Perception versus production of contact-induced cantase in Galician Spanish

Percepción frente a producción de la forma cantase inducida por contacto en el español de Galicia

ANA MARIA ANDERSON
Franklin & Marshall College, Lancaster, PA, EEUU

ABSTRACT. Though the existence of effects of Galician-Castilian contact on patterns of language use in Galicia have been well-documented, the exact scope and consequence of these effects is difficult to determine. This is due in part to the complex attitudes, both positive and negative, toward the Galician language. Because of this, societal reception of influences of Galician on Castilian —such as a heightened use of the verbal morpheme -se— is not easily predictable, even though the form enjoys greater use in Galicia than elsewhere in the Spanish-speaking world. This study explores attitudes of 39 Galicians from the province of Pontevedra toward the production of the -se form, as well as how these attitudes impact participants’ production of the form. Results indicate that the Galician-influenced form -se is more prestigious than the -ra form that is more common outside of Galicia, though whether participants correctly perceive which form is influenced by the minority language is unclear.

Keywords: imperfect subjunctive, language contact, perception, production, Galician Spanish.

RESUMEN. Aunque la existencia de efectos del contacto entre castellano y gallego sobre los esquemas de uso lingüístico en Galicia han sido bien documentados, el ámbito exacto y las consecuencias de esos efectos son difíciles de determinar. Esto se debe en parte a las actitudes complejas, tanto positivas como negativas, acerca de la lengua gallega. A causa de esto, la recepción social de las
influencias del gallego sobre el castellano, como en el incremento de uso del morfema verbal -se, no resulta fácilmente predecible, incluso aunque la dicha forma goce de mayor uso en Galicia que en el resto del mundo de habla hispana. Este estudio explora las actitudes de 39 gallegos de la provincia de Pontevedra hacia la producción de las formas en -se, y también qué impacto tienen estas actitudes en la producción de la forma por los participantes. Los resultados indican que la forma en -se, inducida por el gallego, es más prestigiosa que la forma en -ra, que es más común fuera de Galicia, aunque no está claro si los hablantes perciben adecuadamente qué forma está condicionada por la lengua minoritaria.

*Palabras clave:* imperfecto de subjuntivo, contacto de lenguas, percepción, producción, español de Galicia.

1. **INTRODUCTION**

   In the pre-modern era, Galician was the main language spoken in what is now the Autonomous Community of Galicia, and both the opportunity and the necessity of learning the Castilian language were notably absent outside the largest cities (García 1986; Mariño Paz 1998; Beswick 2002). In the movement toward modernity, however, mobility and social networks began to change (e.g. Britain 2010), and along with a massive migration to urban centers came an increased presence of Castilian in daily life (Bouzada Fernández 2003; Ramallo 2007; Monteagudo 2012). As a result, the Castilian language 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 (e.g. Faginas Souto 2003; Ramallo 2007; Pérez Castillejo 2012; Negro Romero 2013; Silva Valdivia 2013; Pérez Castillejo 2014). The exact scope of the consequences (for either language) are difficult to determine, however, as they extend beyond the linguistic into societal attitudes towards each. Castilian is now the “unmarked” linguistic choice in Galicia outside of the smaller villages, and speakers who switch from speaking predominantly Castilian to predominantly Galician in public —known as neofalantes— are always noticed, and not always in a positive sense, which “den[i]es their use of Galician] the invisibility and anonymity of ‘just talk’” (O’Rourke & Ramallo 2013b: 299). Additional criticism is also often

---

1 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
leveled at these and other speakers of Urban Standard Galician, the standard linguistic variety developed for formal settings and taught in schools, which is often seen as less “authentic” than traditional varieties of the language (Vidal Figueroa 1997; Kováčová-Moman 2007; O’Rourke & Ramallo 2011; 2013b). However, perhaps because speakers of this variety are overwhelmingly young, urban, and middle-class (Costa 2015; O’Rourke & Ramallo 2011), their “inauthentic” 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). These contradictory attitudes, in combination with the drastic increase in formal competence and prestige of the Galician language among the younger generations of speakers (Beswick 2002; Bouzada Fernández 2003), make it anything but clear how perceived influences of Galician on Castilian will be received in the region.

