more conservative confidence level in order to avoid as much as possible erroneously identifying factors as significant (*i.e.* false positives), known in statistics as Type I errors. With a p-value of less than 0.01, fewer than 1 out of every 100 tests should, on average, result in a false positive result for significance.

[11] Pluperfect subjunctive forms *hubiera/hubiese cantado*, are not considered in this thesis, as rates of use of these two forms are quite low in both the present data (fewer than 10 tokens) and overall in Galicia (*e.g.* Rojo & Vázquez Rozas 2014).

(following Rosemeyer & Schwenter, 2017)

- Type of clause: whether the form in question is contained in a noun clause, an adjective clause, an adverbial clause, a hypothetical statement, or a prepositional clause

- Priming: the presence of the same imperfect subjunctive form, of the opposite form, or of neither form preceding the target form (following Rosemeyer & Schwenter, 2017; Pickering & Branigan, 1998; Bock & Griffin, 2000)

- Anteriority: whether or not the form is being used to express an event before another event in the past[12]

- Verb group: the lexical/morphological content of the target form—either *ser*, *tener*, *–ar*, *-er*, or *–ir* verbs (following Hermerén & Lindvall, 1989; Rosemeyer & Schwenter, 2017)

These linguistic variables were included in a logistic regression analysis using R to determine which factors condition the choice of *cantara* or *cantase*. Additionally, the impact of the social variables described below was also considered. This had to be done in a separate analysis, however, as most participants did not produce any variation in their tokens. Thus the percentage of *cantase* production per participant was coded as a dependent variable and a linear regression with ANOVA was run in R using the social factors as independent variables. However, one limitation to the study of grammatical variables in spontaneous speech is the difficulty in obtaining sufficient numbers of tokens to permit a robust statistical analysis (Sankoff, 1982); thus, each of these social and

---

[12] While indicative uses with this temporal reference were excluded as indicated previously, some subjunctive uses of *cantara/cantase* may still have an anterior interpretation, particularly in Galicia, where compound forms such as *hubiera/hubiese cantado* are uncommon (*e.g. "Mi amigo me devolvió el libro la semana pasada, pero habría sido mejor que lo hiciera/hiciese antes"*).

linguistic variables was also analyzed descriptively, and these results are presented in the following chapter.

B. *Aural perception*

Aural perception of *cantase* and *cantara* is examined by way of the pseudo-matched-guise task. For this task, the rating scores given by participants were normalized through the use of z-scores. In using z-scores, each participant's score for each of the 18 characteristics rated was described by its relationship to that participant's average rating; a z-score of zero, for example, indicates a rating right at that participant's average, while a z-score of 0.5 indicates a rating one-half of a standard deviation higher than the average. The advantage of using z-scores over raw scores is that it controls for the possibility that some individuals naturally tend to give higher ratings overall than others. Along with the social variables described below, the matched-guise version presented was included as an independent variable. Additionally, sixteen of the eighteen characteristics from the matched-guise questionnaire were grouped via factor analysis into three categories that can be loosely described as centering around friendship (6 characteristics), capability (6 characteristics), and leadership/charisma (4 characteristics).[13] These first two of these groupings are similar to the categories of personal appeal and capability, respectively, used in Loureiro-Rodriguez, Boggess and Goldsmith (2012), despite the fact that the latter's groupings were created *a priori* rather than through factor analysis. Characteristic group was also considered as an independent

---

[13] Two characteristics were excluded from consideration. Participants consistently rated the speaker's "Similarity to themselves" far lower than any of the other characteristics for both guises, so this characteristic was excluded as an outlier. The other characteristic, *Orgullosa*/"Pride", was excluded due to semantic ambiguity; in Spanish, as in English, pride can be a positive trait, as when an individual takes pride in their accomplishments, or a negative trait, as when an individual refuses to admit their own failings, and the context of the task did not clearly indicate which meaning was intended.

variable. Finally, random effects variables were used to account for the existence of multiple tokens produced by the same participant and responses from multiple participants to the same question. Thus a linear regression with random effects was carried out in R to test for significance of each variable, and both statistical and descriptive results are presented in Chapter 4.

C. *Written perception*

Written perception was analyzed through use of participant responses on the acceptability judgment task. Participant corrections to rejected statements were examined individually and any responses indicating a reason for rejecting the statement other than verb form were excluded. In other words, responses in which the participants modified word choice, word order, or any other part of the sentence beyond the verb forms was not considered.

Two dependent variables, social use and correctness, were considered in this analysis. Both were calculated numerically from responses on the acceptability task by assigning a value of -2 to responses of "*Estoy totalmente en desacuerdo*", 2 to responses of "*Estoy totalmente de acuerdo*", and -1 and 1 to "*Estoy más o menos en descuerdo/de acuero*", respectively. Social use was based on the response to "*Conozco a gente que lo diría así*" and correctness on the response to "*Está bien dicho*". The other question in this task, which targeted the participant's personal use of the form in question, was excluded as pairwise t-tests found that it did not significantly differ from responses to "*Está bien dicho*" (p = 0.56), while both of these questions elicited significantly different responses from the question targeting social use (p < 0.001).

In addition to social variables, the following linguistic variables were considered:

- Verb form in the statement protasis: *cantase*, *cantara*, or conditional

- Verb form in the statement apodosis: conditional or imperfect

- Possibility: whether or not the protasis statement is contrary-to-fact, (following Lavandera, 1975)

- Verb group: whether the verb under consideration is *ser*, *tener*, or a different verb[14]

An ordinal logistic regression with random effects for participant and question including these variables was run in R for each of the dependent variables. Descriptive results for each dependent variable are also found in Chapter 4.

D. *Written production*

Written production was studied based on responses on the fill-in-the-blanks activity, as well as participant corrections on the acceptability judgment task. The dependent variable in each case was whether the verb form produced in the protasis was *cantara* or *cantase;* for the purposes of this study, any responses other than one of these were excluded from consideration.

For the fill-in-the-blank analysis, the following were included as independent variables:

- Immediate prime: *cantase*, *cantara*, or none (if the verb was the first imperfect subjunctive used in the task)

- Recency of prime: the number of words separating the token from the previous instance of imperfect subjunctive, coded as short (3-6 words), medium (11-19

---

[14] Unfortunately, as this variable was added after the initial task was designed and administered, insufficient tokens of some forms existed to break this down further into *–ar, -er*, and *–ir* verbs as in the analysis of spontaneous speech.

words), long (21-25 words) or very long (more than 30 words)

- Prior task prime: which version of the matched-guise task the participant had been presented with

- Verb group: whether the verb in question was a form of *ser, tener*, or was a verb ending in *–ar*, *-er*, or *–ir*

Additionally, both individual and question were included in the analysis as random effects variables.

As mentioned previously, written production was also studied through the corrections produced by participants on the acceptability judgment task. Priming (*cantara*, *cantase*, or conditional), possibility (*irrealis* vs non-*irrealis*), and the verb group being considered (*ser, tener*, or other) were considered as independent variables along with the social variables described in the following section. A logistic regression with random effects for participant was run on this data in R. Though these two measures of written production could not be compared directly, the results of each are presented in Chapter 4 and connections between them are drawn in Chapter 5.

E. Social variables

In addition to the various linguistic variables described above, each statistical and descriptive analysis also included the following social variables identified through the Bilingual Language Profile task:

- Age: participant age at the time of the recording, coded as a continuous variable

- Gender: male or female

- Place of residence: Marín or O Grove

- Level of education: Primary, Secondary, Some undergraduate, Undergraduate,

Graduate[15]

- Initial language: Galician, Castilian, or both (Question 1 on the Bilingual Language Profile)

- School language experience: relative predominance of Galician or Castilian in academic courses, calculated as the difference in number of years spent studying Galician versus Castilian divided by the sum number of years of course work in each, with the result that equal use of both corresponds to a score of zero, while positive scores indicate more use of Galician and negative values indicate more Castilian use (Question 3 on the Bilingual Language Profile)

- Family language experience: relative predominance of Galician or Castilian use with family throughout the participant's lifetime, calculated as the difference in number of years speaking Galician versus Castilian in the family divided by their sum; a value of zero indicates equal use of both languages, while positive values indicate more Galician use and negative results correspond to greater Castilian use (Question 5 on the Bilingual Language Profile)

- Work language experience:[16] relative predominance of Galician or Castilian use throughout the participant's work life, calculated analogously to school and family language experience (Question 6 on the Bilingual Language Profile)

- Current language use: reported percentage of use of Galician in daily activities, calculated by averaging the percentage of Galician use reported in questions 7

---

[15] Reported only descriptively due to some blanks in participant responses

[16] These measures of language experience in school, family, and the workplace are treated as proxies for an analysis of language dominance. Though there are admittedly problems with approaching language dominance in this way, Unsworth (2015) provides compelling evidence that "when language dominance is narrowly defined as relative proficiency, the use of amount of exposure is a valid means of operationalizing language dominance" (p. 173).

through 11 on the Bilingual Language Profile

- Reported language competence: the participant's reported ability in Galician in each of the four modes (speaking, listening, reading, writing) as compared to their reported ability in Castilian (Questions 12 through 15 on the Bilingual Language Profile)[17]

- Language affect: the participant's relative strength of identification with Galician or Castilian, calculated analogously to school, family, and work language experience through averaging responses to Questions 16 through 18 on the Bilingual Language Profile[18]

As discussed previously in this chapter, a total of 39 individuals took part in this study. The distribution of these participants by location, gender, age, and education level is given in Table 3. Eleven of the participants were from the O Grove area, and the remaining 28 were from the area around Marín. Fifteen males and 24 females participated, and eighteen individuals were under 40 while the remaining 21 were over 40 at the time of the study. Participant ages ranged from 18 to 75. Of these participants, one male and two females from Marín were unable to complete all tasks due in two cases to unforeseen interruptions at the end of the group conversation and in one case due to poor vision that only made it possible to complete those tasks which could be done orally (*i.e.* the Bilingual Language Profile, the group conversation, and the matched-guise tasks). Evidently, then, data from these individuals is only included for those tasks which they were able to complete. Finally, three participants reported only having completed primary

---

[17] Because the vast majority of participants rated their abilities in both languages equally, these factors are not included in the statistical analyses. However, trends related to language competence are considered descriptively in Chapter 4.

[18] Reported only descriptively due to blanks in some participant responses

education, fifteen had completed secondary schooling, and six had pursued a graduate degree. In carrying out the statistical analyses the fourteen individuals who had some undergraduate education were divided into two groups based on whether they had completed an undergraduate degree (11 participants) or had not completed their degree (3 individuals).

1. Overview of language experience

Nearly two thirds of participants reported learning both Galician and Spanish simultaneously as their first language, with the other third nearly evenly split between Galician and Spanish as their first acquired language. Approximately the same trend held for both locations studied. Women, however, were more likely than men to report learning Castilian first and, conversely, men were more likely to have Galician as their first language, as can be seen in Figure 2. Additionally, the percentage of participants whose first language is Galician tended to decrease as level of education increased, and conversely an increase in education level also corresponded to an increase in the percentage of participants reporting Castilian as their first language. A pairwise t-test revealed no significant associations between participant age and first language ($p > 0.1$).

Figure 2: First Language by Location, Gender, and Education
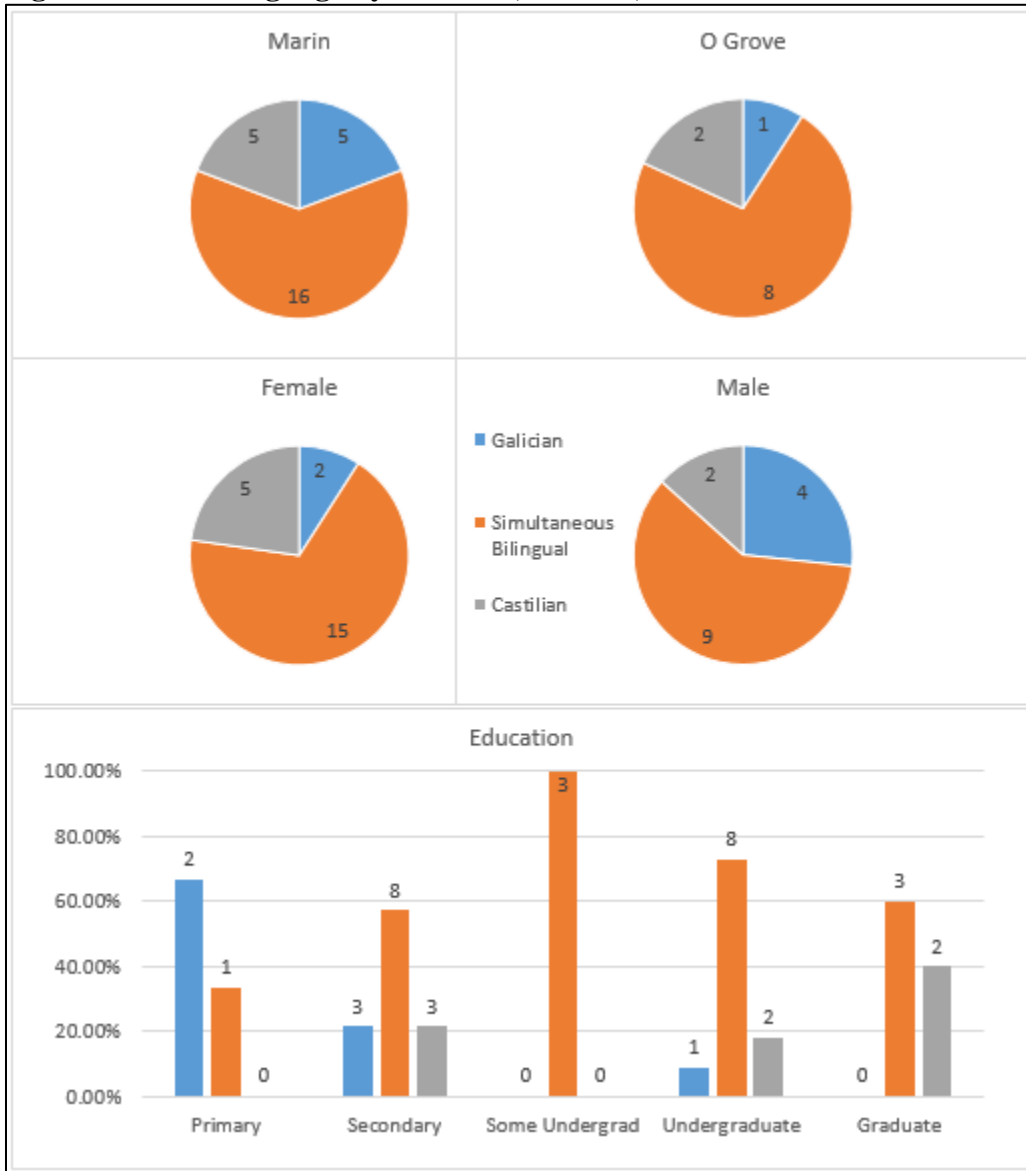
The second question on the Bilingual Language Profile asked about the age at which participants first felt comfortable using each language. A linear regression using age as a numerical dependent variable revealed no significant differences between age of exposure to Castilian and age at which speakers felt comfortable using the language ($R^2 = 0.01$, $F(1,70) = 0.47$, $p = 0.50$) and a difference between the two for Galician that approached significance ($R^2 = 0.06$, $F(1,68) = 5.17$, $p = 0.03$). Because of this only slight difference, the predominance of those who acquired and felt comfortable in both

languages simultaneously, and the very similar nature of the first two questions, first language as revealed in these questions was coded as a categorical variable and the precise ages given were not taken into account.

Questions 3, 5, and 6 dealt with the number of years participants had been exposed to Galician and Castilian in academic coursework, in their families, and in a work environment, respectively. A linear regression on these three categories revealed that responses to these three questions were significantly different for Galician ($R^2 =$ 0.12, $F(2,104) = 8.49$, $p <0.001$) and approached significance for Castilian ($R^2 = 0.05$, $F(2, 107) = 3.78$, $p = 0.03$). A post-hoc Tukey test revealed that the number of years of coursework lagged behind the other two variables for both languages and that the greatest exposure for both languages was in the family. Additionally, a paired t-test revealed that differences between the two languages on these three questions approached significance ($t(36) = 2.49$, $p = 0.02$). Because of this, it was considered important to maintain the information contained in all parts of these three questions. However, to reduce the proliferation of dependent variables, differentials between the two languages were calculated for each question by subtracting the number of years of experience with Castilian from experience with Galician and dividing that difference by the total of both such that a more positive differential indicated greater experience with Galician relative to Castilian. Handling the variables in this way allowed for relative amount of experience to be compared across individuals regardless of their age.

With the variables considered in this way, the average differential among the sample as a whole was -0.23 for language use in education (SD = 0.42), -0.04 for language use in the home (SD= 0.44), and -0.04 for language use at work (SD = 0.31).

These values indicate that the group as a whole has greater experience with Castilian in the classroom and slightly more experience with the same language at home and at work, though in all three cases a balanced history with the two languages is within one standard deviation of the mean.

Location ($t(18) = 1.01$, $p = 0.29$), gender ($t(29) = 1.05$, $p = 0.30$), and education level ($R^2 = 0.01$, $F(4,30) = 1.12$, $p = 0.37$) were not significantly correlated with Galician experience in education, though participant age approached significance as a predictor of this variable ($R^2 = 0.14$, $F(1,34) = 6.57$, $p = 0.01$), with younger participants reporting greater relative school experience in Galician as compared to Castilian (though only in rare cases did the amount of schooling in Galician exceed that in Castilian). Of these same variables, only gender approached significance as a predictor of language experience in the family ($t(35) = -2.06$, $p = 0.05$), as men on average reported greater family use of Galician than did women (means of 0.13 (SD = 0.34) and -0.15 (SD = 0.46), respectively). None of these four variables correlated significantly with language experience in the workplace ($p > 0.1$ in all cases).

Question 4 on the Bilingual Language Profile inquired about the number of years participants had spent living in a country or region where Galician/Castilian is spoken. As mentioned in the previous chapter, all participants had been born in Galicia save four, who had been born to Galician parents and moved back to the region in early childhood. Because of this, responses to Question 4 were categorically at maximum and were therefore not considered in further analysis.

2. Language use

Questions 7 through 11 on the Bilingual Language Profile inquired about the percentage of time that participants use each language with their friends and family, at work or school, when talking to themselves, and when counting. Because the percentages added up to 100%, the percentage of time spent in Galician was sufficient to give information about both languages. A linear regression with question number as the dependent variable found no significant differences between any of these questions ($R^2 = 0.01$, $F(4,180) = 0.61$, $p = 0.66$). Thus responses to these five variables were averaged into one combined language use score. The mean for this score among the sample was 0.39, indicating an average use of Galician 39% of the time, with a standard deviation of 0.30, which means that 65% of the sample reported using Galician between 9% and 69% of the time.

While no significant differences were found between the two locations studied ($t(14) = -1.03$), $p = 0.32$), age ($R^2 = 0.10$, $F(1,35) = 5.02$, $p = 0.03$), gender ($t(22) = 2.27$, $p = 0.03$), and education ($R^2 = 0.10$, $F(1,34) = 4.77$, $p = 0.04$) approached significance as predictors of language use. Specifically, use of Galician decreased from an average high of 80% among those with only primary education to an average low of 22% among those with graduate studies. Males averaged 53% reported use of Galician, while females reported using the language only 30% of the time. Reported use of Galician is decreasing in apparent time among those included in this study, as evidence by its far greater presence in the daily habits of older participants shown in Figure 3.

Figure 3: Reported Galician Use by Age
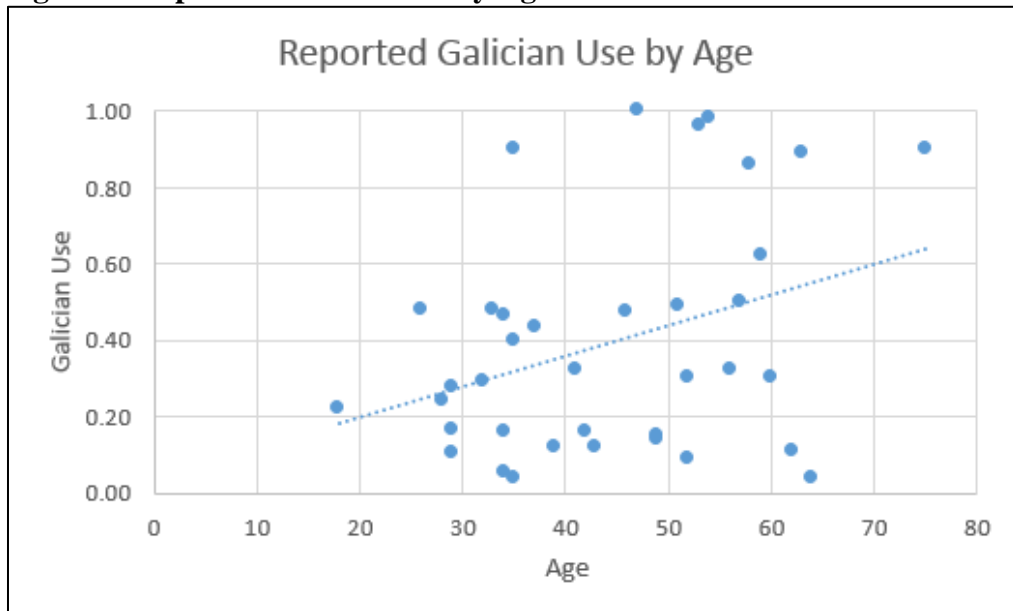
3. Language competence

Questions 12 through 15 focused on participants' reported competence in speaking, understanding, reading, and writing in Galician and Castilian. A series of paired t-tests revealed that participants on average reported greater competence in Castilian across all four modalities, that these differences were significant for reading ($t(34) = 3.22$, $p < 0.01$) and for writing ($t(34) = 4.81$, $p < 0.001$), and that the differences approached significance for speaking ($t(33) = 2.39$, $p = 0.02$) and for listening ($t(34) = 2.71$, $p = 0.01$). To maintain information about these differences in the statistical analysis while reducing the number of needed variables, participants' Galician abilities as relative to their skills in Castilian were determined by calculating the differentials between the two languages for speaking, reading, and writing using a formula analogous to that used for Questions 3, 5, and 6 as described previously. However, as only seven participants reported different levels of listening comprehension between the two languages, values for this variable were coded as either balanced or Castilian-dominant.

Average values indicated greater competence in Castilian among the sample for all four modalities, though the divide in competences was greater for reading (M = -0.10, SD = 0.20) and writing (M = -0.16, SD = 0.23) than for speaking (M = -0.04, SD = 0.32) or listening (M = -0.02, SD = 0.06 before conversion to a categorical variable).

Neither location (t(14) = -0.41, p = 0.62) nor education level (R^2 = 0.04, F(1,32) = 2.21, p = 0.15) were significant predictors of reported speaking competence, while gender approached significance (t(17) = 2.39, p = 0.03), with men on average reporting greater competence in Galician relative to Castilian (M = 0.11, SD = 0.40) than women (M = -0.16, SD = 0.15). Age was a significant predictor of reported speaking ability (R^2 = 0.17, F(1,33) = 8.02, p < 0.01), with greater age corresponding to greater relative ability in Galician. However, as can be seen in Figure 4, this difference is due almost entirely to the existence of three outliers whose reported speaking competence in Castilian was exceptionally low.

Figure 4: Speaking competence differential between Galician and Castilian by Age
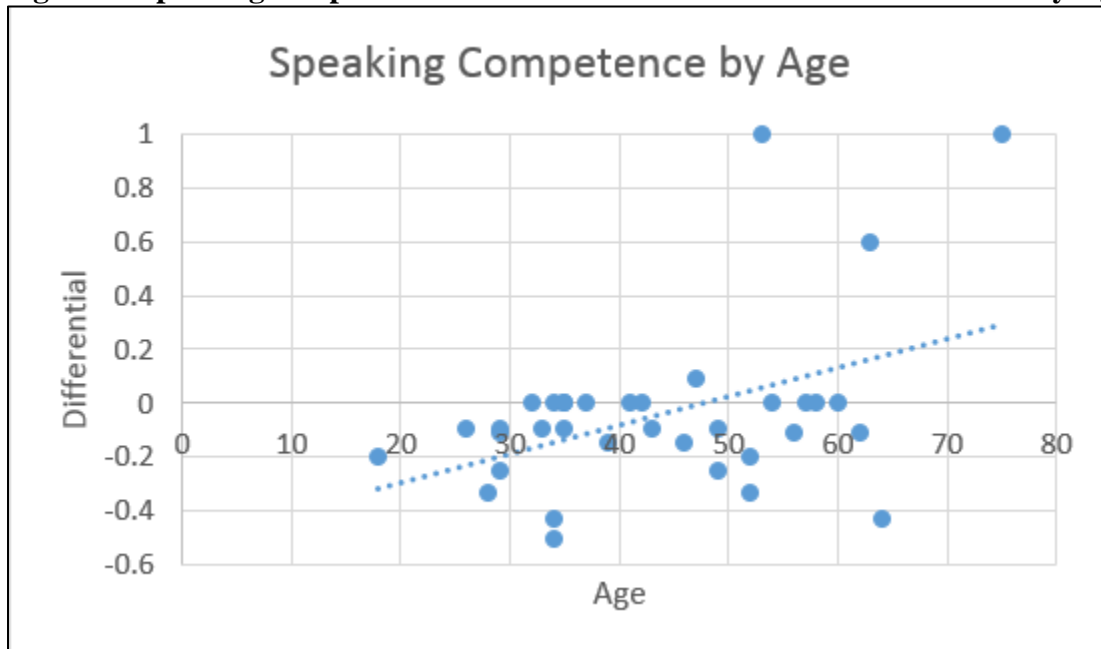
As mentioned previously, only seven individuals reported a non-balanced listening competence, and all of these individuals reported greater abilities in Castilian than in Galician. No patterns were found with respect to gender or location, with precisely 20% of the individuals in each category reporting Castilian dominance in their listening abilities for both variables (*i.e*. gender and location). Similarly, these speakers were distributed throughout all education levels save those who had only completed primary education. A pairwise t-test confirmed that this absence of an observable pattern also held with respect to age ($p = 0.6$). Similarly, neither age ($R^2 = 0.05$, $F(1,33) = 2.94$, $p = 0.10$), gender ($t(20) = -1.30$, $p = 0.21$), education level ($R^2 = 0.03$, $F(1,32) = 1.86$, $p = 0.18$), nor location ($t(32) = -1.31$, $p = 0.20$) were significant predictors of reading competence.

Finally, only age approached significance as a predictor of reported writing ability ($R^2 = 0.12$, $F(1,33) = 5.45$, $p = 0.03$), with written skills in Galician relative to Castilian decreasing among older participants as shown in Figure 5. For each of the other three variables, p-values were greater than 0.1.

Figure 5: Writing competence differential between Galician and Castilian by Age
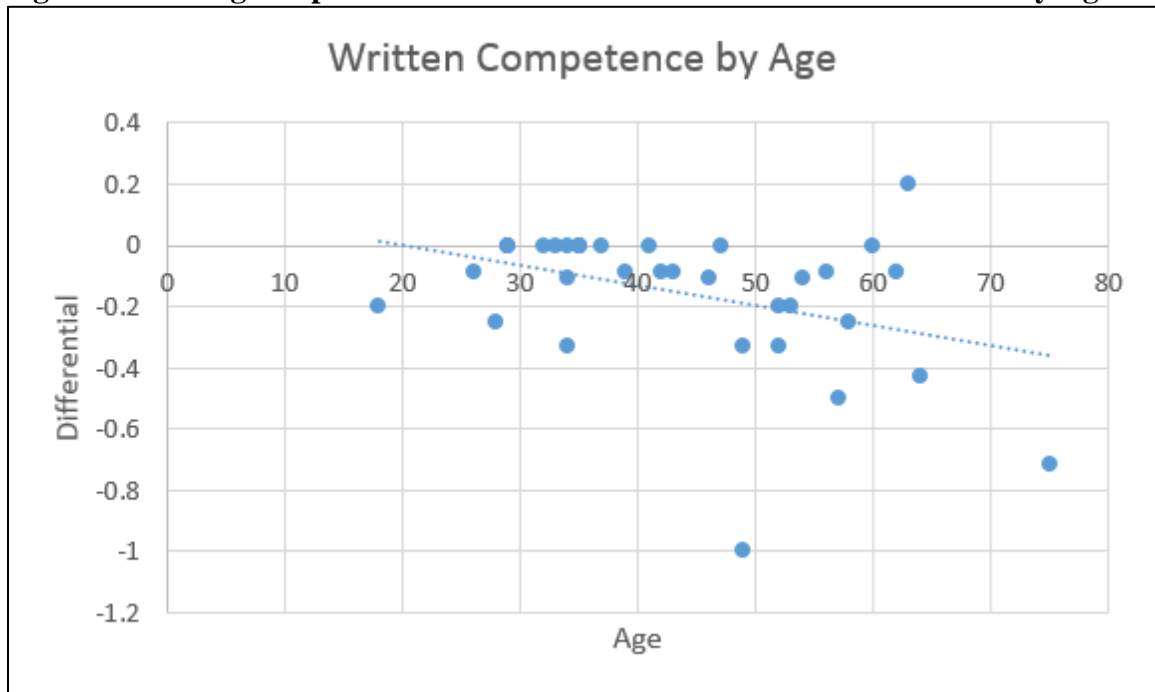
4. Language attitudes

Questions 16 through 19 solicit information about participants' attitudes to each of the region's languages. Of these, Question 19, which inquired about a desire for others to consider the person a native speaker of each language, was omitted. The choice to exclude this question was made because many participants responded disfavorably toward both languages, not because they have negative attitudes toward the languages themselves, but because they indicated not caring what other speakers think of their language use. A linear regression and subsequent post-hoc Tukey test on these responses confirmed that Question 19 was indeed answered differently than the other three questions related to attitudes, particularly for Spanish ($R^2 = 0.06$, $F(3,132) = 3.79$, p = 0.01). Additionally, Question 18 addressed the same issue in a different way by asking whether the individual him/herself wants to speak like a native speaker. Because

responses to this earlier question did not show the same contradictory trends, those responses were the ones included in the present analysis.

As compared to Question 18, neither Question 16, which asked about feeling "like oneself" while speaking each language, nor Question 17, dealing with a feeling of personal identification with each culture, were significantly different for either Castilian or Galician ($p > 0.6$ in all cases). Because of this, responses to these three questions were averaged to give an overall language attitudes score for each language. Additionally, a paired t-test found no significant differences between the two languages on this ($t(33) = -0.40$, $p = 0.7$). Therefore, it should statistically have been sufficient to include the data from only one of the two languages. However, because the question of potential attitudinal differences toward each language is still of practical and theoretical interest, and in order to parallel the decisions made for other social variables, the differential between the two languages was calculated for attitudes as well. Thus only one attitudinal variable is included in subsequent analyses, with positive values indicating more favorable attitudes toward Galician and negative values indicating more favorable attitudes toward Castilian.

Over the entire data set, the mean of the attitudinal variable was 0.02, indicating a very slight overall preference for Galician, with a standard deviation of 0.3. No significant correlations were found between attitude and gender, location, age, or education ($p > 0.16$ in all cases).

Chapter 4: Data Analysis

In this chapter the data obtained as described in Chapter 3 is presented and examined from various perspectives, including both tests for significance and descriptive statistics. The results of each task are considered individually in the order in which they were presented to participants. Broader connections between the results of each task and their significance in the Galician context are drawn in Chapter 5, with specific focus on responding to the research questions which guided this study.

I.   Semi-structured group conversations

As described in Chapter 3, two separate regression models were created to examine spontaneous spoken production. The first model, a logistic regression with random effect for individual, had form produced as the dependent variable and examined linguistic variables associated with each produced imperfect subjunctive form. In total, 130 tokens were extracted from the recordings of the group conversations. These were combined into 30 data points for the examination of social variables, with one data point for each participant who produced an imperfect subjunctive form in their spontaneous speech. These data were fit to a linear regression model with percentage of *cantase* production as the dependent variable.

A. *Linguistic variables associated with oral production*

Of the five linguistic variables included in the statistical model, only Priming approached significance ($F(2,124) = 1.91$, $p = 0.03$). A post-hoc Tukey test revealed that the presence of a *cantase* prime, when compared to the presence of either a *cantara* prime or an unprimed token, significantly conditioned the production of a subsequent *cantase*

form (p < 0.001 in both cases), while *cantara* primes did not correspond to significantly different productions as compared to unprimed tokens (p = 0.07). Despite the lack of significance, however, Figure 6 illustrates that, while *cantara* is the dominant form produced in both contexts, lack of a prime results in roughly 80% *cantara* production, while the presence of a *cantara* prime produces a subsequent *cantara* form at a rate of 95%. Corroborating the statistical results, a *cantase* prime yields a *cantara* token only 25% of the time.[19]

Figure 6: Spoken form production by Priming
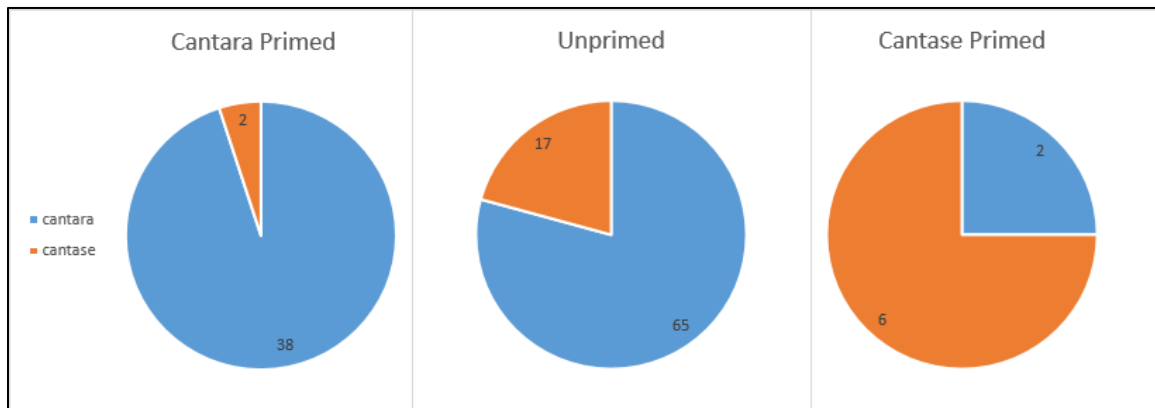
Table 5 below contains the rates of production of each form for the linguistic variables Possibility, Polarity, and Anteriority. Rates of production of *cantase* ranged from 15% to 23% in all cases, and no strong tendencies related to these three independent variables were revealed, though it is difficult to say whether the inclusion of more tokens would reveal more decisive patterns. The strongest association noted is that between *cantara* and anteriority, though again, that association is far from overwhelming.

---

[19] The inclusion of only 8 *cantase*-primed tokens makes it natural to question whether these forms were produced by individuals who categorically produced the *cantase* form. It was subsequently verified that only 2 of the 8 tokens were produced by speakers who never produced *cantara*; thus even without these two tokens the rate of production of *cantase* when primed by the same form is well above the rate when unprimed or when primed by *cantara*.

Table 5: Spoken form production by Possibility, Polarity, and Anteriority

	Possibility				Polarity				Anteriority			
	Contrary		Not Contrary		Negated		Not Negated		Anterior		Not Anterior	
	%	N	%	N	%	N	%	N	%	N	%	N
cantase	0.20	12	0.18	12	0.23	5	0.19	20	0.15	3	0.20	22
cantara	0.80	48	0.82	53	0.77	17	0.81	88	0.85	17	0.80	86

An examination of clause type reveals that adjective clauses and prepositional clauses show almost exclusive *cantara* use in the present data set, while hypothetical statements and noun clauses feature *cantase* roughly a quarter of the time. Adverbial clauses fall in the middle, with 15% containing *cantase*. However, low token numbers over all, particularly for adjective and prepositional clauses, make it necessary to consider these results, shown in Table 6, tentative at best.

Table 6: Spoken form production by Clause Type

	Adjective		Adverbial		Hypothetical		Noun		Preposition	
	%	N	%	N	%	N	%	N	%	N
Cantase	0.08	1	0.15	5	0.24	7	0.24	12	0.00	0
Cantara	0.92	12	0.85	29	0.76	22	0.76	39	1.00	3

Finally, Table 7 reveals that verbs ending in –*ar* and –*ir* were rendered in the *cantase* form nearly a quarter of the time, while the various –*er* verbs, including *ser* and *tener*, were only found in this form at rates of between 13% and 19%.[20]

---

[20] Because the number of *ser* tokens exceeded the number of other –*er* verbs and the number of *tener* tokens was similarly large, these two lexical items were analyzed separately. This also parallels the way verb group was considered in the two written tasks, thereby facilitating Chapter 5's comparison of results across tasks.

Table 7: Spoken form production by Verb Group

	-ar		-er		-ir		ser		tener	
	%	N	%	N	%	N	%	N	%	N
Cantase	0.23	10	0.13	3	0.23	3	0.19	6	0.17	3
Cantara	0.77	34	0.87	21	0.77	10	0.81	25	0.83	15

B. *Social variables associated with oral production*

A linear regression examining the relationship between the rate of production of *cantase* and social variables failed to return any of these variables as significant (p = 0.45). Visually, however, Figure 7 and Figure 8 reveal definite trends with respect to Age and Education. Specifically, no speaker older than age 51 produced any *cantase* forms in their spoken language, while categorical use of *cantara* was spread across the age spectrum. In a possibly related trend, speakers with some higher education experience produced a far greater percentage of *cantase* forms on average than those with only primary or secondary studies. Meanwhile women (M = 0.33, SD = 0.39) and those from O Grove (M = 0.45, SD = 0.38) produced *cantase* at a higher rate than men (M = 0.19, SD = 0.36) or those from Marín (M = 0.21, SD = 0.36).

