

Article

Priming as a Diagnostic of Grammatical Constructions: Second-Person Singular in Chilean Spanish

Matthew Callaghan  and Catherine E. Travis * 

ARC Centre of Excellence for the Dynamics of Language, School of Literature, Languages, and Linguistics, Australian National University, Canberra, ACT 0200, Australia; matthew.callaghan@anu.edu.au

* Correspondence: catherine.travis@anu.edu.au

Abstract: Structural priming has been described as a measure of association between constructions. Here, we apply priming as a diagnostic to assess the status of the Chilean second-person singular (2sg) *voseo*, which exists in variation with the more standard *tuteo*. Despite being the majority variant in informal interactions, Chileans are reported to have little metalinguistic awareness of *voseo* and they avoid the *vos* pronoun, in some cases using the *tú* pronoun with *voseo* verb forms, leading to proposals that *tuteo* and *voseo* are conflated into a single mixed form. The patterning for priming, however, indicates otherwise. Analyses of some 2000 2sg familiar tokens from a corpus of conversational Chilean Spanish reveal that a previous *tuteo* or *voseo* favors the repetition of that same form, indicating that speakers do treat these forms as distinct. We also observe that invariable forms with historically *tuteo* morphology are associated with neither *voseo* nor *tuteo*, while the invariable *voseo* discourse marker *cachái* ‘you know’ retains a weak association with *voseo*. Furthermore, while *tuteo* is favored with a *tú* subject pronoun, this effect does not override the priming effect, evidence that, even with a *tú* pronoun, *voseo* and *tuteo* are distinct constructions in speakers’ representations.

Keywords: structural priming; constructions; metalinguistic awareness; second-person singular; Chilean Spanish; *voseo*



Citation: Callaghan, Matthew, and Catherine E. Travis. 2021. Priming as a Diagnostic of Grammatical Constructions: Second-Person Singular in Chilean Spanish. *Languages* 6: 1. <https://doi.org/10.3390/languages6010001>

Received: 2 September 2020

Accepted: 18 December 2020

Published: 22 December 2020

Publisher’s Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Copyright: © 2020 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

According to a usage-based perspective, grammar is shaped by experience with language, through the repetition and conventionalization of frequently occurring patterns of language use (Bybee 2006, pp. 712–14). That is, speakers create utterances based on a probabilistic grammar that is conditioned by usage. While usage effects can be more global, impacting language change over time, here we examine the local usage effect of structural priming, seen in the tendency to repeat a structure that has been used previously in the same discourse. Specifically, we address questions of grammatical identity in the context of a change in progress in the second-person singular (2sg) in Chilean Spanish, using priming as a means of gauging speaker associations between constructions as evidenced in actual language use.

1.1. Metalinguistic Awareness of *voseo* in Chile

Almost all varieties of Spanish have a distinction between second-person “familiar” and “polite” forms, realized in the singular by *tú* and *usted* in Spain and across much of Latin America (the well-known T-V distinction described in the seminal work by Brown and Gilman 1960). In several Latin American varieties, including Chilean Spanish, there are two 2sg familiar forms, with *tú* existing alongside *vos*. This pronominal variation is accompanied by variation in the verbal paradigms, in what is known as *tuteo* and *voseo*. An example is given in (1), where in the one utterance and with the same verb (*tener* ‘have’), the speaker first uses a *tuteo* verb form and then a *voseo* verb form (captured with subscripts in the translation), both with unexpressed subjects (captured with a $\emptyset$ subscript in the translation).

(1) *cuando **tengas** la posibilidad de conocer,
tenís que ir.*

‘When you-Ø have-TUTEO the chance to get to know ((the place)),
you-Ø have-VOSEO to go.’

(Memories; 475–476; Trinidad)¹

It is widely understood that in late nineteenth-century Chile, *voseo* was a rural phenomenon and a mark of lower socioeconomic status, while *tuteo* was the preferred form of the educated classes and the urban elite (Lenz [1891] 1940, p. 263). Over 100 years ago, the renowned grammarian Andrés Bello rallied strongly against *voseo* in Chile, famously describing it as “*una vulgaridad que debe evitarse*” ‘a vulgarity that must be avoided’ (Bello [1891] 1981, p. 76). By the second half of the twentieth century, a change had been observed in use of the 2sg familiar forms whereby *voseo* verb forms had expanded to the speech of all social classes, including in urban centers, in some instances even coming to occur with a *tú* pronoun (e.g., Torrejón 1986). Recent variationist work has shown the change to be considerably advanced, with rates of *voseo* lying between 70% and 90% today in informal spoken Chilean Spanish between people who are familiar with each other (Callaghan 2020, p. 180; Fernández-Mallat 2018, p. 73; Rivadeneira Valenzuela 2016, p. 101). Despite such high rates in informal interactions, *voseo* is said to retain an element of stigmatization (e.g., Fernández-Mallat 2018, p. 66). This is seen in a number of ways: rates of *voseo* are lower in more formal interactions or between people who do not know each other (e.g., Fernández-Mallat 2018, p. 66); *voseo* is not taught in Chilean schools (e.g., Rivadeneira Valenzuela 2016, p. 90); and in writing, it is largely restricted to interactional genres, such as text messaging and online chats (Cautín-Epifani 2015). This is quite distinct from the situation in neighboring Argentina, where *voseo* is used as a standard form, though with distinct morphology from that used in Chile (Lipski 1994, p. 141).

Chilean speakers appear to have little awareness of their use of *voseo* verb forms. For example, responses to direct questions about *voseo* indicate that many Chileans either do not know what it is, often assuming *voseo* to refer solely to the pronoun *vos*, not the verb form (Huerta Imposti 2011–2012, p. 52; Hummel 2010, p. 112; Stevenson 2007, pp. 163–67), or consider it to be a feature of Argentinian, not Chilean, Spanish (Hummel 2010, p. 110; Stevenson 2007, p. 172; Torrejón 1986, p. 682). While the *voseo* verb forms are in wide use in Chile, the *vos* pronoun is very rare (Callaghan 2020, p. 133; Fernández-Mallat 2018, p. 71; Helincks 2012, p. 193; Rivadeneira Valenzuela 2016, pp. 99–100), so rare that, in some cases, the *voseo* verb form occurs with a *tú* pronoun, as in (2). This phenomenon, known as *voseo mixto verbal* ‘mixed verbal *voseo*’ (henceforth, “mixed *voseo*”) or *voseo culto* ‘educated *voseo*’ (Torrejón 1986), has been said to contribute to “invisibilization” of *voseo* (Huerta Imposti 2011–2012, p. 54), and to the understanding by some Chileans that their use of *voseo* is in fact *tuteo* (Carricaburo 1997, p. 34). Some respondents deny using *voseo* even while using it extensively (and, in some cases, categorically) in the very same discussion (Hummel 2010, pp. 111–12; Stevenson 2007, p. 165).

(2) *Pero **tú tenís** que trabajar po.*

‘But you-TÚ have-VOSEO to work, of course.’

(Cousins; 1245; Carmen)

More indirect research methods provide a similar picture. For example, in one written survey in which participants were asked to select the verb form they would use from sentences illustrating different options (e.g., *¿cómo estás/estái/está?* ‘how are you?’ with *tuteo*, *voseo* and the *usted* form, respectively), reported rates fell drastically short of rates of use in spontaneous speech of the same participants. The overall reported rate of *voseo* (vs. *tuteo* and *usted* forms) was 10% in the survey data, compared with an observed rate

¹ All examples come from the Corpus of Conversational Santiago Spanish (CCSS) (cf. Callaghan 2020). In parentheses following each example are the Transcription name, Intonation Unit number(s) of the example, and the speaker pseudonym. See Appendix F for transcription conventions.

of 54% in the speech data (Bishop and Michnowicz 2010, pp. 419, 424). It appears that, when responding to direct or indirect questions around the 2sg familiar forms, Chileans lean towards what they perceive to be the “correct” form (Bishop and Michnowicz 2010, p. 425). This supports an understanding of *voseo* in terms of covert prestige, as a widely used non-standard form, with a degree of stigma attached to its use (e.g., Fernández-Mallat 2018, p. 76; Stevenson 2007, p. 228).

Given the context of rapid language change, lack of metalinguistic awareness about *voseo*, mixing of the paradigms, and stigmatization, the question arises of whether Chilean Spanish speakers separate *tuteo* and *voseo* in their grammars, or whether the two paradigms have lost their analyzability and become conflated into a single “mixed *voseo*”. Some forty years ago this idea was put forward, and it was proposed that mixed *voseo* might come to replace *tuteo* as the “norma universal de tratamiento de los chilenos cultos en situaciones informales y familiares” ‘universal standard form of address for educated Chileans in informal and familiar situations’ (Torrejón 1986, p. 682).

Nevertheless, there is also some evidence from patterns of spoken language use that Chileans are not oblivious to the existence of *voseo*. For example, one study of the spoken Spanish of Chilean migrants in Montreal found higher rates of *voseo* in interactions with other Chileans compared with conversations with non-Chileans (Fernández-Mallat 2011, p. 47). In this paper, we address the question of Chileans’ awareness of *voseo* by turning to structural priming.

1.2. Structural Priming in Variation and Change

Structural priming, also known as persistence, perseveration, or linguistic parallelism, is the favoring of a previously produced structure over an alternative that has not been recently used. For example, there is a tendency to repeat the same form for future temporal reference in English, such that speakers are more likely to choose *be going to* over *will* if there is an instance of *be going to* in the previous discourse (e.g., Szmrecsanyi 2005). Priming was first observed in spontaneous speech data in early variationist sociolinguistic work, for example for clitic pronouns in Montreal French (*on, tu/vous, nous*) (Sankoff and Laberge 1978); expression of plural -s in the Spanish noun phrase in Puerto Rican Spanish ([s] vs. [h]) (Poplack 1980); and the agentless passive vs. active in English (Weiner and Labov 1983). It has since been observed in laboratory settings, for example for the double object vs. prepositional dative construction, and agentless passive vs. active clauses (e.g., Bock and Griffin 2000; Loebell and Bock 2003), as well as in corpus linguistics, in both speech (e.g., Gries 2005; Jaeger and Snider 2007) and writing (e.g., Szmrecsanyi 2005). The robustness of structural priming effects in naturalistic discourse and laboratory contexts provides overwhelming evidence that “morphological and syntactic variation is controlled by a tendency to preserve parallel structures” (Labov 1994, p. 550).

