



Research Article

Víctor Fernández-Mallat* and David Barrero

Changes and continuities in second person address pronoun usage in Bogotá Spanish

<https://doi.org/10.1515/opli-2022-0241>

received October 13, 2022; accepted June 16, 2023

Abstract: In this article, we provide further evidence that Bogotá Spanish is transitioning from being an extensively *usted*-using variety into one in which *tú* is preferred in informal interaction by analyzing survey data through a quantitative approach, and metalinguistic commentary through a qualitative approach. Our data show that *tú* is mainly thought of as a productive way to convey proximity. At the same time, our data show that, despite this change in second person preference, *usted* and *sumercé* persist in familiar address, albeit at rates considerably lower than *tú*. *Usted* is particularly frequent among males in same-gender dyads because it allows them to avoid the possible connotations of effeminacy that *tú* may have in that specific context. *Sumercé* is frequently selected in addressing older relatives and individuals from the countryside because it is seen as being capable of conveying respect and affection simultaneously. Moreover, *sumercé* is seen as a sign of local identity capable of distinguishing Bogotá Spanish from other national varieties with *vos*, which is marginal in our data. Our findings are best seen through the proposal that address forms may gain specific meanings within their particular context of use, despite having more conventional meanings attached to them.

Keywords: variationist sociolinguistics, sociopragmatics, forms of address, Colombian Spanish

1 Introduction

Colombia is a country of great linguistic diversity. In part, this is because the country's biggest metropolitan areas – Bogotá, Medellín, Cali, and Barranquilla – are in different dialect zones (Montes Giraldo 1982). As a result, each of these metropolitan areas and their surroundings present varying linguistic features that distinguish them from one another. One such feature is the different forms available to speakers to refer to the second person singular (2PS), which is perhaps the field that has received the most attention from scholars to date, and with good reason. Unlike other Spanish-speaking regions where *tú* and *usted* (and their corresponding verbal inflections) are used as ideal exponents of the T/V dichotomy offered by Brown and Gilman (1960) – in which T forms are considered informal and friendly, and V forms are regarded as formal and distant – most of Colombia's varieties include one, if not two, additional variants that in a way compete with *tú* and *usted*. These variants are *vos* and *sumercé*. The former, from Latin *vōs*, long fallen out of favor in Peninsular Spanish but maintained in much of Latin America, has been found to compete with *tú* in the department of Antioquia (Millán 2014, Jang 2015), most of the Cauca region (Newall 2012, Michnowicz and Quintana Sarria 2020), and the department of Nariño (Dos Santos Caliri 2019), to name a few localities. The latter, from the Old Spanish *Vuestra Merced* – a polite address form meaning 'Your Mercy' – has been found to compete with *usted* in most localities on the eastern side of the Central Andes (Bayona 2006, Avendaño de

* **Corresponding author: Víctor Fernández-Mallat**, Department of Spanish and Portuguese, Georgetown University, Washington, United States, e-mail: vf109@georgetown.edu

David Barrero: Departamento de Lingüística y Literatura, Universidad de Antioquia, Medellín, Colombia, e-mail: davidsanbm@gmail.com

Barón 2014, Rojas Cárdenas 2018). Moreover, it has been documented that in some regions *usted* is not only used to signal formality and deference, but it may also be used to signify informality and closeness with interlocutors such as relatives and friends (Rimgaila and Cristina 1966, Uber 1985, Avendaño de Barón 2014).

In the present study, we add to the extant literature on address forms in Colombia by focusing on the Spanish spoken in Bogotá, the largest urban agglomeration of the country and its capital. Bogotá Spanish is often characterized as a heavily *ustedeante* dialect where *tú* and *sumercé* are used to a lesser extent. However, it has recently been suggested that the city's address system is undergoing a series of changes consisting in *tú* gaining ground over *usted* and the gradual demise of *sumercé* (Bayona 2006, Cepeda Ruiz 2014, Lingård 2016; inter alia). Following a variationist approach, in this study we seek not only to assess the current status of these changes, but also to elucidate the interplay of interactional and social factors in determining speakers' preferred forms of address. We do so by means of survey data from 371 speakers of Bogotá Spanish that responded to an online questionnaire about the address forms they would use with varying interlocutors. We supplement our findings with some of these speakers' comments about the value that they associate with each of the address forms of Bogotá Spanish.

In what follows, we start with a general overview of studies on address forms in Colombia. We then survey a selection of studies on Bogotá Spanish to situate our own study within this dialect zone. In the following sections, we introduce the current study, outline our analytical framework, and present the analysis of the data. We conclude with a discussion and implications.

2 Address forms in Colombia

It is well known that 2PS address forms vary widely across the Spanish-speaking world. While some areas show a binary system composed of *tuteo* and *ustedeo* variants (e.g., Spain, most of Mexico, and the Caribbean), other regions show a binary *voseo* vs *ustedeo* system (e.g., Argentina and Nicaragua). A third set of regions, such as Chile, Costa Rica, and Uruguay, have a tripartite system, with *tuteo*, *voseo*, and *ustedeo* all being available to speakers to refer to their interlocutors across a variety of situations (Benavides 2003, Lipski 2004, Carricaburo 2015).

Within all this variation, Colombia represents a particularly interesting case. This is because the country encompasses two of the three systems outlined earlier. The Caribbean region, where Barranquilla is located, has a binary system with *tuteo* and *ustedeo* (Montes Giraldo 2000), whereas the Western region, where both Medellín and Cali can be found, has a tripartite system with *tuteo*, *voseo*, and *ustedeo* (Jang 2015, Michnowicz and Quintana Sarria 2020). Moreover, the Interior Highlands region, where Bogotá is located, currently has a unique tripartite system in which *tuteo* and *ustedeo* are used alongside *sumercedeo* (Bayona 2006), which, as noted earlier, consists in using the pronoun *sumercé* with the verbal forms of *usted*, with which it shares its origins.¹

As might be expected given this level of variation, 2PS address forms in Colombian Spanish have received increased attention from scholars who have set out to study, not only their distribution according to a series of factors (e.g., age, gender, socioeconomic status [SES], setting, etc.), but also their pragmatic import and possible associations with varying social identities. In the Western region, for instance, it has been documented that, to express informality and familiarity, *voseo* is used more frequently than *tuteo*, especially among younger men of the lower and middle classes (Newall 2012, Jang 2015). One explanation for this resides in the fact that in this region the latter is associated with Bogotá, which has caused the former to become a sign of local identity and pride (Weyers 2016, Michnowicz and Quintana Sarria 2020). And because Bogotá is associated with sophistication and refinement, *tuteo* has taken up a nuance of femininity, especially when used by men (Newall 2012, Jang 2015).

¹ For a comprehensive overview of the origins, development, and present-day extension of *sumercedeo*, see Calderón Campos and García-Godoy (2020).