Figure 7: Spoken form production by Age
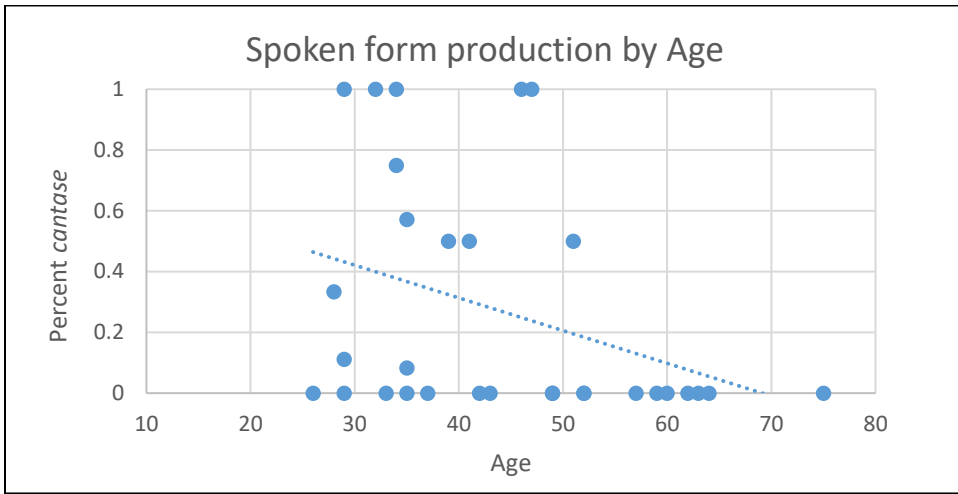
Figure 8: Spoken form production by Education
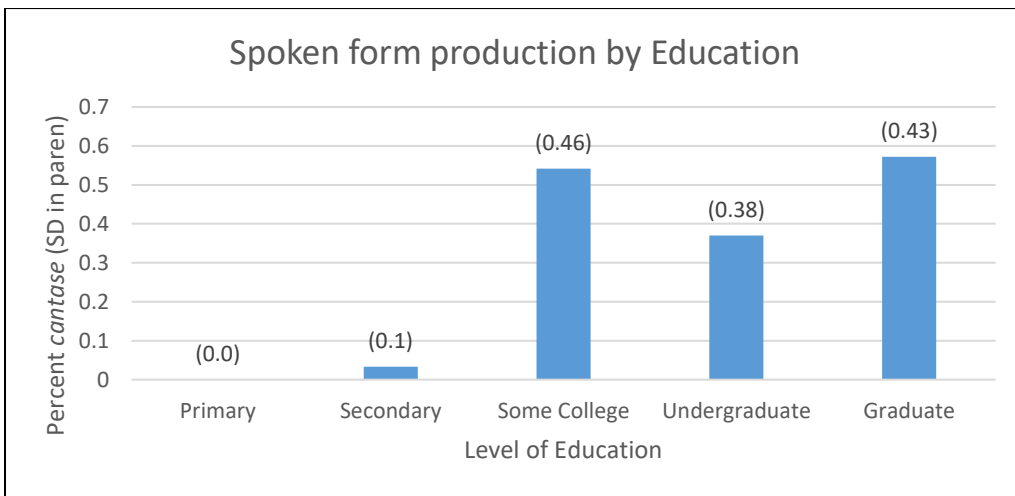
The increasing trend lines in Figure 9 and Figure 10, which would indicate greater production of *cantase* in correlation with increased use of Galician at school and in the home, should be interpreted cautiously due to the heavy concentration of speakers in the center of the language use spectrum and the relative paucity of data points on the extremes. Thus the only statements that can be made with any certainty are that elevated rates of *cantase* use, including categorical productions, were concentrated near the middle of the School Language spectrum, that they were found across the Family

Language spectrum, and that nearly categorical *cantara* use was present for some speakers at all levels.

Figure 9: Spoken form production by School Language

Spoken form production by School Language
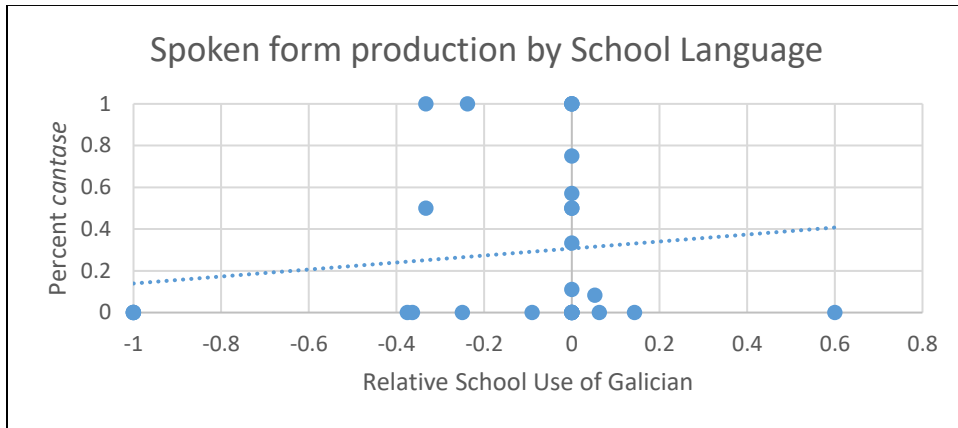
Figure 10: Spoken form production by Family Language
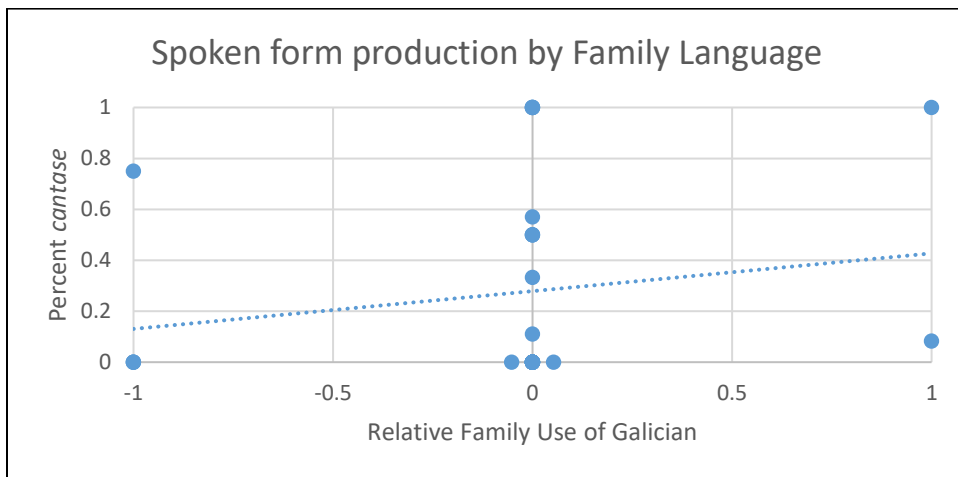
Figure 11 shows that Work Language experience patterns similarly to School Language, with uses of *cantase* limited to those whose work language experience has reportedly been nearly balanced. This trend is also present with respect to initial language: simultaneous bilinguals produced *cantase* on average around 32% of the time (SD = 0.38), while those whose first language was either Galician (M = 0.2, SD = 0.4) or Castilian (M = 0.19, SD = 0.35) used this form less frequently.

Figure 11: Spoken form production by Work Language

Figure 11: Spoken form production by Work Language

Figure 12: Spoken form production by Current Language Use
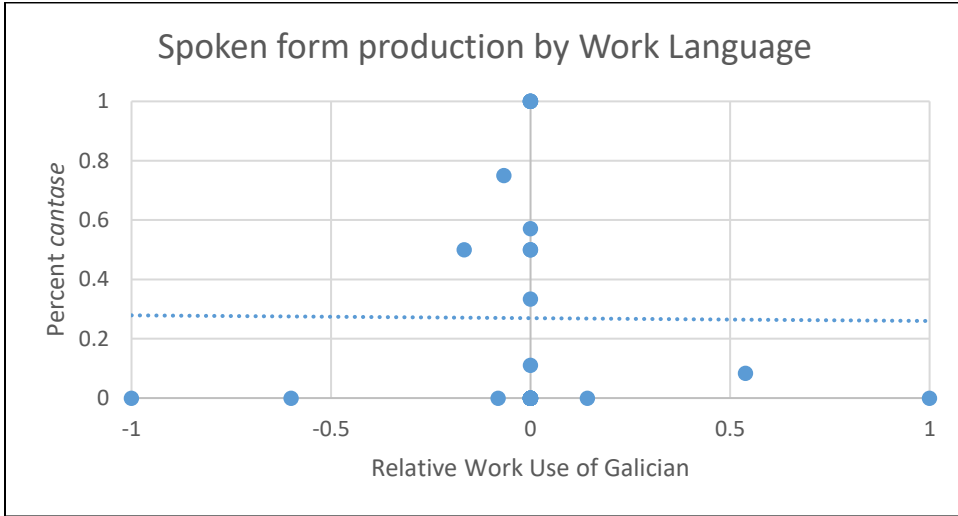
Figure 12: Spoken form production by Current Language Use
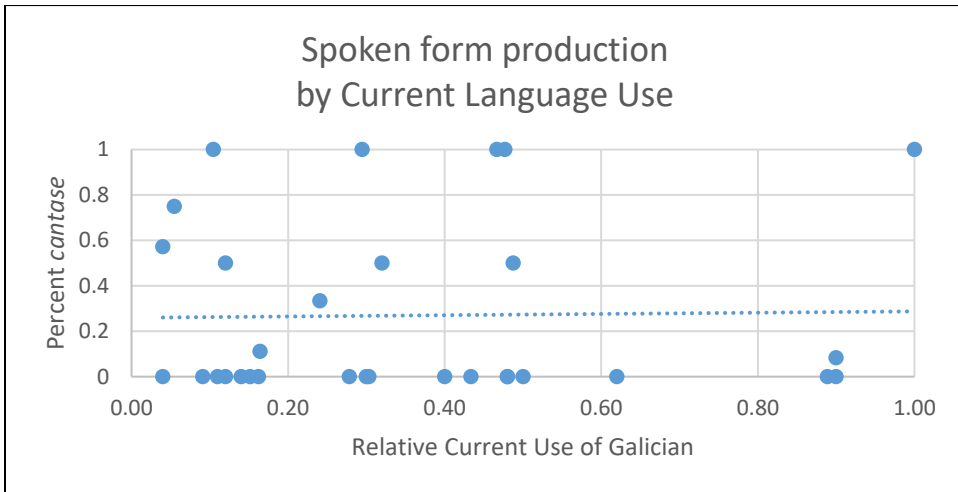
Current language use, shown in Figure 12, reveals no correlation with *cantase* production. Unlike with other variables, this result seems reasonably reliable as data exists across the whole spectrum of current use and categorical uses of both *cantara* and *cantase* are not clustered in one area. Form variation, on the other hand, seems limited primarily to those who use predominantly Castilian.

Figure 13: Spoken form production by Oral Ability

Spoken form production by Oral Ability
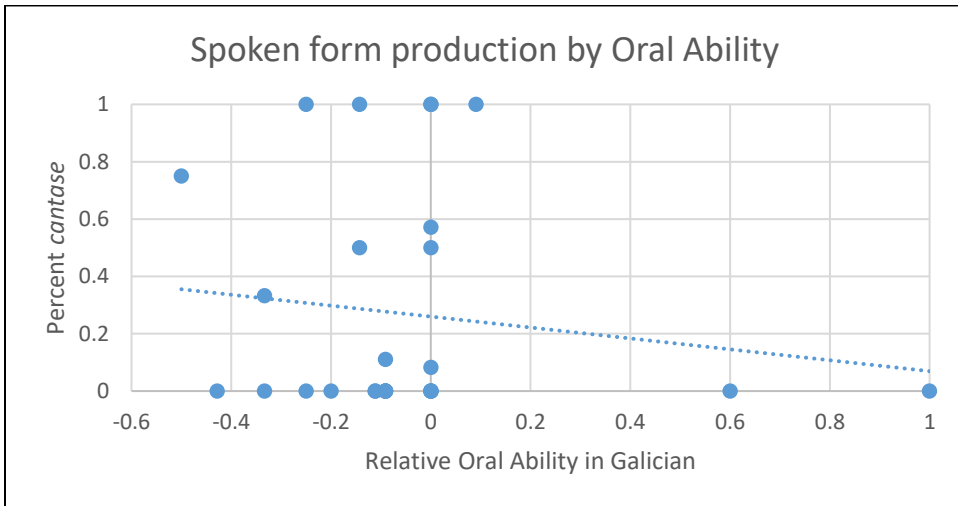
The two speakers who rated themselves as having far greater oral competence in Galician than in Castilian produced no *cantase* tokens; this is likely the reason for the apparent decreasing trend of *cantase* use as Galician spoken competence increases (Figure 13). Across the range in which most data points are contained, however, no pattern emerges. Similarly, with the exception of the one potential outlier in the top right of Figure 14, practically all *cantase* usage, as well as the vast majority of the data points, are concentrated near the point of balanced attitudes toward Galician and Castilian.

Figure 14: Spoken form production by Language Attitudes

Spoken form production by Language Attitudes
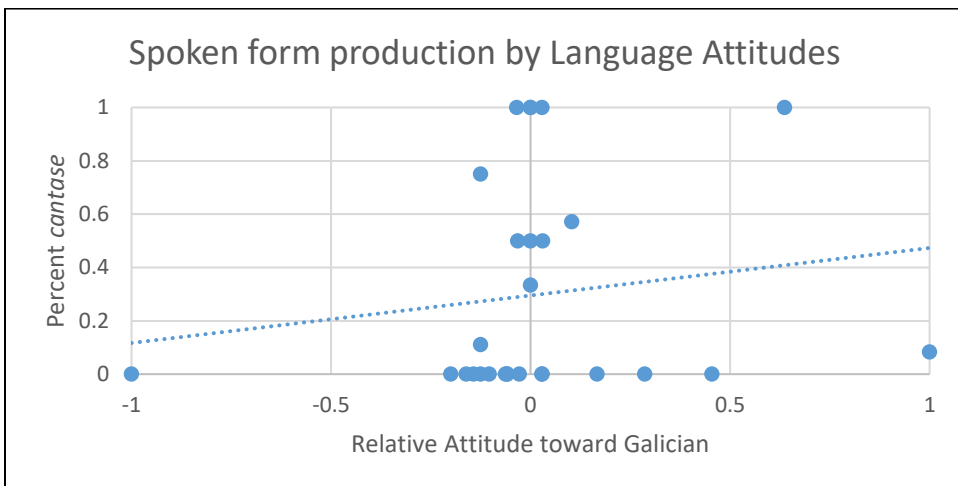
As has been indicated repeatedly throughout this section and in particular with respect to social variables, the results just described herein should be considered tentative at best, due to limitations of the data set. What inferences can safely be drawn are discussed in Chapter 5.

II.   Matched-guise task

This task looked at participant evaluations of a speaker on 18 different characteristics. Two of these characteristics—pride and similarity to the rater—were omitted from the statistical analyses. A factor analysis was run on the remaining sixteen characteristics, which were divided into three groups that can roughly be characterized as based on friendship, leadership/charisma, and capability.[21] The factor loadings and resultant groupings from this analysis can be seen in Table 8.

---

[21] It should be understood that from this point forward references to Friendliness, Charisma, and Capability (or to the adjectives friendly, charismatic, and capable) are meant to refer to the set of characteristics identified by these group headings.

Table 8: Matched-guise characteristics factor analysis: Loadings and communalities[22]

	Friendship	Leadership/Charisma	Capability	Communality
Open		0.30	**0.59**	1.7
Attractive		**0.57**		1.2
Confident		**0.62**	0.42	2.1
Refined			**0.65**	1.4
Fun	**0.92**			1.1
Loyal	**0.44**			2.2
Generous	**0.62**	0.32		1.6
Sense of Humor	**0.67**		0.32	1.5
Intelligent			**0.49**	2.0
Leadership		**0.73**		1.5
Openminded		**0.65**		1.6
Practical	0.44	0.33	**0.54**	2.7
Ambitious			**0.88**	1.1
Nice	**0.75**			1.5
Hardworking	0.58		**0.61**	2.4
Trustworthy	**0.46**		0.31	2.1

Using these groupings as an independent variable, a linear regression with random effects for participant and characteristic was run in R. Of particular interest in this analysis were possible interaction effects between the guise presented to participants and other variables; however, no interaction terms were returned as significant. The only factor which significantly correlated with rating was characteristic grouping ($F(2,583) =$

---

[22] Note: Factor loadings less than 0.3 are suppressed.

10.76, p < 0.01); a post-hoc Tukey test indicated that Capability (M = 0.57, SD = 0.78) differed significantly from both Friendship (M = -0.20, SD = 0.86) and Charisma (M = -0.18, SD = 0.96), which did not differ from each other, as can be seen in Figure 15. The confidence interval for the difference between the Charisma grouping and the Friendship grouping includes zero, indicating no significant difference between the two, while neither of the other two intervals includes zero, corresponding to a significant difference between the means of the factor groups involved. The mean and standard deviation associated with each characteristic group for each guise are presented in Table 9.

Figure 15: 95% confidence interval for difference between characteristic groups
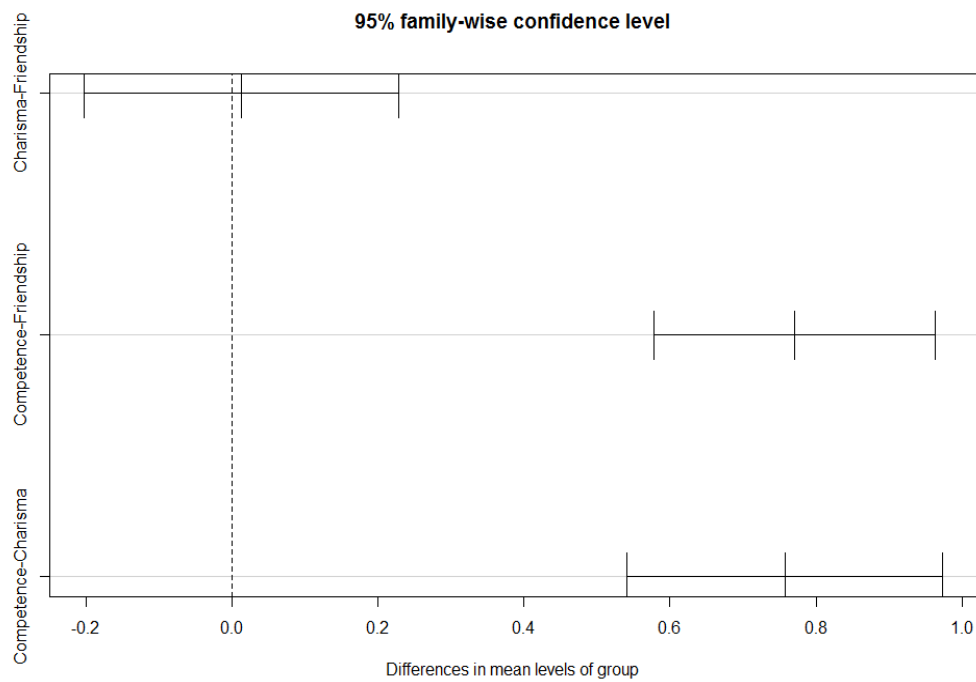
Table 9: Characteristic group means by Guise

	cantase		*cantara*	
	Mean	SD	Mean	SD
Friendliness	-0.19	0.85	-0.20	0.87
Charisma	-0.12	0.93	-0.24	0.99
Capability	0.65	0.74	0.50	0.80

Despite the lack of statistical significance of the other variables considered, the small sample size of this study (N = 586 responses for the matched-guise task) makes it reasonable to consider the possibility that differences might exist that would reach significance were the sample size larger. For this reason, the following subsections provide descriptive results of the social variables examined in the matched-guise task. [23]

A. *Matched-guise results: Gender, Location, Age, and Education*

The use of z-scores as described in Chapter 3, while beneficial in removing extraneous variation due simply to individual bias, makes it unfruitful to compare overall averages across any variable as the process of standardizing scores reduces the overall mean to (near) zero. However, patterns in the distribution of ratings among the three characteristic groupings may still show interesting trends. Average values for each characteristic group for Gender and Location are shown in Table 10. Without exception, participants rated the speaker above average with respect to Capability and below average for Charisma and Friendship. Particularly favoring a high estimation of Capability were women and those from Marín; these same groups also rated the speaker more highly for

---

[23] From this point forward it should be noted than any results not accompanied by a p-value failed to reach significance and thus are descriptive only.

Charisma. Interestingly, however, these groups gave the lowest average ratings for the Friendship grouping.

Table 10: Characteristic group means by Gender and Location

| | Gender | | | | Location | | | |
| | Female | | Male | | Marín | | O Grove | |
	M	SD	M	SD	M	SD	M	SD
Friendship	-0.26	0.90	-0.10	0.80	-0.24	0.85	-0.09	0.88
Charisma	-0.16	0.93	-0.22	1.02	-0.16	0.95	-0.24	1.00
Capability	0.65	0.71	0.46	0.85	0.60	0.76	0.52	0.80

While these general trends are useful to keep in mind, they tell us little about the actual object of study, which in this case is attitudes toward the use of *cantara* and *cantase* forms. To approach this issue, then, it is necessary to look at each of the independent variables not in isolation but as a source of possible interaction with the guise presented. Thus for example the effect of Gender shown in Table 10 can be compared to the joint effects of Gender and Guise in Figure 16. The same directional trend for Capability, with the *cantase* guise seen as more capable than *cantara*, holds for both males and females. However, while men rated both guises equally in terms of Charisma, females saw the *cantase* guise as the more charismatic of the two. Finally, opposing views were found between the genderes as to the Friendliness of the two guises, with men seeing *cantase* as considerably more friendly than *cantara* while women saw it as slightly less so.

Differences in guise evaluation between participants from Marín and O Grove, shown in Figure 17, were minor for perceptions of both Friendliness and Charisma, though the perceived imbalance in Charisma between the two guises was more

pronounced in O Grove. The greatest disagreement lies in their assessment of each guise's Capability, with those from Marín rating the *cantase* guise as the more capable and those from O Grove giving that distinction to the *cantara* guise.

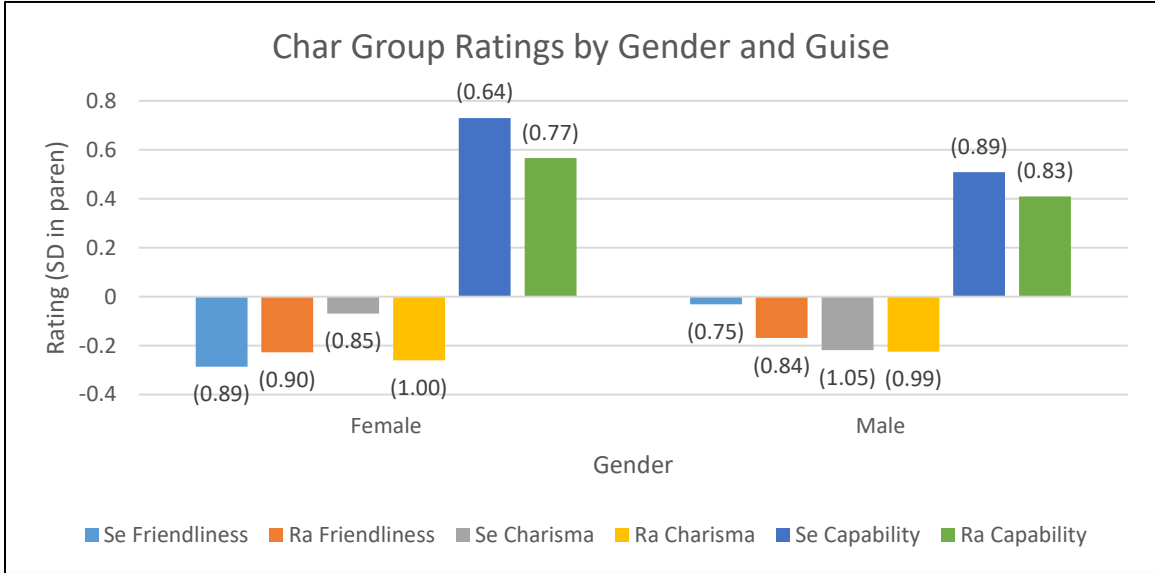**Figure 16: Characteristic group ratings-Interaction between Gender and Guise**

Figure 17: Characteristic group ratings--Interaction between Location and Guise
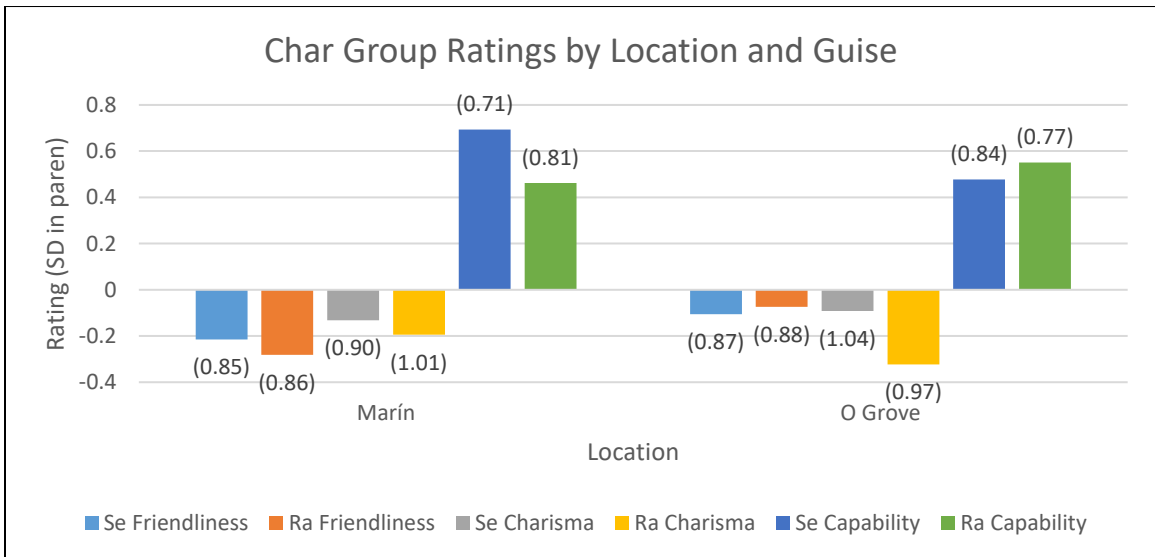
In terms of numeric variables, neither Age nor Education showed any notable patterns other than the aforementioned tendency for Capability to be rated far higher than either of the other two trait groups, as can be seen in Figure 18 and Figure 19.

Figure 18: Characteristic group ratings by Age
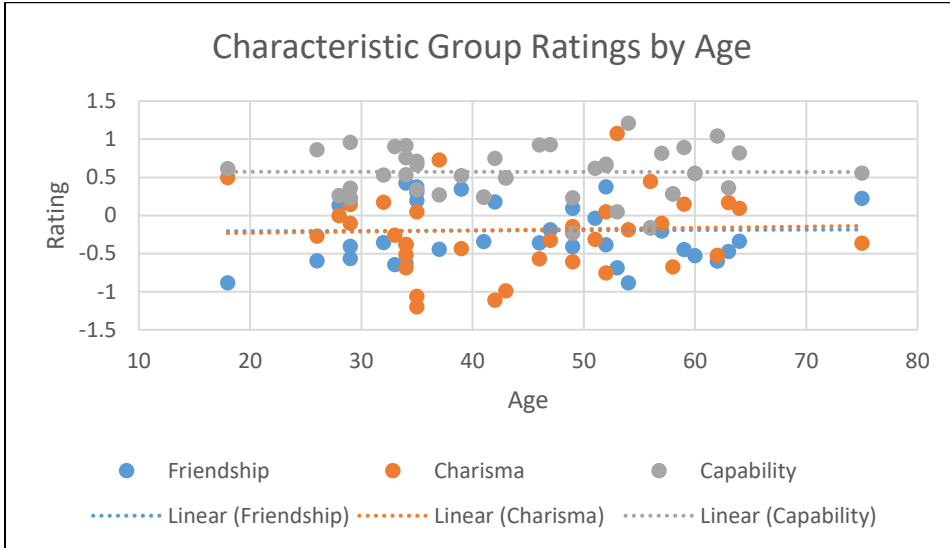
Figure 19: Characteristic group ratings by Education
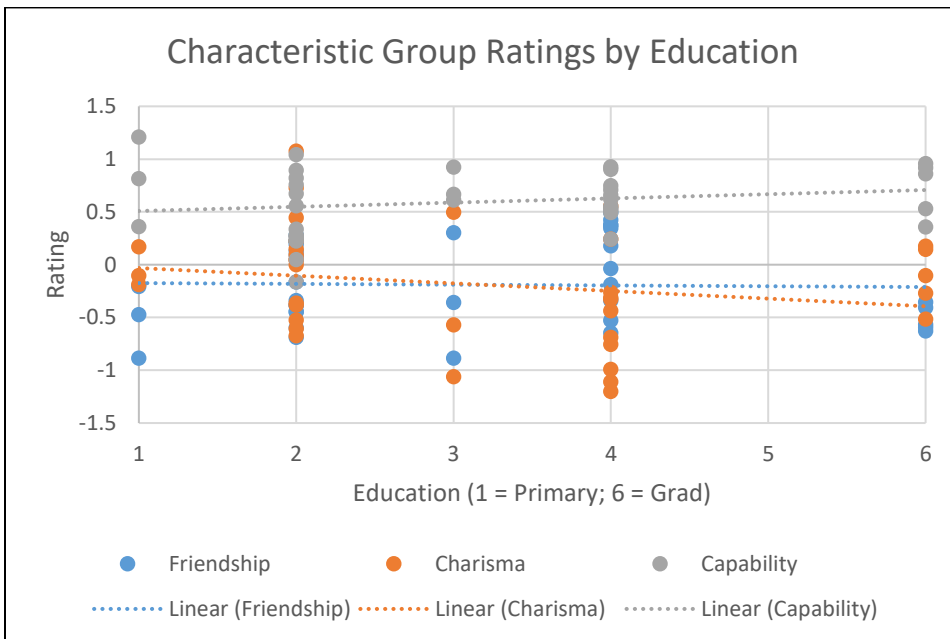
However, in terms of interaction between Age and Guise, *cantase* was seen by older participants as both more capable and less friendly than by younger individuals,

who reported the opposite impressions as can be seen by examining the trend lines in Figure 20. Interestingly, there appears to be a leveling out of perceptions of Charisma in apparent time, as older individuals rated the *cantase* guise as more charismatic, while this distinction was neutralized for the youngest participants.

Figure 20: Characteristic group ratings--Interaction between Age and Guise
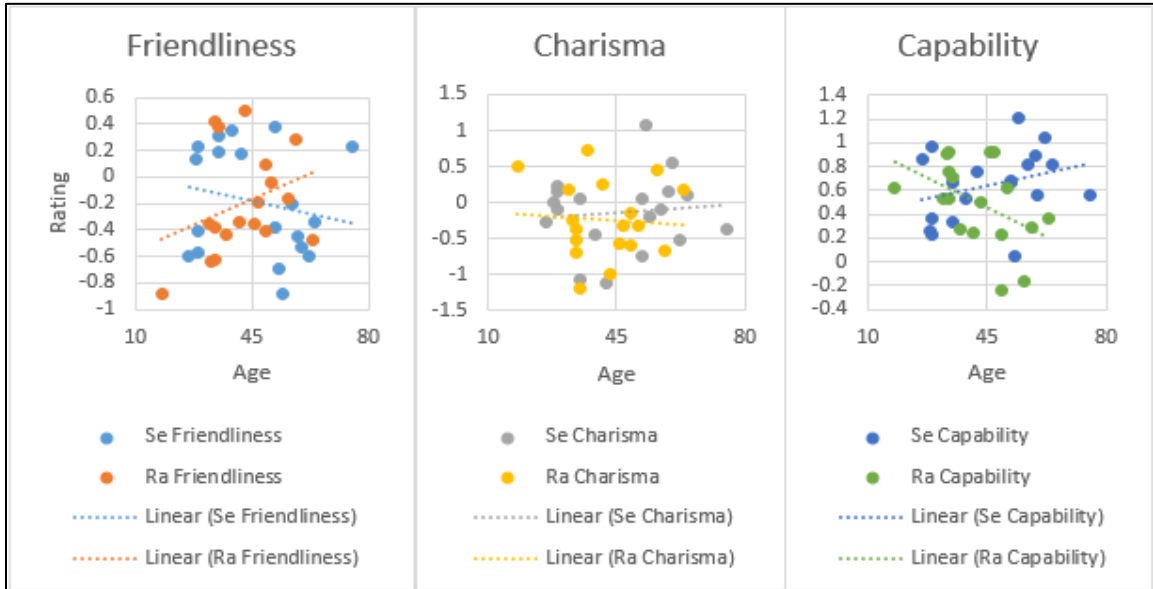
Education level does not appear to have a strong effect on the appraisal of either guise; Figure 21 illustrates that approximately the same patterns are followed by both guises for all three characteristic groups.

Figure 21: Characteristic group ratings--Interaction between Education and Guise

Char Group Ratings by Education and Guise
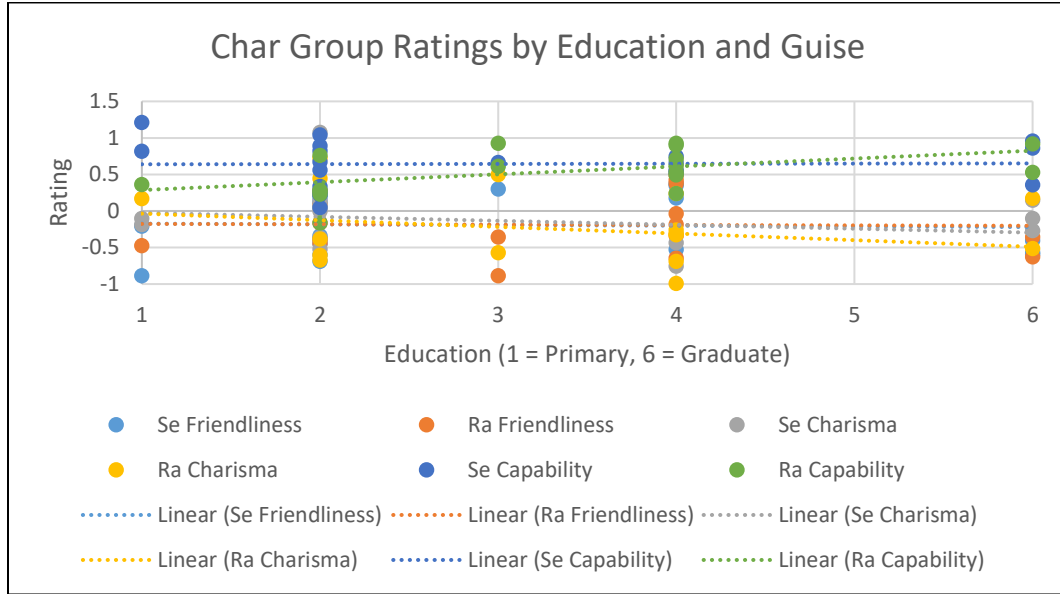
B. Matched-guise results: Initial, School, Family, and Work Language

With respect to initial language, those who first learned Galician rated the speaker as less friendly overall than those who were simultaneous bilinguals or who had Castilian as their first language. Conversely, these same L1 Galician individuals found the speaker more charismatic and more capable than did the other two groups, which did not differ greatly from one another, as shown in Table 11.

Table 11: Characteristic group ratings by Initial Language

	Initial Language					
	Galician		Simultaneous Bilingual		Castilian	
	M	SD	M	SD	M	SD
Friendship	-0.50	0.21	-0.13	0.40	-0.16	0.37
Charisma	0.00	0.41	-0.25	0.53	-0.15	0.48
Capability	0.79	0.35	0.53	0.32	0.54	0.23

Figure 22 reveals some interesting patterns of evaluation based on the participant's first language. Those whose first language was either solely Castilian or solely Galician found the *cantara* guise to be more friendly than the *cantase* guise, while those who were bilingual from the start found *cantase* to be the more friendly of the two, though this unexpected pattern could potentially be due to skewed data, as more individuals fell into the bilingual group (N = 24) than into the Castilian (N = 7) or Galician (N = 6). Less surprising is the fact that those whose first language was Galician considered *cantara* to be more charismatic than *cantase*, as opposed to the contrary viewpoint expressed by bilinguals and particularly by those whose first language was solely Castilian. No strong language-related trends were apparent with respect to estimations of Capability.

Figure 22: Characteristic group ratings--Interaction between First Language and Guise

Figure 23: Characteristic group ratings by School Language
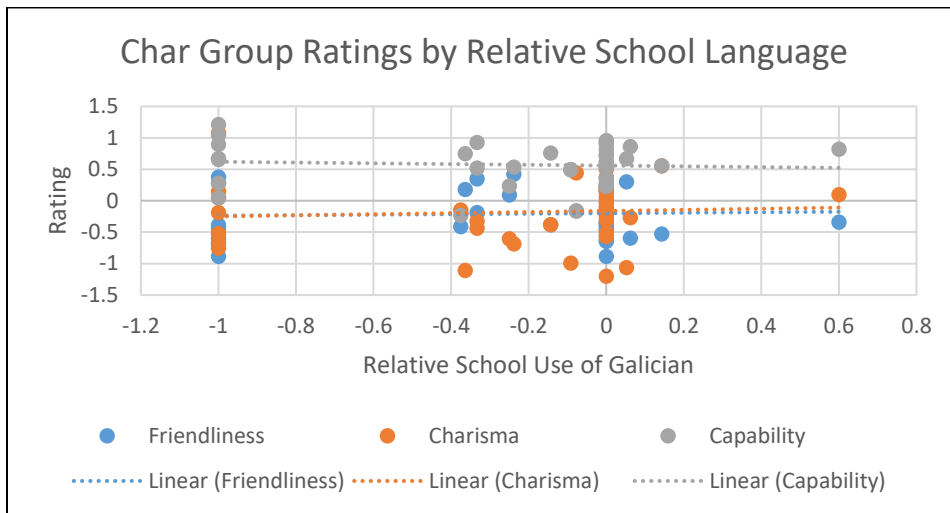
Similarly, no pattern was observed related to relative use of Galician and Castilian at school (Figure 23). Coursework in Galician does, however, appear to have an impact on participant reactions to the guises (Figure 24). Specifically, more schooling in Galician patterns with a more positive view of *cantara* with regards to both Charisma and Capability, while views of *cantase* remain stable or even become more negative. In contrast, the perceived Friendliness of *cantara* decreases sharply with increased school experience in Galician, which goes hand in hand with an increased rating of the Friendliness of the *cantase* guise.