This preservation of parallel structures has also been reported for the Spanish 2sg forms, whereby *tuteo* tends to follow *tuteo*, and *voseo* tends to follow *voseo*. This is illustrated in example (3), where the speakers produce a string of *tuteo*, marked in bold, before a switch to *voseo*, underlined. Priming was found to be the strongest internal factor conditioning the choice between *voseo* and *tuteo* in sociolinguistic interview data from Uruguayan Spanish (Carvalho 2010, p. 17). Further, in a corpus of conversational Spanish from Santiago de Chile, it was observed that the use of *tuteo* was categorical closely following another *tuteo*, and that there was a greater chance of using *voseo* in the context of a preceding *voseo* (Fernández-Mallat 2018, pp. 78–79).

Given the robustness of priming as a factor conditioning variation, this patterning may seem entirely unsurprising. However, priming is not a foregone conclusion for the 2sg familiar forms in Chile, where Chileans claim not to use the *voseo* form, and mix it with a *tú* pronoun. If Chileans genuinely do not perceive *voseo* forms to be distinct from *tuteo*, then we would not expect a previous *tuteo* to favor a subsequent *tuteo* over *voseo*.

(3) Tatiana: cuando **tú te dejas** de acordar.

Ahí realmente *mueres*.

Viviana: ... Sí=,
 pero t- --
 mient- --
si tú quieres estar con alguien.
Estás todo el rato.

Tatiana: [Con el alma].

Viviana: [Porque estái pensando].

Tatiana: Sí po.

Viviana: Cierto.

Tatiana: 'when you_{-TÚ} stop_{-TUTEO} remembering.
 then you_{-Ø} really die_{-TUTEO}.'

Viviana: '... Yes,
 but --
 while- --
 if you_{-TÚ} want_{-TUTEO} to be with someone.
 you_{-Ø} are_{-TUTEO} ((with them)) all the time.'

Tatiana: '[With your soul].'

Viviana: '[Because you_{-Ø} are_{-VOSEO} think]ing.'

Tatiana: 'Yeah.'

Viviana: 'Exactly.'

(Barbecue; 1040–1050; Tatiana, Viviana)

Of particular relevance here are the observations that structural priming occurs between related constructions, and that the strength of priming varies according to the degree of relatedness. For example, structural priming is enhanced by repetition in the prime and target, across verbs (Pickering and Branigan 1998, p. 641; Rosemeyer and Schwenter 2017, p. 27; Szmrecsanyi 2005, p. 133), TAM (Travis 2007, p. 126) or referents (Torres Cacoullos and Travis 2018, p. 91). Similarly, though it occurs both across and within languages, it is stronger within the one language (Torres Cacoullos and Travis 2018, p. 200; Travis et al. 2017, pp. 287–88). Conversely, priming has been shown to be attenuated, or absent, between more weakly related constructions. For example, English future *be going to* (vs. *will*) is primed by previous instances of future *be going to*; it is also primed by instances of motion verb *go*, but to a lesser degree (Szmrecsanyi 2005, p. 139). Research in grammaticalization has reported a diminishing of priming as one form gains autonomy from its source construction. The Spanish progressive *estar* 'be' + VERB-GERUND construction was primed by *estar* + ADJECTIVE and *estar* + LOCATIVE forms when the construction first emerged and was still closely related to these source *estar* constructions. But the priming weakened over time, and was eventually lost, as the progressive construction grammaticalized and gained autonomy from other *estar* constructions (Torres Cacoullos 2015). Finally, variable velar vs. alveolar realizations of English *-ing* in both nominals (e.g., *ceiling* and *morning*) and gerunds (e.g., *working* and *jumping*) are subject to priming, but only within the one morphological category, such that, for example, *working* primes *jumping*, but *ceiling* does not (Tamminga 2016, pp. 348–49). Tamminga regards this lack of priming as an indication that "the prime and target are not representationally related" (2016, p. 337).

From a usage-based perspective, we interpret structural priming as a kind of local entrenchment, whereby the presence of an analogous form in the preceding discourse activates, or provides a temporary boost in the strength of, its mental representation, promoting subsequent use (cf. De Smet 2016, pp. 86–87). Priming thus only occurs across constructions which share a common mental representation, and provides evidence that

speakers perceive two distinct instantiations to be related, that is, exemplars of the same construction. For *voseo* and *tuteo*, this means that if they are conflated, the occurrence of a previous *voseo* or *tuteo* form should not impact the subsequent choice a speaker makes. However, if the rate of *voseo* or *tuteo* rises in the presence of a preceding *voseo* or *tuteo*, respectively, then this would suggest that speakers associate one instance of *tuteo* more closely with another instance of *tuteo* than with an instance of *voseo*, and vice versa; that is, they distinguish between the two paradigms.

Our analyses reveal that priming does play a strong role in conditioning the *tuteo/voseo* variation, leading us to propose that [VERB-TUTEO] and [VERB-VOSEO] exist as two independent constructions. Having established this, we can then use priming as a diagnostic to assess the status of different kinds of constructions, as more or less strongly associated with *tuteo* or *voseo* in speakers' mental representations. We consider two types of non-variable forms, the first a schematic construction deriving from *tuteo* (the Imperative and the Preterit TAMs), and the second a lexically specific construction derived from *voseo* (the highly frequent discourse marker *cachái* 'you know'). Evidence that these non-variable forms have attained some autonomy from their source constructions would be seen if *tuteo* is not favored in the context of a previous Imperative or Preterit, and if *voseo* is not favored in the context of a previous *cachái*. We then turn to the status of "mixed *voseo*" (the use of a *tú* pronoun with a *voseo* verb form) to ask whether this combination might be less sensitive to *voseo* > *voseo* priming, and thus less strongly associated with the *voseo* form than is *voseo* without a *tú* pronoun.

We begin in Section 2 with an overview of the corpus used for this research, and in Section 3 describe the nature of the Chilean *tuteo/voseo* variation. Section 4 then presents the results for priming, before summing up our findings regarding the mental representations of *voseo* in Chilean Spanish in Section 5.

2. Corpus of Conversational Santiago Spanish

While priming has attracted great attention in the lab, given the informal nature of *voseo*, its association with speech, and its stigmatization, the most reliable data in which to accurately observe its patterning is spontaneous speech, for which we use the Corpus of Conversational Santiago Spanish (CCSS) (cf. Callaghan 2020).

The CCSS comprises recordings of 36 residents of Santiago made in 2014 and 2015 by the first author. Participants were recruited from the Province of Santiago de Chile, an area enclosed by mountains to the east, west and north. They were all born and raised in Chile with Chilean parents, were residing in Santiago at the time of the recording, and had not lived outside Santiago for an extended period of time in the last five years—that is, all were *santiaguinos*, or *capitalinos*, as people from Santiago are known in Chile.

The interactional nature of the second person and the association of *voseo* with informal speech between people who know each other render conversation the most appropriate genre in which to study its use. Thus, all recordings are of conversations between friends, family members and colleagues, that took place in the home, car or workplace of one of the participants, that is, the kind of interaction they would be likely to have ordinarily. Participants were told that the focus of the study was the relationship between language and society in Chile, but no further information about any linguistic features of interest was given. They were not assigned topics to discuss, but rather were free to talk about anything that they deemed appropriate. A range of different topics came up, including football, movies, travel, parties, mutual friends, school, work, children, debt collectors, the police, and so on.

Out of a total of some 50 recordings comprising 36 h of speech, 17 were selected for inclusion in the corpus. These were primarily two-party conversations ($n = 12$), but there were also four three-party and one four-party conversation. Recordings were made by 13 research assistants, each of whom was a participant in one or two conversations. The recordings were an average of 50 min long (ranging from 21 to 82 min), and an average of 30 min

were transcribed from each, generally starting from some 15 min into the conversation. This provides a total of nine hours of speech and approximately 110,000 words for analysis.

Table 1 presents the breakdown of the participants for age and gender. Participants range from 20 to 62 years of age, split in this table at 35 years, giving a relatively even distribution in each age group. They were of varying social backgrounds: 22 of the 36 have university education; they have a wide range of occupations, working in areas such as education, health care, management, construction, hospitality, and transport; and they come from different *comunas*, regions of Santiago that have been used to determine socioeconomic class (e.g., [Stevenson 2007](#), p. 233).²

Table 1. Corpus of Conversational Santiago Spanish: participant distribution by age group and gender.

	Males	Females
>35 (36–62 years)	4	13
≤35 (20–35 years)	8	11
	12	24

The data were transcribed by five trained Chilean Research Assistants, who produced time-aligned orthographic transcriptions in F4 ([Dresing and Pehl 2015](#)). The transcription method followed the principles outlined in [Du Du Bois et al. \(1993\)](#), breaking the speech into Intonation Units (IUs), prosodic units generally of two to five words, and rarely more than one clause (though often less, e.g., for single-word IUs consisting of backchannels, response tokens, or discourse markers, such as *sí* ‘yes’, *cierto* ‘exactly’, or *cachái* ‘you know’). A set of orthographic conventions was established, particularly important for the feature under study which occurs predominantly in informal speech and is consequently highly variable in orthographic representation. This is particularly so for syllable-final /s/, which is variably aspirated or elided in Chilean Spanish (cf. [Rivadeneira Valenzuela 2009](#), pp. 119–20). Thus, pronoun *vos* is spelt as *vo*, *voh*, *bo*, or *boh*; this was standardized to *vos* for the transcriptions here. In addition, we follow standard spelling conventions in writing *voseo* forms ending in /ay/ without final /s/ (-ái or -ai) and those ending in /is/ with final /s/ (-ís) (see Table 2).

This conversational corpus, collected by community members and transcribed following defined protocols, provides ideal data to probe the *tuteo/voseo* variation in Chile, and in particular, the impact that priming has on its use.

3. The Chilean Second-Person Singular

Chilean Spanish is characterized by the co-existence of two 2sg familiar forms, both in the pronouns *tú* and *vos*, and in the verb forms, *tuteo* and *voseo*, illustrated in (4) and (5), respectively. These forms can be contrasted with *usted*, the so-called “polite” 2sg form (which takes a third-person singular verb form). There are just 55 occurrences of *usted* verb forms in CCSS, accounting for two percent of all 2sg verbs, almost all of which are either in quoted speech or are to older speakers. Thus, we consider *usted* to be outside the variable context, which we define as second-person singular *familiar* verbal forms.

(4) *Tú tienes ahorros.*

‘You_{-TÚ} have-TUTEO savings.’

(Savings; 1302; Carolina)

(5) *o sea vos no tenís ningún concierto.*

‘In other words you_{-VOS} have-VOSEO no concert.’