As to their pragmatic import, though *tuteo* and *voseo* seem to be unambiguously used to convey informality and familiarity throughout the Western region, in some localities (e.g., Cali, Medellín, and Pasto) *ustedeo* may serve both the purpose of expressing formality and distance along with informality and familiarity; that is, it is used as both a T form and a V form (Millán 2014, Newall 2016, Dos Santos Caliri 2019). It is worth mentioning, however, that the polysemy of *ustedeo* in the Western region is not as widespread as it is in the Interior Highlands region, a point to which we will return in Section 3.

As regards the Caribbean region, it is worth noting that – due to the fact that this zone of the country shows a binary *tuteo* vs *ustedeo* system where *tuteo* serves the purpose of conveying informality and familiarity, and *ustedeo* that of expressing formality and distance (Montes Giraldo 2000) – it is the variety that has received the least attention from scholars, at least with respect to 2PS address. Most of the attention has been focused on either the variable expression of the subject pronoun (Orozco and Guy 2008, Claes 2011, Orozco 2015) or, to a lesser extent, the sociophonetic phenomena typical of Caribbean regions, such as the lenition of coda/s and word-final nasal velarization (Canfield 1988, Zamora Munné and Guitart 1982, File-Muriel 2012).

3 Address forms in Bogotá

As noted in Section 2, both within Colombia and throughout the Spanish-speaking world, Bogotá Spanish is a unique variety in terms of address forms. In part, this is because it is in a region where *sumercé* is used to this day (Calderón Campos and García-Godoy 2020). It is also because in times where most Spanish varieties are considered extensively *tuteantes* or *voseantes*, Bogotá continues to be regarded as heavily *ustedean* (Carricaburo 2015). Moreover, due to large numbers of internal migration – census data show that over the past three decades the city has almost doubled in population (Cortés López 2017) – some studies have suggested that *vos* could be gaining ground (Mestre de Caro 2011).

Given the aforementioned details, it is not surprising that a considerable amount of research on 2PS address in Colombian Spanish has centered on Bogotá, with Uber's (1985) paper being perhaps the first to do so.² In her foundational study, the author argued that one of the reasons why Bogotá Spanish is extensively *ustedean* is that in this dialect *usted* has dual functionality: that of expressing deference toward superiors and strangers, and closeness with friends and relatives. As a result, unlike other Spanish varieties in which speakers generally use *tú* for [–power/+close] relationships and *usted* for [+power/–close] ones, Bogotá Spanish speakers may use *usted* to refer to their superiors, equals, subordinates, relatives, and strangers alike to such a degree that “a visitor to Bogotá might think that *usted* is the only form of address in the singular used there” (Uber 1985, 389). Besides establishing the dual function of *usted* in Bogotá Spanish, other important observations stemming from Uber's (1985) work concern the use of *tú* and *sumercé*. Regarding the former, the author notes that, though clearly less frequent than *usted*, its use was expanding for [–power/+close] relationships, especially among the younger generation and individuals from the upper-middle class. As to the latter, she notes that its use was fairly unpopular among her participants, who thought that it was dying out. Uber concludes that the observed increase in the use of *tú* could mean that *usted* and *sumercé* would be “used less frequently in the future” (1985, 391).

In general, studies following that of Uber (1985) set out to test the idea that, over time, *tú* would gain ground at the expense of both *usted* and *sumercé*, as well as to determine the social distribution of forms of address in Bogotá. Regarding *usted*, studies do not agree on whether this form continues to be preferred over *tú*. For example, Bayona (2006) and Cepeda Ruiz (2017) both report that *usted* persists as the preferred form of Bogotanos. Bayona (2006) found 75% *usted* and 21% *tú* in her interview-based data, while Cepeda Ruiz (2017) reports rates of 66% of *usted* and 34% of *tú* in her survey-based study. In contrast, Mestre de Caro (2011) and

² We are aware that other authors have written about 2PS address in Bogotá Spanish prior to Uber (1985). However, their work was not centered exclusively on this variety. Rather it was aimed at covering 2PS address in the Americas in general (e.g., Kany 1945) or centered on *voseo* in Colombia (e.g., Montes Giraldo 1967).

Lingård (2016) both report that *tú* has surpassed *usted* as the overall preferred address form of Bogotanos. Mestre de Caro (2011) found 42% *tú* and 34% *usted* in her television and radio talk data, while Lingård (2016) reports rates of 52% of *tú* and 44% of *usted* in her survey-based investigation.

Here it is worth recognizing that these studies' findings in terms of frequencies and percentages are not directly comparable due to methodological differences. In particular, whereas some obtained their data from more naturalistic situations (Bayona 2006, Mestre de Caro 2011), others obtained theirs from reported usage (Lingård 2016, Cepeda Ruiz 2017). However, gross comparisons between them suggest that, regardless of whether *tú* was found to be more frequent than *usted* or *vice versa*, there are gradual increases in rates of *tú* alongside gradual decreases in rates of *usted* that largely validate Uber's (1985) observation that the incrementation of *tú* over time could mean that *usted* would be used less frequently in the future.

This is even further supported by some of these authors' examinations of the social factors affecting the choice between *usted* and *tú*. For instance, Bayona (2006) and Lingård (2016) both found that older participants use and report more *usted* than younger participants. Correspondingly, these studies show that younger participants use and report more *tú* than older participants. This indicates a possible change in progress, irrespective of whether *tú* or *usted* is more frequent in general.

The gender of the speaker has also been found to affect people's choice between *usted* and *tú* in an important manner. For example, Lingård (2016) and Cepeda Ruiz (2017) both found that men report more *usted* than women, who correspondingly report more *tú* than men. In Lingård's (2016) data, we can even see that women's rates of *tú* are above the overall rate of use of this form. This suggests that women are well advanced in, even leading, the push toward *tú* observed in Uber (1985).

The gender of the interlocutor is another factor that has been found to affect the selection of *usted* or *tú* in Bogotá Spanish, particularly among men. Cepeda Ruiz (2017), who did a thorough analysis of this variable, found that the preference for *usted* observed among men is further accentuated in same-gender dyads, where they prefer *usted* over *tú* even more markedly than in their overall address selections. According to the author, who complemented her analysis of address selection with participants' metalinguistic comments, this is because men view *tú* as unmasculine. This corresponds with other studies of Colombian Spanish that have found that men tend to avoid *tú*, especially in same-gender dyads (Jang 2015, Newall 2016, Zwisler 2017).

Finally, the SES of individuals has been found to affect the choice between *usted* and *tú*. Cepeda Ruiz (2014), who looked into this by assigning speakers into lower-, middle-, and higher-socioeconomic groups, found that, though all groups favor *usted* over *tú*, the latter form is more prevalent among speakers of higher SES than among those of lower or middle SES. This is in agreement with Uber's (1985) observation that the expansion of *tú* began as a top-down phenomenon.