Figure 24: Characteristic group ratings--Interaction between School Language and Guise

Friendliness | Charisma | Capability

- Se Friendliness
- Ra Friendliness
- Linear (Se Friendliness)
- Linear (Ra Friendliness)

- Se Charisma
- Ra Charisma
- Linear (Se Charisma)
- Linear (Ra Charisma)

- Se Capability
- Ra Capability
- Linear (Se Capability)
- Linear (Ra Capability)
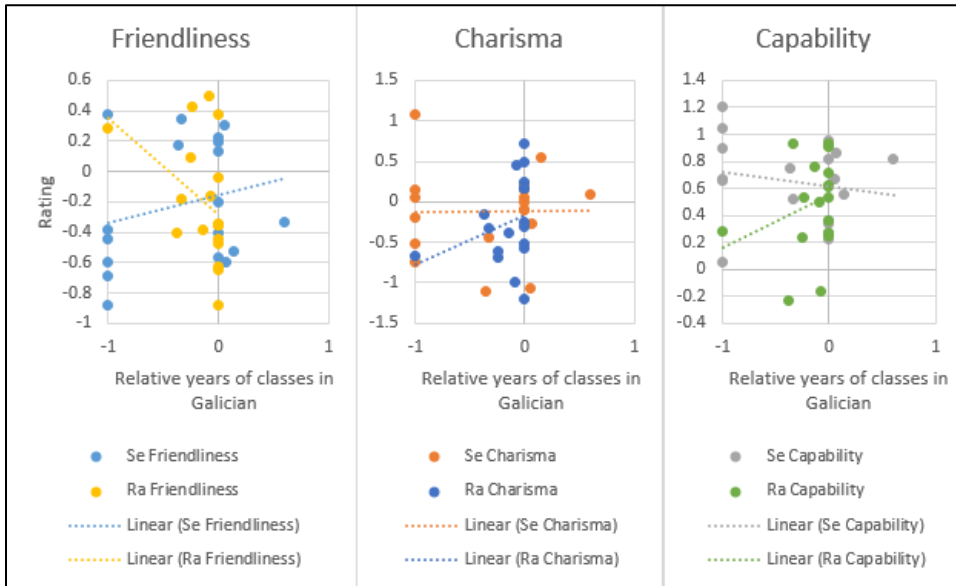
As illustrated by Figure 25, there is a modest tendency for greater relative use of Galician in the workplace to correspond with a higher estimate of the speaker's Capability traits and, conversely, with lower evaluations of the speaker's Friendliness and Charisma.

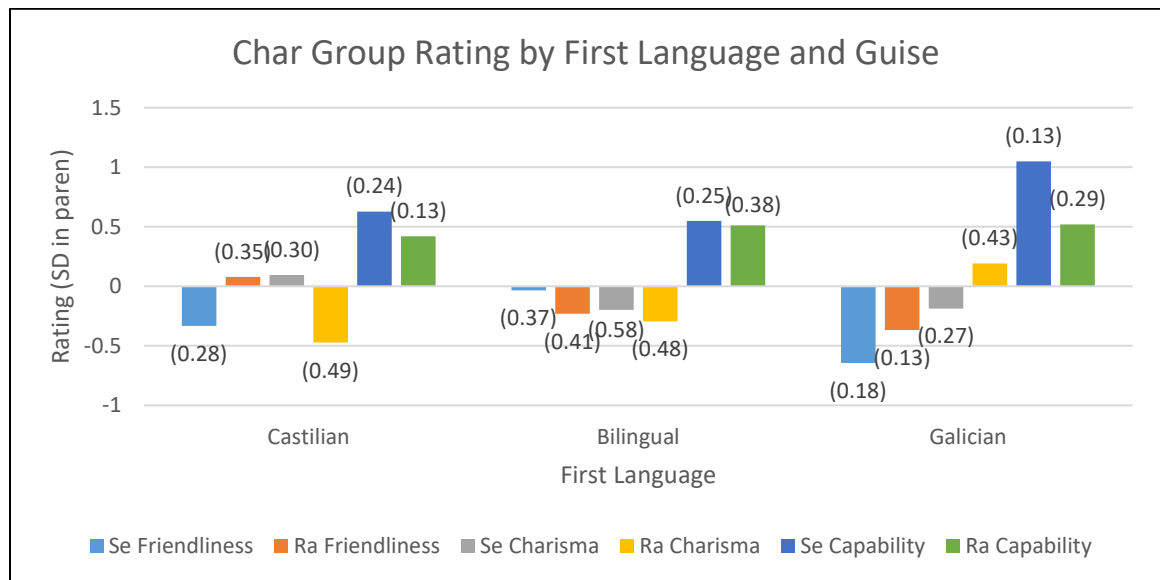**Figure 25: Characteristic group ratings by Work Language**

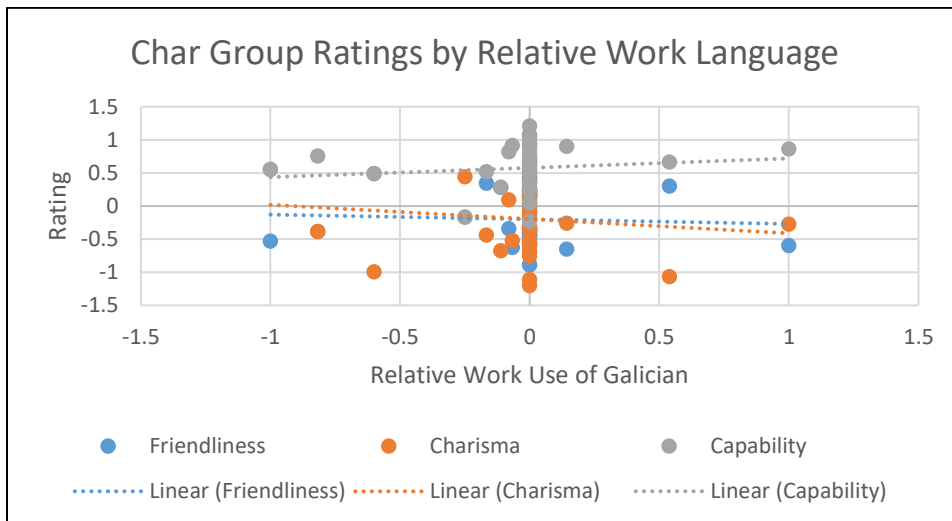Char Group Ratings by Relative Work Language

- Friendliness
- Charisma
- Capability
- Linear (Friendliness)
- Linear (Charisma)
- Linear (Capability)

An examination of the interplay between work language and guise revealed some marked effects, though these should be interpreted with caution since in all cases the range of work use of Galician for those who received the *cantara* guise was more limited than the range for those responding to the *cantase* guise (see Figure 26). Thus it is difficult to know if any trends observed would hold over a more varied range of work language experience. The trends observed, in any case, are for *cantara* to be seen as more friendly and less charismatic by those who use relatively little Galician in the workplace, while those who use the language more often at work see *cantase* as the friendlier but less charismatic form. Conversely, more use of Galician at work corresponds to a higher estimation of the Capability of the recorded *cantase* speaker, while the level of Capability associated with the *cantara* guise remains stable regardless of work language use.

Figure 26: Characteristic group ratings--Interaction between Work Language and Guise
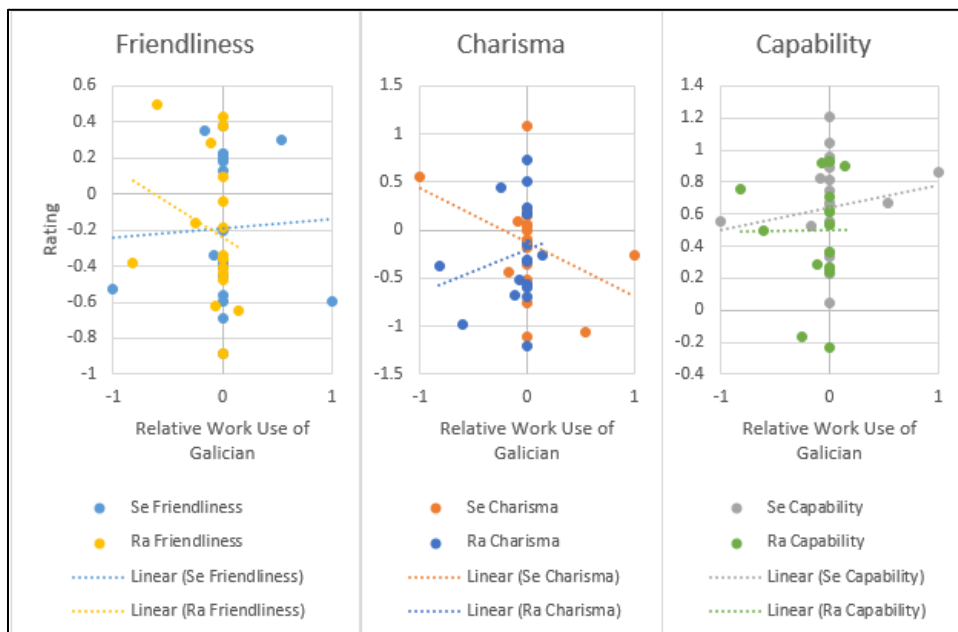
More intriguing is the existence of an interaction between characteristic grouping and family language experience, shown by the crossing trend lines in Figure 27.

Essentially, greater use of Galician in the home correlates with a higher estimation of the Friendliness of the speaker in the recording. At the same time, however, the speaker is considered less charismatic by these same individuals.

Trends by guise (Figure 28) were not as strong. There was in particular no sizable effect on Friendliness and only a slight tendency to view the *cantase* guise as more capable and the *cantara* guise as less capable as Galician use increased. The biggest effects of family language were on Charisma, with *cantase* seen as the more charismatic form by those who have used primarily Castilian in the family and *cantara* seen as more charismatic by those from families who use more Galician.

Figure 27: Characteristic group ratings by Family Language
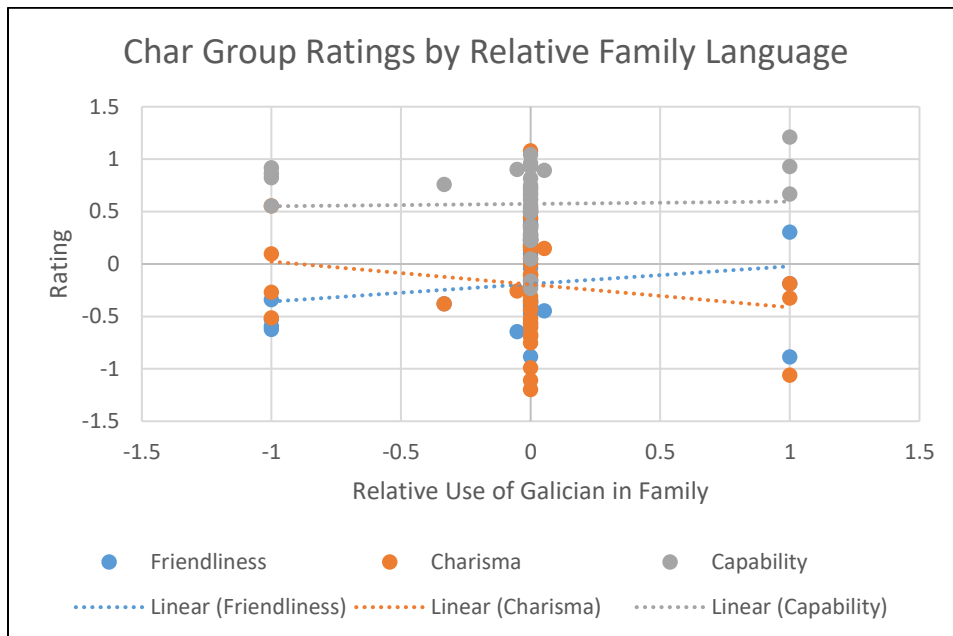
Figure 28: Characteristic group ratings--Interaction between Family Language and Guise
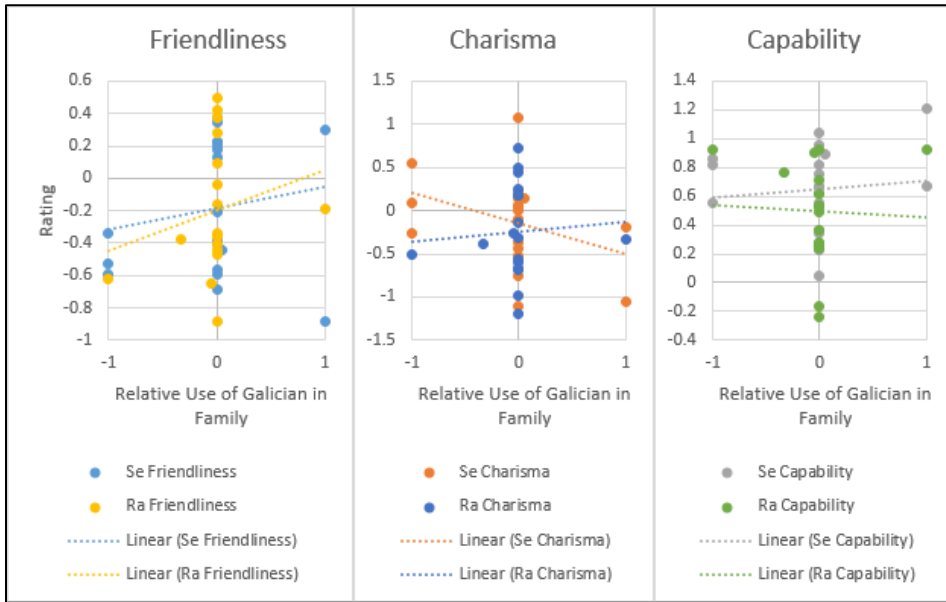
C. Matched-guise results: Current Language Use, Aural Competence, and Attitudes

Finally, as shown in Figure 29 and Figure 30, neither current Galician use nor attitude toward the language appears to affect participant evaluations of the speaker with respect to Friendliness or Charisma. However, the slight increase in the Capability trend line in Figure 30 indicates that more favorable affect toward the Galician language goes hand in hand with more positive evaluations of the recorded speaker's Capability characteristics.

Figure 29: Characteristic group ratings by Current Galician Use

Char Group Rating by Current Language Use
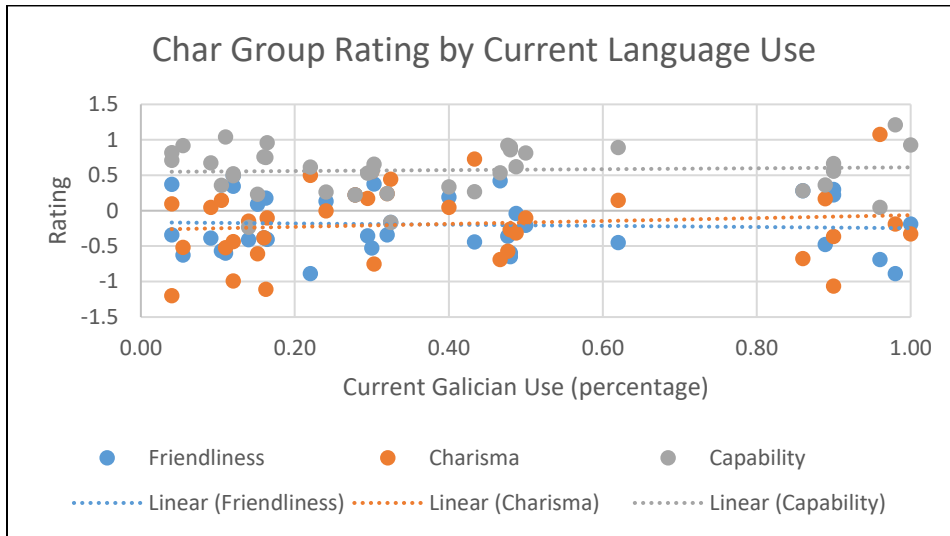
Figure 30: Characteristic group ratings by Attitude toward Galician

Char Group Rating by Attitude toward Galician
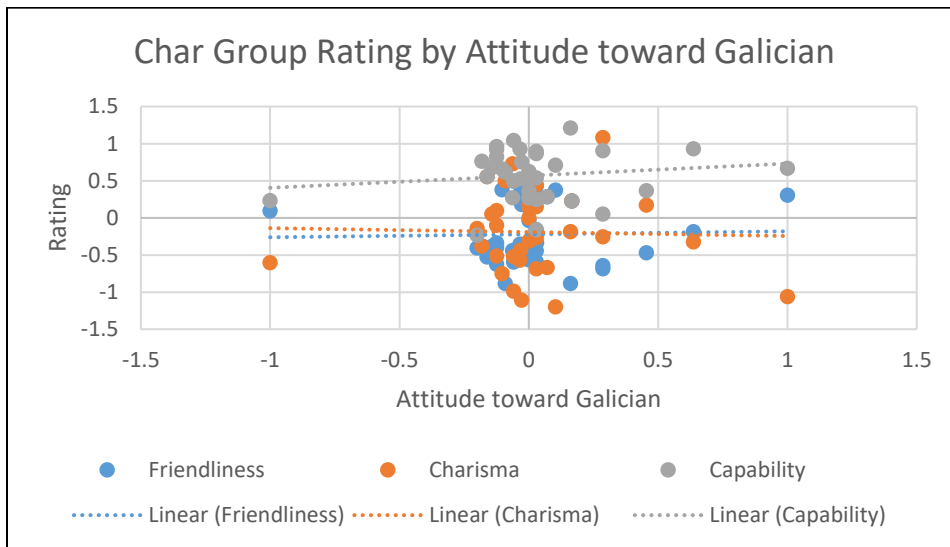
Current language use (Figure 31) had no notable impact on the rating of the two guises. Apparently strong effects are observed for language attitudes, though again, the differing ranges of attitudes for those responding to each guise make definitive conclusions difficult to draw and it is reasonable to infer from Figure 32 that the trend lines may be strongly impacted by data points lying on the extremes of the attitude spectrum. If the trends shown are accurate, however, they indicate that more positive

affect toward Galician corresponds to seeing *cantase* as friendlier, less charismatic, and

slightly less capable while viewing *cantara* as both more capable and slightly more

charismatic, though less friendly.

Figure 31: Characteristic group ratings--Interaction between Current Language Use and Guise
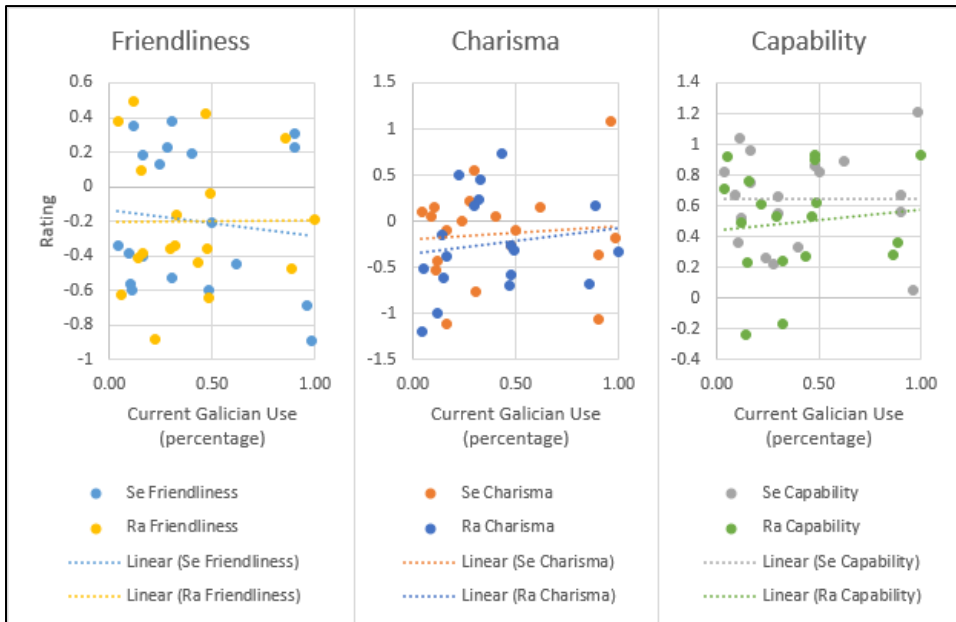
Figure 32: Characteristic group ratings--Interaction between Language Attitudes and Guise
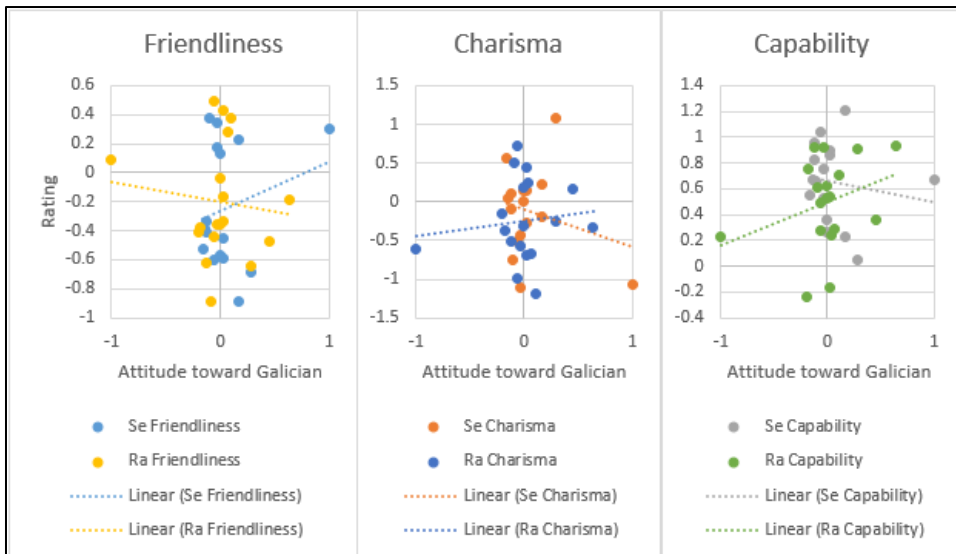
Speakers whose aural abilities were reportedly equally strong in Castilian and Galician generally rated the speaker as more charismatic and more capable, though less friendly, than did their Castilian-dominant counterparts, as shown in Table 12.

Table 12: Characteristic group ratings by Aural Competence

| | Aural Competence | | | |
| | Balanced | | Castilian-Dominant | |
	M	SD	M	SD
Friendship	-0.21	0.85	-0.12	0.86
Charisma	-0.11	1.01	-0.49	0.78
Capability	0.58	0.78	0.49	0.77

Figure 33: Characteristic group ratings--Interaction between Aural Capability and Guise
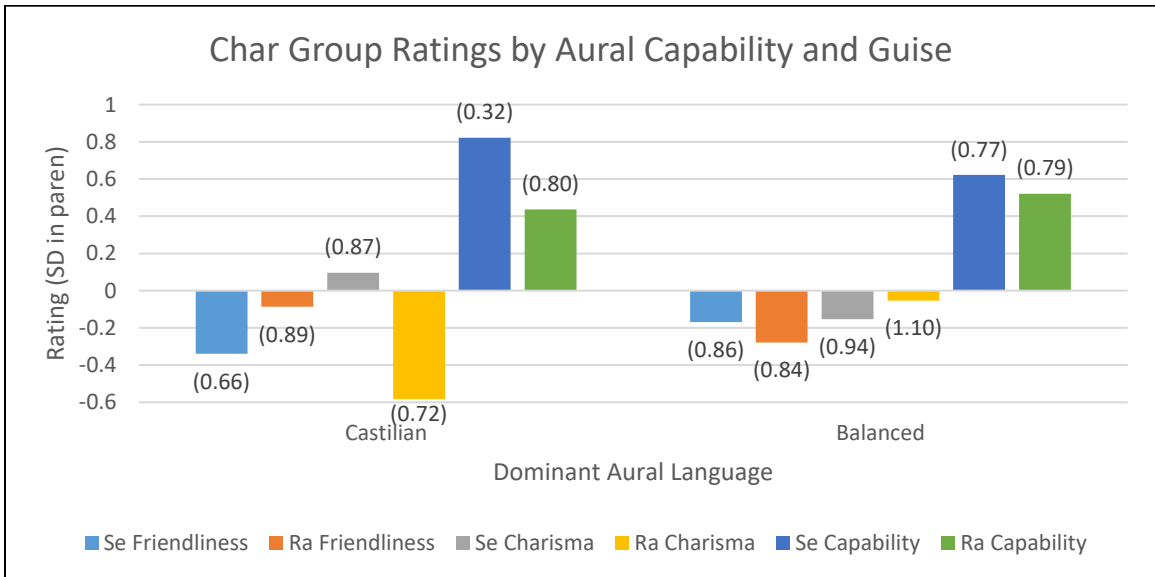
A participant's self-reported aural skills did, however, correspond to differing patterns of evaluation for both Friendliness and Charisma (Figure 33) for the different guises. Those who evaluated their listening comprehension as stronger in Castilian found

the *cantara* guise friendlier than the *cantase* guise, while those whose abilities in the two

languages were reported as balanced rated *cantase* as the slightly friendlier guise. The

cantara guise was also rated as far less charismatic than *cantase* by Castilian-dominant

listeners, while those whose listening abilities in Galician were equally strong considered

it mildly more charismatic. Aural ability did not have a notable effect on evaluations of

guise Capability.

III.   Fill-in-the-blanks task

As described in Chapter 3, the fill-in-the-blanks task consisted of a one-page

written conversation between a doctor and a patient in which all of the conjugated verbs

had been replaced by blanks and an unconjugated verb; participants were asked to fill in

the appropriately conjugated verb form. Though there were 22 blanks in the exercise,

only nine of them were expected to elicit imperfect subjunctive use. The other thirteen

spaces were included to distract from the specific focus of this study. Because all nine

possible contexts for imperfect subjunctive were in the protasis of conditional statements,

it was also frequently the case that these blanks could appropriately be filled out with a

present indicative conjugation. For the purpose of this analysis, only instances in which

the participant actually responded with an imperfect subjunctive form were considered,

regardless of whether such a form *could* have been appropriate. This resulted in a total of

241 tokens being extracted for this task. The dependent variable under consideration was

whether the form produced was *cantara* or *cantase*, which is the categorical variable

included in the logistic regression described hereafter. However, to facilitate graphical

reporting, descriptive statistics will be reported based on the percentage of *cantase*

production.[24] The average *cantase* use over all data points was 0.45 or 45% (SD = 0.50), and if data points were first averaged by participant the average *cantase* use for all participants was 0.42 or 42% (SD = 0.38).

A maximum likelihood logistic regression with random effects for participant and phrase was run in R, returning three variables that approached significance: prime $(F(2,238) = 3.39, p = 0.03)$, Galician use at work $(F(1,240) = 2.09, p = 0.04)$, and current Galician use $(F(1,240) = 3.82, p = 0.04)$. A post-hoc Tukey test revealed that presence of a *cantase* prime as the previous imperfect subjunctive form produced (M = 0.77, SD = 0.42) resulted in significantly more use of the *cantase* form than in unprimed (M = 0.36, SD = 0.48) or *cantara* primed (M = 0.22, SD = 0.42) contexts $(p < 0.001$ for both), while the latter two forms did not differ significantly from one another $(p = 0.20)$.

Figure 34 shows that the tendency is for greater use of Galician in the workplace to correspond to greater written production of *cantase,* though conversely, the percentage of *cantase* use decreases slightly with greater reported current use of Galician, as seen in Figure 35.

---

[24] Percentages are frequently rendered here in decimal form, thus values for *cantara* can be obtained by subtracting the *cantase* value from 1.

Figure 34: FitB form production by Work Language
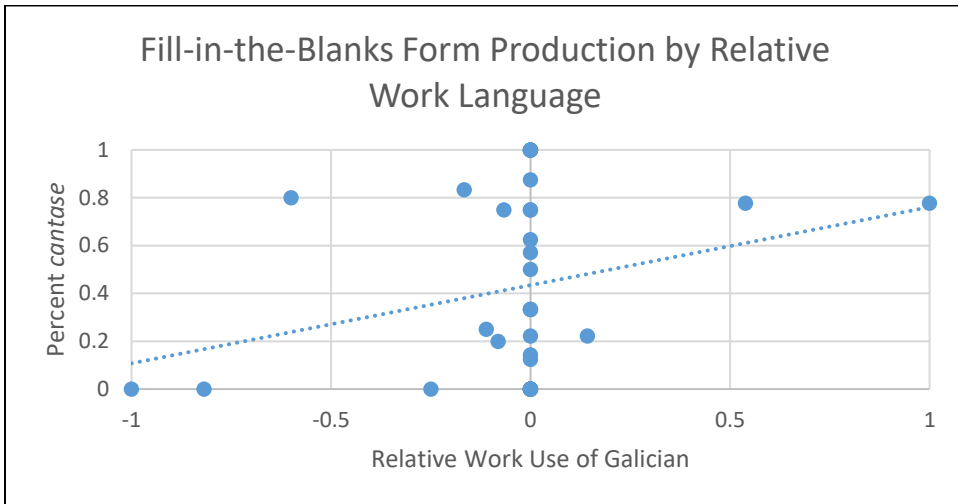
Figure 35: FitB form production by Current Language Use
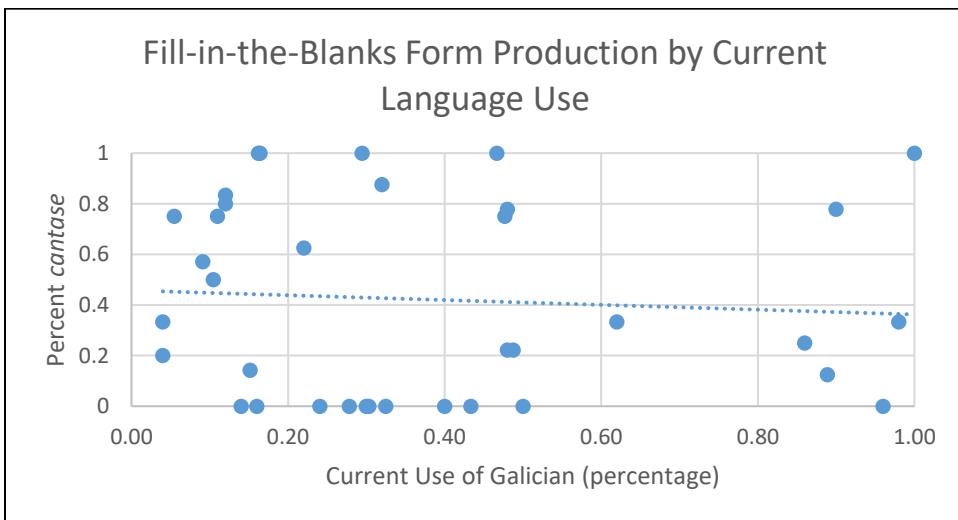
As with the analysis of the matched-guise task carried out in the previous section, the rest of this section will report descriptively on those variables that were not selected as significant by the logistic regression. Related to the priming variable already discussed, distance from the previous prime appears to have some effect on the production of *cantase*, with greater distance corresponding to a decrease in the use of this form (Figure 36). Given that prime itself was a significant factor, it is also logical to consider the effect of a possible interaction between the prime form and its distance from

the target form. However, as can be appreciated in Figure 37, no such interaction is in
evidence. The effect of the prime itself is apparent from the distance between the two
trend lines, but the near parallelism of the two lines indicates that the likelihood of using
cantase decreases uniformly with distance from a previous form, regardless of which
form was primed. No effect was noted due to priming from the guise presented in the
prior task (M = 0.44, SD = 0.50 for the *cantara* guise, and M = 0.46, SD = 0.50 for the
cantase guise).

Figure 36: FitB form production by Recency of Prime

Figure 37: FitB form production--Interaction between Prime and Recency

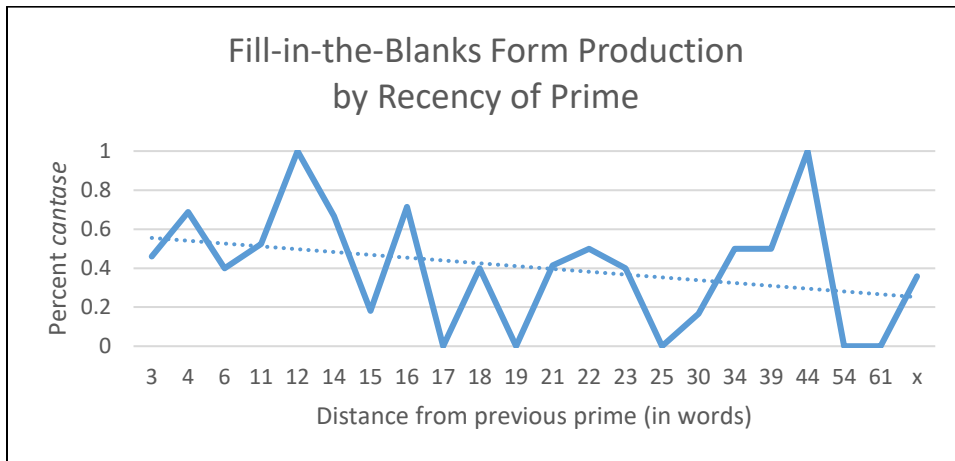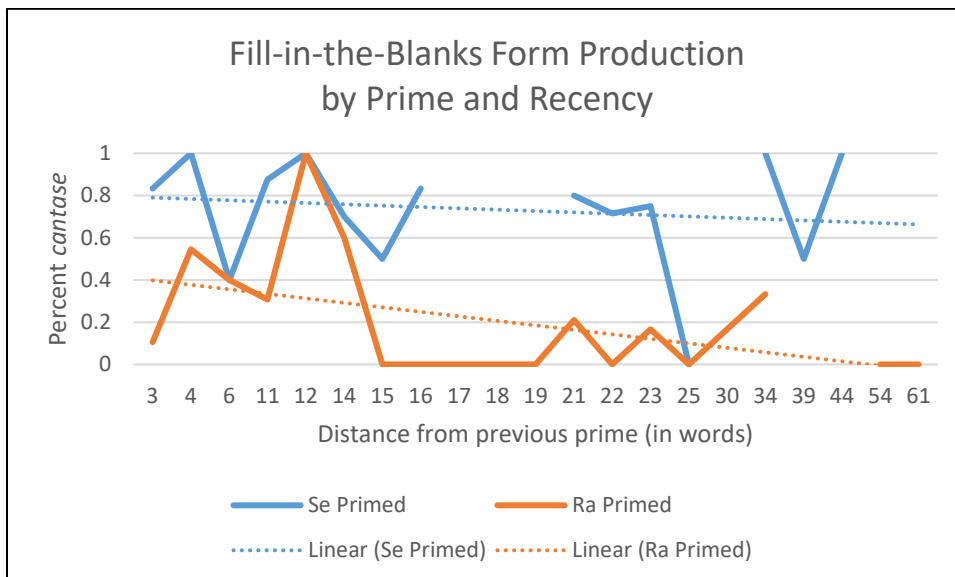The verbs included in this analysis were divided into five groups: *ser*, *tener*, *-ar* verbs, other *-er* verbs, and *–ir* verbs. The percentage of use of the *cantase* form for each group of verbs is shown in Table 13. As stated earlier in this section, the overall percentage of *cantase* use for the entire sample was 0.45; in comparison, *-ar* verbs and *ser* were produced in the *cantase* form more often than average, *-er* verbs and *tener* were produced in this form slightly below the average rate, and *–ir* verbs resulted in a rate of *cantase* use nine percent below the sample average.

Table 13: FitB form production by Verb group

	Percent of *cantase*	Standard Deviation	Number of tokens
-ar verbs	0.50	0.50	70
-er verbs	0.41	0.49	29
-ir verbs	0.36	0.48	55
Ser	0.55	0.50	29
Tener	0.43	0.50	58

Individual participant was, as described previously, included in the logistic regression as a random effect since each participant provided multiple tokens to the data set. However, it is still informative to consider individual production trends. Though there were 39 participants in total, only 36 took part in this task. Of these, eleven produced exclusively the *cantara* form on this task and five produced *cantase* only. Nine produced *cantase* at less than the average rate of 42% for all participants, and the remaining eleven produced *cantase* at an above-average rate.

Table 14 shows the percentage of *cantase* use based on the categorical variables Gender, Location, and Initial Language. The average rate of *cantase* over all participants

was 0.42, and both men and those from Marín tended to lag behind this rate, while women and those from O Grove used *cantase* more frequently than average. There were no differences found based on initial language.

Table 14: FitB form production by Gender, Location, Initial Language

	Gender		Location		Initial Language		
	Female	Male	Marín	O Grove	Galician	Simultaneous Bilingual	Castilian
Percent *cantase*	0.45	0.38	0.35	0.58	0.42	0.42	0.42
Standard Deviation	0.35	0.41	0.36	0.37	0.39	0.35	0.38

With regard to Age, it appears that written use of *cantase* is a form that is strongly disfavored by older participants, though there is much variation across the whole age spectrum (Figure 38). Categorical use of *cantase* is concentrated among speakers in their late 20s to late 40s, while *cantara* is used exclusively by some speakers with ages ranging from the late 20s to 60.

Figure 38: FitB form production by Age

100
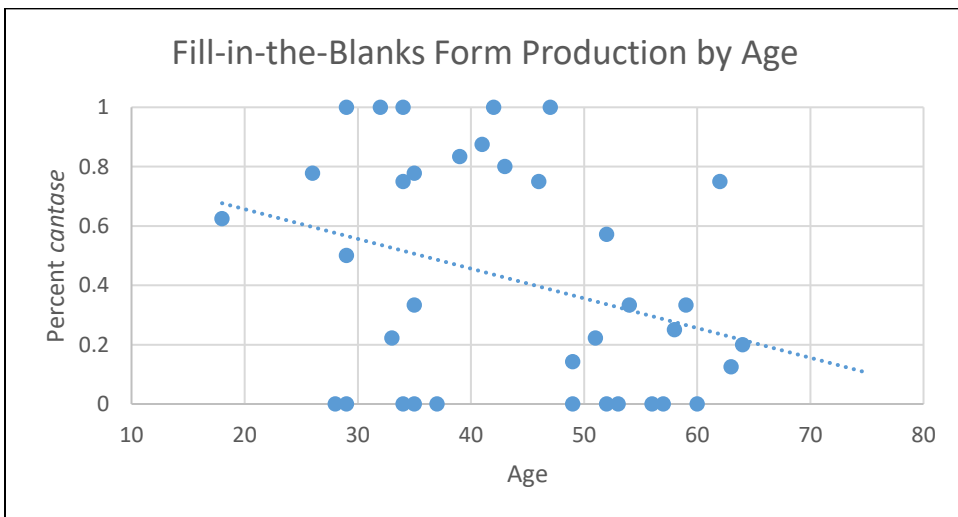
There is a decided tendency for those who have studied at the college level and higher to favor *cantase* use in their writing, while those with only Primary or Secondary education seldom use this form in their writing, as can be seen in Figure 39.