² As the goal of the CCSS was to record informal interaction between people who know each other well, almost all recordings are between people of a similar age and socioeconomic class. This homogeneity does not allow for a rigorous investigation of interlocutor effects. Observations in a subset of the data suggest that males may increase their rate of *tuteo* in interactions with females, thus accommodating to the marginally greater favoring of *tuteo* by females ([Callaghan 2020](#), pp. 235–36).

(Football; 947; Matías)

Chilean Spanish has distinct *voseo* forms throughout most of the verbal paradigm as seen in Table 2, which presents the morphology of the set of TAMs for verbs ending in *-ar* on the left, and *-er* or *-ir* on the right (illustrated for *-er*, which is identical to *-ir* for *voseo*, though not for *tuteo*). As can be seen, there are distinct forms for *tuteo* and *voseo* for all but the Imperative, Preterit and Synthetic Future, which we discuss below.

Tuteo and *voseo* differ in their morpho-phonological shape—*tuteo* is most characterized by /as/ and /es/ endings, and *voseo* by /ái/ and /ís/ endings. This is likely to promote priming within the paradigms, as shared shape renders individual instances of *tuteo* or of *voseo* obviously similar, even for those ignorant of their grammatical status. The *voseo* > *voseo* priming is not, however, dependent on shared morpho-phonological shape, as we show below (Section 4.1).

Table 2. Chilean 2sg verb morphology by tense-aspect-mood (Adapted from Torrejón 1986, p. 678).

	–AR Verbs		–ER/–IR Verbs	
	<i>tuteo</i>	<i>voseo</i>	<i>tuteo</i>	<i>voseo</i>
Present indicative	<i>lleg- as</i>	<i>lleg- ái</i>	<i>com- es</i>	<i>com- ís</i>
Imperfect indicative	<i>lleg- abas</i>	<i>lleg- abai</i>	<i>com- ías</i>	<i>com- íai</i>
Conditional	<i>lleg- arías</i>	<i>lleg- aríai</i>	<i>com- erías</i>	<i>com- eríai</i>
Present subjunctive	<i>lleg- ues</i>	<i>lleg- uís</i>	<i>com- as</i>	<i>com- ái</i>
Imperfect subjunctive	<i>lleg- aras</i>	<i>lleg- arai</i>	<i>com- ieras</i>	<i>com- ierai</i>
Imperative	<i>lleg- a</i>		<i>com- e</i>	
Preterit	<i>lleg- aste</i>		<i>com- iste</i>	
Synthetic Future	<i>lleg- arás</i>		<i>com- erás</i>	

Historically, the Imperative, Preterit and Synthetic Future had distinct *tuteo* and *voseo* forms (e.g., Torrejón 1986, p. 679), but today, only what were the *tuteo* forms remain in wide use, while the *voseo* forms are vanishingly rare. In the CCSS, of 560 Imperative tokens, there are no distinct *voseo* forms; of 167 2sg familiar Preterit tokens, there is only one token of the *-stes* form, which is purported to be historically *voseo* (Rivadeneira Valenzuela 2009; Torrejón 2010)³; and the Synthetic Future does not occur, as the Analytic Future is the preferred form ($n = 86$). As invariable forms, these TAMs fall outside the variable context and must be excluded from analysis (as has been done in previous variationist studies of *voseo*, e.g., Fernández-Mallat 2018, p. 100; Rivadeneira Valenzuela 2016, p. 70).

However, here our interest is in how speakers classify these forms. Do they remain associated with their *tuteo* origins, as an instance of a general [VERB-TUTEO] construction, or are they genuinely neutral, as syncretic forms that collapse the *tuteo/voseo* distinction? As the *tú* pronoun occurs with both *tuteo* and *voseo*, presence of a subject pronoun is not a reliable indication of the form of these verbs, and furthermore, subjects are most commonly left unexpressed, as in example (6) in the Imperative, and (7) in the Preterit. Thus, we appeal to priming to establish the status of these forms: if they indeed share a mental representation with the *tuteo* forms, then we would expect them to favor a subsequent *tuteo*; the lack of such a priming effect would suggest that this association has been lost.

(6) *Consíguete un pololo.*

‘Get-IMPERATIVE-FIXED yourself a boyfriend.’

(Cousins; 1180; Emilia)

(7) *te pusiste huevón ya,*

‘you-Ø already turned-PRTERIT-FIXED into an idiot,’

(Takeaway; 4; Claudia)

³ Though we note that the rarity of final /s/ is to be expected, given the general Chilean phonological patterning of aspiration or elision of syllable-final /s/ (e.g., Rivadeneira Valenzuela 2009, p. 78).

Another non-variable form is the lexically specific, fossilized construction *cachái*, the *voseo* form of the verb *cachar* ‘to understand, get’ (possibly from English ‘to catch’, Rivadeneira Valenzuela 2016, p. 100). In the Present Indicative, *cachar* occurs virtually categorically in the *voseo* form in the CCSS (with just one *cachas* in contrast to 360 *cachái*), consistent with what has been reported in other studies (e.g., Fernández-Mallat 2018, p. 70; Rivadeneira Valenzuela 2016, p. 100; San Martín Núñez 2011, pp. 159–62). Like the invariable TAMs above, then, *cachái* falls outside the variable context. But the question remains, however, of whether speakers perceive *cachái* to be an instance of the [VERB-VOSEO] construction. It is used primarily as a discourse marker, akin to English ‘you know’, as in (8) (322/360). It also occurs as a full verb ($n = 28$), for example, with a direct object, meaning to perceive or realize something (cf. Urzúa-Carmona 2006, pp. 100–5), as in (9), and with a variably expressed pronoun (nearly always *tú*, 10/11 instances).

(8) *tú enrolái como si fueran más delgados,*

cachái?

‘you-_{TÚ} roll-_{VOSEO} ((cigarettes)) as if they were thinner,
you-_Ø know-_{CACHÁI}?’

(A bit of everything; 401–402; Andrea)

(9) *y ahí cachái cómo es el nivel de ellos po huevón.*

‘And that’s when you-_Ø realize-_{CACHÁI} what level that they’re at man.’

(Barcelona; 925; José)

Cachái has emerged within the last fifty years or so, in parallel with the reported rise of *voseo* (e.g., San Martín Núñez 2011, pp. 159–62). It is highly frequent in the speech of young Chileans today, accounting for 94% of all instances of the verb *cachar* in the CCSS, and around 25% of all 2sg familiar forms produced by speakers of 35 years and younger (of a total of 1242 tokens).⁴ For one speaker (Marcela, 31 years old; 123 tokens), *cachái* constitutes 55% of her 2sg familiar tokens. Though *cachái* does occur in discourse of older speakers, it is much rarer, accounting for just 3% of all 2sg familiar forms (of a total of 904).

Do speakers consider *cachái* to be a *voseo* form, that is, a manifestation of the [VERB-VOSEO] construction, or are they unaware of its structure? Although it has a similar phonological shape to the *voseo* Present Indicative, with the /ay/ ending, there is no *tuteo* form to contrast it with, and it has been proposed by some authors that it has lost its status as a verb (e.g., Gille 2015; Mondaca Becerra et al. 2015; Rivadeneira Valenzuela 2016, p. 100). Here, too, priming provides insight into the status of this form: if it is recognized as a *voseo* form, then it should prime a subsequent *voseo*; the lack of such an effect would suggest that it has become fully autonomous and is no longer attached to the *voseo* paradigm.

A final consideration is the impact of a pronoun. As already noted, Spanish has variable subject expression, and thus subject pronouns may be expressed (as *tú* or *vos* in this case) or they may be left unexpressed. The *vos* pronoun, however, shows very minimal use. The CCSS presents a total of 304 tokens of the *tú* subject pronoun, but just 10 of the *vos* subject pronoun.⁵ And while the *vos* subject pronoun does not occur with *tuteo*, the *tú* pronoun occurs with both *tuteo* and *voseo* verb forms, resulting in the occurrence of both [*tú* + VERB-TUTEO] and [*tú* + VERB-VOSEO] constructions. The overwhelmingly favored form with both *tuteo* and *voseo* (accounting for 80% of tokens occurring in the variable context) is with an unexpressed subject. While the use of *voseo* with no pronominal subject has been termed *crypto-voseo*, in the sense that the very existence of a parallel *voseo* paradigm is concealed by the absence of the *vos* pronoun (Lipski 1994, p. 143), it has also been proposed that the “underlying” pronoun in such a context is “always *tú*, as a neutral unmarked form”

⁴ Beyond *cachái*, we find no evidence of lexical effects, though it may simply be that there is insufficient data to test this, as only six verb types have 50 or more tokens with variable *tuteo/voseo* (for discussion, see Callaghan 2020, pp. 216–26).

⁵ The favoring of the *tú* form holds both for subject pronouns (*tú*) and prepositional object pronouns (*ti*); *tú* forms represent 97% (398/412) of all 2sg familiar pronouns in the CCSS.

(Rivadeneira Valenzuela 2016, p. 93). Here again, we use priming to ascertain the status of “mixed *voseo*” [*tú* + VERB-VOSEO], and the impact the presence of a *tú* pronoun has on the representation of a *voseo* verb form.

We have established, then, that instances of the Imperative, Preterit and *cachái* fall outside the variable context, but that tokens with a *tú* pronoun do exhibit variability and thus can be included. Figure 1 depicts the distribution in the data across the non-variable and variable contexts. What is of particular note here is that the non-variable contexts make up over one half of all 2sg familiar instances—the Imperative and Preterit together represent 35%, and *cachái* a further 17%. Within the variable context, the rate of *tuteo* is 20%, in contrast with 80% *voseo*. What factors impact this variation?

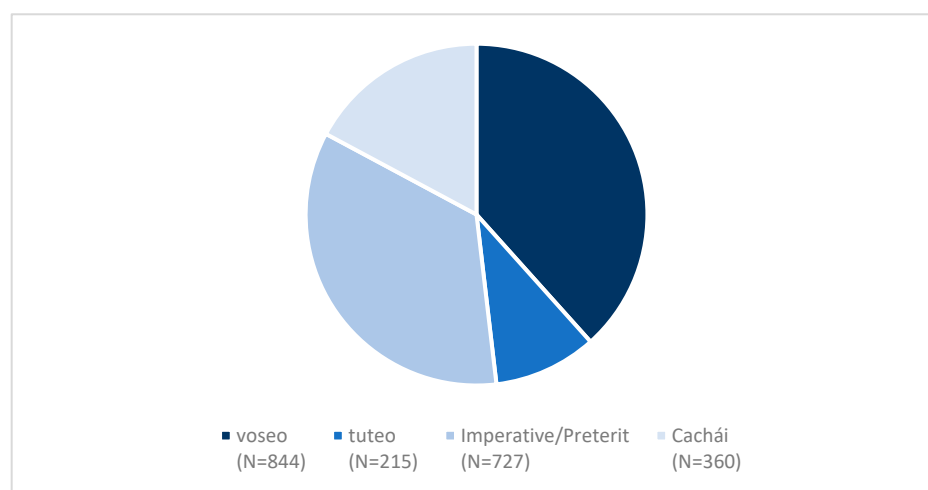


Figure 1. Distribution of the data in and out of the variable context for 2sg familiar verbal forms.