Regarding *sumercé*, studies do not agree on whether this form is set to disappear from Bogotá's address system, as suggested by Bayona (2006) and Mestre de Caro (2011), or whether it will continue to be used for years to come, as argued by Guerrero Rivera (2010) and Cepeda Ruiz (2014). However, with rates of use that do not reach two digits, regardless of whether the data were obtained from more naturalistic situations or survey-based questionnaires, studies agree that *sumercé* is used at rates considerably lower than *tú* and *usted*. They also agree that, similar to *usted*, *sumercé* has a dual function: that of expressing deference toward superiors and strangers, and closeness, even affection, with older relatives (Bayona 2006, Guerrero Rivera 2010).

A further observation regarding *sumercé* is that of Cepeda Ruiz (2014), who noted that, though all socio-economic groups favor *usted* and *tú* over *sumercé*, the latter is more prevalent among speakers of lower and middle SES than among those of higher SES.

Finally, with regard to *vos*, most studies agree that, with rates of use that are either equal to or below 1% – again, regardless of how the data were obtained – this form is marginal in the address system of Bogotá Spanish (Bayona 2006, Cepeda Ruiz 2014, Lingård 2016). Cepeda Ruiz (2014) goes as far as to mention that *vos* is simply not part of this variety's address system. An exception to this can be found in Mestre de Caro (2011) who, on the basis that in her data *vos* is more frequent than *sumercé*, argues that this form may be establishing itself as part of the city's address system due to internal migration from *voseante* regions.

The studies cited earlier have greatly contributed to our understanding of address forms in Colombia in general, and in Bogotá in particular, in addition to establishing the grounds for future research. Nonetheless, in light of the complex picture outlined earlier and the idea that Bogotá's address system may currently be

experiencing important changes, the present study seeks to contribute to this research by answering the following questions:

1. What is the current reported distribution of address forms in Bogotá Spanish, as found in data from 371 online surveys, and does this distribution align with the trends outlined in earlier studies (i.e., overall incrementation in the use of *tú*, decrease in the use of *usted* and *sumercé*, and possible establishment of *vos*)?
2. Taken together, which of the factors individually considered in earlier studies (speaker gender, age, SES, and regional background, as well as interlocutor gender and social role) significantly affect address selection in the Spanish of our Bogotano survey takers, and in what order of importance?
3. What are the meanings (proximity, distance, affection, effeminacy, etc.) that these individuals associate with *tú*, *usted*, *sumercé*, and *vos*, and could these associations help us better understand the distribution of address forms observed in our data?

4 The study

4.1 Survey questionnaire

Data for the present study are drawn from 371 online survey questionnaires with native speakers of Bogotá Spanish that were conducted in 2020. Participants were recruited through the second-named author's social networks in Bogotá – where he spent a significant part of his life – and through subsequent snowball sampling. After obtaining consent and asking participants to complete a short demographic section, the first portion of the survey asked them on 90 different occasions which form they would use between *tú*, *usted*, *sumercé*, and *vos* with 19 different types of interlocutors (child, parent, colleague, boss, service provider, medical doctor, etc.) further differentiated by gender (female or male). The decision to ask participants which pronoun form they would use rather than the verb form was based on the consideration that *usted* and *sumercé* use identical verb forms, and that *tú* and *vos* share many of the same verb forms. Figure 1 shows what each participant saw when asked to choose a form based on different types of interlocutors.

The second portion of the survey asked respondents to answer four open-ended questions concerning the meanings they associate with *tú*, *usted*, *sumercé*, and *vos*. Note that respondents were limited in neither the number of words they could write nor the number of associations they could establish, and that many decided to leave the field empty.

It is important to recognize that data such as that of the first portion of our survey present the disadvantage of not necessarily representing actual language use. Accordingly, there is a risk that “participants may

¿Cómo trata o trataría USTED a las siguientes personas?

(Si no tiene relación con alguna de las personas, por ejemplo, usted no tiene hijos, contemple cómo los trataría hipotéticamente. Seleccione solo una respuesta)

	Sumercé	Usted	Tú	Vos
Papá	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Mamá	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

Figure 1: Screenshot from the main portion of the survey.

under- or overestimate the frequency with which they use a given form” (Michnowicz and Quintana Sarria 2020, 125). Nevertheless, it is well established that the patterns of usage uncovered by survey-based approaches do reflect actual use (Morgan and Schwenter 2016). Moreover, part of the literature we are concerned with based their results on reported use of address forms. In that sense, the current research design serves the purpose of establishing rough comparisons between our data and that of these previous studies and, most importantly, that of determining the constraints significantly affecting address selection among participants.

4.2 Participants

In total, 580 individuals took our survey. After excluding participants who did not meet our selection criteria, 371 surveys remained for analysis. All of the individuals who completed these surveys were adults born in Bogotá, resided in the city at the time of the survey, and had spent the vast majority of their lives there. Participants were grouped based on a series of social factors that are shown in Table 1 alongside their levels and total number per level (not applicable for *participant age*).

As seen in the table, *participant age* was treated as a continuous variable. To facilitate statistical analysis, it was mean centered and *z*-scaled. This common practice enhances the overall goodness of fit and interpretability of the model, particularly when dealing with numeric predictor variables. This is especially crucial if the minimum value of the numeric variable is not zero and a regression model is employed (Drager 2018).

Participant SES was determined using a point-based system similar to that implemented in Otheguy and Zentella (2012). For each participant, a score was calculated based on their educational level, whether they attended a public or a private school, their occupation, their present level of income and social strata as determined by the *Departamento Administrativo Nacional de Estadística* (DANE), and that of each of their parents. While not perfect, the use of this criteria allows for a measurement of SES that is based on each participant’s social background and life trajectory. For instance, most participants in the “lower” group had either a high school education or technical training, an occupation that could be characterized as vocational, a DANE score of 1 or 2, and parents with similar backgrounds. The majority of participants in the middle group were first-generation college students or graduates, were either still in college or employed in their profession, and had a DANE score of 3 or 4. Nearly all participants in the higher group had attained a college education, were either still in college or employed in their profession, had a DANE score of 5 or 6, and had parents with similar backgrounds.

Regarding *participant regional background*, though all participants were born and raised in Bogotá, we wanted to explore the possibility that those with one or both parents from another region of the country could behave differently than participants with a longer history of settlement in the city, particularly in relation to the use of *vos*.