2. **CANTARA/CANTASE AS A SITE OF CONTACT-INDUCED CHANGE**

The focus of this work is on one likely effect of language contact between Galician and Castilian: heightened use of the verbal morpheme –*se*, from the verb form *cantase* (Anderson 2017b). 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 (RAG/ILG 2012). *Cantase*, on the other hand, marks the imperfect subjunctive. Castilian also inherited this structure in its evolution from Latin, but further changes in the language have led to general equivalence of *cantase* and *cantara* in almost all contexts in normative Spanish (Hermerén & Lindvall 1989; Carbonero Cano 1990)². The redundancy of having two practically equivalent forms has led *cantara* to steadily replace *cantase* throughout the Americas (e.g. Lavandera 1975; Moreno de Alba 1978; Navarro 1990; Gutiérrez, 1996; Rojo 1996). Rojo (1996) reported an average use of *cantase* around 18.4% in peninsular Spanish, though this varied from a low of 0% usage in urban Sevilla to highs of 20% or more in other

² 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 *guerría* would be accepted), and the few remaining indicative uses of *cantara*, most commonly found in journalistic language (Carbonero Cano 1990; Rojo & Vázquez Rozas 2014). However, even these exceptional contexts show evidence of overlap between the two forms; Day (2011, in Rosemeyer & Schwenter 2019) and Lunn (1995) report some *cantase* use in modal and politeness verbs, and fully non-subjunctive uses have also been reported (Bejarano 1962, Hermerén 1992). In any case, these exceptions are not considered in the present study.
regions such as Madrid. It has been postulated that *cantase*, while (and perhaps due to being) seldom attested, may be considered more formal and prestigious than *cantara* in multiple varieties of peninsular and American Spanish (Valeš 2004, 2006; Rojo 2011).

It is in Galician Spanish that *cantase* use remains most common, at least in formal tasks. In a written fill-in-the-blanks task, Kempas (2011: 254, 259) found 44.4% use of *cantase*, compared with 22.9% in Spain overall. Similarly, in an oral corpus of interviews of educated speakers of Galician Spanish, Rojo & Vázquez Rozas (2014) found that the *cantase* form was viewed more favorably among men and youth than women and older speakers in general. However, examining individual data led to an important caveat: group response rates may not be a reliable measure of this particular phenomenon in this particular area. They found strong idiosyncratic preferences that transcended age and gender, though *cantara* forms still predominated in their data as in Kempas (2011).

While in Galician both forms may be found expressing irreality in the protasis of *se* clauses (Veiga Rodríguez 1991), the overall frequency of use of these forms is opposite that in Castilian, with *cantase* employed in 82% of imperfect subjunctive contexts and *cantara* employed in 18% of these contexts in the Corpus de Referencia do Galego Actual (Rojo & Vázquez Rozas 2014: 244). This may be related to the perception that use of *cantara* as imperfect subjunctive is a Castilianism (ibid.); indeed, *cantase* is the only form of the imperfect subjunctive provided in the 2012 version of 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)” (RAG/ILG 2012: 109). To the best of my knowledge, no systematic sociolinguistic study on the use of these two forms in Galician exists to date.

Anderson (2017b) is the first study to examine these questions from a language contact perspective. She documented a heightened use of the *cantase* form in both oral and written Galician Spanish when compared to reports from other varieties of Spanish and, based on Silva-Corvalán (1994), concluded that sufficient evidence

---

3 Veiga Rodríguez (1991: 50) clearly argues that it is their expression of irreality, not the subjunctive/indicative opposition, that drives the use of these forms in conditional protases.

4 My thanks to the anonymous reviewer who pointed out that the suggested avoidance of -*ra* as subjunctive did not appear in previous versions of the *Normas*.