Change over Time

The rate of *tuteo* has declined in recent Chilean history, as *voseo* has grown in use. This is evident in the data under study here in comparisons across speakers of different ages. As has been robustly demonstrated in sociolinguistic research, the relative stability of patterns of speech in adults allows for comparisons across age groups to serve as a proxy for language use in different time periods, and differences as indicative of change in apparent time (Sankoff 2006).

The rates of *tuteo* vs. *voseo* by age are presented in Figure 2, which illustrates an overall drop in *tuteo* for younger speakers, supporting the change over time towards *voseo*. Comparing the age groups in Table 1, the rate of use drops from 32% *tuteo* (138/430) for the older speakers to just 12% (77/629) for the younger speakers.

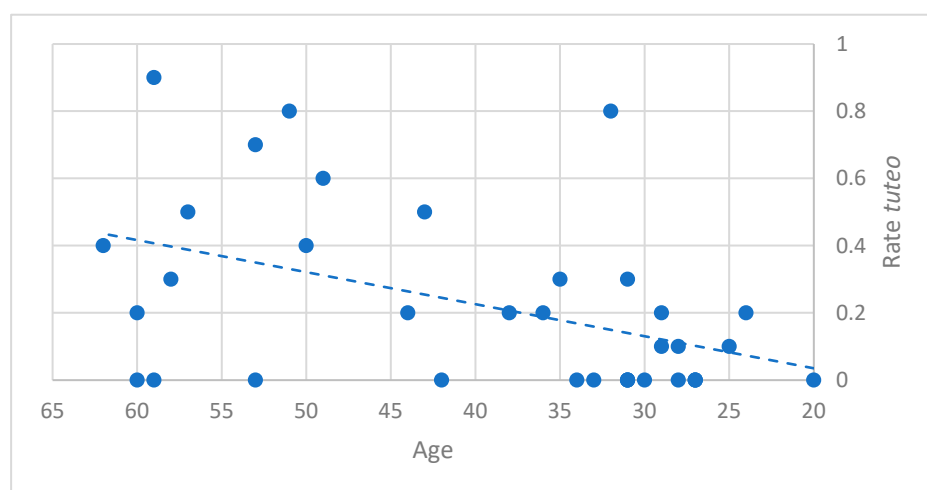


Figure 2. Rates of *tuteo* vs. *voseo* by age, demonstrating lowering rates of use over time.

We also see here a large amount of variability even for speakers of similar ages. This is partly due to a range of conditioning factors, including social (gender and socioeconomic class) and linguistic factors (previous realization, subject pronoun expression, discourse type, clause type, morphological class of verb). Examination of the impact of these factors by Callaghan (2020) revealed wholesale change in the nature of this variation, in that not only has the rate of *tuteo* dropped over time but the conditioning has also changed. Older and younger speakers differ in terms of the set of significant predictors (e.g., discourse type is significant for older speakers only, for whom *tuteo* is disfavored with generic subjects), and in the relative strength of the effect those factors have (e.g., clause type has a stronger effect in younger than older speakers, with a greater disfavoring of *tuteo* in questions). In some cases, they differ in the direction of effect (e.g., *tuteo* is favored by higher social classes for the older age group, but by lower social classes for the younger age group) (cf. Callaghan 2020, p. 195).

One predictor that remains stable across speakers of different ages is previous realization, as a manifestation of priming, which furthermore has the greatest magnitude of effect for both age groups (and overall, see Figure 6 below, Section 4.4). We will therefore now turn to consider this effect in more detail, paying particular attention to what it reveals about speaker awareness of *tuteo/voseo* variation.

4. Priming as a Diagnostic of Associations between *tuteo* and *voseo*

Considering priming to be a measure of speaker associations between forms, priming effects can shed light on whether speakers distinguish between the *tuteo* and *voseo* paradigms, or whether, as a result of mixing, stigmatization and “invisibilization”, they have become conflated in speakers’ variable grammars. We address this question in three ways. We first consider the effect of a variable *tuteo* or *voseo* form in the preceding discourse. We then examine the effect of the presence of a non-variable 2sg familiar form, namely verbs occurring in TAMs that are historically *tuteo* (in the data here, the Imperative and Preterit), and the fossilized *voseo* form *cachái*. For our third set of analyses, we consider the effect of the “mixed *voseo*” and whether the presence of a *tú* pronoun weakens the association between *voseo* forms.

To provide an overall picture of the nature of the priming effect, we describe the general trends through descriptive statistics and visualizations, before testing for statistical significance in those trends with generalized linear mixed effects models and a random forest analysis. For these priming analyses, we exclude tokens for which the previous form or the status of the pronoun (as *tú* or unexpressed) could not be coded due, for example, to unclear speech ($n = 43$); seven instances of a *vos* pronoun in the variable context were also excluded, leaving a total of 1009 tokens for analysis.

4.1. The Impact of the Form of the Previous Realization

We begin by examining the role that the form of previous realization plays in conditioning the *tuteo/voseo* variation. If speakers recognize and negotiate two separate paradigms, then this should be evident in the patterns resulting from priming: a previous *voseo* should favor the repetition of a subsequent *voseo* form, and a previous *tuteo* should favor a subsequent *tuteo*. An absence of priming would support the interpretation that the two paradigms are collapsed, and that speakers do not distinguish between forms.

The sample was coded for the form of the most recent 2sg familiar verb occurring within the previous five IUs, generally, around 20 words. As priming is known to weaken with distance (e.g., Szmrecsanyi 2005, p. 139; Travis 2007, pp. 119–21), this relatively close measure maximizes the possibility of observing an effect. Previous 2sg familiar forms by both the interlocutor and the same speaker were included. Of the 356 instances of a previous *tuteo* or *voseo*, 69 were produced by the interlocutor. These tokens maintain a similar priming effect to those produced by the same speaker, and thus we do not separate out same speaker- from interlocutor-produced primes for the purposes of the analyses.⁶

Five possible previous realizations were coded: a previous *tuteo*, as in (10) (where the prime is underlined, and the target in bold); previous *voseo* (11); previous invariable Imperative or Preterit forms (12); previous *cachái* (13); and no other 2sg familiar token within the previous five IUs.

- (10) *no andabas conmigo cuando **te perdías**.*

'You-Ø weren't-TUTEO with me when you-Ø got lost-TUTEO.'

(Memories; 1261; Trinidad)

- (11) *Sabís la caída que le **vai a dejar** a tu amigo huevona?*

'Do you-Ø know-VOSEO the problems

you-Ø are going-VOSEO to leave your friend idiot?'

(Police; 484; Matilda)

- (12) *Métete cuando --*

*cuando **quieras**.*

You-Ø join in- IMPERATIVE-FIXED when --

whenever you-Ø want-TUTEO.

(Barbeque; 1430–1431; Viviana)

- (13) *Cristián: yo pierdo tiempo con esas huevadas,*

cachái?

Claudia: ***perdís** tiempo.*

Cristián: 'I waste time on that shit,

you-Ø know-CACHÁI?'

Claudia: 'You-Ø waste-VOSEO time.'

(Takeaway; 221–223; Cristián/Claudia)

Table 3 shows the distribution of the data according to these five previous contexts, and the rate of *voseo* vs. *tuteo* in each, which is also depicted in Figure 3. We see in Table 3 that one-third of the tokens have a *tuteo* or *voseo* in the previous five IUs ($n = 356$), and just under one-half have no 2sg familiar token within the previous discourse, providing sufficient data for meaningful comparisons. A small proportion occurred in the context of a previous Imperative or Preterit, or of a previous *cachái*, which we will consider in the following section. For now, we concentrate on the context with a previous (variable)

⁶ The rates of *voseo* in the context of a previous *voseo* produced by the same speaker vs. the interlocutor are the same (92%, 203/220 vs. 56/61, respectively). We cannot compare the rates of *tuteo* in the context of a preceding *tuteo* by speaker vs. interlocutor, as there are only eight tokens occurring in the context of a previous *tuteo* produced by the interlocutor.

voseo or *tuteo*, as compared with no previous mention, the two left-most columns and the right-most column in Figure 3.

Table 3. Rates and relative frequency of *tuteo* vs. *voseo* by previous realization.

	<i>tuteo</i>		<i>voseo</i>		Total
	%	N	%	N	N
Previous <i>tuteo</i>	71%	48	29%	20	68
Previous <i>voseo</i>	8%	22	92%	266	288
No 2sg familiar token in preceding 5 IUs	23%	110	77%	375	485
Previous IMPERATIVE/PRETERIT fixed form	16%	19	84%	102	121
Previous <i>cachái</i>	13%	6	87%	41	47
Total	20%	205	80%	804	1009

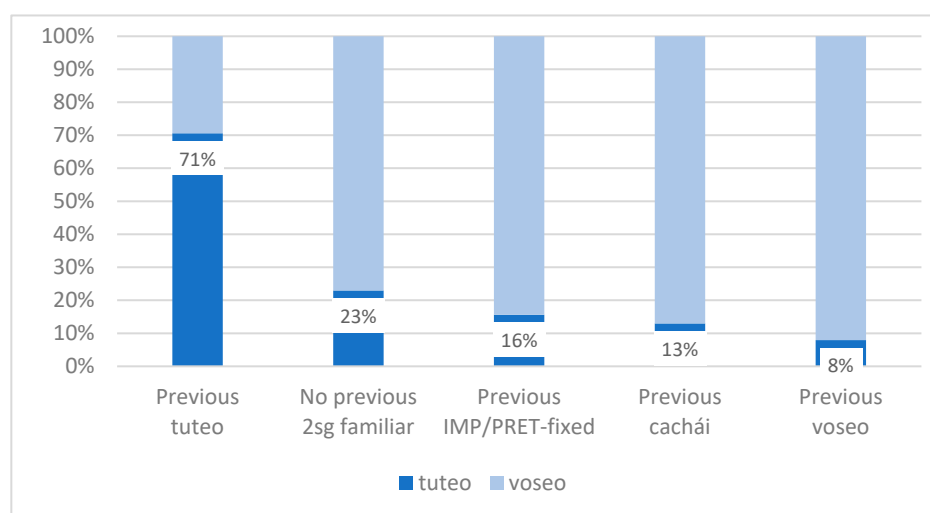


Figure 3. Rates of *tuteo* vs. *voseo* by previous realization.