Figure 39: FitB form production by Education
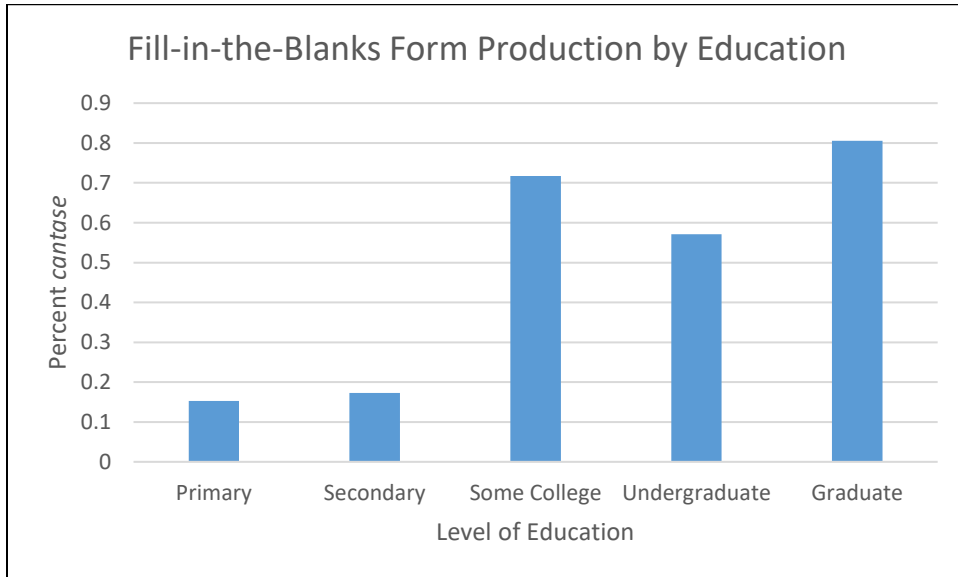
Figure 40: FitB form production by School Language
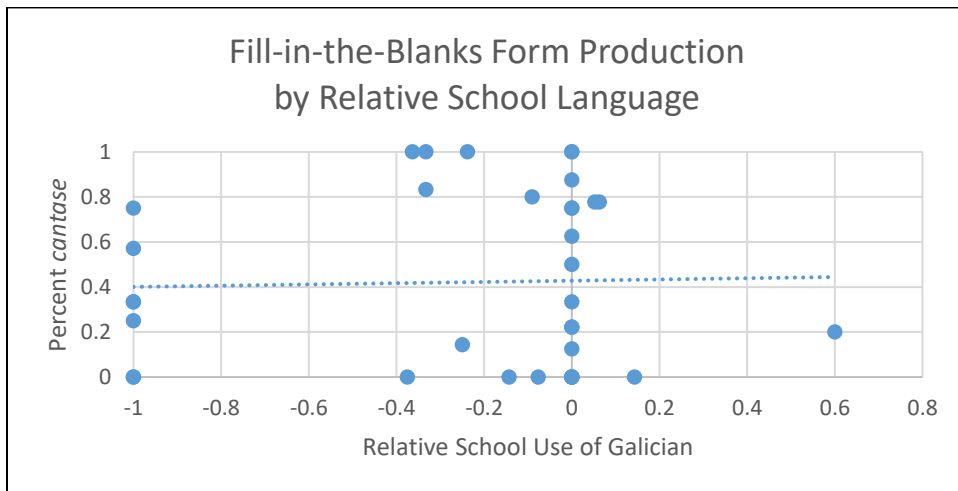
In terms of lifelong language experience, no notable trends emerge with increased use of Galician in the school setting (Figure 40), and categorical uses of one form or the other appear relatively evenly distributed, though those whose school use of Galician and

Castilian was reportedly equal show far more variation in their use of the two forms than do those who predominantly had experience with one language or the other. However, greater use of Galician in the family, illustrated in Figure 41, does bring with it a slight increase in the use of *cantase*.

Figure 41: FitB form production by Family Language
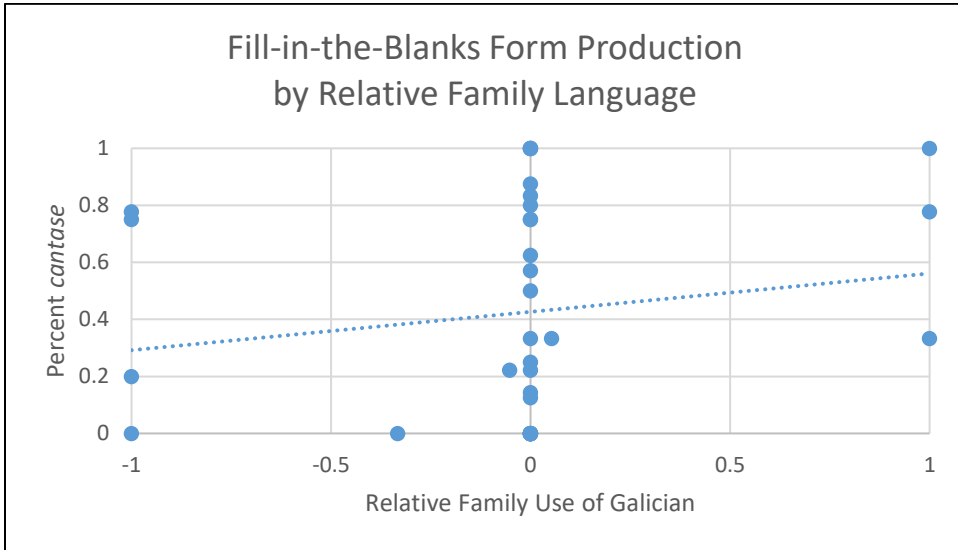
Finally, the greater the participant's reported written competence in Galician relative to Castilian, and the more positive their affect toward the Galician language, the more prevalent their use of *cantase* becomes, as seen in Figure 42 and Figure 43.

Figure 42: FitB form production by Written Competence
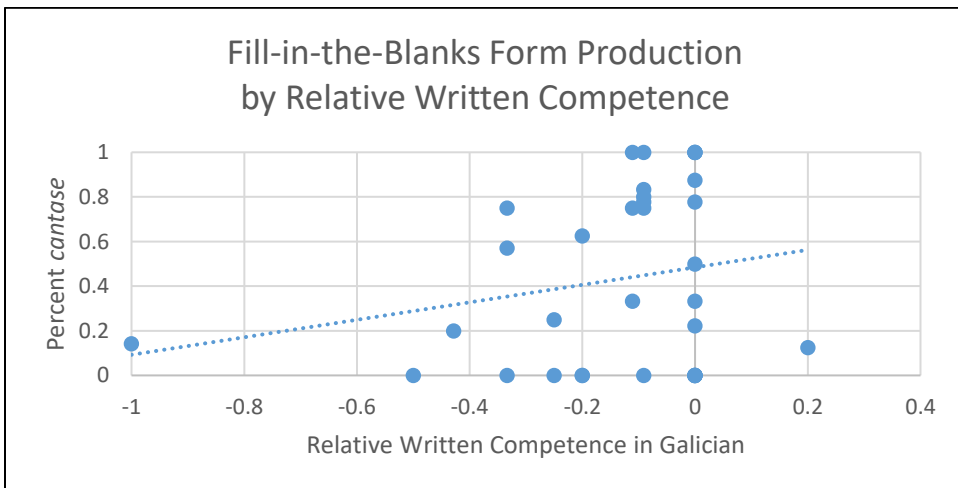
Figure 43: FitB form production by Language Attitudes
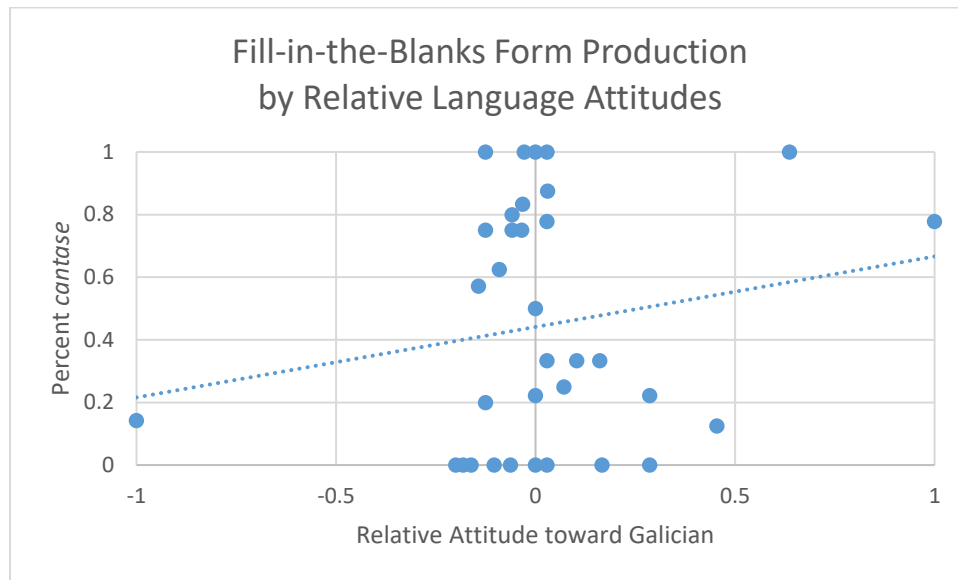
IV.   Acceptability judgment task

The acceptability judgment task was the final task completed by each participant and consisted of two concurrent parts. First, participants evaluated a series of written statements, which provided data about their acceptance of the two imperfect subjunctive forms in written contexts. Second, participants made corrections to any statements which they had previously found to be unacceptable, and these corrections were used to augment the written production data obtained through the fill-in-the-blanks task just analyzed. As these two steps were designed to answer different questions, they are examined separately here.

A.   *Sentence evaluations*

To obtain statistical data about participant evaluations of the phrases in the task, a pair of ordinal logistic regressions with random effects for individual and phrase were run in R, as described in Chapter 3. Two dependent variables—social use of the target form

and correctness—were examined, and a separate regression was run for each. First,
however, a set of pairwise t-tests was run to ascertain whether the inclusion of the
conditional as a possible protasis would skew the results, since this form is not
normatively accepted for use in the protasis of conditional statements. Results of these t-
tests indicated highly significant differences between the conditional and the other two
forms for both dependent variables ($p < 2$ x 10^{-16} in all cases). While these pairwise tests
also found significant differences between *cantara* and *cantase* ($p = 0.001$ for social use
and $p = 0.0005$ for correctness), it was felt that the comparatively weaker significance of
these differences might be overshadowed by the much greater differences with the
conditional form. Because the object of interest of this study is the use of *cantara* and
cantase, the evaluation of statements containing a conditional in the protasis was
considered extraneous; thus these statements ($N = 234$) were excluded from
consideration. The resulting data set consisted of 478 evaluated tokens.

1. Acceptability judgment results: Social use

Running the ordinal logistic regression with social use as the dependent variable
returned Protasis ($p < 0.001$), Apodosis ($p < 0.001$) and Relative School Language ($p =
0.02$) as significant (or nearly significant) predictors of sentence rating (Log Likelihood =
-435.54). Specifically, a *cantase* protasis was rated as more used ($M = 1.16$, $SD = 1.30$)
than a *cantara* protasis ($M == 0.72$, $SD = 1.57$), and the use of the imperfect in the
apodosis of a conditional statement was rejected ($M = -0.16$, $SD = 1.67$), while the use of
conditional in this same context was widely accepted ($M = 0.84$, $SD = 1.59$). Figure 44
illustrates the nearly significant trend for School Language, namely that greater
experience with coursework in Galician corresponds to an elevated evaluation of the

Social Use of conditional statements. Neither of the latter variables interacted

significantly with the verb form in the statement protasis; this can be visually confirmed

by noting the parallel trend lines in Figure 45 and the same downward shift of the

averages from Conditional to Imperfect for both forms in Figure 46.

Figure 44: Acceptability judgment: Social Use by School Language
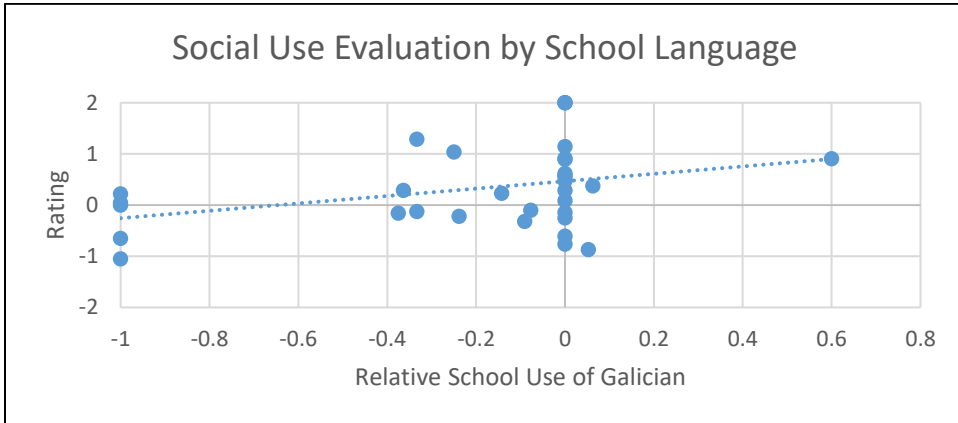
Figure 45: Acceptability judgment: Social Use--Interaction between School Language and Protasis
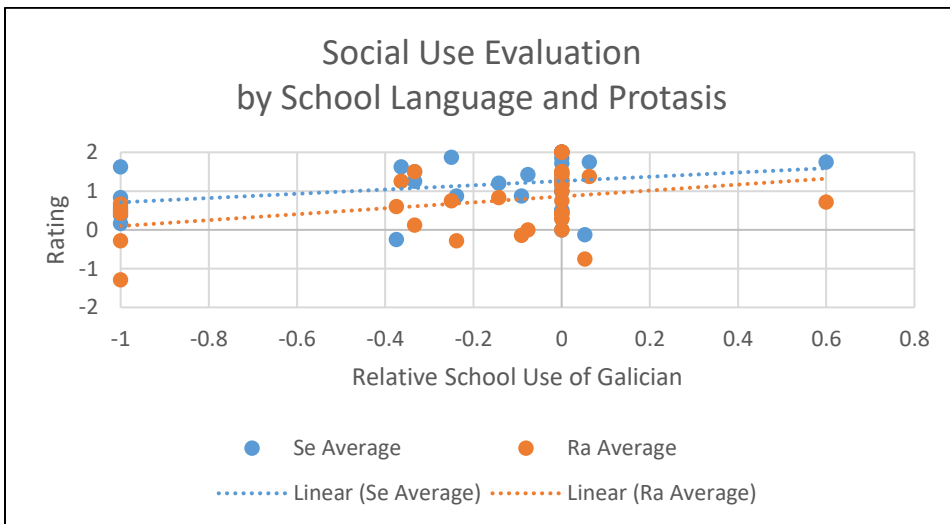
Figure 46: Acceptability judgment: Social Use--Interaction between Apodosis and Protasis

Social Use Evaluation by Apodosis and Protasis
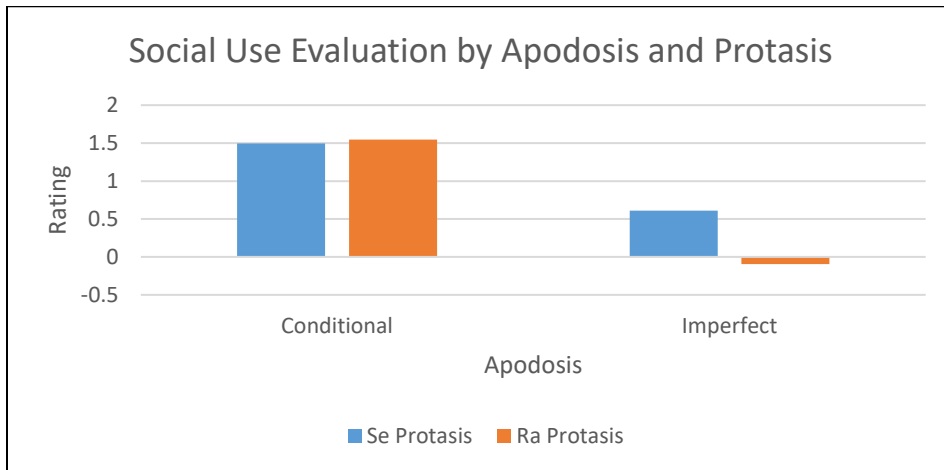
Descriptively, few other variables showed interesting trends, particularly with respect to interactions with Protasis, which is the phenomenon of greatest interest to this study. Level of Education, shown in Figure 47, interacts with protasis in a noticeable way only at the primary education level, where *cantara* was reported to be the more used form in participant social circles than *cantase*. At all other education levels, *cantase* was rated as more frequently used than *cantara.* Figures and tables for the other variables considered can be found in the Appendix.

Figure 47: Acceptability judgment: Social Use--Interaction between Education and Protasis
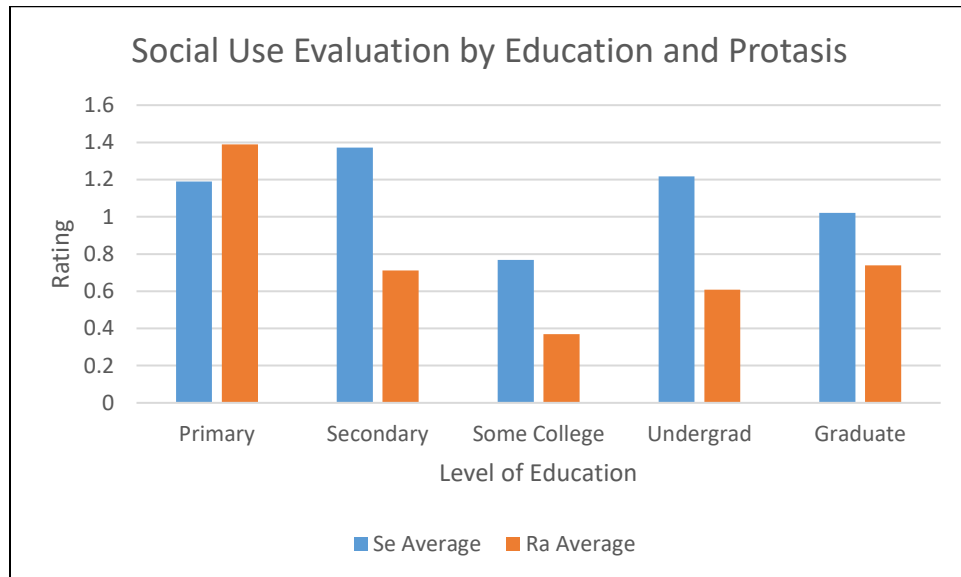
2. Acceptability judgment results: Correctness

An ordered regression with Correctness as the dependent variable identified

Protasis (p < 0.001) and Apodosis (p < 0.001) as significant, while the interactions

between Protasis and Location (p = 0.04) and between Protasis and Possibility (p = 0.04)

approached significance (Log Likelihood = -408.51). Protases in *cantase* tended toward

higher evaluations of correctness overall (M = 0.39, SD = 1.81) than those in *cantara* (M

= -0.11, SD = 1.88), and the conditional (M = 0.30, SD = 1.86) was considered far more

correct on average than the imperfect (M = -1.31, SD = 1.38) in statement apodoses.

Figure 48: Acceptability judgment: Correctness--Interaction between Location and Protasis
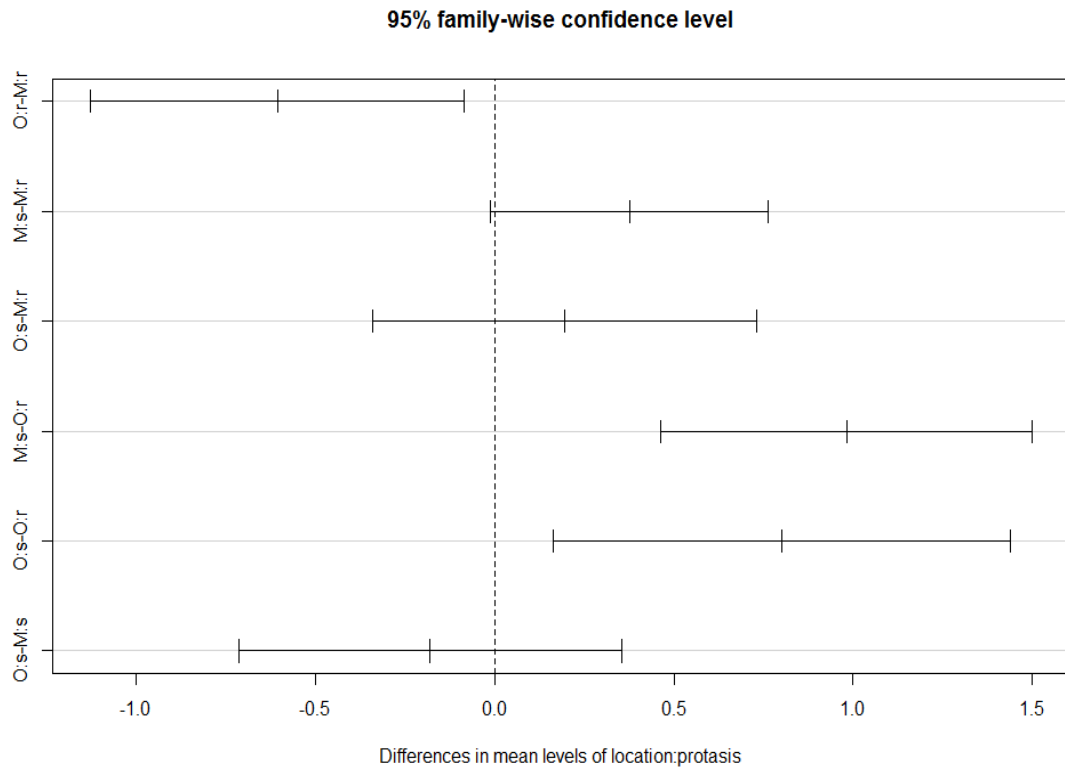
Post-hoc Tukey tests, represented graphically in Figure 48 and Figure 49, indicated that

the *cantara* form in O Grove was considered significantly less correct than both the same

form in Marín and the *cantase* form in both locations, while no other differences were

significant. Similarly, with respect to the interaction between Possibility and Protasis, the

cantase form was considered more correct than the *cantara* form in clearly contrary-to-

fact statements, and contrary-to-fact statements containing *cantase* were considered more

correct than possible statements containing *cantara*.

Figure 49: Acceptability judgment: Correctness--Interaction between Protasis and Possibility
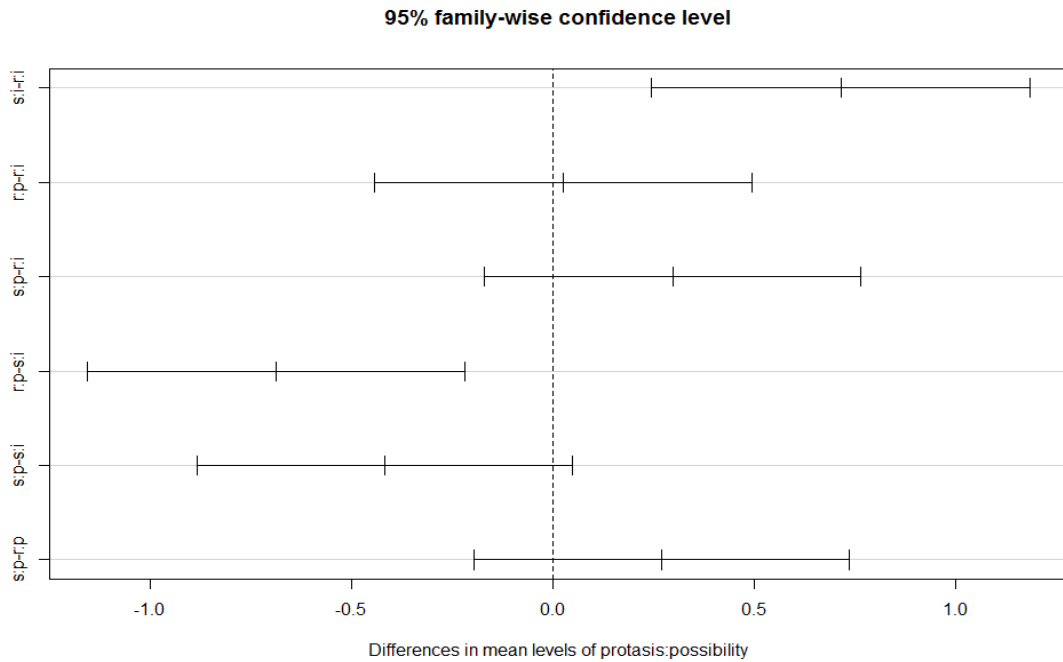
As with Social Use, few non-significant variables demonstrably interacted with protasis form in ratings of Correctness. Education again was one exception, with participants educated only at the primary level considering *cantara* the more correct form, while all other participants gave this distinction to *cantase*, as shown in Figure 50. Additionally, participants from Marín tended toward higher evaluations of correctness overall (M = -0.41, SD = 0.46) than participants from O Grove (M = -0.78, SD = 0.59)

Figure 50: Acceptability judgment: Correctness--Interaction between Education and Protasis

Correctness Ratings by Education and Protasis
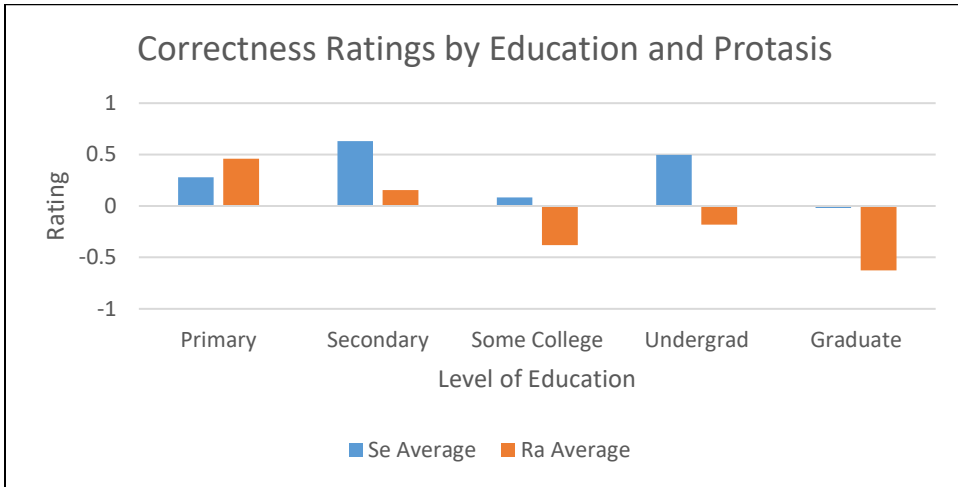
Finally, Language Attitudes did appear to have some relation to the evaluation of the two protasis forms as correct (Figure 51). Specifically, participants with greater relative affect toward Galician tended to rate *cantase* as more correct than *cantara*, while differences in evaluation of each form were neutralized as relative attitude toward Galician declined in favor of Castilian.

Figure 51: Acceptability judgment: Correctness--Interaction between Language Attitudes and Protasis

Correctness Ratings by Language Attitudes and Protasis
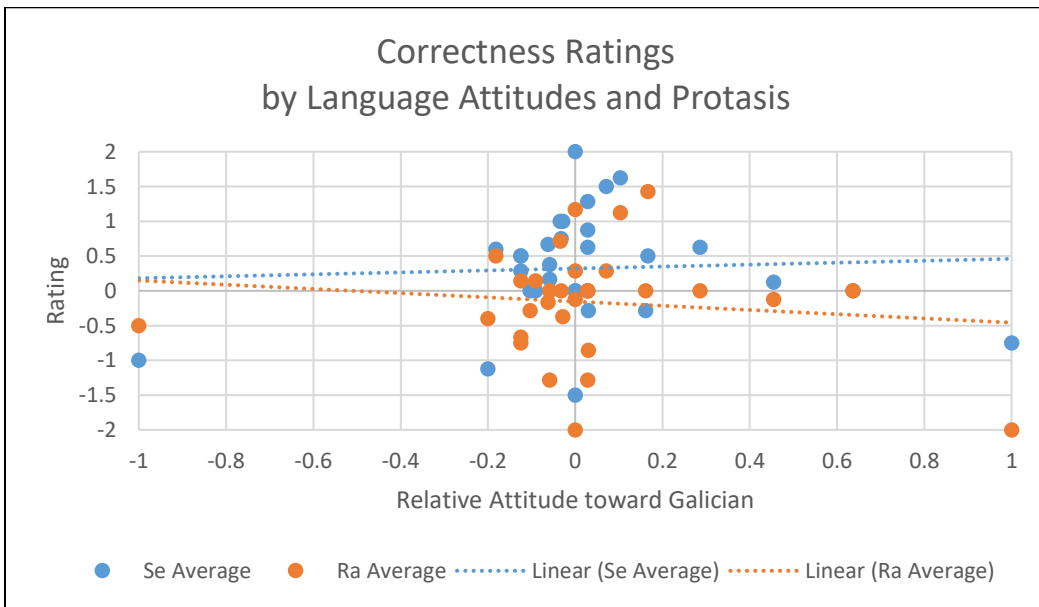
Tables and figures detailing the results of the remaining independent variables are included in the Appendix.

B.   *Participant corrections*

Participant corrections were evaluated statistically through the creation of a logistic regression model with random effect for individual. The dependent variable for this test was the form produced in the corrected protasis. As participants only produced *cantara* or *cantase* in the protases, all 540 tokens obtained were included.

The model produced revealed that Priming was highly significant ($F_{(2, 537)}$ = 41.3, $p < 0.001$) and that Initial Language was also significant ($F_{(2, 537)}$ = 4.15, $p < 0.01$), while Location approached significance ($F_{(1, 537)}$ = 1.40, $p = 0.04$). A post-hoc Tukey test on Prime indicated that differences in the rates of *cantase* production were highly significant for all three primes ($p < 0.001$). Figure 52 illustrates this unsurprising trend; the presence of a *cantara* prime resulted in production of *cantase* only 1/3 of the time, while a Conditional form led to *cantase* in just over half of the cases. A *cantase* prime yielded a rate of use of the *cantase* form over 80%.

Figure 52: Acceptability judgment: Correction form production by Prime
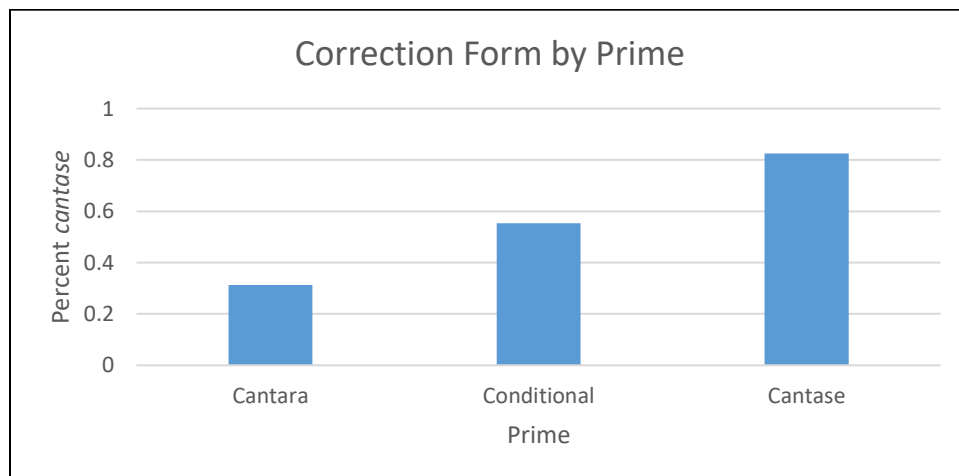
With respect to initial language, a post-hoc Tukey test revealed that only the difference in rate of *cantase* production between L1 Galician speakers and Simultaneous Bilinguals was significant (p < 0.001), with Galician L1ers producing *cantase* at a higher rate. Despite the similar trend between these speakers and Castilian L1ers evident in Figure 53 the difference between these two first languages was not significant; this is likely due to the low number of participants in each of these two initial language groups.

Figure 53: Acceptability judgment: Correction form production by Initial Language
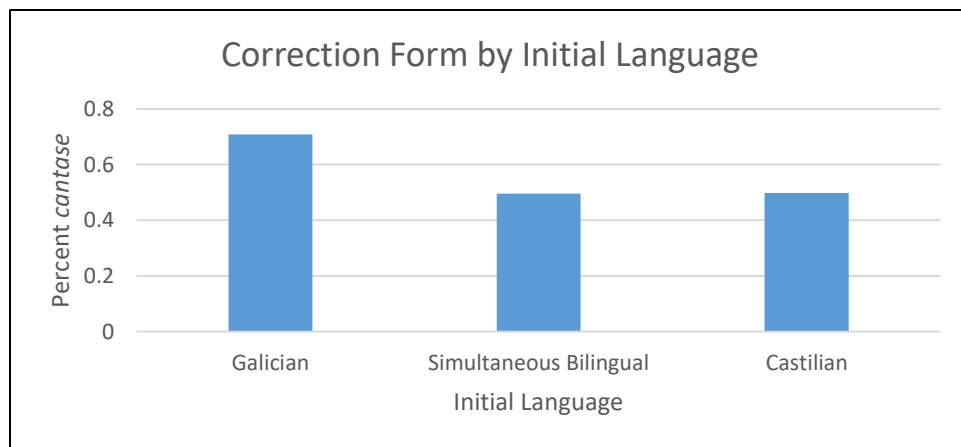
The difference in form production by Location is illustrated in Figure 54; essentially, speakers from O Grove produce *cantase* at a slightly higher rate than those from Marín.

Figure 54: Acceptability judgment: Correction form production by Location
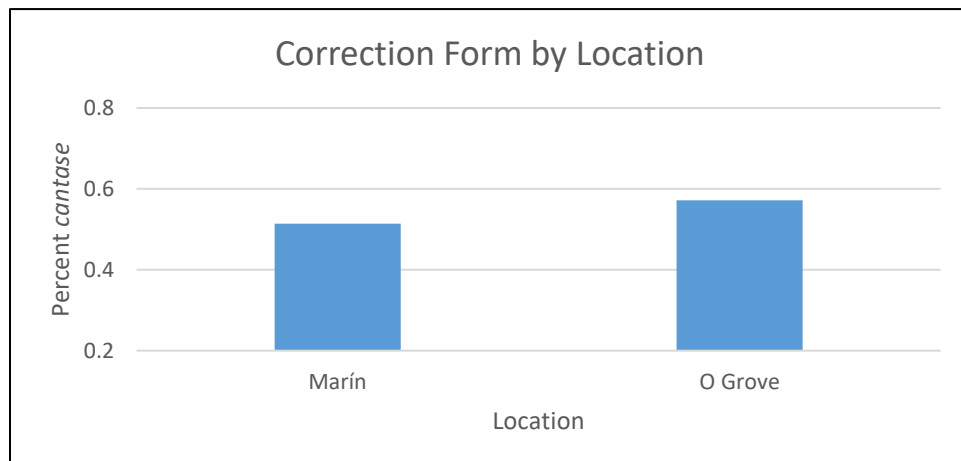
The clear strength of priming as a predictor of form production makes it necessary to consider the other variables in interaction with Prime in addition to independently; without doing so, less robust trends may be obscured. An attempt was made to create a logistic regression model including interaction terms, but the model was returned as unidentifiable. Because of this, the rest of this section descriptively presents the trends present in the independent variables in interaction with Prime, but no claims can be made about the significance (or lack thereof) of such trends.

Figure 55: Acceptability judgment: Correction form production--Interaction between Initial Language and Prime
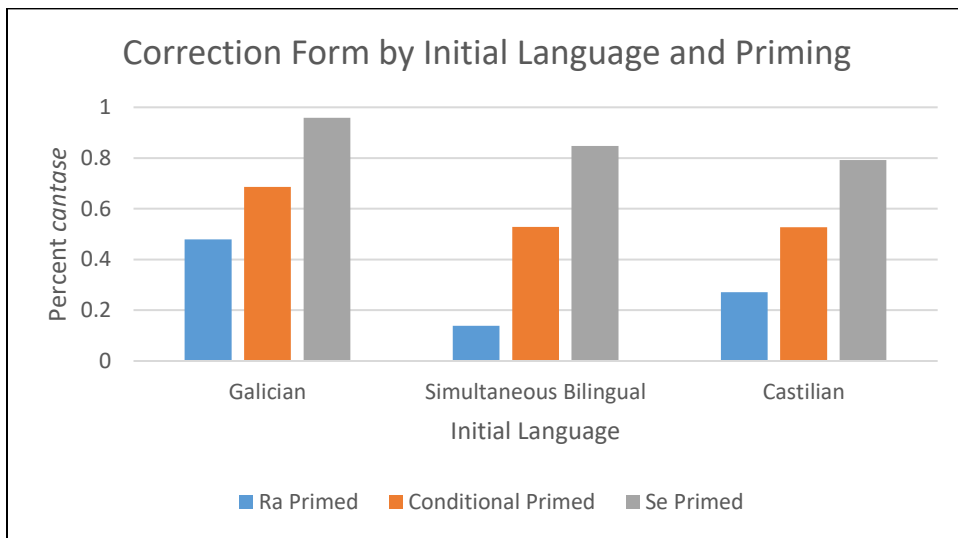
Figure 56: Acceptability judgment: Correction form production--Interaction between Location and Prime
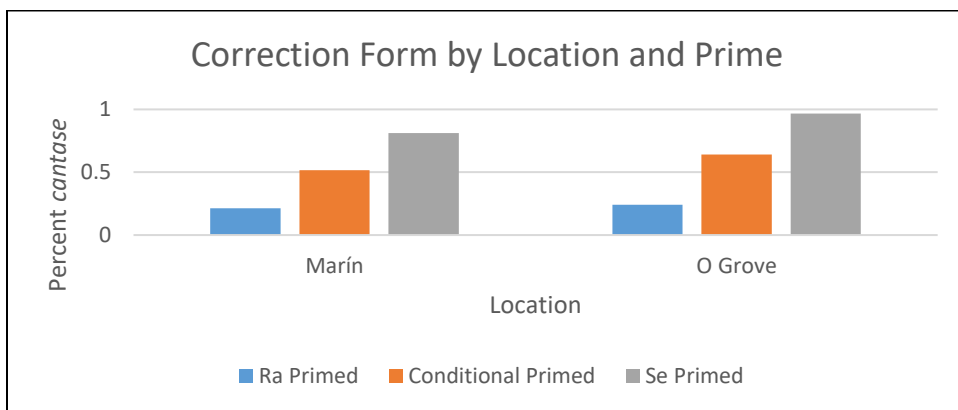
Neither Initial Language nor Location, the two variables selected as individually significant in the original model, show evidence of an interaction with Priming, as the directionality of the difference between Initial Languages (Figure 55) and between Locations (Figure 56) is essentially the same for all three priming conditions. That is, for each kind of prime, O Grove shows a higher rate of *cantase* production than Marín, and similarly, Galician L1ers use *cantase* more frequently than the other two language groups regardless of which prime is being considered.