4.3 Analytical schemata

The analytical framework with which we look at the data drawn from the first portion of the survey is rooted in the tradition of *variationist sociolinguistics* (Labov 1972). One of the main advantages of this approach is

Table 1: Social factors and their levels (not applicable for *participant age*)

Factor	Levels
Participant gender	Woman ($n = 206$); Man ($n = 165$)
Participant age	Median age (range): 31 (18–82)
Participant SES	Lower ($n = 33$); middle ($n = 296$); higher ($n = 42$)
Participant regional background	Both parents from Bogotá ($n = 138$); one parent from Bogotá ($n = 113$); both parents from elsewhere in Colombia ($n = 120$)

that, instead of solely relying on independent raw counts and rates to identify possible differences in language practice between and within varying social groups, it allows for determining whether different linguistic and/or social constraints significantly affect the variability observed in language practice through multifactorial statistical modeling. Furthermore, because this approach assumes that the various constraints or independent variables that may affect a dependent variable (such as address choice) do not necessarily contribute equally to predicting its output, it pays close attention to the hierarchical classification of constraints inferable from the modeling of data in order to rank-order independent variables according to their greater or lesser effect on a given dependent variable. Accordingly, the statistical analysis included a series of mixed-effects logistic regression models, stepwise procedures, and goodness-of-fit measures that were implemented using *lme4* (Bates et al. 2015), *car* (Fox and Weisberg 2019), and *performance* (Lüdtke et al. 2021) in *R* (R Core Team 2021).

Following Chappell's (2018) idea of *motivated comparisons*, models were implemented pitting each form of address against the others. For example, to determine the factors predictive of *tú*, this form was pitted against *usted*, *sumercé*, and *vos*; to establish those predictive of *usted*, this form was pitted against *tú*, *sumercé*, and *vos*; and so on. In the models, the independent variables included *addressee social role*, *addressee gender*, the social factors outlined in Table 1, and interactions between *participant gender* and *participant age*, *participant gender* and *addressee gender*, and *addressee social role* and *addressee gender*. As the models were constructed, predictor variables and interactions were added with a manual stepwise procedure (Crawley 2007). Their contribution to the models' ability to predict the variation observed was determined using the analysis of variance function in *R*'s (R Core Team 2021) *car* package (Fox and Weisberg 2019). The factors that significantly ($\alpha \leq 0.05$) improved the models were retained.

The data drawn from participants' comments regarding the meanings they associate with different address forms were thematically coded and refined to tease out major associations for each address form, following an inductive approach (Hawkins 2018). To verify the reliability of themes, comments were coded by the first-named author and a trained assistant, and the Kappa statistic was used to assess agreement between coders. Agreement was nearly perfect (Landis and Koch, 1977), $\kappa = .93$, 95% confidence interval [.92, .95], and greater than would be expected by chance, $z = 123$, $p < 0.001$. Next, simple descriptive statistics were calculated to determine the top seven most frequent themes associated with *tú*, *usted*, *sumercé*, and *vos*, respectively. This cutoff was selected because, where varying themes were frequent, it allowed for excluding those that did not reach a recurrence level of 10% and, where few themes were frequent, it allowed for visualizing their dominance clearly. Finally, these data were triangulated with the results from the first portion of the survey, based on the idea that using participants' comments in conjunction with evidence of a given feature's distribution allows for potentially enhancing one's analytic claims (Pomerantz 2005).

5 Results

First, Table 2 shows the overall distribution of address forms that the participants selected in the first portion of the survey.

The first observation that can be made from this table is that though our results are not necessarily in agreement with authors claiming that *tú* has replaced *usted* as Bogotanos' favorite address form

Table 2: Overall distribution of address forms

Address form	N	Percentage
Usted	18,203	54.52
Tú	12,199	36.54
Sumercé	2,820	8.44
Vos	167	0.50
Total	33,389	100.00

(Mestre de Caro 2011, Lingård 2016), they do suggest that an overall incrementation in *tú* selection and a correspondent decrease in that of *usted* continues. Compared to Bayona's (2006) distribution of 75% *usted* and 21% *tú* and that of Cepeda Ruiz (2017) of 66% *usted* and 34% *tú*, our data illustrate that trajectory, with the caveat that, unlike the present study and that of Cepeda Ruiz (2017) whose findings are survey-based, Bayona's (2006) data were obtained from interviews.

A second observation regards the selection of *sumercé*. Our results are congruent with previous literature: its rates are considerably lower than *tú* and *usted*. However, they differ from studies like Bayona (2006) and Mestre de Caro (2011) who suggested that this form could disappear from Bogotá's address system. Our rate of 8.44% for *sumercé* being nearly identical to that of 8% reported in Cepeda Ruiz's (2014) survey-based study indicates that, despite being infrequent, the use of this form remains stable among Bogotanos.

The third and last observation to be made from this table is that *vos* selection is practically non-existent in our data, in line with the idea that it is not part of this variety's address system (Cepeda Ruiz 2014). In fact, due to very low counts of *vos*, this form was excluded from the statistical modeling of our data, to which we turn next.

Table 3 provides the results for the constraint ordering of factors significantly affecting 2PS selection in the data. The ordering ranks the constraints according to a statistic, the chi-square (χ^2) value, that quantifies their relative predictive power in the regression: the bigger the χ^2 , the more predictive the constraint. In addition to the χ^2 statistic, the table includes a goodness-of-fit measure (McFadden's R^2) for each of the three models that were fitted to 2PS selection. In each case, its value indicates an acceptable fit of the models to the data (see discussion in Ferguson 2009). Note that the table does not include the factors that were not selected by the statistical modeling of the data – that is, *participant regional background* and the interaction between *participant gender* and *participant age* – and these will not be further discussed. In any case, *participant gender* and *participant age* were found to significantly affect address selection either independently of or in interactions with other factors. Regarding the lack of effect of *participant regional background*, this simply indicates that whether participants have a long history of settlement in the city or not, they share similar address selection norms.

In Table 3, not only do we see that the factors significantly affecting *usted* and *tú* selection are nearly identical, they are also equally ordered in terms of hierarchy. The only difference between these two forms is that *participant SES* was found to affect *tú* but not *usted*. *Sumercé*, similar to *usted* and *tú*, shares the variable of *addressee social role* as the factor with the strongest effect on its selection and the variable of *participant age* as a factor with a less significant influence on this choice. Overall, the table tells us that the selections of *usted* and *tú* are similarly attuned to the same constraints, and that *sumercé* seems to be sensitive to *usted* and *tú* alternations, at least to a certain extent. This is congruent with the idea outlined earlier that *tú* is gaining ground at the expense of both *usted* and *sumercé*.