As can be seen, the rate of *tuteo* is overwhelmingly highest with a *tuteo* in the previous discourse, at 71%, and lowest in the context of a previous *voseo*, at just 8%; in the absence of a prime (with no previous 2sg familiar form), the rate is in between, at 23%. This clearly shows a priming effect, that will be corroborated statistically in Section 4.4 below—*tuteo* is favored in the context of a previous *tuteo*, and is disfavored in its absence, particularly in the context of a previous *voseo*.

We are able to verify that the favoring of *tuteo* following a preceding *tuteo*, and *voseo* following a preceding *voseo*, is a real priming effect, and not simply the inevitable repetition of the same form by speakers with very high rates of one or the other form by comparing the impact of the previous realization across speakers with different baseline rates of *tuteo* (cf. Sankoff and Laberge 1978; Torres Cacoullos and Travis 2018, p. 90). Speakers were binned into four groups according to their overall rate of *tuteo*, separated by 20 percentage points. It was not possible to compare the speakers individually, as few produce enough tokens to reliably identify individual differences—the 36 speakers produce an average of 28 tokens each, with a median of 24, and a range from three to 89. Only five speakers produce over 50 tokens, seven between 30 and 50, and 14 under 20. Fifteen speakers (10 of whom produce under 30 tokens) do not use *tuteo* at all in the data, and they are excluded from these comparisons.⁷

⁷ The speakers who show no variability in the data are, however, included in the other analyses, as there is good reason to assume that they share the same conditioning, and that their lack of *tuteo* is due to a preponderance of *voseo*-favoring contexts in their speech analyzed here (Callaghan 2020, pp. 230–33; on including non-variable speakers in analyses, see Tagliamonte and Baayen 2012, pp. 165–66).

Figure 4 shows the rate of *tuteo* for these groups of speakers when there is no 2sg form in the previous five IUs (the solid, darkest line), in the context of a previous *tuteo* (the line marked with circles) and a previous *voseo* (marked with squares). Here we see that, for both low and high users of *tuteo*, their rate of *tuteo* is higher when there is a *tuteo* in the preceding environment and their rate of *voseo* is higher when there is a *voseo* in the preceding environment; that is, priming holds independently of the rate of *tuteo*. That priming is not limited to any individual speakers is confirmed in the mixed-effect models reported below (Section 4.4), where the priming effect emerges with speaker run as a random effect.

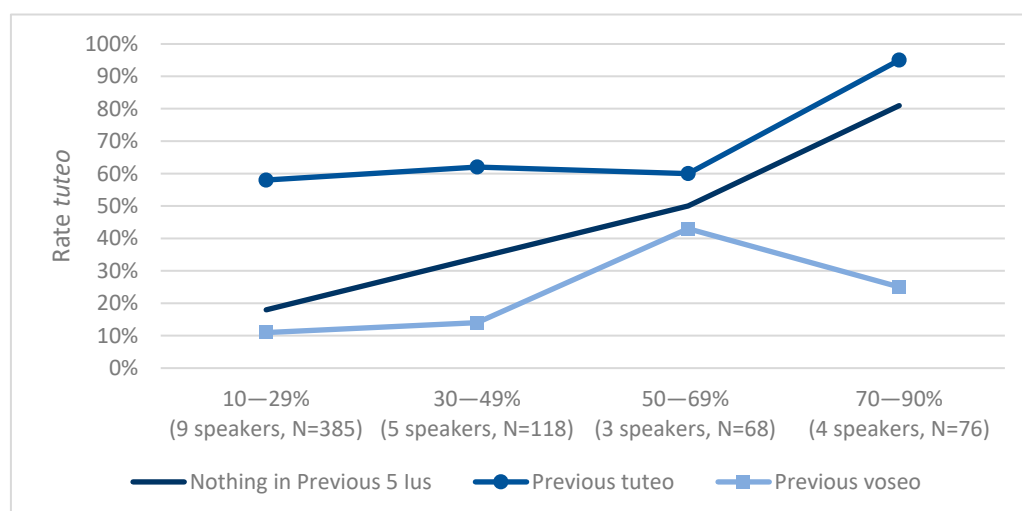


Figure 4. Rates of *tuteo* vs. *voseo* by previous realization comparing speakers binned according to their overall rate of *tuteo*.

A further observation from Figure 4 is that the strongest *tuteo* > *tuteo* priming is found among those with lower overall rates of *tuteo*: for speakers with rates of *tuteo* of under 50% (the first two bins), the rate of *tuteo* in the presence of a previous *tuteo* deviates more from that in the absence of a prime than it does for speakers with rates of over 50%. Conversely, the strongest *voseo* > *voseo* priming occurs amongst those speakers with high baseline rates of *tuteo* (the last bin). Such patterning is consistent with what has been reported in the priming literature, whereby less frequent variants tend to exert a stronger priming effect, under what has been described as “surprisal” (Jaeger and Snider 2007). Thus, while in general, speakers are sensitive to the form of a previous mention, given the predominance of *voseo*, those with lower rates of *tuteo* are even more so. This is corroborated by the findings across age group: younger speakers, who tend to have lower rates of *tuteo* (as seen in Figure 2), exhibit a stronger *tuteo* > *tuteo* priming effect than older speakers (Callaghan 2020, p. 202). From a usage-based perspective, we might interpret this in terms of an unexpected instance being more salient (Jaeger and Weatherholtz 2016), and therefore more readily retrieved and reused in the subsequent discourse.

The priming observed suggests that speakers recognize the [VERB-VOSEO] construction as distinct from the [VERB-TUTEO] construction. But to establish this, we must also demonstrate that priming holds across different morpho-phonological realizations of *voseo*, and is not limited to instances that share the same form. If priming were reliant on similarity of shape, then it should be manifestly stronger when the prime and target share the same form (/ay/ > /ay/ or /is/ > /is/), as in (14) below, than when they do not (/ay/ > /is/ or /is/ > /ay/) as in (11) above. However, although the rate of *voseo* (vs. *tuteo*) is marginally higher when the *voseo* prime and target take the same shape than when they do not (88%, 168/190 vs. 82%, 98/120), this is not significant ($p = 0.13$, Fisher’s Exact

Test).⁸ That different *voseo* morpho-phonological endings prime each other is evidence for *voseo* as an abstract category in speakers' mental representations. The priming here is thus functioning at a schematic level while being marginally strengthened when morphological specificity is shared (cf. Rosemeyer and Schwenter 2017, p. 30; Travis et al. 2017, p. 294).

- (14) *llegái al aeropuerto,*
se te suben dos personas,
y las vai a dejar a Santiago,
 'you-Ø arrive-VOSEO at the airport,
 Two people get in,
 and you-Ø go-VOSEO to drop them in Santiago,'
 (Back to Santiago; 76–78; Ernesto)

4.2. The Status of Non-Variable Constructions: Imperative, Preterit and *cachái*

Having established *tuteo* > *tuteo* and *voseo* > *voseo* priming, we can now use priming to test the status of the non-variable constructions, and whether or not they retain a link with their respective *tuteo*/*voseo* origins. The question we ask is: do speakers associate the historically *tuteo* TAMs (here, the Imperative and Preterit) with the [VERB-TUTEO] construction and the *voseo*-based *cachái* with the [VERB-VOSEO] construction? Evidence of this would be found in a rate of *tuteo* in the context of a previous Imperative or Preterit parallel or similar to that in the context of a previous *tuteo*, or at least a higher rate than when there is no prime. Similarly, for *cachái*, evidence would be a rate of *voseo* in the context of a previous *cachái* parallel or similar to that in the context of a previous *voseo* and higher than that in the absence of a prime.

As we see in Table 3 and Figure 3 above, the rate of *tuteo* in the context of a previous Imperative or Preterit is just 16%. This is less than one quarter of that in the context of a previous *tuteo* (71%), and thus these are clearly treated quite distinctly. It is only slightly lower than in the absence of a prime (23%), and we will see below that this difference is not significant (Table 5, Section 4.4). This indicates that the Imperative and Preterit have shed their historical associations with *tuteo*, and function as syncretic, neutral forms.

For *cachái*, the rate of *tuteo* following a previous *cachái* (13%) is in between that when there is no prime (23%), and when there is a previous *voseo* (8%). The statistical models we report on below (Section 4.4) indicate that, in the data overall, the rate of *tuteo* in the context of a previous *cachái* is not significantly different from that when there is no prime ($p = 0.44$), suggesting that *cachái* may be a neutral form. We also find that the impact of a previous *cachái* is significantly different from that of a previous *voseo* in the data overall ($p = 0.02$). However, this difference is not significant for the young speakers, who are the main users of *cachái* ($p = 0.09$), indicating that for them at least *cachái* retains a degree of association with the *voseo* paradigm.

The classification of these forms is not a minor issue for the analyst, as they account for approximately one-half of all 2sg familiar tokens, as we saw in Figure 1. The patterning we have observed for priming validates the exclusion of both syncretic forms and *cachái* from the *tuteo*/*voseo* variation, and illustrates how priming can help illuminate what counts as an instance of a construction, and so what falls in, and outside of, the variable context (cf. Tammenga 2016).

4.3. The Status of "Mixed *voseo*"

The usage data thus far demonstrate strong awareness of distinct *voseo* vs. *tuteo* paradigms, despite metalinguistic commentary to the contrary. A remaining question is

⁸ Out of a total of 266 *voseo* tokens with a *voseo* prime, stressed /ay/ (45%, $N = 121$) and /is/ (42%, $N = 113$) make up the majority, while unstressed /ay/ (employed in past tenses) accounts for just 12% ($N = 32$) (counting the monosyllabic present indicative forms *vai* 'you go', *dai* 'you give' and *hai* 'you have' as instances of stressed /ay/). Future analyses could compare the relationship between stressed and unstressed /ay/, and the impact of repetition of the same verb

the status of mixed *voseo*, that is, a *voseo* verb form with a *tú* pronoun, seen in (15) (and in (2) above), which we might expect to be less associated with *voseo* than instances that occur without a *tú* pronoun. If this is the case, then there should be a weaker priming effect when the target verb occurs with a *tú* pronoun than without.

(15) *Tú creís que va a entender eso el niño?*

‘Do you_{-TÚ} think_{-VOSEO} the kid is going to understand that?’