1. Correction form production: Gender, Age, Education

No notable trends emerged with regard to participant Gender, either independently or in interaction with Prime. Overall both males and females produced *cantase* in just over half of their corrections (55% and 52%, respectively), and their Primed results behaved as expected, with both genderes overwhelmingly replicating the primed form in their correction. In terms of Age, older individuals generally produced *cantase* in their corrections somewhat less often than younger speakers (Figure 57). This tendency depended greatly, however, on the form provided in the prompt. If primed with a *cantara* or *cantase* form, respondents tended heavily toward maintenance of the primed form regardless of age. When neither of these forms was present as a prime, however, younger participants produced *cantase* nearly as often as their *cantase* primed rates, while older individuals overwhelmingly opted for *cantara* (Figure 58).

Figure 57: Acceptability judgment: Correction form production by Age

Correction Form Production by Age
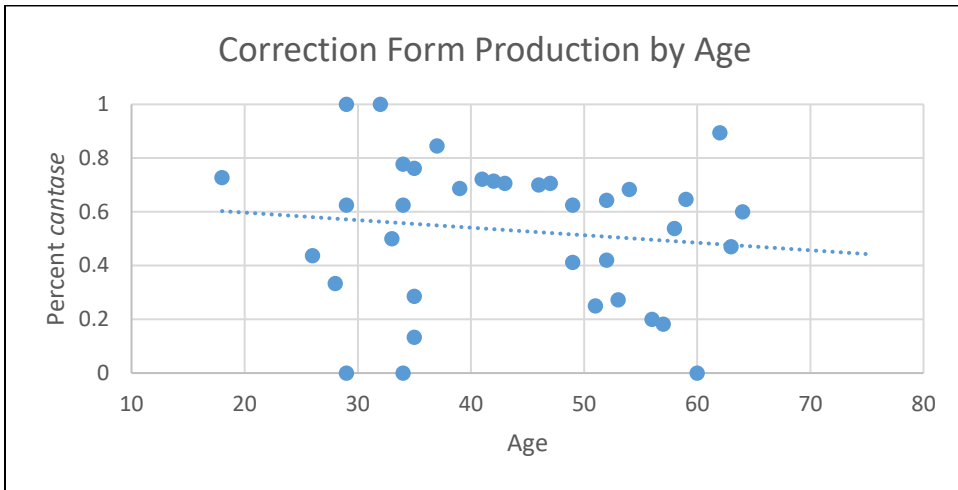
Figure 58: Acceptability judgment: Correction form production--Interaction between Age and Prime

Correction Form Production by Age and Prime
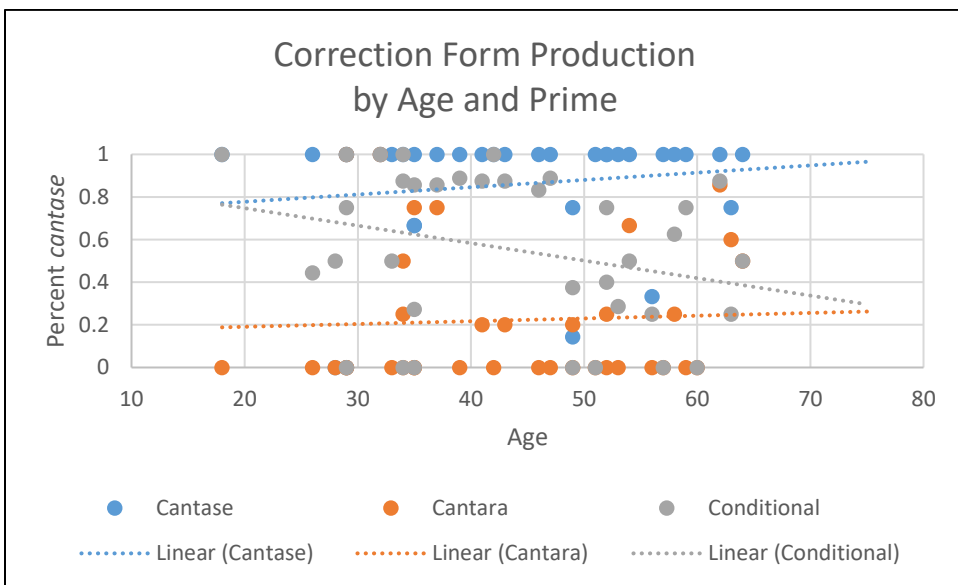
Results for Education are somewhat less clear, though Figure 59 reveals a modest tendency for higher levels of education to correspond to greater *cantase* production. The divide between secondary and post-secondary education is lessened somewhat when a *cantase* prime is present (Figure 60) and it disappears almost entirely in the presence of a *cantara* prime, with participants with undergraduate study showing a strong tendency to

reproduce a primed *cantara.* Only at the level of graduate study does the tendency to favor *cantase* return in the face of a *cantara* prime.

Figure 59: Acceptability judgment: Correction form production by Education

Figure 60: Acceptability judgment: Correction form production--Interaction between Education and Prime
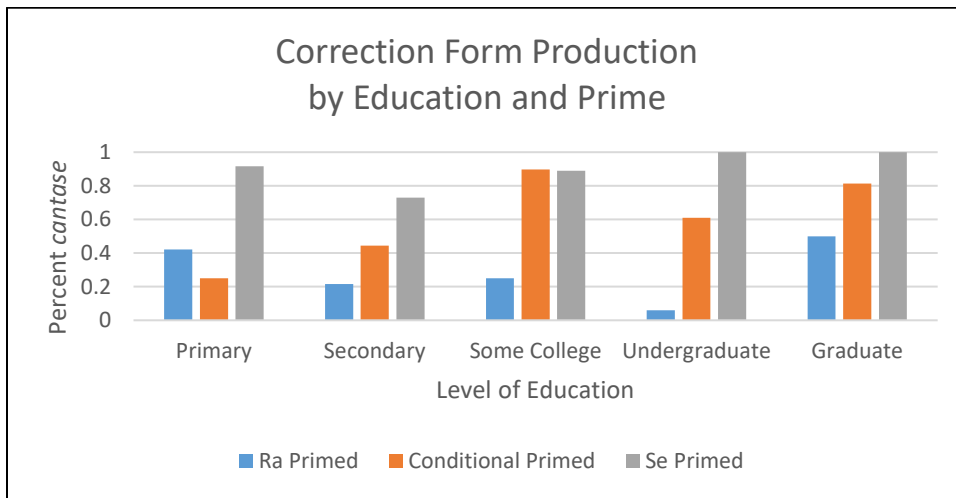
2. Correction form production: Language use variables

Family and Work Language Experience behaved almost identically to one another, showing a tendency for *cantase* production to increase in tandem with greater exposure to Galician regardless of prime (see Figure 61 for Family Language). As was the case with Age, however, these tendencies are due almost entirely to the same trend being observed for conditional-primed tokens; tokens primed with *cantara* or *cantase*

were generally reproduced in the correction (see Figure 62 for Family Language; both analogous figures for Work Language are included in the Appendix).

Figure 61: Acceptability judgment: Correction form production by Family Language Experience
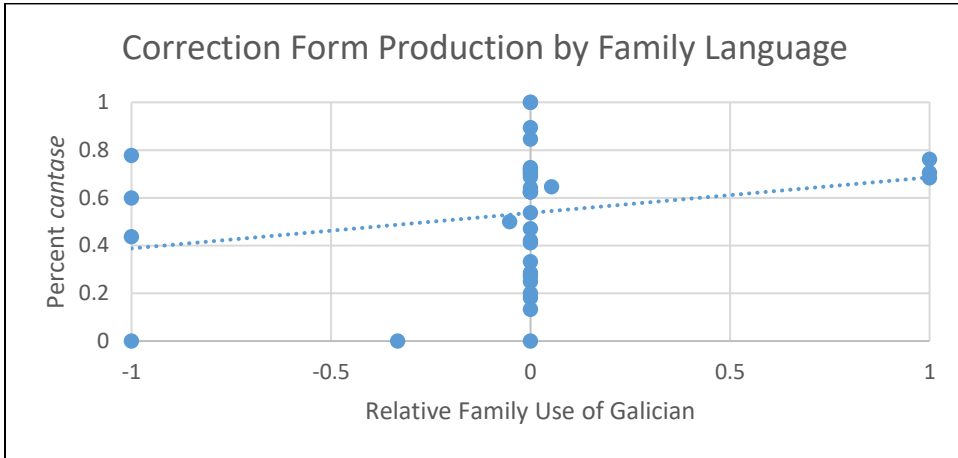
Figure 62: Acceptability judgment: Correction form production--Interaction between Family Language and Prime
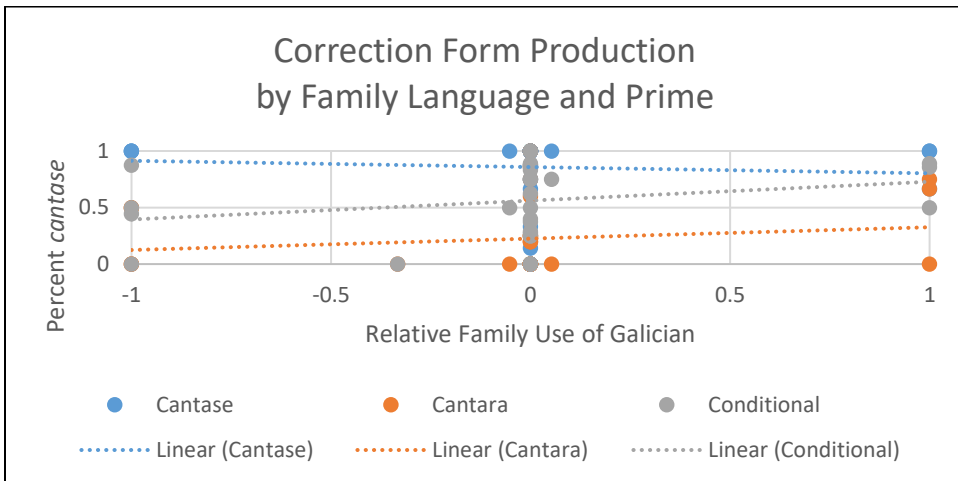
No reliable trends existed, whether independently or in combination with priming, for the remaining language use variables. Figures corresponding to these variables can be found in the Appendix.

3. Correction form production: Apodosis, Possibility, Lexical Item

Neither Apodosis nor Possibility showed independent effects on form production; conditional (M = 0.54, SD = 0.50) and imperfect (M = 0.53, SD = 0.50) apodoses as well as contrary-to-fact (M = 0.54, SD = 0.50) and non-contrary-to-fact (M = 0.55, SD = 0.50) were nearly equal in their slight tendency to favor *cantase* production. However, interesting phenomena were observed when considering these variables in conjunction with Prime.

As can be seen in Figure 63, the conditional form in the apodosis follows the expected priming trend, while the use of imperfect in the apodosis appears to provoke a switch away from the primed form for both *cantara* and *cantase* primes. One possible motivation for a switch to *cantara* to be triggered by the imperfect is the expressly past referential nature of the imperfect form; if *cantara* still retains vestiges of its anterior meaning, it would seem the natural form to use in harmony with other past referent forms. A more plausible explanation, however, is simply that the small number of correction tokens associated with an imperfect apodosis (36 total, of which all but six were paired with a conditional protasis prime) makes these results unreliable (compare to 503 tokens with conditional in the apodosis). This explanation is particularly felicitous in that it also accounts for the unexpected switch from *cantara* to *cantase* in tokens associated with imperfect in the apodosis.

Figure 63: Acceptability judgment: Correction form production--Interaction between Apodosis and Prime

Correction Form Production by Apodosis and Prime
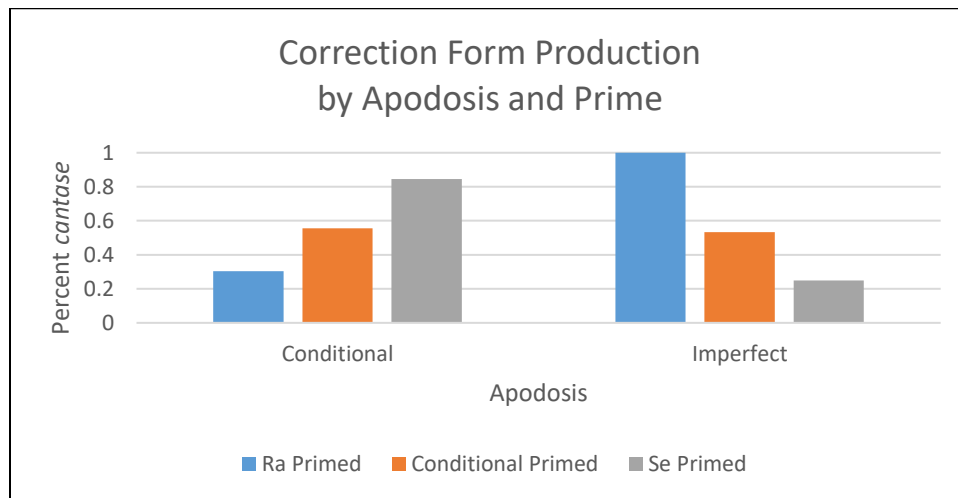
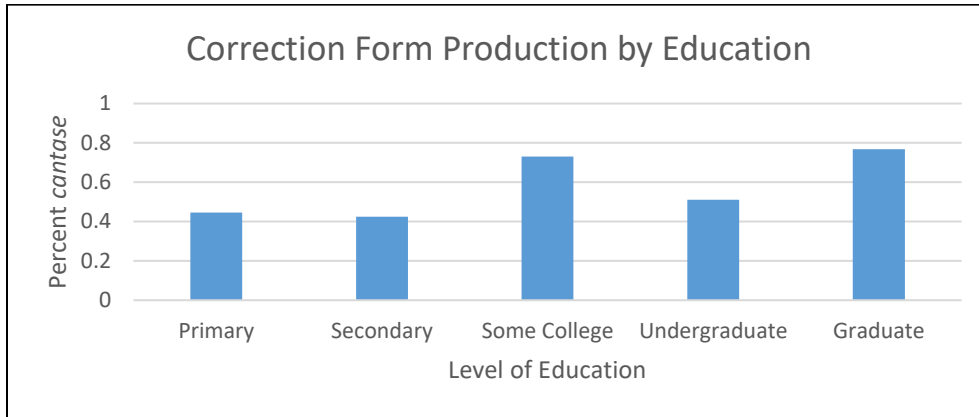**Figure 64: Acceptability judgment: Correction form production--Interaction between Verb Group and Prime**

Correction Form Production by Lexical Item and Prime
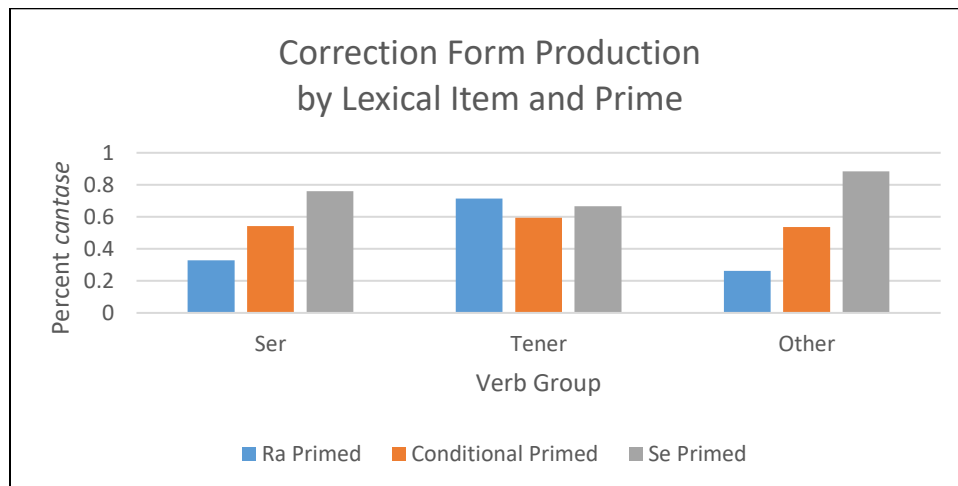
A similar explanation exists for the unusual pattern of production seen for the verb *tener* in Figure 64. While *tener* is independently associated with a slight elevation in *cantase* rate (M = 0.61, SD = 0.49) with respect to *ser* (M = 0.53, SD = 0.50) and other verbs (M = 0.54, SD = 0.50), this association does not explain why the *cantara* form should be so heavily avoided in production for this verb. A closer look at the data reveals that, of the 79 corrections containing the verb *tener* in the protasis, seven them are

cantara-primed and only three are *cantase*-primed, thus making it probable that the

unexpected patterns for *tener* are simply a result of the data being skewed by low

numbers.

Figure 65: Acceptability judgment: Correction form production--Interaction between Possibility and Prime
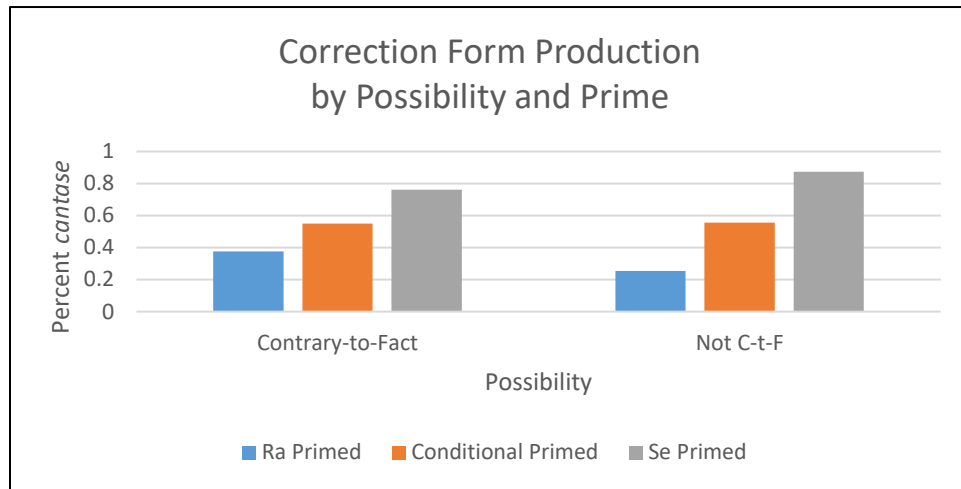
No such explanation, however, accounts for the tendency in Figure 65 for

contrary-to-fact statements to show a switched form (*i.e.* for a *cantara*-prime to produce

cantase or vice-versa) more frequently than non-contrary-to-fact statements. One possible

motive for such a phenomenon could be that the participant is implicitly expressing

rejection of an evidently nonfactual statement through rejection of the verb form

originally associated with it. This explanation, however, is purely speculative and in any

case is unverifiable with the current data set.

As stated earlier in this chapter, in considering the results of a study with a small

sample size, it is important to keep in mind that statistical significance or a lack thereof

may be misleading. It is also quite possible for the inclusion of even a small number of

outliers to create the impression of trends that would be averaged out over a larger and

more representative sample. For these reasons the plausibility of both the statistical and

the descriptive results just presented should be considered in the light of other studies and of general knowledge of both the Galician context and of language contact theory. It is to such considerations that Chapter 5 is dedicated.

Chapter 5: Discussion

The purpose of this chapter is to consider the results presented in Chapter 4 holistically, use these results to attempt to respond to the research questions which motivated the current study, compare outcomes with those of the previous work discussed in Chapter 2, and explore possible motivations for any trends observed. For convenient reference, the research questions that will be discussed in this chapter are repeated below.

Question 1: How is use of *cantara* and *cantase* perceived by speakers of Galician Spanish?

Question 2: Do written and oral production of *cantara* and *cantase* correlate with perception/acceptance of these forms?

Question 3: What social factors, if any, correlate with the use or perception of *cantara* or *cantase* in Galician Spanish?

Question 4: What linguistic factors, if any, condition the choice of *cantara* or *cantase* in Galician Spanish?

Question 5: What relationship, if any, exists between exposure to the Galician language and use of *cantara* and *cantase* in Galician Spanish?

I.   Question 1: Perception of *cantara* and *cantase* in Galicia

Two of the tasks presented in Chapter 3 and analyzed in Chapter 4, the matched-guise task and the acceptability judgment task, were designed to elicit information about perceptions of *cantara* and *cantase* in Galicia. In broad terms, the results of both these tasks indicate that *cantase* appears to enjoy greater prestige in the minds of participants.

These results were particularly strong in the acceptability judgment task, for which significant differences were found between *cantara* and *cantase* ratings for both

Social Use and Correctness. On a Likert scale ranging from -2 to 2, with -2 indicating complete disuse of the form and 2 indicating very frequent use, average ratings for Social Use were 1.16 for *cantase* and 0.71 for *cantara*. Thus, while participants claim to know individuals who would use each of these forms, this claim is stronger for the *cantase* form. On a similar scale, with -2 indicating that a form is completely incorrect and 2 indicating that it is completely correct, *cantase* was considered more correct on average (M = 0.39) than was *cantara* (M = -0.11).

Results of the matched-guise task revealed similar, albeit non-significant, patterns. Despite participants' lack of awareness of the focus on this study, and despite only being provided one guise and therefore having no basis for comparison, the *cantase* guise was considered both more charismatic and more capable on average than the *cantara* guise. The use of z-scores in the analysis of these responses, which normalizes each individual's set of ratings so that the average response of each participant is zero and scores are indicative of how far above or below average a rating is, ensures that the reason for this is not that participants exposed to the *cantase* guise gave higher ratings overall than the other group. In other words, these results were obtained after controlling for possible rater bias.

Because no previous studies exist to my knowledge that address perception of *cantara* and *cantase*, no connections can be drawn to previous results, neither within Galicia nor in other areas of the Hispanic speaking world. However, a plausible hypothesis about why *cantase* may be more positively perceived overall than *cantara* can be drawn from the data in the present study. Because the *cantase* form shares the same function in both Galician and Castilian while the *cantara* form has different functions in

each, it is possible that individuals have a greater uncertainty about when the *cantara*

form may be correctly used in Castilian. This hypothesis is supported by the fact that the

average rating for correctness of the *cantara* form is below zero despite indications that

the form is indeed used by participants' social groups. Additionally, while *cantase* was

favored for the characteristic groups Charisma and Capability, which would be most

associated with professional success and advancement, the differences between the two

forms in terms of Friendliness ratings were negligible. Thus it appears that *cantase* enjoys

greater overt prestige than *cantara* in the region, quite possibly associated with a greater

certainty as to grammatical correctness, while the two are roughly equal in terms of more

intimate, covert prestige where issues of speaking 'correctly' are of less import.

II.   Question 2: Relationship between perception and production

Three tasks were designed to address different aspects of imperfect subjunctive

production: the semi-guided group conversation elicited spoken production of these

forms, while both the fill-in-the blanks activity and the acceptability judgment task

targeted written forms. In the oral data, only 19% of imperfect subjunctive forms (that is,

25 out of 130 total tokens) were in the *cantase* form. This was the only task showing such

a heavy predominance of *cantara* over *cantase*, however, as written production of

cantase reached 42% of all tokens in the fill-in-the-blanks exercise. *Cantase* was also the

form produced in 54% of written corrections, though this percentage varied heavily

depending on the form in the prompt, ranging from a low of 31% *cantase* production

after a *cantara* prime to a high of 83% when *cantase* was in the prompt. Corrections

primed by the conditional, here considered to be a neutral form, featured *cantase* at a rate of 55%.

Given participants' indications in the rating portion of the acceptability judgment task that *cantase* is the more socially used form, these production trends seem somewhat contradictory at first glance. However, there are two important considerations that bear keeping in mind. First, self-reported behaviors such as those elicited through the acceptability judgment task are indicators of explicit participant attitudes but are notoriously unreliable at predicting behaviors (Poplack, 1993). Second, and more satisfactory as an overarching explanation of the data, the tasks employed were designed to vary in both formality and in the amount of attention drawn to the object of study. Thus for example the group conversations were highly informal and did not draw particular attention to any specific element of speech. The fill-in-the-blanks task, however, was more formal in that it was written rather than oral, and it additionally drew attention to verb conjugations. Finally, in correcting statements on the acceptability judgment task, participants engaged in the highly formal task of evaluating and correcting written language and were clearly expected to focus on *cantara* and *cantase* use. This increasing level of formality and focus on form corresponds neatly to an increase in *cantase* production. In this sense, production results do align with the discovery that participants perceive *cantase* as the more overtly prestigious and 'correct' form.

The production results found in this study confirm those of Kempas (2011) with respect to the written use of *cantase*; in carrying out a fill-in-the-blank elicitation exercise throughout Spain, he found rates of *cantase* use in Galicia around 44%, quite similar to the 42% on the corresponding task in the current study. In their study of a corpus of semi-

directed oral interviews of speakers of Galician Spanish, Rojo & Vázquez Rozas (2014) found that just under 25% of imperfect subjunctive forms employed were in the *cantase* form, which also matches up nicely with the 19% rate in group conversations in the present study. Thus it appears that, despite potential limitations associated with a small sample size taken from two specific localities, the trends unveiled in the current body of data may in fact parallel trends for the region at large.

In comparing results to these two previous studies of Galician Spanish, however, it is essential to recognize that both studies found production to be highly varied among individuals, to the point where the differences between members of any one age group, for example, were occasionally as large as the differences between age groups overall. To examine the potential impact of individual variation on the trends just noted, then, Figure 66 and Figure 67 provide a snapshot of individual results for all participants for whom data was available from all tasks. In Figure 66, speaking and fill-in-the-blank (FitB) data is shown as the percentage of imperfect subjunctive produced in the *cantase* form by each individual. Correction data obtained from the acceptability judgment task is separated into three categories based on the form contained in the model sentence, and the data shown for reading indicate the percentage of sentences containing *cantase* for which the protasis was accepted by participants.[25]

---

[25] Included in the reading as 'accepted' are both forms contained in a sentence that was holistically accepted and forms contained in a rejected sentence that were not altered in the produced correction, as the rejection of the sentence in these cases was assumed to be due to a factor other than the protasis form.

Figure 66: Individual variation across tasks--*cantase*
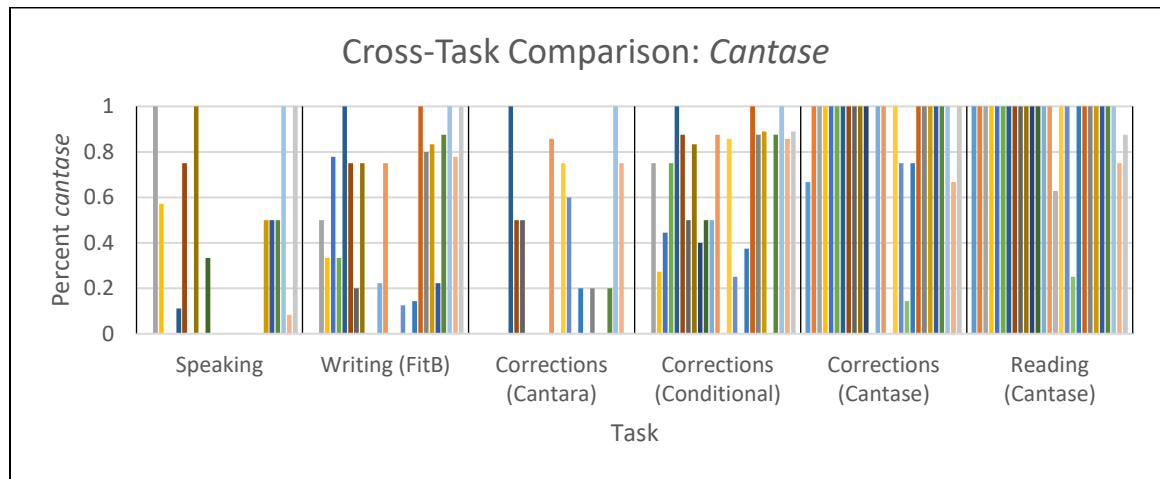
Of particular interest in this graphic are both the relative density (or lack thereof) of bars in the sections corresponding to each task and the quantity of bars that are either fully absent or that reach 100%. Thus for speaking, for example, the relative blankness of the image directly represents the paucity of *cantase* forms produced in the task, while the existence of large gaps with no bars corresponds to those participants who produced no spoken *cantase* forms. Four participants, however, produced exclusively *cantase* forms in their spoken language. The greater number of bars in both the fill-in-the-blanks and the corrections with conditional prime corresponds to more use of the *cantase* form, while the fact that few of these bars reach ceiling indicates that most participants produced both *cantara* and *cantase* forms for these (sub)tasks. Finally, the large proportion of bars that reach 100% in the reading and corrections with *cantase* prime sections reveal that, despite widespread variation in production rates, the *cantase* form is widely accepted by nearly all participants.[26]

---

[26] Although the Corrections (Cantase) data shown in Figure 66 were obtained through a production task, their use as a measure of form perception is justified since the production of a *cantase* form in reformulation of a sentence containing the same form implicitly implies that the participant views the form as correct.

Figure 67 contains similar data to Figure 66, though the values have been inverted to indicate percentages relative to *cantara* rather than *cantase*.[27] Of particular note here is the greater presence of *cantara* in the spoken and fill-in-the-blanks production of many participants, as well as the more sparsely populated sections near the right end of the image. Taken holistically, this indicates that a greater production of *cantara* over *cantase*, particularly in less formal tasks, is indeed the norm not only when averaged across the whole sample but also for the majority of individual participants. With increasing formality, however, comes an increase in *cantase* production, evidenced by the greater density of the Corrections (Conditional) section in Figure 66 as compared to Figure 67. Finally, while both *cantara* and *cantase* strongly tend to be accepted by the majority of participants, the number of individuals for whom this acceptance is at ceiling is greater for *cantase* than *cantara*, again supporting the conclusion drawn on the sample as a whole that *cantase* is the variant more strongly associated with correctness and overt prestige.

---

[27] The consideration of only two forms in this study makes this conversion a simple one: if a participant produced *cantase* 75% of the time on a task, the percentage of *cantara* forms is therefore $1 - 0.75 = 0.25$, or 25%. The only task for which this was not the conversion process was reading, which was calculated analogously to the reading percentages in Figure 66 using sentences containing *cantara* rather than *cantase*. Additionally it should be noted that the positions in the figure of corrections data for *cantase* and *cantara* were reversed so as to display an analogous transition from more intentional to more passive 'corrections' in both figures.

Figure 67: Individual variation across tasks--*cantara*
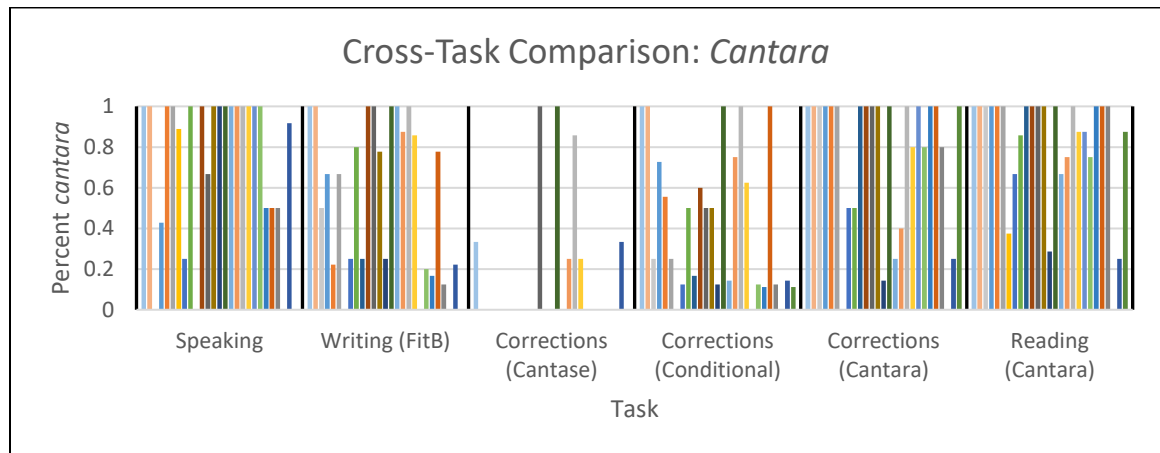
In addition to corroborating the conclusions drawn so far in this chapter, the data on individual behaviors and attitudes in Figure 66 and Figure 67 support the previous assertions of both Kempas (2011) and Rojo and Vázquez Rozas (2014) on the prevalence of individual variation in production. On each production task, each form was categorically used by at least one participant. Much less variation existed with respect to form acceptance, however. This conclusion, unattested to my knowledge in previous literature on the Spanish imperfect subjunctive forms, was not possible to reach based on the data in either Kempas (2011) or Vázquez Rozas (2014), as both studies were limited to an examination of production alone.

III.   Question 3: Social factors related to imperfect subjunctive

Though the impacts of Gender, Age, Location, and Education on imperfect subjunctive perception and production were seldom found to be statistically significant, tendencies with respect to these variables were remarkably uniform across all tasks. Roughly 1 in 3 spoken forms produced by women were *cantase*, while the rate of spoken

use of this form for men was only 19%. Similarly, while both genderes followed the trend observed in the previous section toward increased use of *cantase* on the fill-in-the-blanks task, women again led in production of this form (45% as compared to 38% use of *cantase* by men). No differences between genderes, significant or otherwise, emerged from the acceptability judgment task. The matched-guise task, however, revealed that women viewed the *cantase* form as both more charismatic and less friendly than *cantara*, while men considered *cantase* the more friendly of the two forms. This attitude toward *cantase* on the part of women may serve to reinforce their increased use of the form; women more so than men tend to produce overtly prestigious variants even at the expense of covert prestige (Labov, 1990), and the association of *cantase* with charismatic traits such as leadership and confidence may make women seeking to advance professionally keener to produce this form in their own language.

Part of the overall overt prestige of *cantase* may stem from its association with higher levels of education. Despite failing to reach significance, education appeared to play a large role in favoring the *cantase* form, with individuals with higher education experience producing it far more in both conversation and the Fill-in-the-blanks task than those with only primary or secondary education. Similarly, *cantase* was somewhat favored by more educated speakers in correction productions except in the presence of a *cantara* prime. When such a prime was present, only those with graduate studies showed a tendency to replace it with *cantase,* though these individuals did so roughly 50% of the time (compared to an overall rate of 31% replacement of a *cantara* prime with *cantase*). Finally, all participants save those with primary education alone indicated that *cantase*

was both more correct and more socially prevalent than *cantara*, again reinforcing *cantase* as the form of prestige and correctness.

While ratings of correctness or social use did not appear to be related to participant age, younger participants exhibited the same tendency as women and more educated individuals to favor *cantase* over *cantara* in spoken language, the fill-in-the-blanks task, and conditional-primed corrections, while older participants heavily preferred the *cantara* form in these same contexts. Interestingly, however, it was older speakers in this case who associated the *cantase* guise with the more overtly prestigious characteristic groups Capability and Charisma and saw the form as less friendly. Younger participants, in turn, evaluated *cantase* as the friendlier but less capable of the two forms; both forms were seen as equally charismatic at the younger end of the age spectrum. It is unsurprising that youth, who in general are those with higher levels of education, would favor the prestige form *cantase* taught in schools (Blas Arroyo, 2008). More surprising is that they would associate this form with greater friendliness than *cantara,* which in general appears to be the more covertly prestigious form. It is possible that the increased production of *cantase* by youth in general counteracts to some extent the attitudinal differences toward the two forms; if *cantase* is now the normal form among youth, it may be less associated with unfriendliness than for older generations, though this would not perhaps explain why it would actually surpass *cantara* in ratings of Friendliness rather than simply drawing even. This latter may be due to a single outlier, as the youngest participant also produced by far the lowest rating of *cantara* friendliness (over 0.2 standard deviations lower than the next lowest rating); removing this single data point

brings the trends in Friendliness ratings of the two forms to nearly equal for the youngest remaining participants.

Finally, the results of all three production tasks coincided in indicating that the *cantase* form is produced at a higher rate among participants from O Grove than among those from Marín. In spoken data in particular, *cantase* was produced twice as frequently in O Grove (45%) as in Marín (21%); these rates rose to 58% for the Fill-in-the-blanks activity and 64% for conditional-primed corrections, while Marín continued to lag behind in *cantase* production rates for both tasks (35% and 52%, respectively). Interestingly, while no differences appear to exist between locations with respect to evaluation of guise friendliness or charisma, individuals from Marín associated *cantase* with greater capability, while those from O Grove considered *cantara* the more capable form.