Having determined which factors significantly influence the selection of these different address forms in the data, next we provide an explanation of how each factor value impacts the constrains associated with each

Table 3: Constraint ordering of factors affecting 2PS selection in the data ($N = 33,389$)

Usted		Tú		Sumercé	
Factor	χ^2	Factor	χ^2	Factor	χ^2
Addressee social role	6571.4	Addressee social role	7058.8	Addressee social role	2683.3
Participant gender: Addressee gender	359.9	Participant gender: Addressee gender	446.2	Participant age	15.1
Addressee gender	355.1	Addressee gender	410.7		
Addressee social role: addressee gender	75.8	Addressee social role: addressee gender	73.6		
Participant age	12.3	Participant age	49.3		
Participant gender	3.5	Participant SES	5.9		
		Participant gender	3.5		
$R^2m/R^2c = .36/.59$		$R^2m/R^2c = .41/.65$		$R^2m/R^2c = .17/.74$	
Random factor: participant					

Table 4: Constraint-internal probabilities for *addressee social role: addressee gender* in *usted* and *tú*, and for *addressee social role* in *sumercé* (F = female, M = male)

Usted			Tú			Sumercé		
Levels	P	N	Levels	P	N	Levels	P	N
M medical doctor	.93	964	F partner/spouse	.97	610	Country person	.18	766
M unknown person	.92	948	F child	.96	332	Grandparent	.14	229
M teacher	.89	918	M child	.95	326	Parent	.07	159
M service provider	.87	2,659	M partner/spouse	.94	751	In-law	.05	133
F medical doctor	.87	897	F young relative	.92	313	Auncle	.04	122
F teacher	.82	842	M young relative	.88	297	Service provider	.01	542
F unknown person	.81	837	F friend	.83	837	Child	.01	35
M foreigner	.80	269	F cousin	.75	754	Sibling	.01	96
M boss	.78	793	F colleague	.69	702	Teacher	.007	121
F service provider	.78	2,412	F parent	.64	221	Unknown person	.007	120
F foreigner	.70	243	F sibling	.63	438	Medical doctor	.006	110
F boss	.66	701	M friend	.58	636	Boss	.005	97
M in-law	.62	216	F unknown kid	.58	206	Colleague	.003	74
M country person	.61	638	M cousin	.55	611	Friend	.003	76
F country person	.54	595	M sibling	.53	397	Cousin	.003	68
M colleague	.50	546	M unknown kid	.53	198	Foreigner	.002	15
F in-law	.48	182	M parent	.49	187	Partner/spouse	.001	29
M unknown kid	.43	164	F auncle	.48	179	Unknown kid	.001	14
M auncle	.40	156	M colleague	.43	522	Young relative	.001	14
M cousin	.38	451	F grandparent	.42	162			
F unknown kid	.38	156	M auncle	.34	149			
M sibling	.36	287	M grandparent	.33	145			
M friend	.35	423	F boss	.27	364			
F auncle	.30	131	F foreigner	.26	118			
F sibling	.28	251	F in-law	.25	115			
F colleague	.27	370	M boss	.16	269			
M grandparent	.25	112	M foreigner	.16	92			
M parent	.24	109	M in-law	.15	87			
F cousin	.21	311	F service provider	.13	650			
F grandparent	.18	91	F teacher	.12	208			
F friend	.14	231	F unknown person	.12	213			
F parent	.11	64	F medical doctor	.09	162			
M young relative	.11	64	F country person	.07	136			
F young relative	.07	46	M service provider	.06	393			
M partner/spouse	.04	41	M teacher	.06	136			
F partner/spouse	.03	43	M country person	.04	90			
M child	.03	24	M medical doctor	.04	92			
F child	.02	18	M unknown person	.04	101			
Range	.91		Range	.93		Range	.18	

address form. This analysis will enable us to determine the direction of the effects imposed by these constraints. It is important to highlight that, in order to avoid redundancy and capture the combined influence of variables involved in interactions in the outputs of the models, we will report on these variables within their interrelationships with one another. Therefore, when examining *usted* and *tú*, we will evaluate the effects of each individual value within the *addressee social role* variable alongside the effects of each individual level within the *addressee gender* variable. Similarly, we will assess the effects of each individual value within the *addressee gender* variable in conjunction with the effects of each individual level within the *participant gender* variable.

5.1 Interaction between addressee social role and addressee gender

Table 4 shows the effect of the interaction between *addressee social role* and *addressee gender* on the selection of *usted* and *tú*, and the effect of *addressee social role* on that of *sumercé*. In this and the following tables describing the effects of categorical variables, we include model-calculated probabilities (P) and raw counts (N) for each of the variables' levels (and/or their interaction), as well as the range of variation of the variable (or interaction) for each address form.³

In this table it can be seen that, far from the idea that a visitor to Bogotá would think that *usted* is the only address form used there (Uber 1985), this visitor would be faced with the reality that this form is in fact favored only with [+power/–close] interlocutors, such as medical doctors, service providers, and strangers. In line with the observation that *tú* has gradually gained ground over *usted* (Bayona 2006, Lingård 2016, Cepeda Ruiz 2017), this visitor would be able to see that the former is at present the most likely to be selected for [–power/+close] relationships, such as those that one nurtures with friends, siblings, and partners. Finally, and in agreement with the idea that the use of *sumercé* is fairly marginal, the visitor would notice that this form is reserved for sporadic use with very specific types of interlocutors, namely individuals from the countryside and older relatives.

What the interaction between *addressee social role* and *addressee gender* adds to our understanding of these relationships in the data is that, whether favored or not, *usted* is consistently more likely to be selected with male interlocutors of any social role than with female interlocutors. Correspondingly, and in line with the idea that *tú* may be seen as unmasculine (Jang 2015, Newall 2016, Zwisler 2017), this form is invariably more likely to be selected with female interlocutors than with male ones. The fact that *sumercé* is unaffected by the gender of the interlocutor is further proof that, overall, this form is independent from *usted* and *tú* alternations.

5.2 Interaction between participant gender and addressee gender

Proceeding to the interaction between *participant gender* and *addressee gender*, Table 5 shows its effect on the selection of *usted* and *tú*.

From this table, we can see that male interlocutors elicit more *usted* than female interlocutors, whereas the latter tend to elicit more *tú* than the former. This is in line with the observations from Section 5.1. However, the effect that *participant gender* has on this association suggests that it is strongly mediated by the participants' gender identity. Most of the observed effect seems to be driven by the fact that men are particularly unlikely, even reluctant, to select *tú* in same-gender dyads, where they clearly favor the selection of *usted*. Though this appears to be in line with the aforementioned idea that *tú* has connotations of effeminacy, the fact that men do not seem to be equally reluctant to select *tú* with female interlocutors adds to our understanding of this indexicality by showing that it is contextually bound to male, same-gender dyads only. In other words, while male participants seem more likely than female participants to select *usted*, who correspondingly appear to be likelier than their male counterparts to select *tú*, this difference ceases to be clearly visible when the gender of the addressee is considered.

5.3 Participant age

Table 6 displays the results for the effect of *participant age* on 2PS selection. In contrast to the previous tables outlining the effects of categorical variables, this table provides model estimates (E) and raw counts (N) for

³ The model-calculated probabilities, as implied by the name, were derived from the output of the models. To accomplish this, we utilized *emmeans* (Lenth 2020) in *R* (R Core Team 2021).