5 While language contact had previously been suggested as a likely influence on the frequency of these verb forms in Galicia (see, for example, De Sterck 2000), Anderson (2017b) was to my knowledge the first to attempt to verify this influence empirically.
exists to support a contact-based explanation for her findings. The present study uses the data from Anderson (2017b) to explore two related questions:

1. How do speakers of Galician Spanish perceive use of the forms *cantara* and *cantase*?
2. What relationship exists between speaker perception and production of these forms?

Given the conflicting ideological views surrounding Galician language use, these questions have important implications, and it is to such considerations that the rest of this work is dedicated.

### 3. METHODOLOGY

#### 3.1 Participants

This study incorporates data from 39 individuals living in the municipalities of O Grove and Marín in the Galician province of Pontevedra from a very young age.

Data were collected during the summer of 2015. Ages of participants at the time of data collection ranged from 18 to 75, and education ranged from primary only to completion of a graduate degree. The breakdown of participants by gender, age, and education level is shown in Table 1.

<table>
<thead>
<tr>
<th>Gender</th>
<th>Male</th>
<th>Female</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age</td>
<td>Under 40</td>
<td>40+</td>
</tr>
<tr>
<td>Education</td>
<td>P</td>
<td>S</td>
</tr>
<tr>
<td>Number</td>
<td>3</td>
<td>3</td>
</tr>
</tbody>
</table>

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

**Table 1.** Breakdown of participants by gender, age, and education

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. Due to the ideological

---

6. Variation in form production could be expected in both O Grove and Marin, which roughly correspond to sites of variation P13 and P18, respectively, on Map 23 of the Atlas Lingüístico Galego (García *et alii* 1990: 86-87).

7. 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.

8. Only 38 individuals are included in the table, as one older male participant did not report his education level.
complications involved in asking neofalantes to participate in oral tasks in Spanish, no neofalantes were included in this study.

3.2 Tasks

Participants in this study completed a total of five tasks: a bilingual language profile, a semi-structured group conversation involving 2-5 participants, a pseudo-matched-guise task, a fill-in-the blanks task, and an acceptability rating task. All tasks except the group conversation were completed individually on paper during the same session.
as the conversation. The tasks were carried out in the order listed in order to maintain participants’ unawareness of the focus of the study for as long as possible. Of course, all participants gave consent for their data and responses to be used for linguistic research.

### 3.2.1 Bilingual language profile

This task, designed to obtain information regarding participants’ demographic information, language practices, and self-reported language attitudes, consisted of a 5-page background questionnaire adapted from Birdsong, Gertken & Amengual (2012) to the Galician context. Because the present article does not focus on the sociolinguistic results of the study, this task will not be discussed further.\(^9\)

### 3.2.2 Semi-structured group conversation

In this task, groups of two to five participants were recorded in an hour-long conversation with the researcher about a variety of daily topics, including family, work, and current social issues such as emigration. These conversations were carried out in groups rather than individually with the interviewer in order to elicit the informal, vernacular speech styles that are the object of variationist sociolinguistic inquiry (Tagliamonte 2006). 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 (Labov 1972).

### 3.2.3 Pseudo-matched-guise task

This task accessed aural perception of the *cantara* and *cantase* forms.\(^{10}\) In a true matched-guise task (i.e. Lambert *et alii* 1960), speakers who are able to use two different languages or dialects are recorded producing the same text in both linguistic

---

\(^{9}\) Full results and analysis of this task and its relationship to the others can be found in Anderson (2017b).

\(^{10}\) 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.
varieties. The present study follows Silva-Corvalán (1984) in adopting a modified version of this task suitable to studying differences in the verb form used within one linguistic variety. A native speaker of both Galician and Spanish, 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 in Spanish 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 listened to one of the two versions of the recording and was asked to individually 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 was the matched-guise questionnaire designed by the *Seminario de Sociolingüística de la Real Academia Galega* research team (Fernández Salgado et alii 2002), 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.

### 3.2.4 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*, adapted from a Spanish language textbook activity, by filling in the blanks with a conjugated form of a verb given in parenthesis\(^ {11} \). This activity consisted of total of 22 blanks to fill, of which 9 were *si*-clauses meant to elicit an imperfect subjunctive form and the remaining 13 were meant to distract the participant from identifying the specific focus of this study.