Given that the other social variables just discussed, and particularly gender and education, appear to indicate that *cantase* is the more overtly prestigious form, it seems somewhat contradictory that this same form would be favored in O Grove, a municipality which is still heavily linked to traditional trades such as fishing and also farther from the city of Pontevedra, thereby limiting mobility practices (*e.g.* Britain, 2010) of its inhabitants. Unfortunately, the explanation for this unexpected trend likely lies in a limitation of the data set, namely, the overrepresentation of more highly educated speakers from O Grove as compared to Marín. Data from the 2001 census from the Instituto Galego de Estatística indicate that Marín is a slightly more educated populace on average than O Grove; however, 73% of participants from O Grove had at least some university education, as compared to 44% of participants from Marín. It appears likely,

then, that the preference for *cantase* among participants from O Grove is essentially a reproduction of the preference for *cantase* among more educated participants.

The results of this study with respect to social variables do not consistently align with those of any previous study, whether within Galicia or elsewhere. Lavandera (1975) associated *cantase* with men, the middle aged, and non-college-educated speakers in Buenos Aires, all of which are in direct opposition to the results presented here. Navarro (1990) found that *cantase* was correlated with older, more educated, and wealthier individuals in Venezuela; while the current data does not permit a comparison of socioeconomic class based on wealth, only the relationship between education and *cantase* can be shown to hold true for the current data set. Within Spain, Kempas (2011) found that it was predominantly men who used the *cantase* form, and this result was backed up by Rojo and Vázquez Rozas (2014)'s finding that men and youth favored the *cantase* form within Galicia. While the current results support the latter finding with respect to age, the discrepancy between these two studies and the present analysis with respect to gender is intriguing. On the one hand, both the current study and Rojo and Vázquez Rozas (2014) have relatively low token counts for oral production, making statistical analysis potentially unreliable, and in any case gender was not a significant correlate of spoken production in the present study while Rojo and Vázquez Rozas (2014) did not appear to submit their data to statistical testing. On the other hand, for men to use *cantase* more frequently in spoken language does not necessarily contradict the present finding that men evaluate *cantase* as friendlier than *cantara*. Finally, given the differing attitudes between men and women toward the two forms uncovered in the present study, it is possible that different results may also be related to different speech

contexts. While the current study observed spoken speech in the context of informal group conversations, very little detail is given in Rojo and Vázquez Rozas (2014) on the nature of the interviews they examined.

IV. Question 4: Linguistic correlates of imperfect subjunctive form

Seven linguistic variables were considered in this study, of which three—possibility, priming, and lexical item—were applicable to multiple tasks. Polarity, Clause type, and Anteriority were only considered with respect to oral production. The small number of oral tokens makes speculation about these three variables highly tentative, particularly since the differences between, for example, negated statements and non-negated statements were slight in comparison to the large differences between *cantara* and *cantase* production overall in this task. However, average rates of use indicated a very slight preference for using *cantase* with negated statements—logical given that *cantase* was also favored in contrary-to-fact statements as will be discussed shortly—and a very slight dispreference for this form to express anteriority, which is also unsurprising if *cantara* in the region's Castilian still holds some association with its simple pluperfect use in Galician. This goes hand in hand with one possible explanation for the association between imperfect apodosis forms and *cantara* in the acceptability judgment task, as increased salience of past reference may make the *cantara* form more adequate for expression of this nuance, though as stated in Chapter 4 it is equally likely that this association is merely the result of few corrections containing the imperfect in the apodosis.

Anecdotally, the anterior association of *cantara* does appear to exist for some individuals, as indicated by the following unsolicited comment from a male participant in Marín:

> *A veces que uno vaya hablar se extraña por el uso sobre todo del*
>
> *pluscuamperfecto (?) y de repente empiezo a decir e:m estuviera o pasara o*
>
> *cantara o- (.) en lugar de había cantado. Y ya te dicen "¿Cómo? Espera que no*
>
> *pillo yo lo que estás diciendo." Pero yo no soy consciente de que lo hago.* [MM8,
>
> recording MF8M8, 10:41]

However, as the small number of existing tokens of *cantara* with pluperfect indicative reference were excluded from consideration in the current study, no further detail on the frequency or on factors associated with this type of use is available.

In terms of clause type, oral data revealed elevated rates of *cantase* use in both noun clauses and hypothetical statements, despite the fact in the latter case that this was the very context in which the *cantara* form is thought to have initially encroached on the imperfect subjunctive (see, for example, Veiga (1996)) and thus could reasonably be expected to be among the most advanced contexts for the disappearance of the *cantase* form. At least two possible explanations for the retention of *cantase* in these statements exist. First, unlike other dependent clause types that can differ greatly in form, hypothetical constructions follow a generally formulaic pattern of "*si* + imperfect subjunctive, conditional" whose regularity may promote the maintenance of traditional forms disfavored in less rigid constructions.[28] Second, and equally probable given social

---

[28] Evidently other verb tense combinations are possible and frequent with this type of construction, but all have in common that there are relatively rigid patterns of which forms can be combined and in which contexts.

changes in the region in the last 40 years, any in-progress evolution with respect to *cantara* and *cantase* use in Galicia may no longer be the result of the same phenomena that led *cantara* to gain its subjunctive associations and *cantase* to lose ground in other parts of the Spanish-speaking world. Rather, impetus for current developments (or the lack thereof) in Galician Spanish may arise from contact with Galician; this possibility will be explored further in Chapter 6. In either case, it is possible that the association of *cantase* with hypothetical clauses may contribute to the tendency for this form to be more prevalent in tasks other than the group conversation (*i.e.* Figure 66), as each of the other tasks were created specifically to examine these clauses.

Contradictory results emerged with respect to association of *cantase* with any specific verb group. In oral production, on the one hand, *cantase* was favored in –*ar* and –*ir* verbs and disfavored for *ser*, *tener*, and other –*er* verbs.[29] In the fill-in-the-blanks exercise, on the other hand, -*ir* disfavored *cantase* production, which was favored for –*ar* and *ser*. No trend was found in either part of the acceptability judgment task. While not conducive to providing a firm argument about association of either form with specific verb groups or with high versus low frequency items, contradictory results about the verbs commonly paired with either form also exist in the literature (*cf.* Hermerén & Lindvall, 1989; Rosemeyer & Schwenter, 2017). Thus it appears that lexical item, if it conditions use of *cantara* and *cantase* at all, may do so only weakly.

Clearer indications were found for an association of *cantase* with contrary-to-fact statements; this connection was only tentative in the spoken data, but approached

---

[29] Here "favored" should not be taken to mean that *cantase* was produced more than *cantara*, but rather that *cantase* was produced at a higher rate (*e.g.* 23% for –*ar* verbs) than its overall average across all forms (19%).

significance on the evaluative portion of the acceptability judgment task. This supports

the assertion by Bolinger that "-*se* implies 'remoteness, detachment, hypothesis, lack of

interest, vagueness, greater unlikelihood,' while *–ra* brings everything into relatively

sharper focus" (1956, p. 346). Additionally, an association of unlikelihood/impossibility

with *cantase*, combined with the "Principle of Distance" discussed by Silva-Corvalán

(1984), would account for the diminishing use of *cantase* in both less formal language

within Galicia and in all contexts throughout the Spanish-speaking world. This also

aligns with arguments about forms referring to unreal contexts in Lavandera (1975),

though both she and Silva-Corvalán (1984) made their arguments with respect to the

extension of the conditional at the expense of *cantara* rather than with regard to the two

imperfect subjunctive forms.

Perhaps the most influential linguistic factor in conditioning form choice is

Priming, which either was significant or approached significance in each of the three

tasks in which it was considered. Essentially, the existence of either *cantara* or *cantase* in

the context preceding an imperfect subjunctive form heavily favored replication of the

primed form, while the absence of any prime (or the presence of the conditional in a

protasis) was relatively neutral in conditioning form choice. Despite Rosemeyer and

Schwenter (2017)'s finding that priming had a greater effect on *cantase* than on *cantara*,

which they attributed to *cantase*'s tendency toward obsolescence, the present study found

strong effects for priming on both forms. However, a greater distance from the prime did

result in a reduced effect of the prime for *cantase*, in line with the former study's

findings. Additionally, the exceptional strength of priming in conditioning form choice in

corrections on the acceptability judgment task can be explained with reference to the

finding of Pickering and Branigan (1998) that lexical repetition, an inherent aspect of the acceptability judgment task from the present study, can greatly strengthen the influence of structural priming. In sum, priming is an essential variable to consider in future studies of imperfect subjunctive production and may even be a confounding factor in some studies if not properly controlled for. On the extreme end of the spectrum, however, no effect of guise form, considered as a possible prime for the fill-in-the-blanks task, was encountered in this study, despite findings by Bock and Griffin (2000) that priming effects can endure even over substantial intervening material.

Finally, it is essential to recognize that the gold standard for sociolinguistic data is to examine natural language use in context. This was attempted as far as possible through the semi-directed group conversations. However, as mentioned previously, obtaining sufficient numbers of tokens from spontaneous speech presents a particular challenge in the analysis of syntactic variation (Sankoff, 1982). Ideally, in addition to the broad strokes of analysis already presented, a satisfactory exploration of the difference in *cantara* and *cantase* would involve an in-depth discursive analysis of the spontaneous speech elicited in an attempt to tease out any nuances that may condition form choice. However, in the end this wasn't feasible with the current data set. Of the 39 individuals who took part in the study, only eight produced both forms in their spoken language. A contrastive analysis of form meaning is only informative for those participants who produced both forms and thus may have communicated some difference of meaning through form choice; thus the sample size over which such an analysis could be carried out is greatly reduced. In addition, of these eight participants, only five produced more than two imperfect subjunctive tokens in total, and only one participant produced more

than one token of each form. Because of this, while acknowledging that a detailed analysis of this kind is essential for a full understanding of imperfect subjunctive use in Galicia, such an analysis is a challenge to be met by future research endeavors.

V.   Question 5: Influence of Galician language experience

In general, it is difficult to identify reliable results with respect to the language experience variables explored in this study due to the extremely high prevalence of balance between Castilian and Galician across all domains save current language use, which did not correlate noticeably with the dependent variable on any task. Without more information about behaviors of individuals on the extremes of language experience, the few results herein must be tentative at best. That said, elevated production of *cantase,* particularly in spoken language, seemed in general to be concentrated near the point of balance between the two languages when any trend existed at all.

In extremely general terms, greater experience with Galician appeared to go hand in hand with a slight preference for *cantase* production in writing. On the fill-in-the-blanks task, Galician use at work and in the family, written competence in the language, and more positive attitudes toward Galician relative to Castilian all resulted in slightly elevated *cantase* use. In production of corrected statements on the acceptability judgment task, having Galician as the first language favored *cantase* production across all three priming conditions. Additionally, increased experience with Galician in the family or at work, while showing no trends for *cantase-* or *cantara*-primed tokens, did correlate with an increase in *cantase* production with conditional primes.

With respect to perception, those with greater relative experience in Galician in various settings, as well as those with greater oral capacity and more positive affect toward the language generally rated *cantase* as more friendly but less charismatic and capable than their Castilian-dominant counterparts. The only trend that emerged with respect to these variables on the acceptability judgment task was a correlation between more positive attitudes toward Galician and an increased perception of *cantase* as a correct form.

Again, the indication here of a possible relationship between the Galician language and more favorable production and perception of the *cantase* form should be considered tenuous due to the limited amount of variation among the sample population with respect to these variables. It does appear plausible, however, that the association of *cantase* with Galician may be indicative of a greater uncertainty for bilinguals in navigating appropriate use of the *cantara* form in their two languages, particularly as research has shown that bilinguals have both of their languages (and thus in this context both competing functions of the *cantara* form) activated when producing and processing either language (Kroll & Bialystok, 2013).

Having discussed the results of the first five research questions, I now turn to consideration of language contact as a possible explanation for these results to conclude this study.

Chapter 6: Conclusion

Having applied the data presented in Chapter 4 to answer five research questions specific to imperfect subjunctive use in Galicia in Chapter 5, I now attempt to tie these findings together through an examination of their implications for a language contact argument. I end this chapter with a discussion of the present work's contributions to the field of language contact, an admission of some of its limitations, and suggestions for future work in this area.

I.   Results in the light of language contact

Identifying effects of language contact, complex enough even in cases of non-related languages, is anything but trivial when the two languages in question are typologically similar and are directly descended from the same language. The following two conditions, proposed by Silva-Corvalán (1994), were discussed in Chapter 2 as possible indications of transfer between languages:

- When two languages X and Y share a form that is structurally similar but with different functions in X and Y, the function of the form in language X may become associated with the parallel form in language Y, even if language Y already has a different form with the same function.

- When two forms in language Y are in competition, a contact-variety of language Y may have more frequent use of the form that is most similar to that of language X, as compared to a possibly categorical or highly preferred alternate form in non-contact varieties.

More specific to the context of the imperfect subjunctive in Galicia, these

conditions could be read as follows:

1. The *cantara* form in the Spanish of Galicia may take on the perfective meaning of the Galician form.

2. The *cantase* form of imperfect subjunctive, far less frequent than the *cantara* form in most varieties of Spanish, may be more common in Galicia.

While the focus of this thesis was on the second possibility, it is interesting to note in passing that there is some indication that the first also holds for some speakers, and that additionally this association of the *cantara* form with anteriority does not appear to be limited to the subjunctive. Whether these anterior associations are the result of the influence of Galician on the Spanish of the region or whether this association is an anachronism held over from Latin is a question certainly worthy of pursuit, albeit one that lies beyond the scope of this thesis.

As to whether the second of the two conditions is attested in the present data, the answer is a resounding affirmative. This study found a rate of 19% use of *cantase* in oral production. While this may not seem elevated when compared to the 18% rate cited by Rojo (1996) or the 23% rate in Kempas (2011), both of these studies focused solely on written language. In that sense rates in Galicia do appear to be elevated, with 42% on the fill-in-the-blanks activity (in line with results from Kempas (2011) for Galicia alone) and 55% of corrections to conditional-primed tokens in the acceptability judgment task in the *cantase* form. Additionally, acceptability of the *cantase* form is even higher than its production rate for individuals at all points of the social variable spectrum. Thus it appears highly likely that language contact does indeed have some effect on the variety of Spanish spoken in Galicia.

However, the question remains whether this influence from Galician is in the form of a perpetual conserving effect, that is, a slowing of the tendency toward obsolescence of the *cantase* form attested in most other varieties of Spanish that would possibly have been in play for centuries, or whether there is evidence of a more recent impact of the Galician language. Here, evidence suggests that both may be contributing factors. The all-around acceptance of *cantase* as an imperfect subjunctive form regardless of participant age, education level, or language experience indicates that the form endures from earlier forms of the language rather than being a recent innovation.

On the other hand, if the use of *cantase* were solely a conservative vestige of older language forms, it would be expected that its prevalence in produced language would continue to diminish, however slightly, in apparent time, in parallel to trends in other parts of the Spanish-speaking world such as Buenos Aires (Lavandera, 1975) and Venezuela (Navarro, 1990). This expectation does not hold on the current data set, however, as it was younger speakers who produced *cantase* most frequently, while older individuals preferred *cantara*. Additionally, despite the problems discussed in the previous chapter with respect to language experience variables, increased preference for *cantase* as associated with Galician language experience only appeared to hold approximately to the point of balance between Galician and Castilian, with few reliable trends noted for those with relatively little Castilian experience. Were sustained language contact alone an explanation for an increased use of *cantase*, we might not only expect the form to be more prevalent among older speakers, but also among those with increased exposure to the Galician language. Neither of these appears to be the case based on current data.

Instead, and somewhat at odds with what one would expect from an obsolescing form, the previous chapter indicated that women, youth, and the educated are those who show preference for the *cantase* form, from which it is possible to infer that the form may be resurgent even if it was at one point diminishing in Galician Spanish.[30]

To provide a possible explanation for a potential resurgence of *cantase*, particularly among youth and the educated, it is useful to take into account that education was not highly prioritized in the Franco era, as illustrated anecdotally by the following exchange between myself and two participants:

> MF10: Ir a la escuela suponía aprender solo las cuatro reglas básicas. Eh leer, sumar, restar, y dividir. Y ya no era necesario nada más. Entonces ya cuando un niño o una niña llegaba a esas cuatro reglas […] Ya no era necesario que fuera más al colegio.
> […]
> MM9: Después a trabajar ya. […] Porque España (.) era mucho mano de obra que hay que- Franco quería manos de obra no quería estudios. Manos de obra. Manos de obra.
> I: Claro. Es más difícil controlar a la gente educada.
> […]
> MF10: Claro. Es que eso es el problema. El problema es que
> MM9: Los estudios-
> MF10: No le interesaba en absoluto que-
> MM9: Que la- que la gente supiera.
> MF10: Que la gente supiera. [MF91011M9 25:00]

Thus increased levels of formal education in both languages, but particularly in Galician, followed after the end of the dictatorship in 1975. With this came increased instruction in the 'correct' way to speak both languages, and the overlap of the function of the *cantase* form in Castilian and Galician may have reinforced its use in both languages. Anecdotally this appears to be the case for some individuals, as a contact of mine from

---

[30] This would also explain the discrepancies between results of this study and those from Rosemeyer and Schwenter (2017) in terms of priming; these authors argue that a larger priming effect exists for the *cantase* form than for *cantara* because the former is obsolescing, while the present study found equally strong priming effects for both forms.

the region reported being advised in school to avoid *cantara* completely so as to not mistakenly employ it as pluperfect indicative in Castilian.

As to why this resurgence should center primarily on contrary-to-fact forms and hypothetical statements, the explanation is similar. It is precisely at the higher levels of education where students are most often exposed to hypothetical and contrary-to-fact thought processes, and if the use of the *cantase* form is promoted at essentially the same time as critical higher-order thinking is emphasized, it seems natural that the one should reinforce the other. This reinforcement also makes sense from a cognition standpoint, as Matras (2010) argues that grammar forms such as modality that correspond to complex thought processes are cognitively vulnerable and as such may be particularly susceptible to transfer effects. Thus, in an effort to reduce the processing load associated with discussing hypothetical and unreal content, participants may prefer use of a form that does not add to their cognitive uncertainty. Finally, recurring to changing education patterns as a likely primary cause of increased *cantase* use also supports the strong tendency documented in this study for this form to be increasingly accepted and employed at the expense of the *cantara* form in more formal tasks that involve greater attention to form.

In summary, while contact with Galician undoubtedly has played a role in the maintained acceptance of the *cantase* form that has become widely obsolete in other parts of the Spanish-speaking world, contact alone is insufficient to explain the demographic trends associated with its use. When considered in conjunction with social changes in the linguistic and educational norms of the Galician community, however, a reasonably coherent 'big picture' comes into view.

II.   Contributions, limitations, and directions for future study

Within the field of language contact more broadly, the present study serves to reemphasize the importance of considering social and cultural context in addition to linguistic features of the languages. It appears that changes in mobility practices (*i.e.* Britain, 2010) can lead to changes in "habit[s], routin[es], convention[s] and institution[s] of human practices", by extension producing consequences related to the situation-transcending properties of language (Linell, 2009, p. 50). This consideration of social practices, naturally important for any sociolinguistic study of languages in contact, is clearly even more essential when the languages in question are closely related, as linguistic arguments alone are insufficient to distinguish between internal (drift) and external (contact) motivations for language change.

While the present thesis has attempted to contribute to current understanding not only of imperfect subjunctive use in Galicia but also to the study of contact between related languages more broadly, much remains to be done in both. With respect to the former, a more comprehensive corpus of oral data is clearly needed to inform an in-depth discursive analysis on the potential implications of form choice. Additionally, a study of both perception and production that involves a random sample of participants with more diverse linguistic experiences and would enable clearer conclusions to be drawn about social factors in general and about the potential impact of exposure to Galician outside the classroom in particular. Finally, in order to truly paint a complete picture of imperfect subjunctive patterns in the region, a sociolinguistic study on the distribution of these forms in Galician itself is badly needed.

In theoretical terms, this study is able to contribute a statement about the general importance of the extralinguistic societal context in potentially conditioning linguistic practices. However, it remains unclear if similar importance would be found were the forms under consideration not so intimately linked to higher-level thinking and education; in this sense, examining a variety of linguistic forms of varying complexities acquired at different stages of development would be informative. Finally, while not an explicit focus of this study, examining the linguistic practices of *neofalantes*—"New Speakers"—of minority languages in comparison to more mainstream speakers could lead to fascinating information about how language attitudes enhance or inhibit the effects of social context on linguistic production. It is my hope that the current study serves as a base for such considerations in the future.

References

Asratián, A. (2007). Variación –ra/-se en el español hablado en Caracas. *Boletín de lingüística XIX*(27), 5-41.

Beltrán Lloris, F. (2004). El latín en la Hispania romana: una perspectiva histórica. In Cano Aguilar (Ed), *Historia de la lengua española* (pp. 83-106). Barcelona: Ariel.

Becker, M. G. (2008). From temporal to modal. Divergent fates of the temporal Latin synthetic pluperfect in Spanish and Portuguese. In Ulrich Detges and Richard Waltereit (eds.): *The Paradox of Grammatical Change. Perspectives from Romance* (pp. 147-179). Amsterdam: John Benjamins.

Bejarano, V. (1962). Sobre las dos formas del imperfecto de subjuntivo y el empleo de la forma en –se con valor de indicativo. *Strenae. Estudios de filología e historia dedicados al profesor Manuel García Blando (Acta Salmanticensia, 16)*, 78-86.

Beswick, J. (2002). Galician language planning and implications for regional identity: restoration or elimination?. *National Identities, 4*(3), 257-271. doi: 10.1080/1460894022000026123

Beswick, J. (2010). Linguistic ideology in Galician factual radio broadcasts: Towards the castilianisation of Galician pronunciation traits?. *Spanish at work: Analysing institutional discourse across the Spanish-speaking world. London, Palgrave Macmillan*, 35-49.

Beswick, J. (2014). Borders within borders: Contexts of language use and local identity configuration in southern Galicia. In D. Watt & C. Llamas (Eds.), *Language, borders and identity* (pp. 105-117). Edinburgh: Edinburgh University Press.

Birdsong, D., Gertken, L.M., & Amengual, M. (2012). *Bilingual language profile: An easy-to-use instrument to assess bilingualism*. COERLL, University of Texas at Austin. https://sites.la.utexas.edu/bilingual

Blas Arroyo, J.L. (2008). *Sociolingüística del español.* Madrid: Cátedra.

Bock, K. & Griffin, Z.M. (2000). The persistence of structural priming: Transient activation or implicit learning? *Journal of experimental psychology, 129*(2), 177-192.

Bolinger, D.L. (1956). Subjunctive –ra and –se: "Free variation"?. *Hispania, 39*(3), 345-349.

Bouzada-Fernández, X. M. (2003). Change of values and future of the Galician language. *Sociolinguistic Studies*, *4*(1), 321-341. doi: 10.1558/sols.v4i1.321

Bowern, C. (2010). Fieldwork in contact situations. In R. Hickey (Ed.), *The handbook of language contact* (pp. 340-357). Oxford: Wiley-Blackwell.

Britain, D. (2010). Contact and dialectology. In R. Hickey (Ed.), *The handbook of language contact* (pp. 208-229). Oxford: Wiley-Blackwell.

Buzelin Haro, C.M. (2010). Prótasis de imperfect de subjuntivo en oraciones condicionales de textos de Córdoba de los siglos XVI-XVII e inicios del XXI. In V. M. Castel and L. Cubo de Severino (eds): *La renovación de la palabra en el bicentenario de la Argentina. Los colores de la mirada lingüística* (pp. 245-251). Mendoza: Editorial FFyL.

Carbolová, K. (2010). La situación sociolingüística de Galicia. (Master's thesis). Masaryk University, Brno, Czech Republic.

Carbonero Cano, P. (1990). Usos de las formas verbales en –ra y en –se en el habla de Sevilla (nivel popular). In P. Carbonero Cano & M. T. Palet Plaja (Eds.), *Habla de Sevilla y hablas americanas* (pp. 45-58). Sevilla: Universidad de Sevilla.

Costa, J. (2015). New speakers, new language: on being a legitimate speaker of a minority language in Provence. *International journal of the sociology of language 231*, 127-145.

Croft, W. (2010). *Explaining language change: An evolutionary approach.* Harlow: Longman.

Day, M. (2011). *Variation in the use of the –ra and –se forms of the imperfect subjuncive in Modern Spoken Peninsular Spanish.* NWAV 40, Georgetown University.

del Valle, J. (2000). Monoglossic policies for a heteroglossic culture: Misinterpreted multilingualism in modern Galicia. *Language & Communication 20*, 105-132. doi: 10.1016/S0271-5309(99)00021-X

Domínguez-Seco, L. (2002). Social prestige and linguistic identity. On the ideological conditions behind the standardisation of Galician. *Sociolinguistic Studies*, *3*(2), 207-228. doi: 10.1558/sols.v3i2.207

Epps, P., Huehnergard, J., & Pat-El, N. (2013). Introduction: Contact among genetically related languages. *Journal of language contact, 6*(2013), 209-219.

Espinoza, A. (1930). The use of the conditional for the subjunctive in Castilian popular speech. *Modern Philology* 27, 445-449.

Even-Zohar, I. (2008). Culture planning, cohesion, and the making and maintenance of entities. *Benjamins Translation Library 75*, 277-292. doi: 10.1075/btl.75.22eve

Faginas Souto, S. (2001). A interferencia fonética no español da Coruña: a vocal [o] tónica. *Actas do I simposio internacional sobre o bilingüismo* (pp. 686-698). Vigo: Universidad de Vigo.

García de Cortázar y Ruiz de Aguirre, J. (2004). Resistencia frente al Islam, reconquista y

repoblación en los reinos hispanocristianos (años 711-1212). In Cano Aguilar (Ed), *Historia de la lengua española* (pp. 239-258). Barcelona: Ariel.

García, C. (1986). El castellano en Galicia. En M. Alvar, M. Etxebarria, C. García & M. Marsá (eds.): *El castellano actual en las comunidades bilingües de España*. Salamanca: Junta de Castilla y León, 49-64.

García de Diego, V. (1951). *Gramática histórica española.* Madrid: Gredos.

Gili Gaya, S. (1973). *Curso superior de sintaxis española.* Barcelona: Bibliograf.

González González, M. (1985). La recuperación del gallego. *Revista de filología románica, 3*(1985), 101-120.

González González, M., Rodríguez Neira, M., Fernández Salgado, A., Loredo Gutiérrez, X., & Suárez Fernández, I. (2007). *Mapa sociolingüístico de Galicia 2004. Volume I. Lingua inicial e competencia lingüística en Galicia.* A Coruña: Real Academia Galega.

González González, M., Rodríguez Neira, M., Fernández Salgado, A., Loredo Gutiérrez, X., & Suárez Fernández, I. (2008). *Mapa sociolingüístico de Galicia 2004. Volume II. Usos lingüísticos en Galicia.* A Coruña: Real Academia Galega.

Gutiérrez, M. J. (1996). Tendencias y alternancias en la expresión de condicionalidad en el español hablado en Houston. *Hispania, 79*(3), 567-577.

Hermerén, I., & Lindvall, L. (1989). La distribución de las formas en ra y se en un texto español moderno. *Moderna sprak,* 83, 34-42.

Hornberger, N. H. & King, K. A. (1996). Language revitalization in the Andes: Can the schools reverse language shift? *Journal of Multilingual and Multicultural Development*, 17, 427-41.

Kempas, I. (2011). Sobre la variación en el marco de la libre elección entre cantara y cantase en el español peninsular. *Moenia 17*(2011), 243-264.

King, K.A. & Hermes, M. (2014). *Why is this so hard?* Ideologies of endangerment, passive language learning approaches, and Ojibwe in the U.S. *Journal of Language, Identity and Education*,*13* (4), 268-282.

Kovácová-Moman, V. (2007). *A mudança da língua usual nos novos locutores de galego – neofalantes*. Santiago de Compostela: GZe-editora.

Kroll, J.F. & Bialystok, E. (2013). Understanding the consequences of bilingualism for language processing and cognition. *Journal of cognitive psychology, 25*(5), 497-514.

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

Press.

Labov, W. (1990). The intersection of gender and social class in the course of linguistic change. *Language variation and change, 2*(1990), 205-254.

Lambert, W.E., Hodgson, R.C., Gardner, R.C., & Fillenbaum, S. (1960). Evaluational reactions to spoken languages. *The journal of abnormal and social psychology, 60*(1), 44-51.

Lavandera, B. R. (1975). *Linguistic structure and sociolinguistic conditioning in the use of verbal endings in si-clauses (Buenos Aires Spanish)* (Doctoral dissertation, Graduate School of Arts and Sciences, University of Pennsylvania).

Linell, P. (2009). *Rethinking language, mind, and world dialogically: Interactual and contextual theories of human sense-making.* Charlotte, NC: Information Age Publishing, Inc.

Lloyd, P. (1987). *From Latin to Spanish.* Philadelphia: American Philosophical Society.

Loureiro-Rodríguez, V., Boggess, M.M., & Goldsmith, A. (2012). Language attitudes in Galicia: Using the matched-guise test among high school students. *Journal of multilingual and multicultural development, 34*(2), 136-153.

Lunn, P. V. (1995). The evaluative function of the Spanish subjunctive. In L. Bybee & S. Fleischman (Eds.), *Modality in grammar and discourse* (pp. 429-449). Amsterdam: John Benjamins.

Luquet, G. (1988). *Systematique historique du mode subjonctif espagnol.* Paris: Klincksieck.

Marín, F. (1979). Observaciones sobre las construcciones condicionales en la historia de la lengua española. *Nueva revista de filología hispánica 28*: 86-105.

Matras, Y. (2010). Contact, convergence, and typology. In R. Hickey (Ed.), The *handbook of language contact* (pp. 66-85). Oxford: Wiley-Blackwell.

Monteagudo, H. (2012). *Facer país co idioma: Sentido da normalización linguística.* Santiago de Compostela: Real Academia Galega.

Moreno de Alba, J.G. (1978). *Valores de las formas verbales en el español de México.* México: UNAM.

Moscovici, S. (1976). *Social influence and social change.* London: Academic Press.

Muysken, P. (1981). Halfway between Quechua and Spanish: The case for relexification. In A.R. Highfield & A. Valdman (Eds.) *Historicity and variation in creole studies* (p. 52-78). Ann Arbor: Karoma Publishers.

Navarro, M. (1990). La alternancia-ra/-se y-ra/-ría en el habla de Valencia (Venezuela). *Thesaurus-Instituto Caro y Cuervo*, *45*(2), 481-488.

N'Diaye, G. (1970). *Structure du dialecte basque de maya.* The Hague: Mouton.

Negro Romero, M. (2013). Contacto galego-castelán e cambio no léxico do corpo humano. In E. Gugenberger, H. Monteagudo, & G. Rei-Doval (eds.), *Contacto de linguas, hibrididade, cambio: contextos, procesos e consecuencias* (221-244). Santiago de Compostela: Consello da Cultura Galega.

Nowikov, N. (1984). El valor doble de la forma en –SE en el español peninsular y americano. *Ibero-americana Pragensia* 18, 61-66.

O'Rourke, B., & Ramallo, F. (2011). "The native-non-native dichotomy in minority language contexts. Comparisons between Irish and Galician". *Language Problems and Language Planning*. Vol. 35, no. 2, pp. 139–159.

O'Rourke, B., & Ramallo, F. (2013a). "A mina variedade é defectuosa": a lexitimidade social das neofalas. *Estudos de lingüística galega 5*, 89-103.

O'Rourke, B., & Ramallo, F. (2013b). Competing ideologies of linguistic authority amongst new speakers in contemporary Galicia. *Language in society 42*, 287-305.

O'Rourke, B., & Ramallo, F. (2015). Neofalantes as an active minority: understanding language practices and motivations for change amongst new speakers of Galician. *International journal of the sociology of language 231*, 147-165.

Ortega, A., Amorrortu, E., Goirigolzarri, J., Urla, J., & Uranga, B. (2014). New Basque speakers: Linguistic identity and legitimacy. *Digithum 16*, 47-58.

Pat-El, N. (2013). Contact or inheritance? Criteria for distinguishing internal and external change in genetically related languages. *Journal of language contact, 6*(2013), 313-328.

Pérez Castillejo, S. (2012). Estudio sociofonético de los tonemas de las interrogativas absolutas en el castellano de Galicia. *Estudios de lingüística*, (26), 235-268.

Pérez Castillejo, S. (2014). *La entonación del español de Galicia desde una perspectiva sociofonética* (Unpublished doctoral dissertation). University of Minnesota, Minneapolis, MN.

Pérez Torres, R. (2014). *Aproximación a la distribución diacrónica de las formas de pretérito imperfecto de subjuntivo (-ra y –se) en el español de España*. (Thesis). Universidade de Santiago de Compostela.

Pickering, M.J. & Branigan, H.P. (1998). The representation of verbs: Evidence from syntactic priming in language production. *Journal of memory and language*, *39*, 633-651.

Pollán, C. (2001). The expression of pragmatic values by means of verbal morphology: A variationist study. *Language variation and change* 13(1), 59-89.

Poplack, S. (1993). Variation theory and language contact. In D. Preston (Ed.), *American dialect research* (pp. 251-286). Philadelphia: John Benjamins.

Porcar Miralles, M. (1993). *La oración condicional: La evolución de los esquemas verbales condicionales desde el latín al español actual.* Castellón de la Plana: Universitat Jaume I.

Rabanal, M. (1967). *Hablas hispánicas. Temas gallegos y leoneses.* Madrid: Alcalá.

RAG / ILG. (2012). *Normas ortográficas e morfolóxicas do idioma galego* (23rd ed.). Santiago de Compostela: Real Academia Galega/Instituto da Lingua Galega.

Ramallo, F. (2007). Sociolinguistics of Spanish in Galicia. *International journal of the Sociology of Language*, *2007*(184), 21-36.

Ramallo, F. (2013). Neofalantismo. In E. Gugenberger, H. Monteagudo, & G. Rei-Doval (eds.), *Contacto de linguas, hibrididade, cambio: contextos, procesos e consecuencias* (245-258). Santiago de Compostela: Consello da Cultura Galega.

Rei-Doval, G. (2013). Purismo e control normativo na lingua galega: análise crítica dunha proposta actual. In E. Gugenberger, H. Monteagudo, & G. Rei-Doval (eds.), *Contacto de linguas, hibrididade, cambio: contextos, procesos e consecuencias* (259-286). Santiago de Compostela: Consello da Cultura Galega.

Rojo, G. (1996). Sobre la distribución de las formas llegara y llegase en español actual. *Casado Velarde (1996: II, 677-691)*.

Rojo, G. & Montero Cartelle, E. (1983). *La evolución de los esquemas condicionales (Potenciales e irreales desde el poema del Cid hasta 1400).* Santiago de Compostela: Universidade de Santiago de Compostela.

Rojo, G. & Vázquez Rozas, V. (2014). Sobre las formas en -ra en el español de Galicia. In A. Enrique-Arias, M.J. Gutiérrez, A. Landa y F. Ocampo (eds.): *Perspectives in the study of Spanish Language Variation. Papers in Honor of Carmen Silva-Corvalán* (pp. 237-270). Santiago de Compostela: Universidade de Santiago de Compostela.

Rosemeyer, M. & Schwenter, S. A. (2017). Entrenchment and persistence in language change: The Spanish past subjunctive. *Corpus Linguistics and Linguistic Theory 13*(1). Advance online publication. doi: 10.1515/cllt-2016-0047

Sankoff, D. (1982). Sociolinguistic method and linguistic theory. *Logic, methodology, philosophy of science, 6*, 677-689.

Silva-Corvalán, C. (1982). Conditional for subjunctive in Old Castile. *Proceedings of the*

Eighth Annual Meeting of the Berkeley Linguistics Society, 87-96.

Silva-Corvalán, C. (1984). The social profile of a syntactic-semantic variable: Three verb forms in Old Castile. *Hispania 67(4),* 594-601.

Silva-Corvalán, C. (1994). *Language contact and change: Spanish in Los Angeles.* Oxford: Clarendon Press.

Silva Valdivia, B. (2013). Galego e castelán: entre o contacto e a converxencia. In E. Gugenberger, H. Monteagudo, & G. Rei-Doval (eds.), *Contacto de linguas, hibrididade, cambio: contextos, procesos e consecuencias* (287-314). Santiago de Compostela: Consello da Cultura Galega.

Tagliamonte, S.A. (2006). *Analyzing sociolinguistic variation.* Cambridge: Cambridge University Press.

Thomason, S. G., & Kaufman, T. (1988). *Language contact, creolization, and genetic linguistics*. University of California Press.

Thomason, S. (2001). *Language contact: An introduction*. Georgetown University Press.

Thomason, S. (2010). Contact explanations in linguistics. In R. Hickey (Ed.), The *handbook of language contact* (pp. 29-47). Oxford: Wiley-Blackwell.

Trudgill, P. (1986). *Dialects in contact*. Oxford: Basil Blackwell Ltd.

Unsworth, S. (2015). Amount of exposure as a proxy for dominance in bilingual language acquisition. In C. Silva-Corvalán & J. Treffers-Daller (Eds.), *Language dominance in bilinguals: Issues of measurement and operationalization* (pp. 156-173). Cambridge: Cambridge University Press.

van Coetsem, F. (1988). *Loan phonology and the two transfer types in language contact.* Dordrecht: Foris Publications.

van Coetsem, F. (2000). *A general and unified theory of the transmission process in language contact* (Vol. 19). Winter.

Vann, R. E. (2002). Linguistic ideology in Spain's ivory tower: (Not) analyzing Catalan Spanish. *Multilingui, 21*(2-3), 227-246. doi: 10.1515/mult.2002.010

Veiga, A. (1996). *La forma verbal española "cantara" en su diacronía.* Santiago de Compostela: Servicio de Publicaciones e Intercambio Científico.

Vidal Figueiroa, T.V. (1997). Estructuras fonéticas de tres dialectos de Vigo. *Verba, 24*, 313-332.

Villalobos, C.M. (1997). La oración condicional: Del Latin al Español. *Analecta Malacitana, XX*(2): 447-474.

Vincent, N. (1982). The development of the auxiliaries habere and esse in Romance. In Nigel Vincent and Martin Harris (eds.): *Studies in the Romance verb* (pp. 71-96). London: Croom Helm Ltd.

Vladu, A. I. (2011). La normalización del Gallego—desafíos y perspectivas para el futuro. *Limbi şi Literaturi Straine,* 37-44.

Weinreich, U. (1953). *Languages in contact: Findings and problems* (No. 1). Walter de Gruyter.