(Family; 361; Sara)

For a first view, we look to patterns of subject expression. While theoretically there are three options for 2sg subject realization (*tú*, *vos*, $\emptyset$), due to the rarity of the *vos* pronoun (just 10 tokens in the CCSS), only two are fully exploited by speakers, *tú* or an unexpressed subject. Subject expression can therefore be used as a proxy for the strength of association between the pronoun and verb: if speakers do not distinguish between the two paradigms, then we would expect the rate of occurrence with a *tú* pronoun to be similar across *voseo* and *tuteo*. Table 4 compares the rates of subject pronoun expression with each verb form. As can be seen, there is a higher rate of pronominal expression with *tuteo* than *voseo* (31% vs. 18%), indicating that, though speakers do use a *tú* pronoun with *voseo* verb forms, they are less likely to do so than with a *tuteo* verb form. The favoring effect of a *tú* pronoun on rate of *tuteo* also emerges as significant in the statistical model (see Table 5, Section 4.4). Furthermore, mixed *voseo* accounts for under 15% of the data (147/1009), suggesting that this is a relatively minor phenomenon, despite the reported predominance of this form in the literature (e.g., Torrejón 1986, p. 682).

Table 4. Rates of subject expression (*tú* vs. $\emptyset$) for *tuteo* vs. *voseo*.

	% <i>tú</i>	N	% $\emptyset$	N
<i>Tuteo</i>	31%	63	69%	142
<i>Voseo</i>	18%	147	82%	657
Total	21%	210	79%	799

It is important to note that multiple factors condition patterns of subject expression (cf., Torres Cacoullós and Travis 2018, chp. 5), such that care needs to be taken in comparing overall rate differences. A preponderance of tokens in contexts favorable to pronominal expression (e.g., non-coreferential contexts) may result in a higher rate of expression. Such analyses are beyond the scope of this paper, and we now turn to priming as a further measure of the status of mixed *voseo*.

Figure 5 shows the rate of *tuteo* across three priming contexts (previous *tuteo*, no 2sg familiar token in the previous five IUs, previous *voseo*), with a *tú* pronoun (darker column, on the left of each pair) and with no subject pronoun (lighter column on the right). First, note that, in the absence of a prime, we observe the predicted effect: a higher rate of *tuteo* with a *tú* pronoun than with no pronoun. But we also observe a nearly identical priming effect with and without a *tú* pronoun: with a *tú* pronoun, the rate of *tuteo* rises from 13% with a *voseo* previous mention to 33% with no previous mention, to 71% with a *tuteo* previous mention; with no subject pronoun, it rises from 7% to 19% to 70%. That is, whether or not the target occurs with a *tú* pronoun, speakers are sensitive to the overall tendency to repeat the previous form.

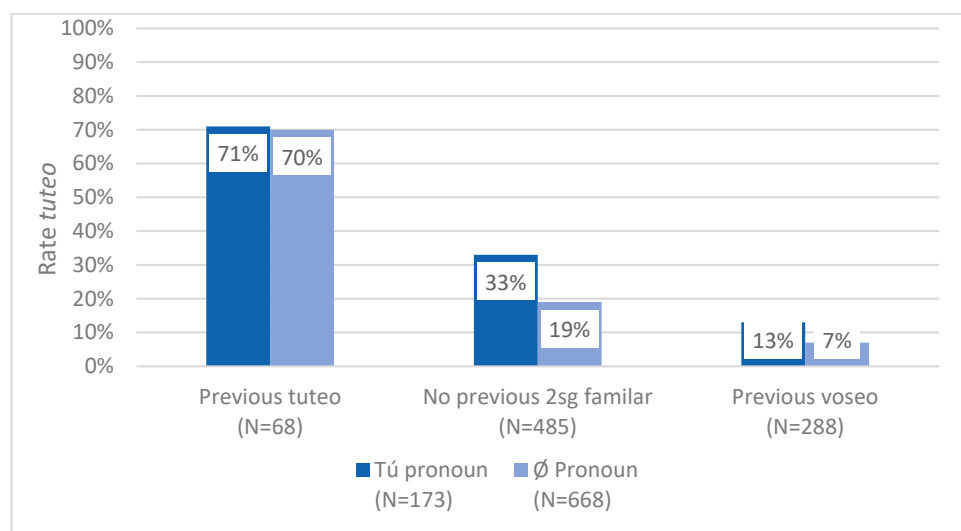


Figure 5. Rates of *tuteo* vs. *voseo* by previous realization and subject expression (*tú* vs. $\emptyset$).

It is particularly telling that, in the context of a previous *tuteo*, the presence of a pronoun has no impact on the rate of *tuteo*; that is, the previous *tuteo* overrides the pronoun effect. Although the presence of a *tú* pronoun does impact the rate of *tuteo* in the context of a previous *voseo*, the effect of a previous *voseo* is still felt in this context, where the rate of *tuteo* is barely one-third that in the absence of a prime (13% vs. 33%). The favoring of *tuteo* by a *tú* pronoun is confirmed in the statistical models presented below (Table 5, Section 4.4), which also reveal that there is no significant interaction between the presence of a pronoun and the previous realization, indicating that each holds independently of the other.

The avoidance of the *vos* pronoun, the use of a *tú* pronoun with a *voseo* verb form, and reports that Chileans claim not to use *voseo* because they associate it exclusively with the pronoun may seem to support an interpretation of pronouns as more salient in speakers' minds than verb forms. This is precisely what has been proposed in the literature, in accordance with the notion that the greater saliency of the *vos* pronoun as a lexical item has contributed to it becoming stigmatized, and therefore avoided, while *voseo* verb forms fly under the radar, allowing for their expansion in use (Bertolotti 2015, p. 19; Huerta Imposti 2011–2012, p. 52; Hummel 2010, p. 112; Stevenson 2007, p. 167). This is not, however, what we observe in usage patterns, where priming of verb forms is upheld even in the presence of a *tú* pronoun. Though mixed *voseo* has been used as evidence of the loss of awareness of the *voseo* paradigm (Carricaburo 1997, p. 34; Huerta Imposti 2011–2012, p. 54), what the priming data show is that even mixed *voseo* is treated by speakers as an instance of the [VERB-VOSEO] construction.

These results also allow us to respond to the suggestion that there exists an underlying *tú* pronoun for all instances of 2sg familiar verbs with an unexpressed subject (e.g., Rivadeneira Valenzuela 2016, p. 93). The priming patterns we observe do not support this interpretation, but rather indicate that speakers create their utterances with *tuteo* and *voseo*, and with or without pronouns, based on probabilistic usage-based factors, influenced by, among other things, what precedes in the discourse.

4.4. Statistical Analyses

Having observed the general patterns, we now test the statistical significance of these differences, using generalized linear mixed effects models with the `glmer()` function in R (Bates et al. 2019; R Development Core Team 2019). Models were fit with verb form (*tuteo/voseo*) as the dependent variable, and previous realization, presence of a *tú* pronoun, and age as independent variables. For previous realization, the absence of a prime was used as the reference level; an unexpressed subject was the reference level for pronoun; and age was modeled continuously. Three- and two-way interactions between previous

realization, pronoun and age were tested, and then pruned from the model, as none were significant. To take into account differences between speakers and verb types, speaker and verb were included as random intercepts. Table 5 presents the final model summary.

Table 5. Output of a generalized linear mixed effects model predicting *tuteo*.

	Estimate	Std. Error	Z	p
(Intercept)	5.86752	1.16632	5.031	<0.001
Previous <i>tuteo</i>	−1.53467	0.33822	−4.538	<0.001
Previous <i>voseo</i>	0.87474	0.28351	3.085	<0.01
Previous syncretic	0.37496	0.32207	1.164	=0.24
Previous <i>cachái</i>	−0.40438	0.52527	−0.770	=0.44
Pronoun <i>tú</i>	−0.90796	0.23260	−3.903	<0.001
Age	−0.09075	0.02627	−3.454	<0.001

Overall *tuteo* 20% (205/1009); negative coefficients are associated with a higher rate of *tuteo*. For 36 speakers, variance = 2.7 (SD = 1.6) and for 134 verb types, variance = 0.1 (SD = 0.32).

First, this model corroborates the overall priming effect we saw in Table 3 and Figure 3: compared with when there is no prime in the previous discourse, the rate of *tuteo* is significantly lower in the context of a previous *voseo* and significantly higher in the context of a previous *tuteo*. Thus, *tuteo* primes a subsequent *tuteo* and *voseo* primes a subsequent *voseo*, confirming that these speakers do recognize a relationship between distinct verbs produced in the *voseo* vs. *tuteo* forms. Despite a lack of explicit metalinguistic awareness and mixing of *voseo* verb forms with a *tú* pronoun, in actual usage, speakers keep the two paradigms separate.

What of previous syncretic forms and *cachái*? Table 5 indicates that neither is significantly different from contexts where there is no prime, in accordance with our observations above. Releveled models run with a previous *tuteo* as the reference level indicate that we are significantly less likely to get *tuteo* following syncretic forms than following *tuteo* ($\beta = 1.91, p < 0.001$), supporting the notion that these TAMs have lost their association with *tuteo*. *Cachái* is slightly different. Although overall, we are more likely to get a subsequent *voseo* following another *voseo* than following *cachái* ($\beta = 1.28, p < 0.05$), the effect is not significant for the younger speakers, the main *cachái* users ($\beta = 1.19, p < 0.09$), suggesting that the association between *cachái* and *voseo* has not been entirely lost. That is, *cachái* may not be a central exemplar of the [VERB-VOSEO] construction, but it does retain a link to it.

The model in Table 5 also indicates that the priming effect is not an artefact of speakers with very high rates of *voseo*, as the impact of a previous *voseo* or *tuteo* holds even when we take account of individual speakers' preference by including a random intercept for speaker in the model. To further test this, we ran another model identical to the above, but with the addition of an interaction between rate of *tuteo* and previous realization, and no interaction was found between a previous *voseo* and rate of *tuteo* ($\beta = 1.40, p = 0.24$), nor a previous *tuteo* and rate of *tuteo* ($\beta = 2.23, p = 0.10$).

Finally, the model in Table 5 also tests the impact of a *tú* pronoun, which favors *tuteo*, as we saw in Table 4 and Figure 5. The model shows that the impact of a previous *tuteo* is stronger than that of a *tú* pronoun (with z scores of −4.54 and −3.90 respectively). The same is not so for a previous *voseo*, which has a slightly lower z score than that of the pronoun (3.10). Nevertheless, that *voseo* priming holds in the presence of a *tú* pronoun is evident in the fact that an interaction between the presence of a pronoun and previous *voseo* fails to reach significance ($\beta = 0.11, p = 0.87$)—priming and pronoun presence are therefore indeed independent effects.