\(^ {11} \) 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.
3.2.5 Acceptability rating 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 33 written statements (24 target statements and 9 distractors), which provided data about their acceptance of the two imperfect subjunctive forms in written contexts. The composition of the 24 conditional statements is summarized in Table 2.

<table>
<thead>
<tr>
<th>Protasis</th>
<th>Contrary-to-Fact</th>
<th>Not Contrary-to-Fact</th>
<th>Totals</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>-ra</td>
<td>-se</td>
<td>conditional</td>
</tr>
<tr>
<td>Apodosis</td>
<td>Conditional</td>
<td>2</td>
<td>2</td>
</tr>
<tr>
<td></td>
<td>Imperfect</td>
<td>2</td>
<td>2</td>
</tr>
<tr>
<td>Totals</td>
<td>4</td>
<td>4</td>
<td>4</td>
</tr>
</tbody>
</table>

Table 2. Composition of conditional statements in acceptability task

Given these sentences, participants were asked to rate a series of three statements related to each on a four-point descriptive scale. The series of statements rated for each sentence, respectively addressing Social Use, Individual Use, and Correctness, was:

1. Conozco a gente que lo diría así.
2. Yo lo diría así.
3. Está bien dicho.

For each of these statements, participants rated their reaction to the sentence by choosing between Estoy muy en desacuerdo, Estoy más o menos en desacuerdo, Estoy más o menos de acuerdo, and Estoy muy de acuerdo. Participants were instructed to rate the sentences based on their grammatical correctness rather than their propositional content.

---

12 A native speaker of Galician Spanish verified the grammaticality (apart from verb form) of all statements. A prior version of this task was piloted in Anderson (2017a).
13 e.g. Si yo fuera tú...
14 e.g. Si yo ganase la loteria...
In the second step of this task, participants made corrections to any statements which they had previously found to be unacceptable by reformulating the phrase so that it would be correct and/or something they themselves would say. This allowed for verification that rejections were related exclusively to verb form, and the proposed corrections supplement the written data from the fill-in-the-blank task discussed previously.

4. RESULTS

As stated previously, the main question 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. 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 is analyzed individually, with larger connections examined descriptively.

Because the focus of the present article is on the overall relationship between perception and production rather than on the contributions of social or linguistic factors, results presented here are limited to those directly relevant to this focus. For a more comprehensive breakdown of results, see Anderson (2017b).

4.1 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.

---

15 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 the larger 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.

16 Pluperfect subjunctive forms *hubiera/hubiese cantado*, are not considered in this analysis, 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).
In total, 130 tokens produced by 31 individuals were extracted from the recordings of the group conversations (the remaining 8 participants did not produce these forms). The forms produced were combined into 31 data points, with one data point indicating the percentage of cantase production for each participant who produced an imperfect subjunctive form in their spontaneous speech. A total of 25 of these tokens, or 19% overall, were produced in the cantase form. The average of participant rates of use of cantase was 26.9% (S.D. 0.38), though production ranged from 0% among some participants to 100% among others\(^1\).

4.2 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. Sixteen of the eighteen characteristics\(^1\) from the matched-guise questionnaire were grouped via factor analysis into three categories that can be loosely described as centering around friendliness (6 characteristics), capability (6 characteristics), and leadership/charisma (4 characteristics)\(^1\). These first two of these groupings are similar to the categories of personal appeal and capability, respectively, used in Loureiro-Rodriguez, Boggess & Goldsmith (2012), despite the fact that the latter’s groupings were created \textit{a priori} rather than through factor analysis. The factor loadings and resultant groupings from this analysis can be seen in Table 3.

\(^1\) Individual production data are included in Figures 3 and 4.

\(^1\) 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, \textit{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.

\(^1\) 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.
A linear regression with random effects for participant and characteristic was run in R. While the differences between guises did not reach significance, 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, results are examined descriptively here.

Without exception, participants rated the speaker in both guises above average with respect to Capability and below average for Charisma and Friendliness. However, while both guises were rated similarly in terms of friendliness, the canta se guise was seen as more charismatic and more capable than the cantara guise, as shown in Table 4.