Williams, A. (1982). The use of the-ra and-se forms of the past subjunctive in Navarre. *Hispania 65(1)*, 89-93.

Woolard, K. A. (2008). Language and identity choice in Catalonia: The interplay of contrasting ideologies of linguistic authority. In K. Süselback, U. Mühlschlegel, & P. Masson (Eds.), *Lengua, nación e identidad: La regulación del plurilingüismo en España y América Latina*, (p. 303-323). Madrid: Iberoamericana.

Wright, L. (1932). *The –ra verb form in Spain.* Berkeley: University of California Press.

Appendix: Tasks

I.   Bilingual Language Profile: Español-Gallego

Nos gustaría pedir su ayuda para contestar a las siguientes preguntas sobre su historial lingüístico, uso, actitudes y competencia. La encuesta contiene 19 preguntas y le llevará menos de 10 minutos para completar. Esto no es una prueba, por tanto no hay respuestas correctas ni incorrectas. Por favor conteste cada pregunta y responda con sinceridad, ya que solamente así se podrá garantizar el éxito de esta investigación. Muchas gracias por su ayuda.

I.   Información biográfica

Nombre _____   Fecha de hoy _____/_____/_____

Edad _____   Hombre/Mujer___   Lugar de residencia actual: _____

Nivel más alto de formación académica:

     Menos de la escuela secundaria   Escuela Secundaria

     Un poco de universidad   Universidad (diplomatura, licenciatura)

     Un poco de escuela graduada   Máster

     Doctorado   Otro: _____

Lugar de nacimiento: _____   Profesión/Trabajo: _____

II.   Historial lingüístico

En esta sección, nos gustaría que contestara algunas preguntas sobre su historial lingüístico marcando la casilla apropiada.

1. ¿A qué edad empezó a aprender las siguientes lenguas?

    Español

☐		☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐
Desde el Nacimiento		1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20+

    Gallego

☐		☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐
Desde el Nacimiento		1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20+

2. ¿A qué edad empezó a sentirse cómodo usando las siguientes lenguas?

    Español

☐		☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐
Tan pronto como recuerdo	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20+	

    Gallego

☐		☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐
Tan pronto como recuerdo	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20+	

3. ¿Cuántos años de clases (gramática, historia, matemáticas, etc..) ha tenido en las siguientes lenguas (desde la escuela primaria a la universidad)?

    Español

☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐
0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20+

    Gallego

☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐
0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20+

4. ¿Cuántos años ha pasado en un país/región donde se hablan las siguientes lenguas?

    Español

☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐
0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20+

    Gallego

☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐
0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20+

5. ¿Cuántos años ha pasado en familia hablando las siguientes lenguas?

    Español

☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐
0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20+

    Gallego

☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐
0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20+

6. ¿Cuántos años ha pasado en un ambiente de trabajo hablando las siguientes lenguas?

    Español

☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐
0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20+

    Gallego

☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐
0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20+

III.   Uso de lenguas

En esta sección, nos gustaría que contestara algunas preguntas sobre su uso de lenguas marcando la casilla apropiada. El uso total de todas las lenguas en cada pregunta debe llegar al 100%.

7. En una semana normal, ¿qué porcentaje del tiempo usa las siguientes lenguas con sus amigos?

Español ☐ ☐ ☐ ☐ ☐ ☐ ☐ ☐ ☐ ☐ ☐
0% 10% 20% 30% 40% 50% 60% 70% 80% 90% 100%

Gallego ☐ ☐ ☐ ☐ ☐ ☐ ☐ ☐ ☐ ☐ ☐
0% 10% 20% 30% 40% 50% 60% 70% 80% 90% 100%

Otras lenguas ☐ ☐ ☐ ☐ ☐ ☐ ☐ ☐ ☐ ☐ ☐
0% 10% 20% 30% 40% 50% 60% 70% 80% 90% 100%

8. En una semana normal, ¿qué porcentaje del tiempo usa las siguientes lenguas con su familia?

Español ☐ ☐ ☐ ☐ ☐ ☐ ☐ ☐ ☐ ☐ ☐
0% 10% 20% 30% 40% 50% 60% 70% 80% 90% 100%

Gallego ☐ ☐ ☐ ☐ ☐ ☐ ☐ ☐ ☐ ☐ ☐
0% 10% 20% 30% 40% 50% 60% 70% 80% 90% 100%

Otras lenguas ☐ ☐ ☐ ☐ ☐ ☐ ☐ ☐ ☐ ☐ ☐
0% 10% 20% 30% 40% 50% 60% 70% 80% 90% 100%

9. En una semana normal, ¿qué porcentaje del tiempo usa las siguientes lenguas en la escuela/el trabajo?

Español ☐ ☐ ☐ ☐ ☐ ☐ ☐ ☐ ☐ ☐ ☐
0% 10% 20% 30% 40% 50% 60% 70% 80% 90% 100%

Gallego ☐ ☐ ☐ ☐ ☐ ☐ ☐ ☐ ☐ ☐ ☐
0% 10% 20% 30% 40% 50% 60% 70% 80% 90% 100%

Otras lenguas ☐ ☐ ☐ ☐ ☐ ☐ ☐ ☐ ☐ ☐ ☐
0% 10% 20% 30% 40% 50% 60% 70% 80% 90% 100%

10. Cuando se habla a usted mismo, ¿con qué frecuencia se habla a sí mismo en las siguientes lenguas?

Español ☐ ☐ ☐ ☐ ☐ ☐ ☐ ☐ ☐ ☐ ☐
0% 10% 20% 30% 40% 50% 60% 70% 80% 90% 100%

Gallego ☐ ☐ ☐ ☐ ☐ ☐ ☐ ☐ ☐ ☐ ☐
0% 10% 20% 30% 40% 50% 60% 70% 80% 90% 100%

Otras lenguas ☐ ☐ ☐ ☐ ☐ ☐ ☐ ☐ ☐ ☐ ☐
0% 10% 20% 30% 40% 50% 60% 70% 80% 90% 100%

11. Cuando hace cálculos contando, ¿con qué frecuencia cuenta en las siguientes lenguas?

Español ☐ ☐ ☐ ☐ ☐ ☐ ☐ ☐ ☐ ☐ ☐
0% 10% 20% 30% 40% 50% 60% 70% 80% 90% 100%

Gallego ☐ ☐ ☐ ☐ ☐ ☐ ☐ ☐ ☐ ☐ ☐
0% 10% 20% 30% 40% 50% 60% 70% 80% 90% 100%

Otras lenguas ☐ ☐ ☐ ☐ ☐ ☐ ☐ ☐ ☐ ☐ ☐
0% 10% 20% 30% 40% 50% 60% 70% 80% 90% 100%

IV. Competencia

En esta sección, nos gustaría que considerara su competenccia de lengua marcando la casilla de 0 a 6.

0=no muy bien   *6=muy bien*

12. a. ¿Cómo habla en Español? □ 0 □ 1 □ 2 □ 3 □ 4 □ 5 □ 6

b. ¿Cómo habla en Gallego? □ 0 □ 1 □ 2 □ 3 □ 4 □ 5 □ 6

13. a. ¿Cómo entiende en Español? □ 0 □ 1 □ 2 □ 3 □ 4 □ 5 □ 6

b. ¿Cómo entiende el Gallego? □ 0 □ 1 □ 2 □ 3 □ 4 □ 5 □ 6

14. a. ¿Cómo lee en Español? □ 0 □ 1 □ 2 □ 3 □ 4 □ 5 □ 6

b. ¿Cómo lee en Gallego? □ 0 □ 1 □ 2 □ 3 □ 4 □ 5 □ 6

15. a. ¿Cómo escribe en Español? □ 0 □ 1 □ 2 □ 3 □ 4 □ 5 □ 6

b. ¿Cómo escribe en Gallego? □ 0 □ 1 □ 2 □ 3 □ 4 □ 5 □ 6

V.   Actitudes

En esta sección, nos gustaría que contestara a las siguientes afirmaciones sobre actitudes lingüísticas marcando las casillas de 0 a 6.

0=no estoy de acuerdo   *6=estoy de acuerdo*

16. a. Me siento "yo mismo" cuando hablo en Español.

☐ 0   ☐ 1   ☐ 2   ☐ 3   ☐ 4   ☐ 5   ☐ 6

b. Me siento "yo mismo" cuando hablo en Gallego.
☐ 0   ☐ 1   ☐ 2   ☐ 3   ☐ 4   ☐ 5   ☐ 6

17. a. Me identifico con una cultura Hispanohablante.
☐ 0   ☐ 1   ☐ 2   ☐ 3   ☐ 4   ☐ 5   ☐ 6

b. Me identifico con una cultura Gallegohablante.
☐ 0   ☐ 1   ☐ 2   ☐ 3   ☐ 4   ☐ 5   ☐ 6

18. a. Es importante para mí usar/llegar a usar Español como un hablante nativo.
☐ 0   ☐ 1   ☐ 2   ☐ 3   ☐ 4   ☐ 5   ☐ 6

b. Es importante para mí usar/llegar a usar Gallego como un hablante nativo.
☐ 0   ☐ 1   ☐ 2   ☐ 3   ☐ 4   ☐ 5   ☐ 6

19. a. Quiero que los demás piensen que soy un hablante nativo de Español.
☐ 0   ☐ 1   ☐ 2   ☐ 3   ☐ 4   ☐ 5   ☐ 6

b. Quiero que los demás piensen que soy un hablante nativo de Gallego.
☐ 0   ☐ 1   ☐ 2   ☐ 3   ☐ 4   ☐ 5   ☐ 6

II.   Interview Protocol

Juntos

- ¿Cómo os conocisteis? ¿Cuál fue la primera impresión que tuvisteis el uno del otro?
- ¿Qué tipo de cosas soléis hacer juntos?
- ¿Qué es algo que os gustaría hacer juntos, pero que nunca habéis hecho? ¿Por qué no lo habéis hecho? ¿Bajo qué condiciones lo haríais?
- ¿Podéis contarme una historia graciosa sobre la otra persona?
- ¿Qué es una cosa que creéis que la otra persona nunca haría? (a la otra persona) ¿Bajo qué condiciones lo harías?
- ¿Cuál es el momento más memorable que habéis pasado juntos?

Escuela

- En la escuela primaria, ¿cómo era un día típico?
- En general, ¿te gustó la escuela? ¿Por qué? ¿Qué podría haberlo hecho mejor?
- ¿Cuál es tu memoria favorita de la escuela primaria?
- ¿Alguna vez sufriste abuso o intimidación en la escuela? ¿Viste a alguien que lo sufrió? ¿Puedes contarme la historia?
- ¿Qué tipo de juegos jugabas con los otros niños en la escuela?
- ¿Alguna vez copiaste, o tuviste experiencia con alguien que copiaba? ¿Si le hubieras pillado a un amigo copiando, qué habrías hecho? (Imagínate que pillaste a un amigo copiando…qué habrías hecho?) ¿Qué habría hecho la profesora?
- ¿Alguna vez tuviste problemas disciplinarios en la escuela? ¿Cómo habrían reaccionado tus padres?
- ¿Cómo es la escuela diferente hoy en día? ¿Qué crees que les sorprendería a los niños de hoy si pudieran viajar en el tiempo para visitar tu escuela primaria?
- ¿Crees que hay problemas en el sistema educativo aquí? ¿Qué podría ayudar a solucionar estos problemas?

Familia

- ¿Cómo era tu familia cuando eras pequeño? ¿En qué era típica, y en qué era única? ¿Cómo era la relación con tu familia extendida? ¿Qué cosas hacíais juntos como familia?
- ¿Qué tipo de trabajo hacían tus padres? ¿Cúales eran los papeles de los hombres/mujeres en la familia?
- ¿Crees que tus padres eran estrictos? ¿Por qué crees así?
- ¿Me cuentas la historia de una vez que te metiste en problemas en casa?
- ¿Alguna vez pensaste en escaparte de la casa? ¿Por qué? ¿Qué habrían hecho tus padres si hubieras amenazado con hacerlo?

161

- ¿Cómo habrían reaccionado tus padres si tu o alguno de tus hermanos hubieran invitado a casa un amigo de otra raza? ¿Que hablara otra lengua?
- ¿Cómo habrían reaccionado tus abuelos?
- ¿Cómo reaccionarían/habrían reaccionado tus padres si tu o alguno de tus hermanos llevaran/hubieran llevado a casa un novio de otra raza? ¿Que hablara otra lengua? ¿Una pareja del mismo gendero?
- ¿Cómo reaccionarían tus abuelos?

Individuo

- ¿Alguna vez pensaste que ibas a morir? ¿Me cuentas la historia?
- ¿Qué es la cosa más miedosa que te ha pasado?
- ¿Alguna vez experimentaste algo inexplicable? ¿Me cuentas la historia?
- ¿Te consideras una persona feliz? ¿Bajo qué condiciones crees que podrías sentirte más feliz de lo que eres ahora?
- ¿Cómo crees que tu vida será en diez años? ¿Qué será igual/diferente? ¿Qué tiene que pasar entre hoy y aquel entonces para que se cumpla esa visión? ¿Qué podría complicar tus planes?

Identidad/Asuntos regionales

- ¿Conoces a alguien que ha emigrado a otro país? ¿Algunos de tus antepasados emigraron? ¿Por qué? ¿Bajo qué condiciones crees que habrían quedado en Galicia? ¿Bajo qué condiciones volverían?
- ¿Alguna vez pensaste en emigrar? ¿Por qué? ¿Bajo qué condiciones emigrarías? y ¿a dónde irías?
- ¿Cuáles son las causas principales de la emigración histórica? ¿Qué se podría haber hecho para evitar la necesidad de tanta emigración? ¿Qué se podría hacer para evitarla en el futuro?
- ¿Cuál(es) lengua(s) son necesarias para que una persona se considera gallego? ¿Es posible ser gallego sin hablar castellano? ¿Sin hablar gallego?
- ¿Por qué crees que los jóvenes hablan gallego menos hoy? ¿Es algo que se debe remediar? ¿Bajo qué condiciones se podría invertir esta tendencia?
- ¿Cómo ves el gallego "estándar"? ¿Es necesario? ¿Es auténtico? ¿Cómo se ve a la gente que sólo habla gallego estándar y no una variedad tradicional?
- He leído sobre gente que se criaron en castellano pero que han cambiado y ahora hablan casi totalmente en gallego; se denominan "neofalantes". ¿Por qué crees que hacen eso? ¿Crees que más gente lo debe hacer? ¿Bajo qué condiciones pensarías tú en hacerte neofalante?

III. Matched-Guise Text Transcriptions

Version 1: *Cantara*

A: Bueno, a ver, te considerarías una persona feliz?

M: Ah, yo creo que sí, bueno, puedo tener otras cosas, pero sí que…sí que me considero una persona feliz.

A: Y por qué?

M: Ah…bueno porque…bueno estoy en otro país, entonces sí, siempre se echa un poco de menos a la familia. Si estuviera mi familia aquí pos estaría más feliz. Ahm, también, sí echo de menos España a veces, la comida sobre todo,

A: Pues claro!

M: Sí, la comida, y…y sí, allí también tengo a todos mis amigos entonces si estuvieran aquí pues mejor. Si pudiera trasladar la universidad, porque por otra parte me gusta…estoy contenta aquí, me gusta este sitio, me gusta también un poco…bueno, a veces tengo quejas de, de la, de los Estados Unidos, ¿no? Pero, pero sí me gusta.

A: Y…y como podría ser para que te gustaran más los Estados Unidos?

M: Ah, pues bueno, que viniera mi familia, que viniera toda mi familia y mis amigos, pero lo supongo imposible, creo que sería más fácil trasladar la universidad a España.

A: Jajaja ah que sí

M: Sí, porque claro, las oportunidades de trabajo en España no son buenas, pues, si tuviera este trabajo en España…porque aunque termine aquí, luego si me voy a España no voy a tener buenas oportunidades. Entonces eso, si se trasladara esta universidad, también mi novio, aunque supongo improbable porque no habla español, pero sí, trasladar la universidad y y en general el ambiente, el mercado laboral también, si hubiera mejor mercado laboral en España, sí, estaría bien. Porque la verdad es que al final me gustaría más vivir en España que a lo mejor aquí. Sí, en general, aunque bueno, tampoco me importaría vivir aquí y viajar a España durante las navidades o en el verano. Bueno, y también, bueno, si tuviera mejor sueldo, sería bueno. Porque tener mejor sueldo, ah, y una buena casa también…ahora tengo que buscar casa, y sería más fácil si tuviera más dinero y no tuviera que buscar tanto.

Version 2: *Cantase*

A: Bueno, a ver, te considerarías una persona feliz?

M: Ah, yo creo que sí, bueno, puedo tener otras cosas, pero sí que…sí que me considero una persona feliz.

A: Y por qué?

M: Ah…bueno porque…bueno estoy en otro país, entonces sí, siempre se echa un poco de menos a la familia. Si estuviese mi familia aquí pos estaría más feliz. Ahm, también, sí echo de menos España a veces, la comida sobre todo,

A: Pues claro!

M: Sí, la comida, y…y sí, allí también tengo a todos mis amigos entonces si estuviesen aquí pues mejor. Si pudiese trasladar la universidad, porque por otra parte me gusta…estoy contenta aquí, me gusta este sitio, me gusta también un poco…bueno, a veces tengo quejas de, de la, de los Estados Unidos, ¿no? Pero, pero sí me gusta.

A: Y…y como podría ser para que te gustasen más los Estados Unidos?

M: Ah, pues bueno, que viniese mi familia, que viniese toda mi familia y mis amigos, pero lo supongo imposible, creo que sería más fácil trasladar la universidad a España.

A: Jajaja ah que sí

M: Sí, porque claro, las oportunidades de trabajo en España no son buenas, pues, si tuviese este trabajo en España…porque aunque termine aquí, luego si me voy a España no voy a tener buenas oportunidades. Entonces eso, si se trasladase esta universidad, también mi novio, aunque supongo improbable porque no habla español, pero sí, trasladar la universidad y y en general el ambiente, el mercado laboral también, si hubiese mejor mercado laboral en España, sí, estaría bien. Porque la verdad es que al final me gustaría más vivir en España que a lo mejor aquí. Sí, en general, aunque bueno, tampoco me importaría vivir aquí y viajar a España durante las navidades o en el verano. Bueno, y también, bueno, si tuviese mejor sueldo, sería bueno. Porque tener mejor sueldo, ah, y una buena casa también…ahora tengo que buscar casa, y sería más fácil si tuviese más dinero y no tuviese que buscar tanto.

III.   Matched-Guise Questionnaire
CUESTIONARIO

NOMBRE:...

La persona que habla,	**Muy poco**					**Mucho**
Te parece inteligente	1	2	3	4	5	6
Te parece simpática	1	2	3	4	5	6
Te parece una persona culta	1	2	3	4	5	6
Te parece atractiva físicamente	1	2	3	4	5	6
Te parece digna de confianza	1	2	3	4	5	6
Te parece que tiene sentido del humor	1	2	3	4	5	6
Te parece que tiene interés en progresar	1	2	3	4	5	6
Te parece una persona abierta	1	2	3	4	5	6
Te parece que tiene confianza en si misma	1	2	3	4	5	6
Te parece que acepta nuevos usos y costumbres	1	2	3	4	5	6
Te parece una persona generosa	1	2	3	4	5	6
Te parece que es capaz de dirigir	1	2	3	4	5	6
Te parece una persona divertida	1	2	3	4	5	6
Te parece una persona trabajadora	1	2	3	4	5	6
Te parece orgullosa	1	2	3	4	5	6
Te parece fiel	1	2	3	4	5	6
Te parece una persona práctica	1	2	3	4	5	6
Te parece una persona similar a ti	1	2	3	4	5	6

¿A qué tipo de trabajo crees que se dedica?...

IV.   Fill-in-the-Blank Task

Nombre: _____

En esta actividad, se te pide completar la conversación rellenando cada espacio con una forma apropiada del verbo en paréntesis.

DON ISMAEL: No me _____ (sentir) bien, doctora. ¡Ay,

si _____ (tener) más energía!

DRA. SÁNCHEZ: Si usted _____ (tomar) estas vitaminas y _____

(realizar) más ejercicio, se _____ (sentir) mejor,

don Ismael.

DON ISMAEL: Pero doctora, las vitaminas _____ (ser) caras. Si

_____ (tener) el dinero para comprar pastillas ya me

las _____ (comprar). Y hacer ejercicio

_____ (ser) aburrido. Si _____ (vivir) más

cerca del gimnasio, lo _____ (hacer), pero…

DRA. SÁNCHEZ: Entiendo que _____ (ser) difícil, don Ismael. Pero ¿qué

_____ (hacer) su esposa si algo le _____ (pasar) a

usted? Si _____ (seguir) mis consejos,

_____ (ser) mucho más feliz y su esposa no

_____ (temer) por su salud.

DON ISMAEL: Usted _____ (tener) razón. Si _____ (poder)

seguir sus consejos me _____ (sentir) mejor.

_____ (tratar) de hacerlo. ¡Ay, si

_____ (ser) más joven!

V. Acceptability Judgment Task

Nombre: _____

En esta actividad, se te pide evaluar la aceptabilidad de una serie de frases. Es importante notar que no se está preguntando acerca del contenido de la frase, sino de las formas gramaticales que se usan para expresarla. Es decir, la frase "Don Quixote es mi mejor amigo", por ejemplo, sólo se debe marcar como inadecuado si implica extrañeza gramatical decirlo así, no porque estás o no de acuerdo con la proposición expresada.

1. Llegué al restaurante tarde, pero mi amiga todavía no salió de casa.

	Estoy muy en desacuerdo	Estoy más o menos en desacuerdo	Estoy más o menos de acuerdo	Estoy muy de acuerdo
A. Conozco a gente que lo diría así.	☐	☐	☐	☐
B. Yo lo diría así.	☐	☐	☐	☐
C. Está bien dicho.	☐	☐	☐	☐

Si contestaste que no está bien dicho o que no la dirías así, ¿Cómo se debe decir? o ¿Cómo lo dirías tú?

2. Si yo fuera tú, estudiaría más.

	Estoy muy en desacuerdo	Estoy más o menos en desacuerdo	Estoy más o menos de acuerdo	Estoy muy de acuerdo
A. Conozco a gente que lo diría así.	☐	☐	☐	☐
B. Yo lo diría así.	☐	☐	☐	☐
C. Está bien dicho.	☐	☐	☐	☐

Si contestaste que no está bien dicho o que no la dirías así, ¿Cómo se debe decir? o ¿Cómo lo dirías tú?

3. Si yo ganase la lotería, me compraba un coche nuevo.

	Estoy muy en desacuerdo	Estoy más o menos en desacuerdo	Estoy más o menos de acuerdo	Estoy muy de acuerdo
A. Conozco a gente que lo diría así.	☐	☐	☐	☐
B. Yo lo diría así.	☐	☐	☐	☐
C. Está bien dicho.	☐	☐	☐	☐

Si contestaste que no está bien dicho o que no la dirías así, ¿Cómo se debe decir? o ¿Cómo lo dirías tú?

4. Antes de mi nacimiento, mis padres ya escogieran mi nombre.

	Estoy muy en desacuerdo	Estoy más o menos en desacuerdo	Estoy más o menos de acuerdo	Estoy muy de acuerdo
A. Conozco a gente que lo diría así.	☐	☐	☐	☐
B. Yo lo diría así.	☐	☐	☐	☐
C. Está bien dicho.	☐	☐	☐	☐

Si contestaste que no está bien dicho o que no la dirías así, ¿Cómo se debe decir? o ¿Cómo lo dirías tú?

5. Si serías más joven, tenías menos estrés.

	Estoy muy en desacuerdo	Estoy más o menos en desacuerdo	Estoy más o menos de acuerdo	Estoy muy de acuerdo
A. Conozco a gente que lo diría así.	☐	☐	☐	☐
B. Yo lo diría así.	☐	☐	☐	☐
C. Está bien dicho.	☐	☐	☐	☐

Si contestaste que no está bien dicho o que no la dirías así, ¿Cómo se debe decir? o ¿Cómo lo dirías tú?

6. Si los jóvenes fueran a la universidad, ganarían más dinero.

	Estoy muy en desacuerdo	Estoy más o menos en desacuerdo	Estoy más o menos de acuerdo	Estoy muy de acuerdo
A. Conozco a gente que lo diría así.	☐	☐	☐	☐
B. Yo lo diría así.	☐	☐	☐	☐
C. Está bien dicho.	☐	☐	☐	☐

Si contestaste que no está bien dicho o que no la dirías así, ¿Cómo se debe decir? o ¿Cómo lo dirías tú?

7. Si él fuera mujer, mostraba más sus emociones.

	Estoy muy en desacuerdo	Estoy más o menos en desacuerdo	Estoy más o menos de acuerdo	Estoy muy de acuerdo
A. Conozco a gente que lo diría así.	☐	☐	☐	☐
B. Yo lo diría así.	☐	☐	☐	☐
C. Está bien dicho.	☐	☐	☐	☐

Si contestaste que no está bien dicho o que no la dirías así, ¿Cómo se debe decir? o ¿Cómo lo dirías tú?

8. Si mi amiga consiguiese otro trabajo, se preocuparía menos.

	Estoy muy en desacuerdo	Estoy más o menos en desacuerdo	Estoy más o menos de acuerdo	Estoy muy de acuerdo
A. Conozco a gente que lo diría así.	☐	☐	☐	☐
B. Yo lo diría así.	☐	☐	☐	☐
C. Está bien dicho.	☐	☐	☐	☐

Si contestaste que no está bien dicho o que no la dirías así, ¿Cómo se debe decir? o ¿Cómo lo dirías tú?

9. Cuando le pregunté a mi amigo qué quería comer, él ya había llamado para pedir una pizza.

	Estoy muy en desacuerdo	Estoy más o menos en desacuerdo	Estoy más o menos de acuerdo	Estoy muy de acuerdo
A. Conozco a gente que lo diría así.	☐	☐	☐	☐
B. Yo lo diría así.	☐	☐	☐	☐
C. Está bien dicho.	☐	☐	☐	☐

Si contestaste que no está bien dicho o que no la dirías así, ¿Cómo se debe decir? o ¿Cómo lo dirías tú?

10. Si mis amigos y yo seríamos alemanes, aprenderíamos el inglés más fácilmente.

	Estoy muy en desacuerdo	Estoy más o menos en desacuerdo	Estoy más o menos de acuerdo	Estoy muy de acuerdo
A. Conozco a gente que lo diría así.	☐	☐	☐	☐
B. Yo lo diría así.	☐	☐	☐	☐
C. Está bien dicho.	☐	☐	☐	☐

Si contestaste que no está bien dicho o que no la dirías así, ¿Cómo se debe decir? o ¿Cómo lo dirías tú?

11. Si ellos preguntasen, encontraban opciones.

	Estoy muy en desacuerdo	Estoy más o menos en desacuerdo	Estoy más o menos de acuerdo	Estoy muy de acuerdo
A. Conozco a gente que lo diría así.	☐	☐	☐	☐
B. Yo lo diría así.	☐	☐	☐	☐
C. Está bien dicho.	☐	☐	☐	☐

Si contestaste que no está bien dicho o que no la dirías así, ¿Cómo se debe decir? o ¿Cómo lo dirías tú?

12. Si tuviera una madre diferente, yo sería una persona totalmente distinta.

	Estoy muy en desacuerdo	Estoy más o menos en desacuerdo	Estoy más o menos de acuerdo	Estoy muy de acuerdo
A. Conozco a gente que lo diría así.	☐	☐	☐	☐
B. Yo lo diría así.	☐	☐	☐	☐
C. Está bien dicho.	☐	☐	☐	☐

Si contestaste que no está bien dicho o que no la dirías así, ¿Cómo se debe decir? o ¿Cómo lo dirías tú?

13. Llegué al restaurante tarde, pero mi amiga todavía no había salido de casa.

	Estoy muy en desacuerdo	Estoy más o menos en desacuerdo	Estoy más o menos de acuerdo	Estoy muy de acuerdo
A. Conozco a gente que lo diría así.	☐	☐	☐	☐
B. Yo lo diría así.	☐	☐	☐	☐
C. Está bien dicho.	☐	☐	☐	☐

Si contestaste que no está bien dicho o que no la dirías así, ¿Cómo se debe decir? o ¿Cómo lo dirías tú?

14. Si mis padres tuviesen el dinero, pasarían una semana en Roma.

	Estoy muy en desacuerdo	Estoy más o menos en desacuerdo	Estoy más o menos de acuerdo	Estoy muy de acuerdo
A. Conozco a gente que lo diría así.	☐	☐	☐	☐
B. Yo lo diría así.	☐	☐	☐	☐
C. Está bien dicho.	☐	☐	☐	☐

Si contestaste que no está bien dicho o que no la dirías así, ¿Cómo se debe decir? o ¿Cómo lo dirías tú?

15. Antes de mi nacimiento, mis padres ya escogieron mi nombre.

	Estoy muy en desacuerdo	Estoy más o menos en desacuerdo	Estoy más o menos de acuerdo	Estoy muy de acuerdo
A. Conozco a gente que lo diría así.	☐	☐	☐	☐
B. Yo lo diría así.	☐	☐	☐	☐
C. Está bien dicho.	☐	☐	☐	☐

Si contestaste que no está bien dicho o que no la dirías así, ¿Cómo se debe decir? o ¿Cómo lo dirías tú?

16. Si mis profesores fuesen analfabetos, tendrían trabajos diferentes.

	Estoy muy en desacuerdo	Estoy más o menos en desacuerdo	Estoy más o menos de acuerdo	Estoy muy de acuerdo
D. Conozco a gente que lo diría así.	☐	☐	☐	☐
E. Yo lo diría así.	☐	☐	☐	☐
F. Está bien dicho.	☐	☐	☐	☐

Si contestaste que no está bien dicho o que no la dirías así, ¿Cómo se debe decir? o ¿Cómo lo dirías tú?

17. Si los jóvenes se quejaran menos, estaban más felices.

	Estoy muy en desacuerdo	Estoy más o menos en desacuerdo	Estoy más o menos de acuerdo	Estoy muy de acuerdo
A. Conozco a gente que lo diría así.	☐	☐	☐	☐
B. Yo lo diría así.	☐	☐	☐	☐
C. Está bien dicho.	☐	☐	☐	☐

Si contestaste que no está bien dicho o que no la dirías así, ¿Cómo se debe decir? o ¿Cómo lo dirías tú?

18. Si yo fuese tú, estudiaba más.

	Estoy muy en desacuerdo	Estoy más o menos en desacuerdo	Estoy más o menos de acuerdo	Estoy muy de acuerdo
A. Conozco a gente que lo diría así.	☐	☐	☐	☐
B. Yo lo diría así.	☐	☐	☐	☐
C. Está bien dicho.	☐	☐	☐	☐

Si contestaste que no está bien dicho o que no la dirías así, ¿Cómo se debe decir? o ¿Cómo lo dirías tú?

19. Si yo ganaría la lotería, me compraría un coche nuevo.

	Estoy muy en desacuerdo	Estoy más o menos en desacuerdo	Estoy más o menos de acuerdo	Estoy muy de acuerdo
A. Conozco a gente que lo diría así.	☐	☐	☐	☐
B. Yo lo diría así.	☐	☐	☐	☐
C. Está bien dicho.	☐	☐	☐	☐

Si contestaste que no está bien dicho o que no la dirías así, ¿Cómo se debe decir? o ¿Cómo lo dirías tú?

20. Cuando le pregunté a mi amigo qué quería comer, él ya llamara para pedir una pizza.

	Estoy muy en desacuerdo	Estoy más o menos en desacuerdo	Estoy más o menos de acuerdo	Estoy muy de acuerdo
D. Conozco a gente que lo diría así.	☐	☐	☐	☐
E. Yo lo diría así.	☐	☐	☐	☐
F. Está bien dicho.	☐	☐	☐	☐

Si contestaste que no está bien dicho o que no la dirías así, ¿Cómo se debe decir? o ¿Cómo lo dirías tú?

21. Si fueses más joven, tendrías menos estrés.

	Estoy muy en desacuerdo	Estoy más o menos en desacuerdo	Estoy más o menos de acuerdo	Estoy muy de acuerdo
A. Conozco a gente que lo diría así.	☐	☐	☐	☐
B. Yo lo diría así.	☐	☐	☐	☐
C. Está bien dicho.	☐	☐	☐	☐

Si contestaste que no está bien dicho o que no la dirías así, ¿Cómo se debe decir? o ¿Cómo lo dirías tú?

22. Si los jóvenes irían a la universidad, ganaban más dinero.

	Estoy muy en desacuerdo	Estoy más o menos en desacuerdo	Estoy más o menos de acuerdo	Estoy muy de acuerdo
A. Conozco a gente que lo diría así.	☐	☐	☐	☐
B. Yo lo diría así.	☐	☐	☐	☐
C. Está bien dicho.	☐	☐	☐	☐

Si contestaste que no está bien dicho o que no la dirías así, ¿Cómo se debe decir? o ¿Cómo lo dirías tú?

23. Cuando le pregunté a mi amigo qué quería comer, él ya llamó para pedir una pizza.

	Estoy muy en desacuerdo	Estoy más o menos en desacuerdo	Estoy más o menos de acuerdo	Estoy muy de acuerdo
A. Conozco a gente que lo diría así.	☐	☐	☐	☐
B. Yo lo diría así.	☐	☐	☐	☐
C. Está bien dicho.	☐	☐	☐	☐

Si contestaste que no está bien dicho o que no la dirías así, ¿Cómo se debe decir? o ¿Cómo lo dirías tú?

24. Si él sería mujer, mostraría más sus emociones.

	Estoy muy en desacuerdo	Estoy más o menos en desacuerdo	Estoy más o menos de acuerdo	Estoy muy de acuerdo
D. Conozco a gente que lo diría así.	☐	☐	☐	☐
E. Yo lo diría así.	☐	☐	☐	☐
F. Está bien dicho.	☐	☐	☐	☐

Si contestaste que no está bien dicho o que no la dirías así, ¿Cómo se debe decir? o ¿Cómo lo dirías tú?

25. Llegué al restaurante tarde, pero mi amiga todavía no saliera de casa.

	Estoy muy en desacuerdo	Estoy más o menos en desacuerdo	Estoy más o menos de acuerdo	Estoy muy de acuerdo
A. Conozco a gente que lo diría así.	☐	☐	☐	☐
B. Yo lo diría así.	☐	☐	☐	☐
C. Está bien dicho.	☐	☐	☐	☐

Si contestaste que no está bien dicho o que no la dirías así, ¿Cómo se debe decir? o ¿Cómo lo dirías tú?

26. Si mi amiga consiguiera otro trabajo, se preocupaba menos.

	Estoy muy en desacuerdo	Estoy más o menos en desacuerdo	Estoy más o menos de acuerdo	Estoy muy de acuerdo
D. Conozco a gente que lo diría así.	☐	☐	☐	☐
E. Yo lo diría así.	☐	☐	☐	☐
F. Está bien dicho.	☐	☐	☐	☐

Si contestaste que no está bien dicho o que no la dirías así, ¿Cómo se debe decir? o ¿Cómo lo dirías tú?

27. Si mis amigos y yo fuésemos alemanes, aprendíamos el inglés más fácilmente.

	Estoy muy en desacuerdo	Estoy más o menos en desacuerdo	Estoy más o menos de acuerdo	Estoy muy de acuerdo
A. Conozco a gente que lo diría así.	☐	☐	☐	☐
B. Yo lo diría así.	☐	☐	☐	☐
C. Está bien dicho.	☐	☐	☐	☐

Si contestaste que no está bien dicho o que no la dirías así, ¿Cómo se debe decir? o ¿Cómo lo dirías tú?

28. Si ellos preguntaran, encontrarían opciones.

	Estoy muy en desacuerdo	Estoy más o menos en desacuerdo	Estoy más o menos de acuerdo	Estoy muy de acuerdo
A. Conozco a gente que lo diría así.	☐	☐	☐	☐
B. Yo lo diría así.	☐	☐	☐	☐
C. Está bien dicho.	☐	☐	☐	☐

Si contestaste que no está bien dicho o que no la dirías así, ¿Cómo se debe decir? o ¿Cómo lo dirías tú?

29. Si tendría una madre diferente, yo era una persona totalmente distinta.

	Estoy muy en desacuerdo	Estoy más o menos en desacuerdo	Estoy más o menos de acuerdo	Estoy muy de acuerdo
A. Conozco a gente que lo diría así.	☐	☐	☐	☐
B. Yo lo diría así.	☐	☐	☐	☐
C. Está bien dicho.	☐	☐	☐	☐

Si contestaste que no está bien dicho o que no la dirías así, ¿Cómo se debe decir? o ¿Cómo lo dirías tú?

30. Antes de mi nacimiento, mis padres ya habían escogido mi nombre.

	Estoy muy en desacuerdo	Estoy más o menos en desacuerdo	Estoy más o menos de acuerdo	Estoy muy de acuerdo
D. Conozco a gente que lo diría así.	☐	☐	☐	☐
E. Yo lo diría así.	☐	☐	☐	☐
F. Está bien dicho.	☐	☐	☐	☐

Si contestaste que no está bien dicho o que no la dirías así, ¿Cómo se debe decir? o ¿Cómo lo dirías tú?

31. Si mis padres tendrían el dinero, pasaban una semana en Roma.

	Estoy muy en desacuerdo	Estoy más o menos en desacuerdo	Estoy más o menos de acuerdo	Estoy muy de acuerdo
A. Conozco a gente que lo diría así.	☐	☐	☐	☐
B. Yo lo diría así.	☐	☐	☐	☐
C. Está bien dicho.	☐	☐	☐	☐

Si contestaste que no está bien dicho o que no la dirías así, ¿Cómo se debe decir? o ¿Cómo lo dirías tú?

32. Si mis profesores fueran analfabetos, tenían trabajos diferentes.

	Estoy muy en desacuerdo	Estoy más o menos en desacuerdo	Estoy más o menos de acuerdo	Estoy muy de acuerdo
A. Conozco a gente que lo diría así.	☐	☐	☐	☐
B. Yo lo diría así.	☐	☐	☐	☐
C. Está bien dicho.	☐	☐	☐	☐

Si contestaste que no está bien dicho o que no la dirías así, ¿Cómo se debe decir? o ¿Cómo lo dirías tú?