Table 5: Constraint-internal probabilities for *participant gender: addressee gender* in *usted* and *tú*

Usted			Tú		
Levels	<i>P</i>	<i>N</i>	Levels	<i>P</i>	<i>N</i>
Male Male	.60	4,804	Male Female	.53	3,833
Female Male	.38	3,612	Female Male	.46	3,518
Female Female	.38	4,809	Female Female	.45	3,202
Male Female	.32	4,978	Male Male	.22	1,646
Range	.28		Range	.31	

Table 6: Effect of *participant age* in *usted*, *tú*, and *sumercé*

Usted		Tú		Sumercé	
<i>I</i>		<i>I</i>		<i>I</i>	
-3.645		2.581		-5.197	
<i>E</i>	<i>N</i>	<i>E</i>	<i>N</i>	<i>E</i>	<i>N</i>
0.260	18,203	-0.586	12,199	0.584	2,820

each address form, along with their corresponding estimated intercept coefficients (*I*). This allows for calculating the percentage of change in the intercept of each address form for every one-unit increase in *participant age* using the anti-logit function in *R* (R Core Team 2021). It is important to note, however, that since this variable was mean centered and z-scaled, a one-unit difference corresponds to a one standard deviation difference.

Two observations can be drawn from Table 6. First, there is an apparent time decline in both *usted* and *sumercé*, with the probability of selection of these forms predicted to increase by 6.5 and 14.2%, respectively, for every one-unit increase in *participant age*. Second, there is an apparent time increase in *tú*, with the probability of selection of this form predicted to decrease by 14.26% for every one-unit increase in *participant age*. These effects, which are further illustrated in Figure 2, are in line with the idea that *usted* and *sumercé* are being ousted from Bogotá's address system gradually over time and that the incrementation of *tú* is part of an ongoing change driven by younger generations (Uber 1985, Bayona 2006, Lingård 2016).

5.4 Participant SES

Let us now move to the last constraint affecting the selection of a 2PS variant in the data: that of the effect of *participant SES* on the selection of *tú* presented in Table 7.

In this table we see that, as expected from previous studies (Uber 1985, Cepeda Ruiz 2014), the shift toward higher rates of *tú* is mostly driven by participants of higher SES, who are significantly more likely than both individuals of lower and middle SES to select this form. As such, these results suggest that the expansion of *tú* seems to be well advanced and largely accepted by individuals of higher social strata, but that this is not necessarily the case among members of the lower and middle socioeconomic groups, a point to which we will return in Section 6. Next, let us look at the meanings that participants associate with the four address forms available to speakers in Colombian Spanish and how these associations could enhance our understanding of the results presented thus far.

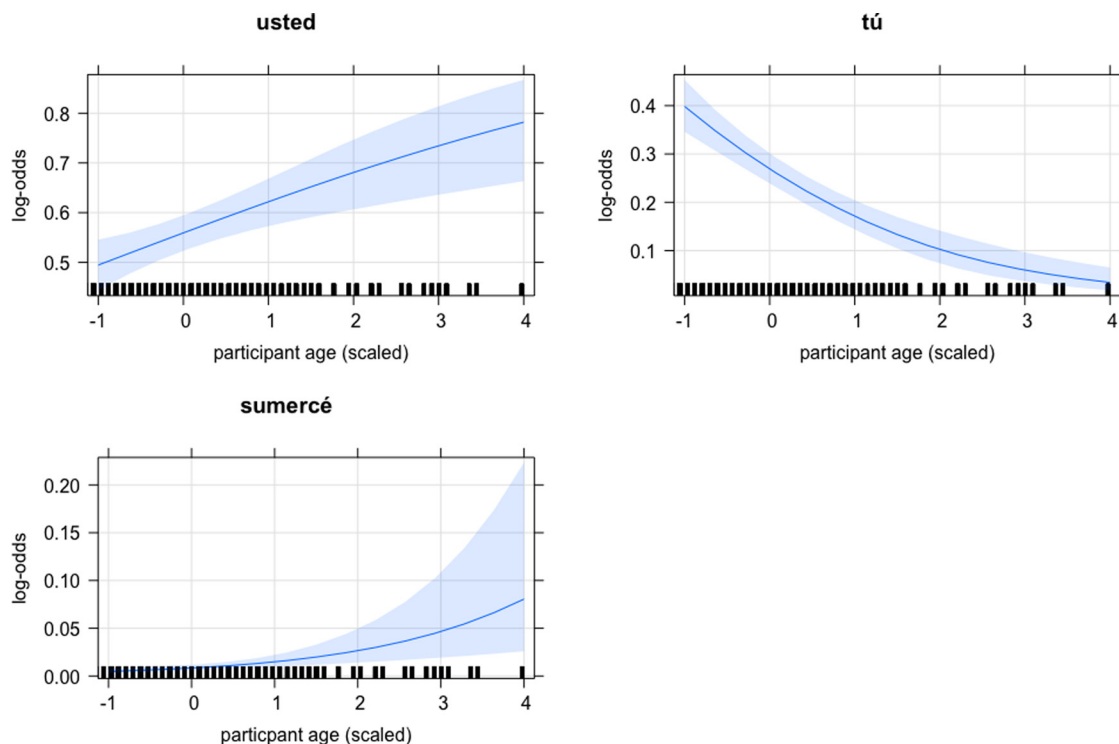


Figure 2: Effect plots of participant age on *usted*, *tú*, and *sumercé*.

Table 7: Constraint-internal probabilities for *participant SES* in *tú*

	Tú	
Levels	<i>P</i>	<i>N</i>
Higher	.52	1,481
Middle	.38	9,635
Lower	.33	1,083
Range	.19	

5.5 Main themes associated with varying address forms

Figure 3 presents the top seven most frequent themes that were elicited from participants' comments when asked about the meanings they associate with *usted*. In this and the following figures, the plus symbol beside a theme indicates that more than one topic was coded from a participant's comments.

The themes from Figure 3 demonstrate that *usted* is overwhelmingly associated with proximity, distance, and respect, in line with Uber's (1985) idea that this form has a dual T and V function in Bogotá Spanish. This is illustrated both by the fact that these are the three most recurring standalone themes in the data ($n = 86$, $n = 71$, and $n = 15$) and by their repetition as the third ($n = 49$), fifth ($n = 11$), and seventh ($n = 9$) most frequent themes, alongside the acknowledgement from participants that the form has multiple meanings.