Table 3. Matched-guise characteristics factor analysis: loadings and communalities

<table>
<thead>
<tr>
<th>Characteristic</th>
<th>Friendliness</th>
<th>Leadership/Charisma</th>
<th>Capability</th>
<th>Communalities</th>
</tr>
</thead>
<tbody>
<tr>
<td>Open</td>
<td>0.30</td>
<td></td>
<td>0.59</td>
<td>1.7</td>
</tr>
<tr>
<td>Attractive</td>
<td>0.57</td>
<td></td>
<td></td>
<td>1.2</td>
</tr>
<tr>
<td>Confident</td>
<td>0.62</td>
<td></td>
<td>0.42</td>
<td>2.1</td>
</tr>
<tr>
<td>Refined</td>
<td></td>
<td></td>
<td>0.65</td>
<td>1.4</td>
</tr>
<tr>
<td>Fun</td>
<td>0.92</td>
<td></td>
<td></td>
<td>1.1</td>
</tr>
<tr>
<td>Loyal</td>
<td>0.44</td>
<td></td>
<td></td>
<td>2.2</td>
</tr>
<tr>
<td>Generous</td>
<td>0.62</td>
<td></td>
<td>0.32</td>
<td>1.6</td>
</tr>
<tr>
<td>Sense of Humor</td>
<td>0.67</td>
<td></td>
<td>0.32</td>
<td>1.5</td>
</tr>
<tr>
<td>Intelligent</td>
<td></td>
<td></td>
<td>0.49</td>
<td>2.0</td>
</tr>
<tr>
<td>Leadership</td>
<td>0.73</td>
<td></td>
<td></td>
<td>1.5</td>
</tr>
<tr>
<td>Openminded</td>
<td>0.65</td>
<td></td>
<td></td>
<td>1.6</td>
</tr>
<tr>
<td>Practical</td>
<td>0.44</td>
<td></td>
<td>0.33</td>
<td>0.54</td>
</tr>
<tr>
<td>Ambitious</td>
<td>0.75</td>
<td></td>
<td></td>
<td>1.1</td>
</tr>
<tr>
<td>Nice</td>
<td></td>
<td></td>
<td>0.88</td>
<td></td>
</tr>
<tr>
<td>Hardworking</td>
<td>0.58</td>
<td></td>
<td>0.61</td>
<td>2.4</td>
</tr>
<tr>
<td>Trustworthy</td>
<td>0.46</td>
<td></td>
<td>0.31</td>
<td>2.1</td>
</tr>
</tbody>
</table>

Factor loadings less than 0.3 are suppressed.

Table 4. Characteristic group means by Guise

<table>
<thead>
<tr>
<th>Characteristic</th>
<th>Cantase</th>
<th>SD</th>
<th>Cantara</th>
<th>Mean</th>
<th>SD</th>
</tr>
</thead>
<tbody>
<tr>
<td>Friendliness</td>
<td>-0.19</td>
<td>0.85</td>
<td>-0.20</td>
<td>0.87</td>
<td></td>
</tr>
<tr>
<td>Charisma</td>
<td>-0.12</td>
<td>0.93</td>
<td>-0.24</td>
<td>0.99</td>
<td></td>
</tr>
<tr>
<td>Capability</td>
<td>0.65</td>
<td>0.74</td>
<td>0.50</td>
<td>0.80</td>
<td></td>
</tr>
</tbody>
</table>
Perception versus production of contact-induced cantase in Galician Spanish

4.3 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 were not considered.

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 (\(p < 2 \times 10^{-16}\) 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), 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.

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 (SD = 0.61) and 0.71 for cantara (SD = 0.83)\(^{21}\). 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 = 0.38, SD = 1.22) than was cantara (M = -0.09, SD = 1.25)\(^{22}\).

---

\(^{21}\) These high standard deviations were expected given the task design, as some statements were fully correct and thus received high ratings, while some statements were incorrect due to the inclusion of the imperfect in the statement apodosis. However, because the number of phrases containing imperfect apodosis forms was equal for both cantara and cantase protases, the elevated standard deviation values do not detract from the validity of conclusions drawn from the corresponding means.