4.5. The Effect of Priming in Conjunction with Other Conditioning Factors

We mentioned above (Section 3) that multiple factors condition the choice of *tuteo* over *voseo*. Up to now, we have focused on a subset of these that are directly relevant to priming. Prior work conducting regression analyses that include a full set of linguistic

and social predictors has found priming not only to operate alongside these other factors, but to be the strongest predictor of this variation (Callaghan 2020, p. 197). In order to confirm this here, we conduct one final set of analyses, including the identical dataset to that employed above ($n = 1009$), and adding three new predictors: discourse type (reported speech, generic subjects, specific subjects); clause type (questions, main clause declaratives, subordinate clauses), and gender.

To determine the impact of priming while taking into account a full range of predictors, we conduct a conditional random forest analysis. Random forests are built from multiple conditional inference trees, a statistical approach that makes recursive binary splits in the data, according to the strength of the predictors in each subsequent subset of the data. Random forests measure the overall importance of each predictor included in the model by averaging the results across multiple conditional inference trees (here, 1000), each based on a randomly generated subset of the data (Tagliamonte and Baayen 2012, pp. 159–60). Such a model “ensures that the evaluation of a variable’s importance takes into consideration its behavior in relation to other variables in its ranking” (Schnell and Barth 2018, p. 64). The result of this analysis, obtained using the `cree()` function from the “party” package for R (Hothorn et al. 2006), is presented in Figure 6.

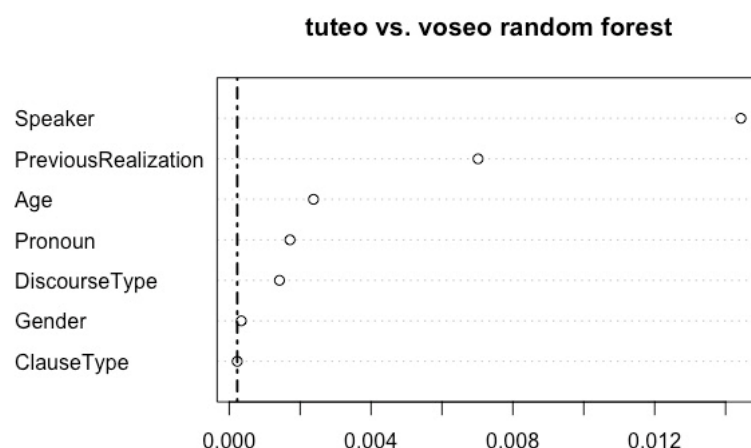


Figure 6. Variable importance ranking from random forest for *tuteo* vs. *voseo*.

As can be seen, the strongest effect is that of individual speaker, as is common in such analyses (Tagliamonte and Baayen 2012, p. 162).⁹ Priming, however, is the next strongest predictor, and is substantially stronger than any other. Age, pronoun and discourse type are ranked next, followed by gender with a marginal effect, and clause type is found not to be worth further consideration. Thus, priming is not only upheld when considering a full set of predictors but it exerts the strongest effect.

5. Conclusions

We have sought here to gain access to information about speaker representations through the study of actual usage. While perception tasks and surveys are often employed to extract attitudinal information, speakers’ own intuitions and judgements do not necessarily produce reliable data for stigmatized variables (Sankoff 1988, pp. 145–46), as has been shown to be the case for *voseo* use (Bishop and Michnowicz 2010). Here we have demonstrated how spontaneous speech data can offer insights into speakers’ implicit understanding of constructions, examined in terms of degrees of association: because priming occurs across related constructions, the relative strength of priming provides an indication of the strength of the relationship.

⁹ The strong effect for speaker is not due to those speakers who use only *voseo* in the data, as a similarly strong effect was found in a random forest generated with data that excluded those speakers.

The fact that speakers' choice between *voseo* and *tuteo* is influenced by the previous 2sg familiar form used in the discourse provides strong evidence that they do distinguish between these paradigms; that is, they recognize, and keep separate, [VERB-TUTEO] and [VERB-VOSEO] constructions. For the non-variable forms, however, we found different results. The Imperative and Preterit, despite deriving historically from *tuteo*, appear to no longer be associated with it; when occurring in the previous discourse, their impact is no different from that of the absence of a prime, and it is significantly different from that of a previous *tuteo*. *Cachái*, on the other hand, appears not to have wholly lost its association with the [VERB-VOSEO] construction, as, for young speakers at least (the main *cachái* users), the impact of a previous *cachái* in the discourse is not significantly different from that of a previous *voseo*. This illustrates that mental representations can change over time, and that non-variable, or fossilized, forms may not be homogeneous in their degree of autonomy from their source constructions (cf. Bybee 2006, p. 715).

As we have seen, *voseo* and *tuteo* take distinct morpho-phonological forms (e.g., *estás* vs. *estái* 'you are-TUTEO/VOSEO', *tienes* vs. *tenís* 'you have-TUTEO/VOSEO'), and it is highly likely that this aids in keeping these paradigms apart. But we have also seen that *voseo* > *voseo* priming is not dependent on repetition of the same morpho-phonological shape, demonstrating the existence of a schematic [VERB-VOSEO] construction. Furthermore, the priming of *voseo* verb forms is retained even with a *tú* pronoun in the so-called mixed *voseo*. Thus, though overt expression of a *tú* pronoun favors *tuteo* over *voseo*, it does not override the impact of the previous form, indicating that [*tú* + VERB-VOSEO] is still an exemplar of the more general [VERB-VOSEO] construction.

We can capture the mental representations evidenced here with the set of associations depicted in Figure 7, with [Ø + VERB-VOSEO] (the majority variant) at its center: the *tú* pronoun is associated with both *tuteo* and *voseo*, but more strongly with the former; the *vos* pronoun, though minimally used, is associated with *voseo* verb forms; the syncretic Imperative and Preterit forms are associated with neither construction; and *cachái* remains (weakly) associated with *voseo*.

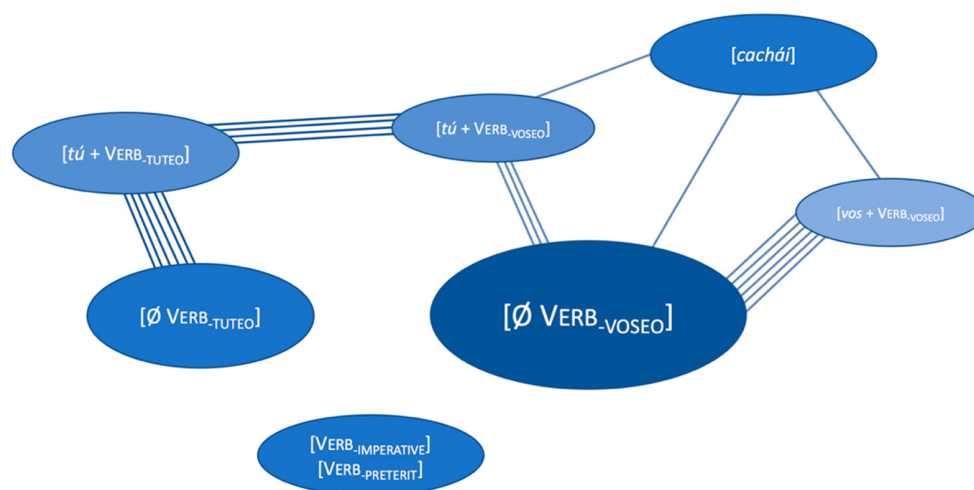


Figure 7. Mental representation of *tuteo/voseo* constructions.

This picture does not seem to be consistent with Chileans' metalinguistic awareness of these forms, which revolves around the *vos* pronoun, the subject of much commentary and stigmatization, while the *tuteo/voseo* verb forms pass unnoticed. Nor is it wholly consistent with much of the literature on this topic, that has assumed the existence of an underlying form for each syncretic form, and for instances with an unexpressed pronoun. The results presented here are, however, consistent with a usage-based understanding of grammar, according to which speakers do not construct their speech around abstract underlying forms, but rather as part of their constantly unfolding experience with language. Further,

this can function at a highly local level, where one factor constraining variant choice is the form that has been used previously, as evidenced in structural priming.

Author Contributions: This paper derives from a Ph.D. thesis by M.C., completed in 2020 at the Australian National University under the supervision of C.E.T. Both authors contributed equally to the writing of this paper. All authors have read and agreed to the published version of the manuscript.

Funding: This research was supported by an Australian Government Research Training Program (RTP) Ph.D. scholarship awarded to Matthew Callaghan, and by the ARC Centre of Excellence for the Dynamics of Language (CE140100041).

Institutional Review Board Statement: The study was conducted according to the guidelines of the Declaration of Helsinki, and approved by the Human Research Ethics Committee of the Australian National University (Protocol #2014-413, approved on 29th October 2014).

Informed Consent Statement: Informed consent was obtained from all subjects involved in the study.

Data Availability Statement: The data presented in this study may be made available on request from the corresponding authors. The data are not publicly available to accord with the informed consent guidelines provided to the participants.

Acknowledgments: We are grateful to the anonymous reviewers and to the editors of this volume, Esther Brown and Javier Rivas, for valuable feedback on this paper, and to Danielle Barth for advice on the statistical analysis.

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

Appendix F. Transcription Conventions

Table A6. Transcription Conventions (Du Bois et al. 1993).