33. Si los jóvenes se quejarían menos, estarían más felices.

	Estoy muy en desacuerdo	Estoy más o menos en desacuerdo	Estoy más o menos de acuerdo	Estoy muy de acuerdo
A. Conozco a gente que lo diría así.	☐	☐	☐	☐
B. Yo lo diría así.	☐	☐	☐	☐
C. Está bien dicho.	☐	☐	☐	☐

Si contestaste que no está bien dicho o que no la dirías así, ¿Cómo se debe decir? o ¿Cómo lo dirías tú?

Appendix: Statistical tests, figures, and tables

I. Spontaneous speech

Table 15: Spontaneous speech logistic regression with random effect for Individual--ANOVA table

	Df	Sum Sq	F value
Prime*	2	3.82	1.91
Verb Group (n.s.)	3	1.74	0.58
Possibility (n.s.)	1	0.27	0.27
Polarity (n.s.)	1	0.43	0.43
Anteriority (n.s.)	1	0.30	0.30
Clause Type (n.s.)	3	4.29	1.43
Marginal R^2 = 0.20; Conditional R^2 = 0.86; Log Likelihood = -38.5 (for the associated Logistic Regression)			
* = approaches significance, (n.s.) = non-significant			

Table 16: Spontaneous speech linear regression--ANOVA table

	Df	Sum Sq	F value
Gender (n.s.)	1	0.08	0.53
Age (n.s.)	1	0.57	3.68
Location (n.s.)	1	0.21	1.35
Initial Language (n.s.)	2	0.26	0.84
School Language (n.s.)	1	0.17	1.08
Family Language (n.s.)	1	0.01	0.04
Work Language (n.s.)	1	0.14	0.92
Current Language (n.s.)	1	0.00	0.00
R^2 = 0.01; $F_{(9,19)}$ = 1.03 (for the associated Linear Regression)			
(n.s.) = non-significant			

II. Matched-guise task

Table 17: Matched-guise linear regression with random effects for Individual and Characteristic--ANOVA table

	Df	Sum Sq	F value
Guise (n.s.)	1	1.00	1.45
Characteristic Group**	2	14.88	10.76
Gender (n.s.)	1	0.14	0.20
Age (n.s.)	1	0.01	0.01
Location (n.s.)	1	0.15	0.22
Initial Language (n.s.)	2	0.17	0.12
School Language (n.s.)	1	0.23	0.34
Family Language (n.s.)	1	0.04	0.06
Work Language (n.s.)	1	0.14	0.20
Current Language Use (n.s.)	1	0.06	0.08
Guise*Char Group (n.s.)	2	0.85	0.61
Guise*Gender (n.s.)	1	0.02	0.03
Guise*Age (n.s.)	1	0.14	0.20
Guise*Location (n.s.)	1	0.17	0.24
Guise*L1 (n.s.)	2	0.25	0.18
Guise*School Lang (n.s.)	1	0.00	0.00
Guise*Family Lang (n.s.)	1	0.07	0.10
Guise*Work Lang (n.s.)	1	0.03	0.05
Guise*Current Lang (n.s.)	1	0.46	0.66
Marginal R^2 = 0.16; Conditional R^2 = 0.25			
** = significant, (n.s.) = non-significant			

III.   Fill-in-the-blanks task

Table 18: FitB logistic regression with random effects for Individual and Phrase-- ANOVA table

	Df	Sum Sq	F value
Prime*	2	6.79	3.39
Recency (n.s.)	3	1.21	0.40
Prior Task Prime (n.s.)	1	0.01	0.01
Verb Group (n.s.)	4	8.63	2.16
Gender (n.s.)	1	0.69	0.69
Age (n.s.)	1	1.15	1.15
Location (n.s.)	1	3.19	3.19
Initial Language (n.s.)	2	1.89	0.95
School Language (n.s.)	1	0.30	0.30
Family Language (n.s.)	1	0.06	0.06
Work Language*	1	2.09	2.09
Current Language Use*	1	3.83	3.83
Written Competence (n.s.)	1	0.83	0.83
Language Attitudes (n.s.)	1	0.29	0.29
Marginal R^2 = 0.39; Conditional R^2 = 0.63; Log Likelihood = -112.0 (for the associated Logistic Regression)			
* = approaches significance, (n.s.) = non-significant			

IV. Acceptability judgment task

Table 19: Social Use ordinal logistic regression with random effects for Individual and Phrase--Analysis of Deviance table

	LR Chi Sq	Df	P value
Protasis**	11.77	1	0.00
Gender (n.s.)	0.68	1	0.41
Age (n.s.)	0.02	1	0.88
Location (n.s.)	0.21	1	0.65
Apodosis**	37.81	1	0.00
Possibility (n.s.)	0.44	1	0.51
Initial Language (n.s.)	2.78	2	0.25
School Language*	5.91	1	0.02
Family Language (n.s.)	3.48	1	0.06
Work Language (n.s.)	0.05	1	0.83
Current Language Use (n.s.)	1.48	1	0.22
Protasis*Gender (n.s.)	0.72	1	0.40
Protasis*Age (n.s.)	0.04	1	0.84
Protasis*Location (n.s.)	0.30	1	0.58
-Protasis*Apodosis (n.s.)	0.33	1	0.57
Protasis*Possibility (n.s.)	0.01	1	0.94
Protasis*L1 (n.s.)	3.90	2	0.14
Protasis*School Lang (n.s.)	1.63	1	0.20
Protasis*Family Lang (n.s.)	1.22	1	0.27
Protasis*Work Lang (n.s.)	0.61	1	0.43
Protasis*Current Lang (n.s.)	0.12	1	0.73
$R^2 = 0.07$; Log Likelihood = -435.54 (for the associated Ordinal Logistic Regression)			
** = significant, * = approaches significance, (n.s.) = non-significant			

Table 20: Correctness ordinal logistic regression with random effects for Individual and Phrase--Analysis of Deviance table

	LR Chi Sq	Df	P value
Protasis**	12.45	1	0.00
Gender (n.s.)	0.01	1	0.94
Age (n.s.)	0.05	1	0.83
Location (n.s.)	3.61	1	0.06
Apodosis**	44.20	1	0.00
Possibility (n.s.)	1.29	1	0.26
Initial Language (n.s.)	0.79	2	0.67
School Language (n.s.)	0.00	1	0.99
Family Language (n.s.)	1.38	1	0.24
Work Language (n.s.)	1.63	1	0.20
Current Language Use (n.s.)	1.29	1	0.26
Protasis*Gender (n.s.)	1.22	1	0.27
Protasis*Age (n.s.)	0.68	1	0.41
Protasis*Location*	4.32	1	0.04
Protasis*Apodosis (n.s.)	0.28	1	0.60
Protasis*Possibility*	4.24	1	0.04
Protasis*L1 (n.s.)	0.15	2	0.93
Protasis*School Lang (n.s.)	0.61	1	0.43
Protasis*Family Lang (n.s.)	0.68	1	0.41
Protasis*Work Lang (n.s.)	0.34	1	0.56
Protasis*Current Lang (n.s.)	0.11	1	0.74
$R^2 = 0.07$; Log Likelihood = -408.51 (for the associated Ordinal Logistic Regression)			
** = significant, * = approaches significance, (n.s.) = non-significant			

Table 21: Corrections logistic regression with random effect for Individual--ANOVA table

	Df	Sum Sq	F value
Prime**	2	82.62	41.31
Verb Group (n.s.)	2	1.37	0.69
Possibility (n.s.)	1	0.02	0.02
Gender (n.s.)	1	0.15	0.15
Age (n.s.)	1	0.73	0.73
Location*	1	1.40	1.40
Initial Language**	2	8.30	4.15
School Language (n.s.)	1	0.45	0.45
Family Language (n.s.)	1	0.00	0.00
Work Language (n.s.)	1	1.35	1.35
Current Language Use (n.s.)	1	2.76	2.76
Marginal R^2 = 0.37; Conditional R^2 = 0.64; Log Likelihood = -261.1 (for the associated Logistic Regression)			
** = significant, * = approaches significance, (n.s.) = non-significant			

Table 22: Social Use ratings--Interaction between Protasis and Gender, Location, L1

	cantase		*cantara*	
	M	SD	M	SD
Female	1.15	0.57	0.73	0.61
Male	1.16	0.80	0.68	0.99
Marín	1.20	0.68	0.78	0.77
O Grove	1.03	0.63	0.52	0.78
Castilian L1	1.46	0.56	0.69	0.51
Simultaneous Bilingual	1.06	0.66	0.66	0.71
Galician L1	1.18	0.70	0.90	1.14

Figure 68: Acceptability judgment: Social Use--Interaction between Age and Protasis
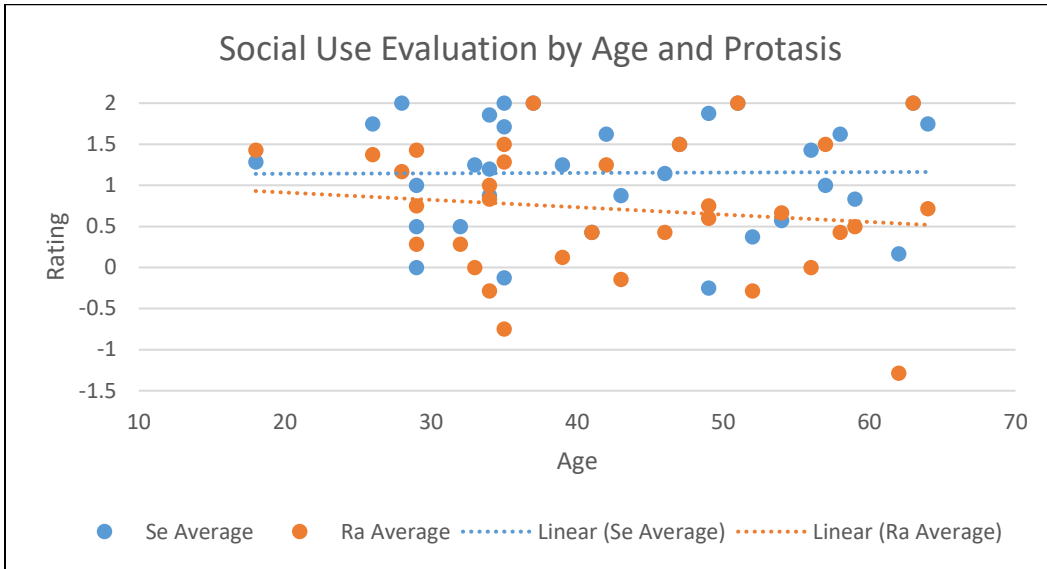

Social Use Evaluation by Age and Protasis

Figure 69: Acceptability judgment: Social Use--Interaction between Family Language and Protasis

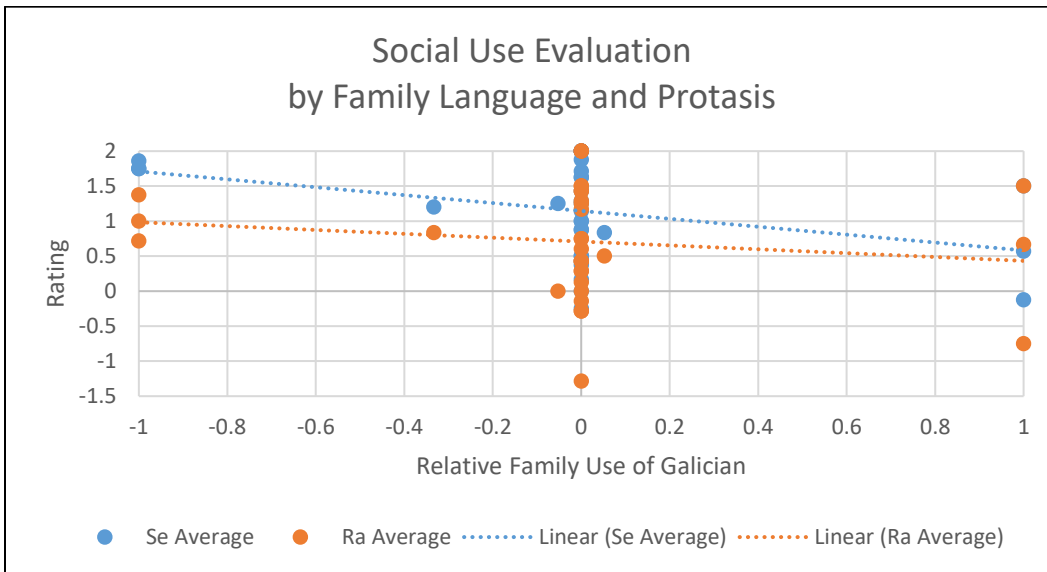Social Use Evaluation by Family Language and Protasis

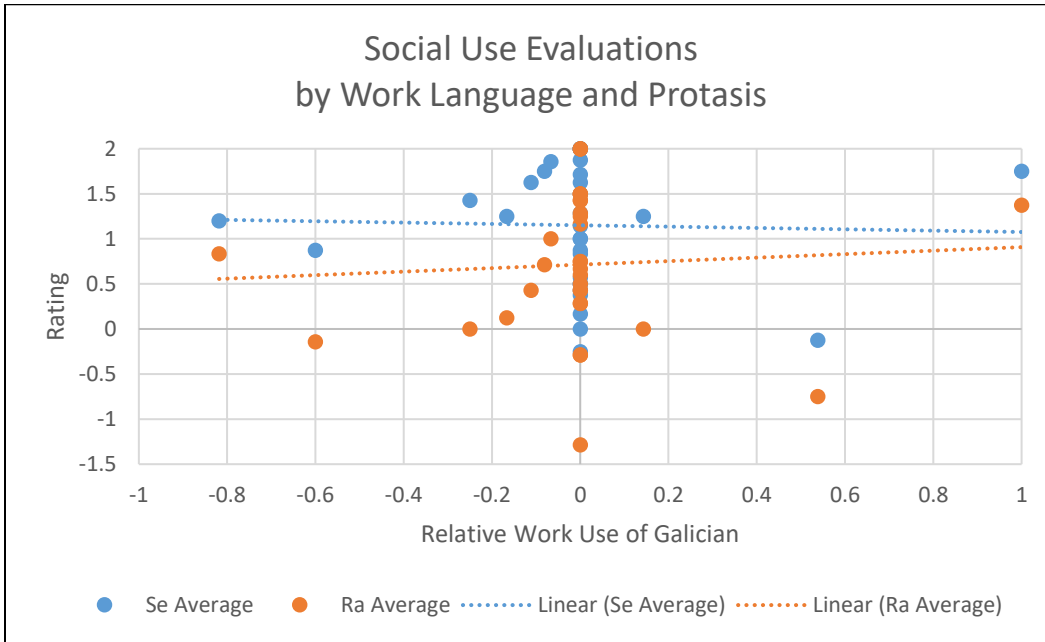**Figure 70: Acceptability judgment: Social Use--Interaction between Work Language and Protasis**

Social Use Evaluations by Work Language and Protasis

Figure 71: Acceptability judgment: Social Use--Interaction between Current Language and Protasis

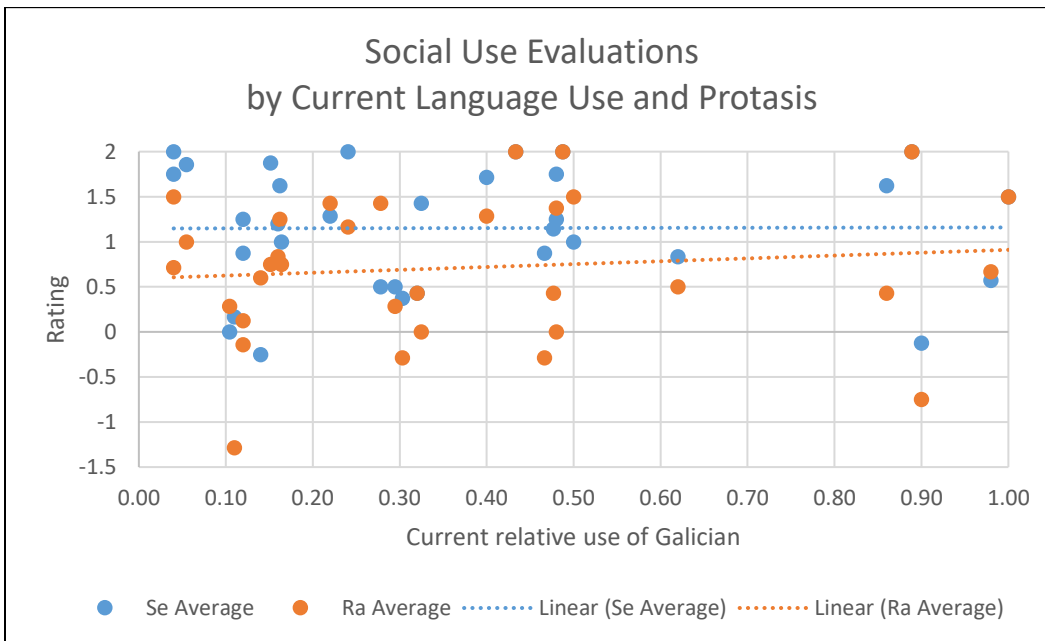Social Use Evaluations by Current Language Use and Protasis

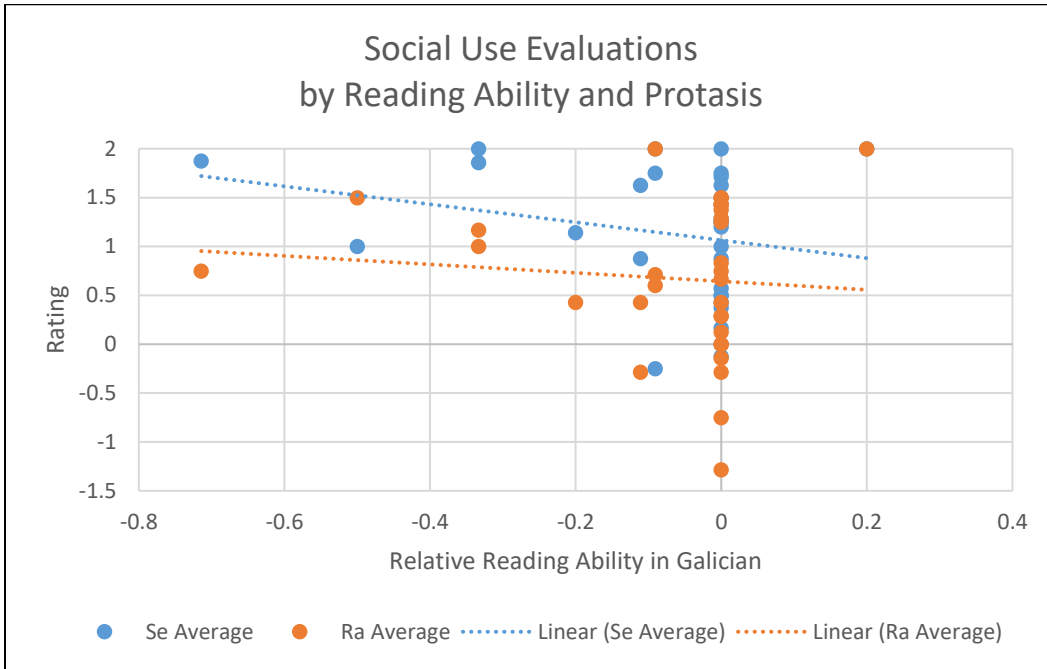**Figure 72: Acceptability judgment: Social Use--Interaction between Reading Ability and Protasis**

Social Use Evaluations
by Reading Ability and Protasis

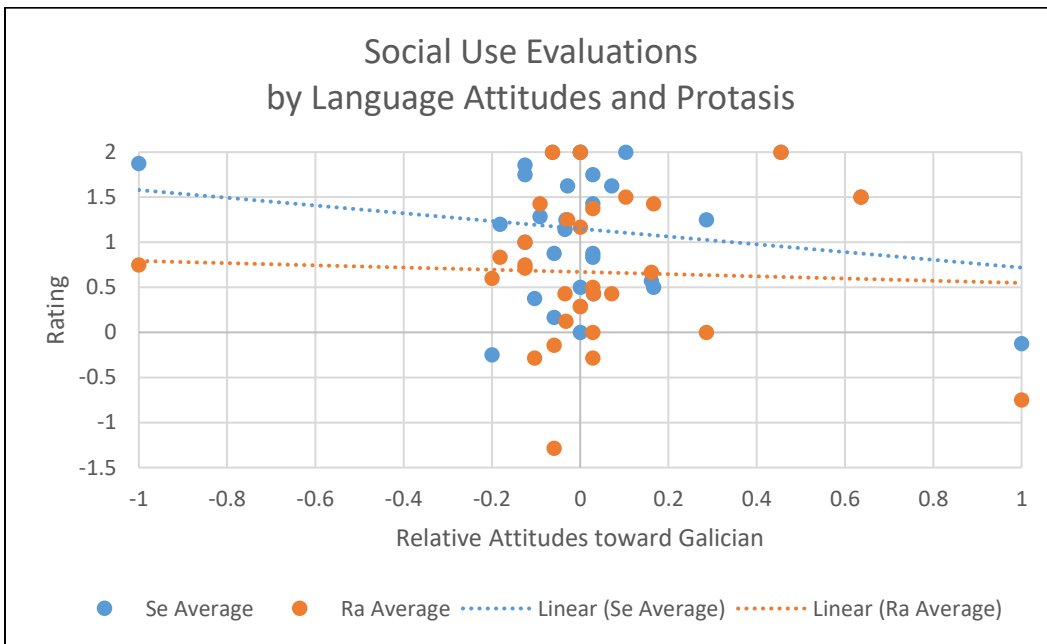**Figure 73: Acceptability judgment: Social Use--Interaction between Language Attitudes and Protasis**

Social Use Evaluations
by Language Attitudes and Protasis

Table 23: Acceptability judgment: Social Use--Interactions between Protasis and Possibility, Verb Group

	cantase		cantara	
	M	SD	M	SD
Contrary-to-Fact	1.16	0.49	0.75	0.88
Not C-t-F	1.17	0.70	0.70	0.77
ser	1.28	0.0	0.48	0.87
tener	---	---	1.55	0
other	0.97	0.70	0.70	0.77

Table 24: Acceptability judgment: Correctness--Interaction between Protasis and Gender, L1

	cantase		cantara	
	M	SD	M	SD
Female	0.36	0.66	-0.01	0.72
Male	0.35	0.89	-0.25	0.90
Castilian L1	0.17	1.14	-0.20	0.95
Simultaneous Bilingual	0.48	0.68	-0.03	0.83
Galician L1	0.10	0.29	-0.06	0.46

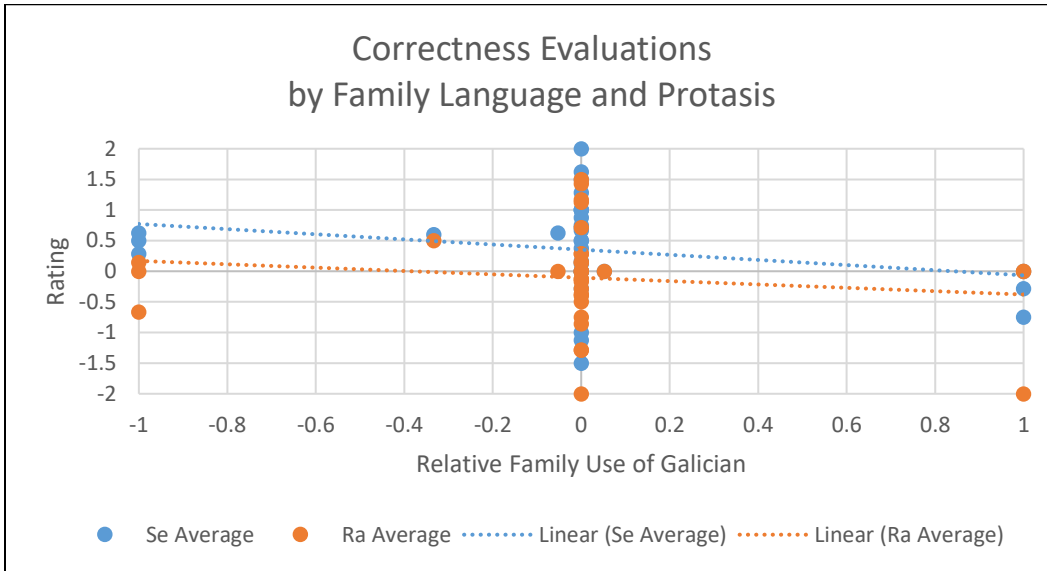**Figure 74: Acceptability judgment: Correctness--Interaction between Age and Protasis**

Correctness Evaluations by Age and Protasis

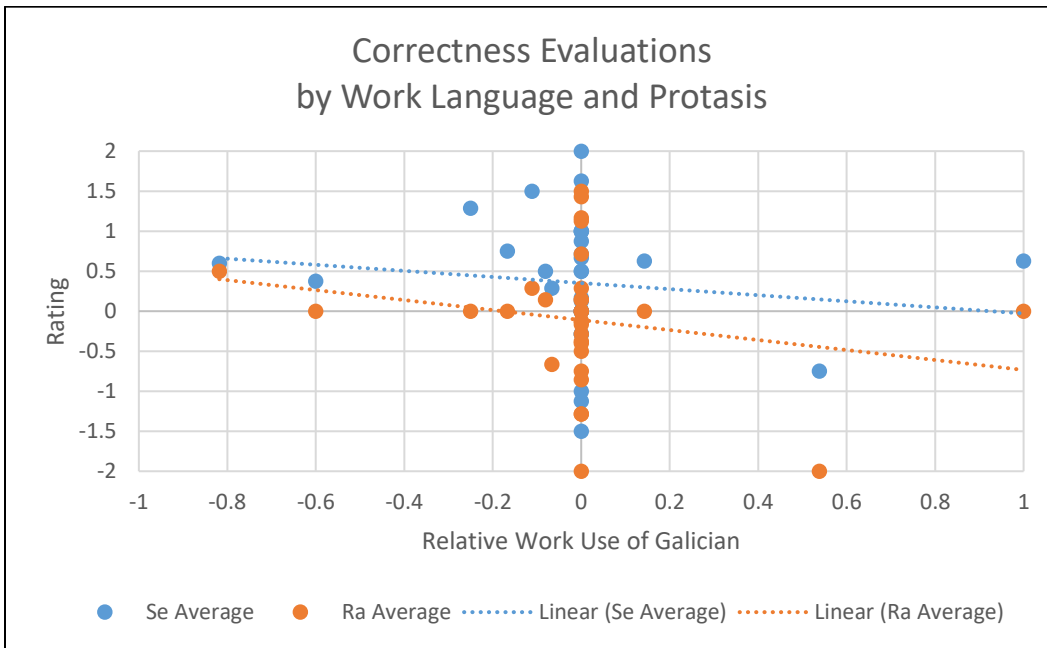**Figure 75: Acceptability judgment: Correctness--Interaction between School Language and Protasis**

Correctness Evaluations by School Language and Protasis

Figure 76: Acceptability judgment: Correctness--Interaction between Family Language and Protasis

Correctness Evaluations by Family Language and Protasis

Figure 77: Acceptability judgment: Correctness--Interaction between Work Language and Protasis

Correctness Evaluations by Work Language and Protasis

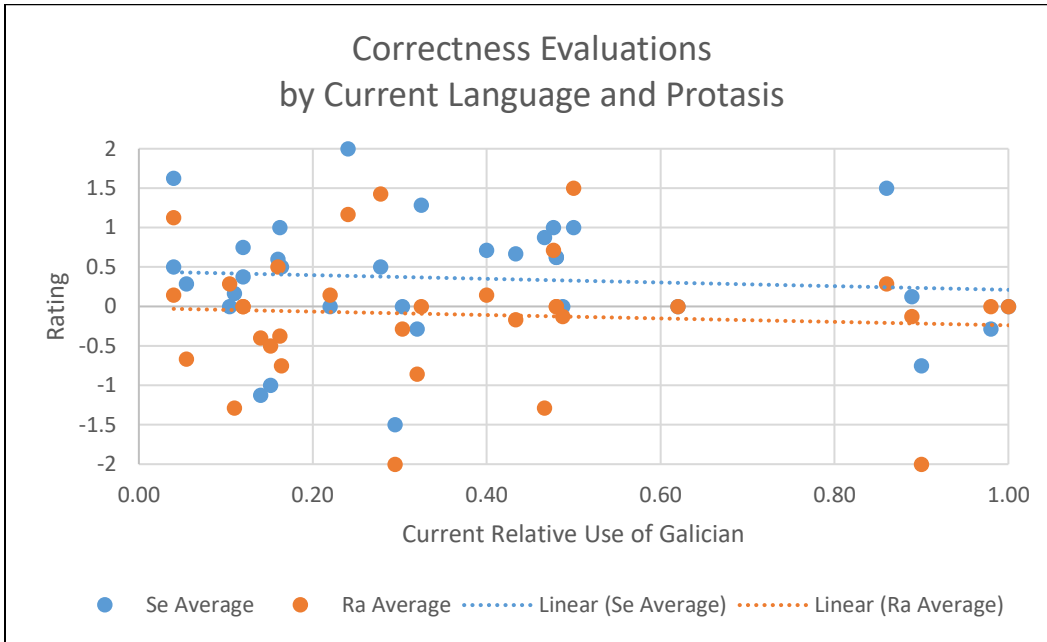**Figure 78: Acceptability judgment: Correctness--Interaction between Current Language and Protasis**

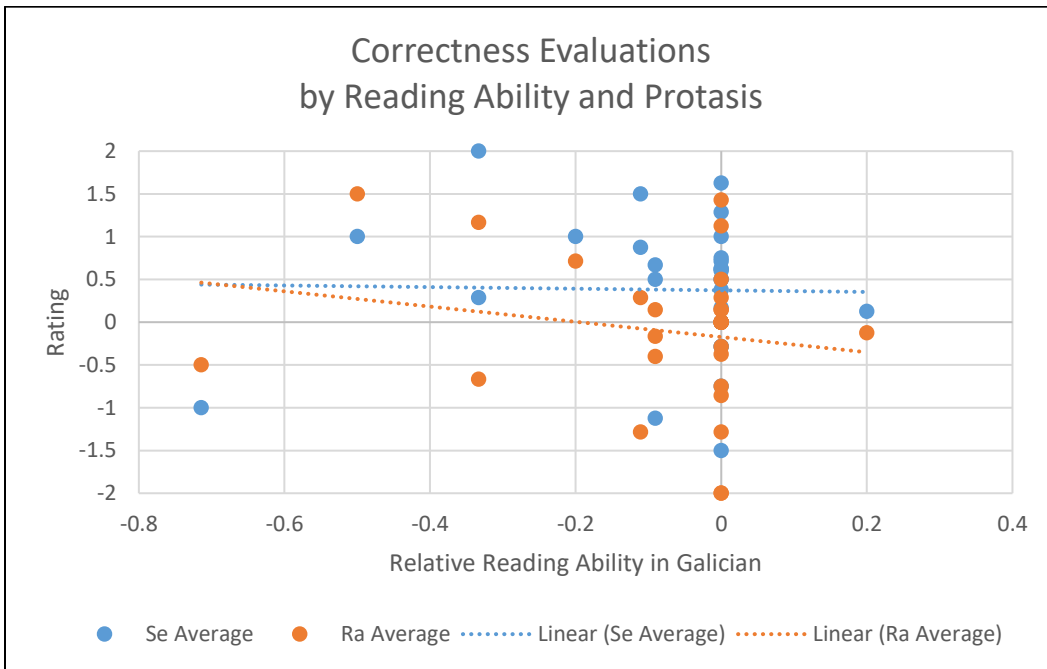**Figure 79: Acceptability judgment: Correctness--Interaction between Reading Ability and Protasis**

Figure 80: Acceptability judgment: Correctness--Interaction between Verb Group and Protasis

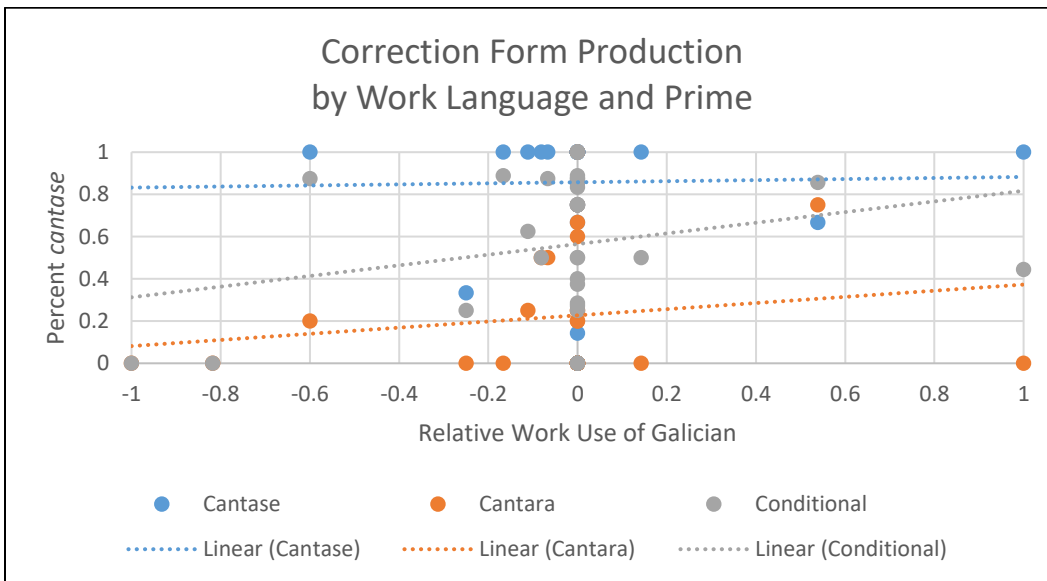|	*cantase*		*cantara*	
	M	SD	M	SD
ser	0.84	0.98	-0.53	1.22
tener	---	---	1.10	0
other	-0.38	1.21	-0.05	1.22

Figure 81: Acceptability judgment: Correction form production--Interaction between Work Language and Prime

Figure 82: Acceptability judgment: Correction form production--Interaction between Current Language and Prime
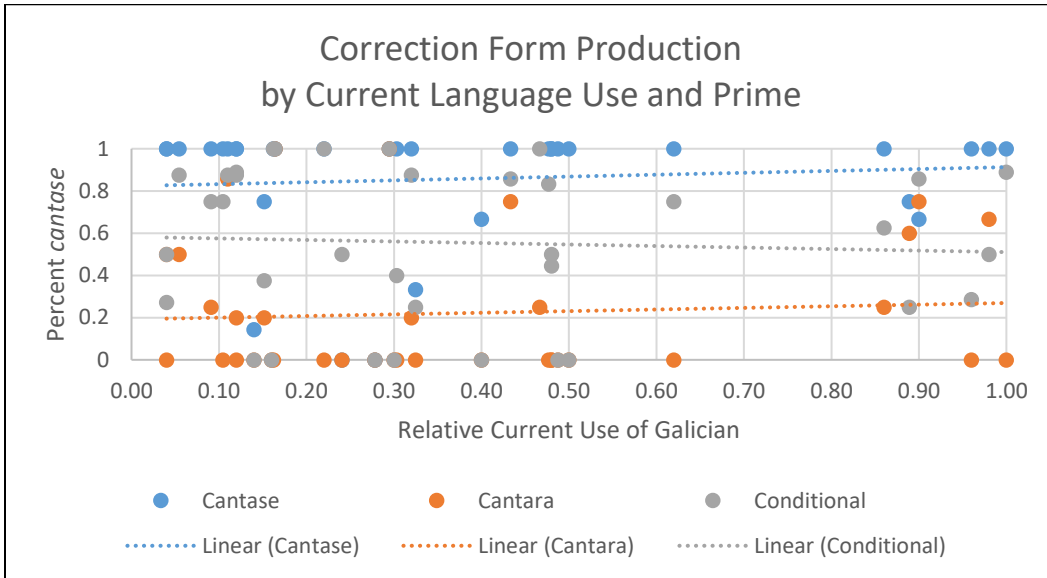

Correction Form Production by Current Language Use and Prime

Figure 83: Acceptability judgment: Correction form production--Interaction between Written Competence and Prime

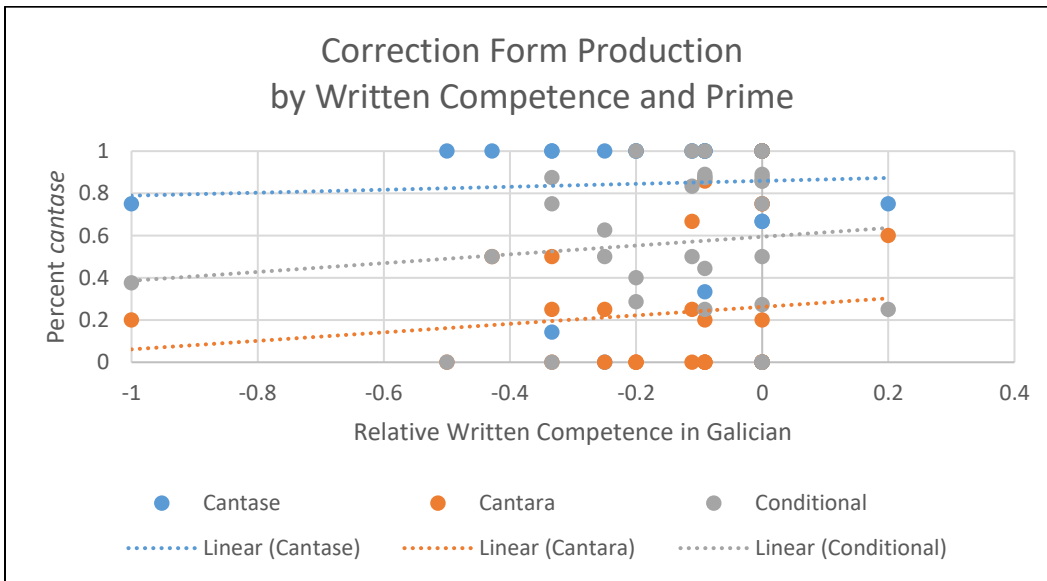Correction Form Production by Written Competence and Prime

Figure 84: Acceptability judgment: Correction form production--Interaction between Language Attitudes and Prime