\(^{22}\) Because there was no significant difference between Correctness and Individual use, only the former is considered here.
4.4 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, regardless of whether such a form could have been appropriate. This resulted in a total of 241 tokens being extracted for this task. 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).

As mentioned previously, written production was also studied through the corrections produced by participants on the acceptability judgment task. As participants only produced *cantara* or *cantase* in their corrected protases, all 540 tokens obtained were included. Of these tokens, 54% were in the *cantase* form.

However, this base percentage rate does not reveal the whole picture. 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. Among other results, the model produced revealed that Priming from the form in the rated sentence was highly significant ($F(2, 537) = 41.3, p < 0.001$) in conditioning form production. A post-hoc Tukey test indicated that differences in the rates of *cantase* production were highly significant for all three primes ($p < 0.001$). Figure 2 illustrates this unsurprising trend; the presence of a *cantara* prime in the rated prompt 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 2. Acceptability judgement: correction form production by prime
5. DISCUSSION

5.1 Perception of cantara and cantase in Galicia

Two of the tasks just discussed, 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. 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 in Galicia, no connections can be drawn to previous results in the region. Valeš (2006) did examine perception of these forms in Granada, however, and comparison with this study is illuminating. Valeš employed a forced-choice task to examine explicit perceptions of these forms, finding that focusing the attention of respondents on the verb form greatly increased the selection of cantase as compared to spontaneous production. This is similar to the results of the present study, though in Valeš (2006) cantase never fully overtook cantara even in the most formal, focused tasks, while in the present study cantase did surpass cantara in unprimed contexts and the former more than quadrupled the latter’s rate when primed with cantase.

While Valeš (2006) correctly highlights that increased formal use of cantase corresponds to higher overt prestige than usage data may indicate, denoting cantase as the prestige form in Granada may be problematic when in no case — neither in this study nor in Valeš (2004) — does its use surpass that of cantara. The somewhat more dramatic results in the present study, however, point to similar conclusions. 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* (at least in the areas studied), 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.

5.2 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 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.

---

23 This hypothesis does not explain the similarity of results with Valeš (2006), which is a topic that deserves further study. However, as indicated in the following section, a possible interaction between formality and location when comparing the two studies implies that somewhat different phenomena may be at play in producing the results found.

24 Rosemeyer & Schwenter (2019) found that priming has a greater impact for obsolescing forms than for mainstream forms. That priming has a strong effect for both imperfect subjunctive forms, then, implies that the *cantase* form may not be obsolescing in (this part of) Galicia, despite its disuse in other parts of the Spanish-speaking world.
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. 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.

In comparison to Valeš (2006), who used similar interview and fill-in-the-blanks tasks as well as the previously discussed forced-choice task, results of the present study are once again similar but indicative of the increased presence of cantase in the region studied as compared to Granada. However, as shown in Figure 3, not only is cantase use heightened overall in the present data, but the difference between the data sets also increases from 6% in informal speech to roughly 15% in the most formal contexts considered. It appears, then, that the impact of increased task formality (and therefore of the overt prestige of cantase) may be even greater in the region under consideration than in Granada.

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: 249) 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

25 While the forced-choice task in Valdeš (2006) is not precisely equivalent to any of the tasks in the present study, it comes quite close to the task of correcting sentences with cantaria in the protasis (in other words, those “unprimed” by either cantara or cantase), since in practice participants in the present study only ever produced cantara or cantase in their corrections and were thus, in a sense, “choosing” between the two forms.
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 4 and Figure 5 provide a snapshot of individual results for all participants for whom data was available from all tasks. In Figure 4, 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.\(^\text{26}\)

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

\[^{26}\text{Included in the reading as ‘accepted’ are both forms contained in a sentence that was holistically accepted and forms contained in a rejected sentence that were not altered in the produced correction, as the rejection of the sentence in these cases was assumed to be due to a factor other than the protasis form.}\]
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.