As to how these findings may enhance our understanding of the distribution of address forms previously observed, the fact that participants associate *usted* with the expression of proximity to such a great extent likely explains why this form has not been completely displaced by *tú* in [-power/+ close] relationships, even though the preference for the latter in this type of dyad is obvious from the results presented earlier. In other words, the fact that participants strongly associate *usted* with the expression of proximity – at the same time that they acknowledge other values for this form – could well explain why the shift from *tú* to *usted* initially

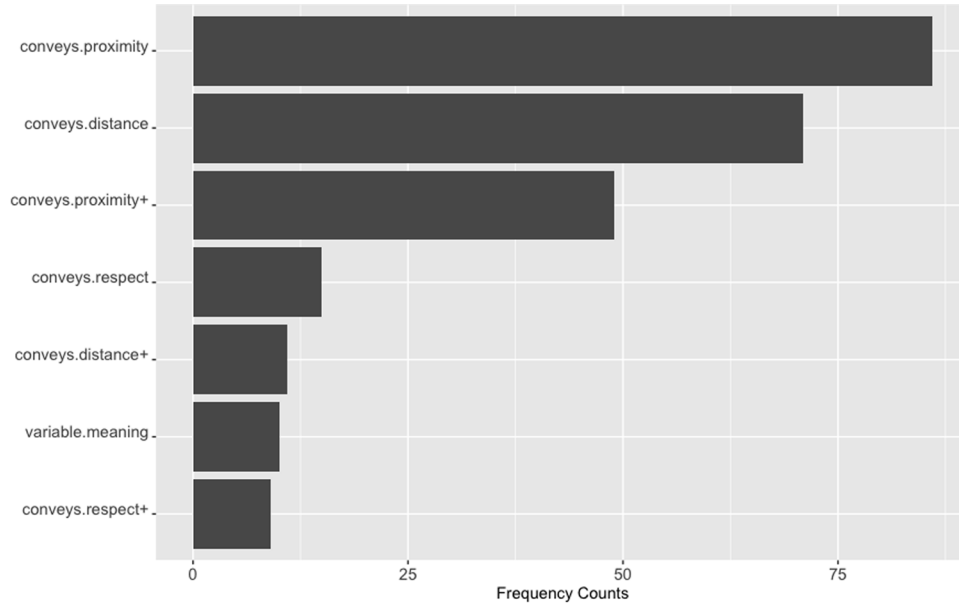


Figure 3: Top seven most frequent themes for *usted* (N = 273).

reported by Uber (1985) nearly four decades ago, and attested here and in previous studies, does not seem to reach completion.

In considering the *tú* form, Figure 4 displays the top seven most frequent themes that were coded from participants’ comments when asked about the values they associate with it.

Similar to *usted*, we can see that the main association of *tú* is with proximity, which is both the most frequent theme in the data (n = 84) and the fourth most frequent one overall (n = 24), alongside the acknowledgment from participants that it may have additional connotations. Unlike *usted*, however, the idea that this form can also convey the opposite notions of distance and deference is absent from the data, suggesting that *tú* lacks the duality of *usted* and is mostly exclusive to familiar settings. The fact that the second most frequent

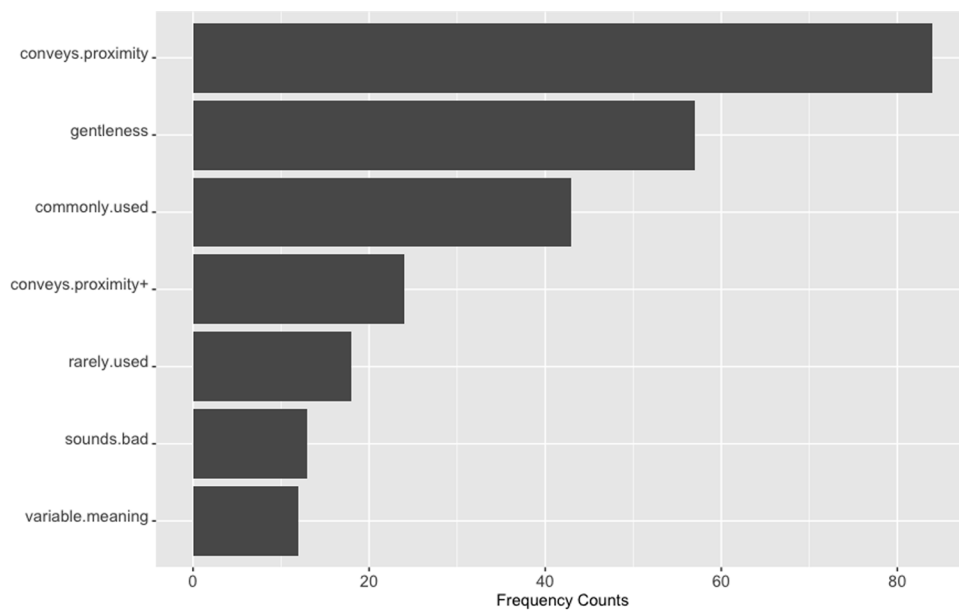


Figure 4: Top seven most frequent themes for *tú* (N = 282).

theme is that of gentleness ($n = 57$) is not unconnected to this, as it makes sense that one reserves delicate treatment for [-power/+close] interlocutors like relatives and friends.

Regarding the ways in which these findings could help us better understand the results documented in Section 5, it is important to note that some of the comments coded under “gentleness” included references to the gender-specificity of *tú* observed earlier. For instance, one participant noted that he thinks of this form as a way to *expresarse delicadamente ... con casi todas las mujeres y con familiares más cercanos* ‘gently express yourself ... with most women and your closest family members’, and another pointed out that, though she thinks of it as a productive way to indicate delicate treatment across varying situations, *[sus] amigos y [su] novio solo usteen a sus amigos hombres [porque] si comenzaran a tutear tal vez sonaría “gei” [sic] [her] friends and [her] boyfriend only use usted with their male friends [because] if they started to use tú it might sound gay*. This helps to further understand why, despite repeated evidence that *tú* is gaining ground over *usted* in [-power/+close] contexts, the former has not yet ousted the latter, at least in man-to-man dyads, where its use could be seen as contravening the conventions of a largely heteronormative society. It could also explain the contradiction that emerges from its association with a feature that is both commonly ($n = 43$) and rarely ($n = 18$) used, that is variable in meaning ($n = 12$), and that may sound bad ($n = 13$). This is because while many participants think of *tú* as a productive form to convey familiarity, some think of it as a feature to be actively avoided, especially in men-to-men dyads, due to its connotations of effeminacy.

Next, Figure 5 shows the top seven most frequent themes that were obtained from participants’ comments when asked about the meanings they associate with *sumercé*.

Similar to *usted*, the themes from Figure 5 demonstrate that *sumercé* is polysemic and thought of as both a T and a V form. This is illustrated by its association with both respect and proximity, which emerged as recurring themes on their own ($n = 25$ and $n = 20$, respectively) and in combination with other topics ($n = 32$ and $n = 14$, respectively). Interestingly, though, its major association is with affection ($n = 72$ and $n = 39$), which, in terms of contributing to a better understanding of the results presented herein, fits with the observation made earlier that this is a form that participants mostly reserve for older relatives. This is because within a prototypical Colombian household these types of interlocutors are likely to elicit both deference and familiarity from other family members, at the same time that they are deserving of their affection. The fact that *sumercé* was found to be even more likely with individuals from the countryside, however, remains largely unexplained by the participants’ comments.