New line	new Intonation Unit	...	medium to long pause (>0.3 s)
.	final intonation contour	-	truncated word
,	continuing intonation contour	=	lengthened syllable
?	appeal intonation contour	[]	overlapped speech
--	truncated intonation contour	(())	transcriber's comment

References

- Bates, Douglas, Martin Mächler, Ben Bolker, Steve Walker, Rune Haubo Bojesen Christensen, Henrik Singmann, Bin Dai, Fabian Scheipl, Gabor Grothendieck, Peter Green, and John Fox. 2019. lme4: Linear Mixed-Effects Models Using 'Eigen' and S4. R package. Version 1.1-21. Available online: <https://cran.r-project.org/web/packages/lme4/index.html> (accessed on 21 December 2020).
- Bello, Andrés. 1981. *Gramática de la Lengua Castellana: Destinada al Uso de los Americanos*. Caracas: La Casa de Bello. First published 1891.
- Bertolotti, Virginia. 2015. A Mí de Vos no Me Trata ni Usted ni Nadie. *Sistemas e Historia de las Formas de Tratamiento en la Lengua Española en América*. Ciudad de México: Universidad Nacional Autónoma de México.
- Bishop, Kelley, and Jim Michnowicz. 2010. Forms of Address in Chilean Spanish. *Hispania* 93: 413–29.
- Bock, Kathryn, and Zenzi M. Griffin. 2000. The persistence of structural priming: Transient activation or implicit learning? *Journal of Experimental Psychology: General* 129: 177–92. [CrossRef]
- Brown, Roger, and Albert Gilman. 1960. The pronouns of power and solidarity. In *Style in Language*. Edited by Thomas A. Sebeok. Cambridge: MIT Press, pp. 253–76.
- Bybee, Joan L. 2006. From usage to grammar: The mind's response to repetition. *Language* 82: 711–33. [CrossRef]
- Callaghan, Matthew. 2020. I'm Talking tú vos: A Comparative Study of Morphosyntactic Variation and Change in the Chilean Second-Person Singular. Ph.D. thesis, The Australian National University, Canberra, Australia. [CrossRef]
- Carricaburo, Norma. 1997. *Las Fórmulas de Tratamiento en el Español Actual*. Madrid: Arco Libros.
- Carvalho, Ana María. 2010. ¿Eres de la frontera o sos de la capital? Variation and alternation of second-person verbal forms in Uruguayan border Spanish. *Southwest Journal of Linguistics* 29: 1–23.
- Cautín-Epifani, Violeta. 2015. Formas de Tratamiento en Interacciones Verbales Escritas en la 'Biografía Facebook' de Hablantes de la Provincia de Iquique. Ph.D. thesis, Pontificia Universidad Católica de Valparaíso, Valparaíso, Chile.
- De Smet, Hendrik. 2016. How gradual change progresses: The interaction between convention and innovation. *Language Variation and Change* 28: 83–102. [CrossRef]
- Dresing, Thorsten, and Thorsten Pehl. 2015. F4transkript. Marburg, Germany. dr. Dresing and Pehl GmbH. Available online: <https://www.audiotranskription.de/english/> (accessed on 20 January 2015).

- Du Bois, John W., Stephan Schuetze-Coburn, Susanna Cumming, and Danae Paolino. 1993. Outline of discourse transcription. In *Talking Data: Transcription and Coding in Discourse Research*. Edited by Jane A. Edwards and Martin D. Lampert. Hilldale: Lawrence Erlbaum, pp. 45–89.
- Fernández-Mallat, Víctor. 2011. El ‘voseo mixto verbal’ de hablantes chilenos en Montreal: Estudio de caso en un contexto de contacto dialectal. *Boletín de Filología* 46: 35–58.
- Fernández-Mallat, Víctor. 2018. Alternancia y variación de formas verbales tuteantes y voseantes en el español de santiaguinos: Estudio de caso basado en un corpus conversacional. *Boletín de Filología* 53: 63–82.
- Gille, Johan. 2015. On the development of the Chilean Spanish discourse marker *cachái*. *Revue Romane. Langue et littérature. International Journal of Romance Languages and Literatures* 50: 3–29. [CrossRef]
- Gries, Stefan T. 2005. Syntactic priming: A corpus-based approach. *Journal of Psycholinguistic Research* 34: 365–99. [CrossRef]
- Helincks, Kris. 2012. La variación social y estilística del voseo chileno en diferentes géneros televisivos. *Revista Internacional de Lingüística Iberoamericana* 19: 185–211.
- Hothorn, Torsten, Kirk Hornik, and Achim Zeileis. 2006. Unbiased recursive partitioning: A conditional inference framework. *Journal of Computational and Graphical Statistics* 15: 651–74. [CrossRef]
- Huerta Imposti, M. Cristina. 2011–2012. Percepción del fenómeno del voseo verbal o morfológico en Chile. *Nueva Revista del Pacífico* 56–57: 37–56.
- Hummel, Martin. 2010. Reflexiones metodológicas y teóricas sobre el estudio de las formas de tratamiento en el mundo hispanohablante, a partir de una investigación en Santiago de Chile. In *Formas y Fórmulas de Tratamiento en el Mundo Hispánico*. Edited by Martin Hummel, Bettina Kluge and María Eugenia Vázquez Laslop. México: El Colegio de México, pp. 101–62.
- Jaeger, T. Florian, and Neal Snider. 2007. Implicit learning and syntactic persistence: Surprisal and cumulativity. *University of Rochester Working Papers in the Language Sciences* 3: 26–44.
- Jaeger, T. Florian, and Kodi Weatherholtz. 2016. What the heck is salience? How predictive language processing contributes to sociolinguistic perception. *Frontiers in Psychology* 7: 1115. [CrossRef] [PubMed]
- Labov, William. 1994. *Principles of Linguistic Change: Internal Factors*. Oxford: Basil Blackwell.
- Lenz, Rodolfo. 1940. Sobre la morfología del español de América. In *El Español en Chile: Trabajos de Rodolfo Lenz, Andrés Bello y Rodolfo Oroz*. Edited by Amado Alonso and Raimundo Lida. Buenos Aires: Facultad de Filosofía y Letras de la Universidad de Buenos Aires, pp. 259–68. First published 1891.
- Lipski, John M. 1994. *Latin American Spanish*. London and New York: Longman.
- Loebell, Helga, and Kathryn J. Bock. 2003. Structural priming across languages. *Linguistics* 41: 791–824. [CrossRef]
- Mondaca Becerra, Lissette Andrea, Andrea Patricia Méndez Carrasco, and Marcela Rivadeneira Valenzuela. 2015. “No es muletilla, es marcador, ¿cachái?”: Análisis de la función pragmática del marcador discursivo conversacional *cachái* en el español de Chile. *Literatura y Lingüística* 32: 233–58. [CrossRef]
- Pickering, Martin J., and Holly P. Branigan. 1998. The representation of verbs: Evidence from syntactic priming in language production. *Journal of Memory and Language* 39: 633–51. [CrossRef]
- Poplack, Shana. 1980. The notion of the plural in Puerto Rican Spanish: Competing constraints on (s) deletion. In *Locating Language in Time and Space*. Edited by William Labov. New York: Academic Press, pp. 55–67.
- R Development Core Team. 2019. *R: A Language and Environment for Statistical Computing*. Vienna: R Foundation for Statistical Computing, Available online: <http://www.R-project.org> (accessed on 21 December 2020).
- Rivadeneira Valenzuela, Marcela. 2009. El Voseo en Medios de Comunicación de Chile. Descripción y Análisis de la Variación Dialectal y Funcional. Ph.D. thesis, Universitat Pompeu Fabra, Barcelona, Spain.
- Rivadeneira Valenzuela, Marcela. 2016. Sociolinguistic variation and change in Chilean voseo. In *Forms of Address in the Spanish of the Americas*. Edited by María Irene Moyna and Susana Rivera-Mills. Philadelphia: John Benjamins, pp. 87–117.
- Rosemeyer, Malte, and Scott A. Schwenter. 2017. Entrenchment and persistence in language change: The Spanish past subjunctive. *Corpus Linguistics and Linguistic Theory* 15: 167–204.
- San Martín Núñez, Abelardo. 2011. Los marcadores interrogativos de control de contacto en el corpus PRESEEA de Santiago de Chile. *Boletín de Filología* 46: 135–66.
- Sankoff, David. 1988. Sociolinguistics and syntactic variation. In *Linguistics: The Cambridge Survey (Volume 4, Language: The Socio-Cultural Context)*. Edited by Frederick J. Newmeyer. Cambridge: Cambridge University Press, pp. 140–61.
- Sankoff, David, and Suzanne Laberge. 1978. Statistical dependence among successive occurrences of a variable in discourse. In *Linguistic Variation: Models and Methods*. Edited by David Sankoff. New York: Academic Press, pp. 119–26.
- Sankoff, Gillian. 2006. Age: Apparent time and real time. In *Encyclopedia of Language and Linguistics*, 2nd ed. Edited by Keith Brown. Oxford: Elsevier, pp. 110–16.
- Schnell, Stefan, and Danielle Barth. 2018. Discourse motivations for pronominal and zero objects across registers in Vera’a. *Language Variation and Change* 30: 51–81. [CrossRef]
- Stevenson, Jeffrey Lee. 2007. The Sociolinguistic Variables of Chilean Voseo. Ph.D. dissertation, University of Washington, Seattle, WA, USA.
- Szmrecsanyi, Benedikt. 2005. Language users as creatures of habit: A corpus-based analysis of persistence in spoken English. *Corpus Linguistics and Linguistic Theory* 1: 113–50. [CrossRef]

- Tagliamonte, Sali A., and Harald R. Baayen. 2012. Models, forests, and trees of York English: *Was/were* variation as a case study for statistical practice. *Language Variation and Change* 24: 135–78. [\[CrossRef\]](#)
- Tamminga, Meredith. 2016. Persistence in phonological and morphological variation. *Language Variation and Change* 28: 335–56. [\[CrossRef\]](#)
- Torrejón, Alfredo. 1986. Acerca del voseo culto de Chile. *Hispania* 69: 677–83.
- Torrejón, Alfredo. 2010. Nuevas observaciones sobre el voseo en el español de Chile. In *Formas y Fórmulas de Tratamiento en el Mundo Hispánico*. Edited by Martin Hummel, Bettina Kluge and María Eugenia Vázquez Laslop. México: El Colegio de México, pp. 755–70.
- Torres Cacoullos, Rena. 2015. Gradual loss of analyzability: Diachronic priming effects. In *Variation in Language: System- and Usage-Based Approaches*. Edited by Aria Adli, Marco García García and Göz Kaufmann. Berlin: De Gruyter, pp. 265–88.
- Torres Cacoullos, Rena, and Catherine E. Travis. 2018. *Bilingualism in the Community: Code-Switching and Grammars in Contact*. Cambridge: Cambridge University Press.
- Travis, Catherine E. 2007. Genre effects on subject expression in Spanish: Priming in narrative and conversation. *Language Variation and Change* 19: 101–35. [\[CrossRef\]](#)
- Travis, Catherine E., Rena Torres Cacoullos, and Evan Kidd. 2017. Cross-language priming: A view from bilingual speech. *Bilingualism: Language and Cognition* 20: 283–98. [\[CrossRef\]](#)
- Urzúa-Carmona, Paula. 2006. El verbo “cachar” en el español coloquial de Chile. *Onomázein* 1: 97–107.
- Weiner, E. Judith, and William Labov. 1983. Constraints on the agentless passive. *Journal of Linguistics* 19: 29–58. [\[CrossRef\]](#)