Figure 5 contains similar data to Figure 4, though the values have been inverted to indicate percentages relative to *cantara* rather than *cantase*. 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.

---

27 Although the Corrections (*cantase*) data shown in Figure 4 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.

28 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 4 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.
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 4 as compared to Figure 5. 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.

![Figure 5. Individual variation across tasks — *cantara*](image)

In addition to corroborating the conclusions drawn so far in this study, the data on individual behaviors and attitudes in Figure 4 and Figure 5 support the previous assertions of both Kempas (2011) and Rojo & 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 Rojo & Vázquez Rozas (2014), as both studies were limited to an examination of production alone.

6. CONCLUSION

As indicated in the previous section, *cantara* and *cantase* are both widely accepted among participants in this study as appropriate in the protasis of conditional statements, and neither appears to index greater covert prestige than the other. However, the picture changes in terms of overt prestige, with *cantase* perceived as both more
charismatic and more capable. Similarly, the *cantase* form enjoys greater use in more formal, form-focused tasks than in more spontaneous settings, in line with results from Valeš (2006). In contrast to the latter study, however, the present results indicate that *cantase* production in the most formal tasks does in fact exceed *cantara* production, clearly marking *cantase* as the overtly prestigious variant in the region studied here.

One question that remains unanswered is why *cantase* should enjoy more overt prestige than the *cantara* form that is popular elsewhere in the Spanish-speaking world. This is a particularly intriguing question given that the frequency of this form is attributable to contact with Galician (Anderson, 2017b). From this standpoint, correct use of *cantara* as imperfect subjunctive could be seen as a rejection of association with the historically lower-status Galician language. Why, then, would speakers overtly embrace *cantase* and reject *cantara*?

One possible contributing factor is widespread bilingual education throughout the region. Increased levels of formal education in both regional languages, but particularly in Galician, followed after the end of the Franco dictatorship in 1975 and the passage of the Law of Linguistic Normalization in 1983. 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. 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. Anecdotally this appears to be the case for some individuals, as a contact of mine from the region reported being advised in school to avoid *cantara* completely so as to not mistakenly employ it as pluperfect indicative in Castilian. While anecdotal evidence is far from proof, this raises the possibility that it is the use of *cantara* rather than *cantase* that is perceived as a marker of the Galician language. In other words, it is possible that speakers may privilege the actual Galician-influenced form while assigning less prestige to the form they (mistakenly) perceive as indicative of this influence.

The results of this study bring to light the importance of questioning assumptions that speakers accurately perceive the impact of contact on their various languages. Since participants were selected from only two municipalities, however, more data is needed to clarify this issue. A study of both perception and production that involves a random sample of participants with more diverse linguistic experiences and from throughout the Galician territory would enable clearer conclusions in general; this is
also important to be able to test the hypothesis about the impact of education against
the potential impact of exposure to Galician outside the classroom. Such a study
would also ideally include perceptions of Galician-language influence on each form,
something that the present study did not include. Finally, in order to paint a complete
picture of imperfect subjunctive patterns in the region, a sociolinguistic study on the
distribution of these forms in Galician itself is needed.

REFERENCES

ANDERSON, A. M. (2017a): “On the acceptability and use of -ra and -se in conditio-
languages2040021

ANDERSON, A. M. (2017b): Receding or resurgent? On the use of cantase (and cantara)
in Galician Spanish. Unpublished doctoral dissertation. University of Minnesota,
Minneapolis, MN.

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 filologia e historia
dedicados al profesor Manuel García Blanco (Acta Salmanticensia 16), pp. 77-86.

BESWICK, J. (2002): “Galician language planning and implications for regional identity:
org/10.1080/1460894022000026123

an easy-to-use instrument to assess bilingualism. COERLL, University of Texas
at Austin. https://sites.la.utexas.edu/bilingual

sols.v4i1.321

org/10.1002/9781444318159.ch10

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.):

COSTA, J. (2015): “New speakers, new language: on being a legitimate speaker of
a minority language in Provence”, International Journal of the Sociology of


Perception versus production of contact-induced cantase in Galician Spanish