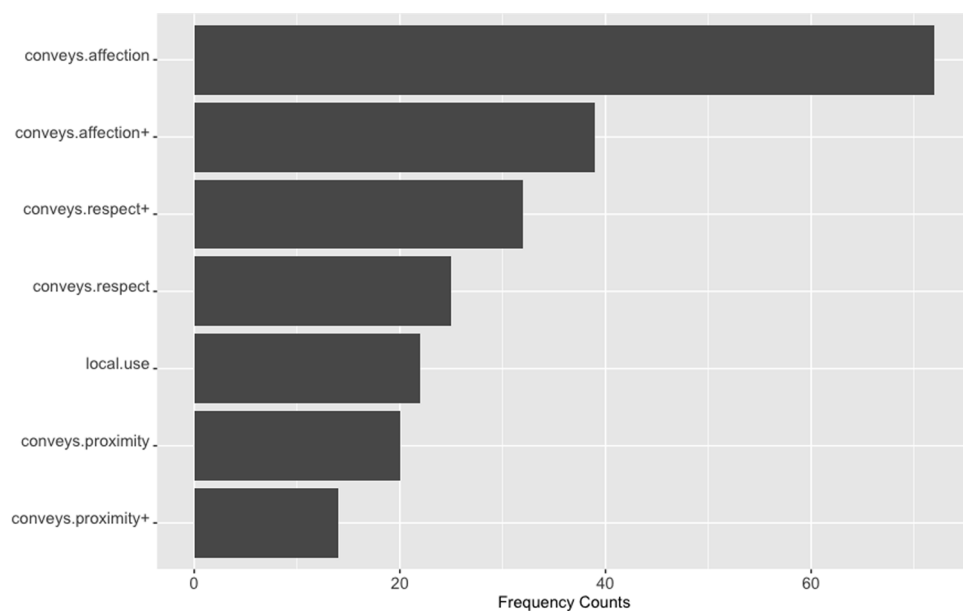


Figure 5: Top seven most frequent themes for *sumercé* ($N = 293$).

Another insight gained from triangulating participants' comments with the results from the first portion of the survey concerns the fact that the local nature of *sumercé* emerged as the fifth most frequent theme ($n = 22$) in participants' comments. We argue that its association with the local context could explain why, despite being used at very low rates, it remains part of Bogotanos' address system. This is because, in a way, it marks local identity, thus helping to distinguish the local variety from that of other regions of the country.

Finally, Figure 6 displays the top seven most frequent themes that were elicited from participants' comments when asked about the values they associate with *vos*.

From this figure, we can observe that *vos* is overwhelmingly thought of as not belonging to Bogotá's address system. Nearly nine in ten participants ($n = 299$) who commented about this form mentioned something to the effect of *el "vos" esta [sic] más asociado a otras regiones* 'using *vos* is associated more with other regions'. All remaining associations can be said to be marginal, with the exception of "internal migration" ($n = 39$), which consists of comments where participants mentioned associating *vos* with the presence of migrants in the city. In any case, this sufficiently explains the extremely low counts for *vos* found in our data.

6 Discussion and conclusion

Three main questions guided this study. The first asked whether our data could provide further evidence for change in the address system of Bogotá Spanish, particularly as regards the alternation between *usted* and *tú*. Though the evidence presented herein does not support the hypothesis that, since Uber's (1985) study, the latter has replaced the former as Bogotanos' most frequently used address form overall (Mestre de Caro 2011, Lingård 2016), it does support the idea that *tú* has consistently gained ground over *usted* (Bayona 2006, Cepeda Ruiz 2017). As a result, it is currently favored over *usted* in familiar settings. This finding is in agreement with previously established tendencies, not only throughout the Spanish-speaking world (Carricaburo 2015), but also in other T/V languages, like French, where according to Clyne et al. "*tu* is the default form of address ... with family and close friends" (2009, 43).

Consideration of this study's second research question – which asked what factors significantly affect address selection in Bogotá Spanish – further supports the change-in-progress hypothesis. This is because the gender and age of the participants were both found to similarly affect address selection (though, in line with

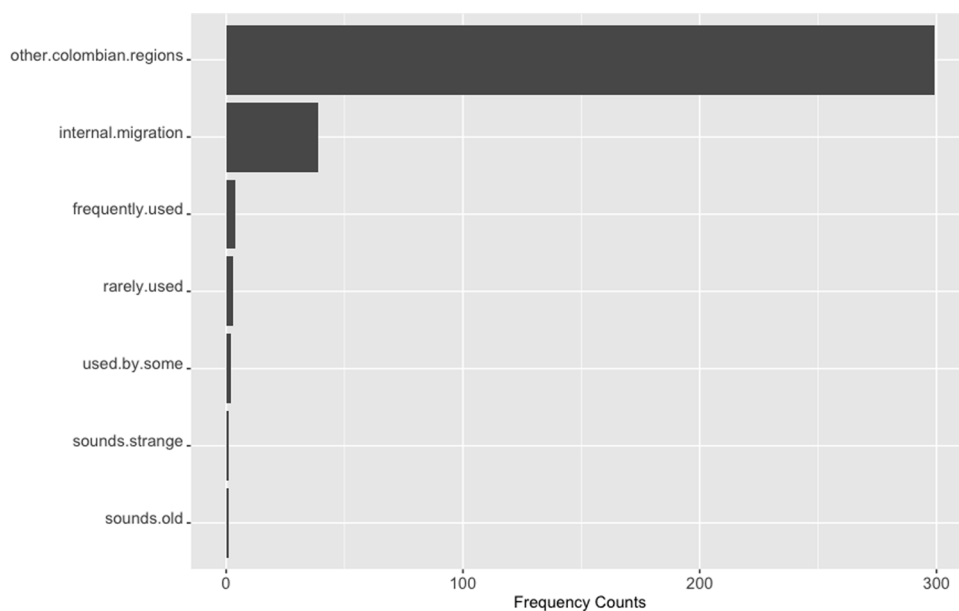


Figure 6: Top seven most frequent themes for *vos* ($N = 349$).

the pragmatic nature of address choice, the strongest factor in determining reported address forms is by far the social role of interlocutors). Specifically, it was found that the shift toward a greater use of *tú* at the expense of *usted* seems to be led by members of the younger generation and women. This is widely accepted as evidence that change is underway (Díaz-Campos 2014).

However, it is important to note that, while there is a push toward an increased use of *tú* with [–power/+close] interlocutors, and that this push seems to be led by younger individuals and women, there are also some factors that appear to be impeding the completion of this shift. One such factor is the gender of the addressee. In interaction with the social role of the addressee, it was seen that irrespective of whether *tú* is favored or not, male interlocutors consistently elicit less *tú* and more *usted*. In interaction with participant gender, it was seen that men actively avoid *tú* in favor of *usted* in same-gender dyads, while they seem to be equally – if not more – prone than women to use it in opposite-gender dyads. An additional factor affecting the shift toward *tú* is the SES of the participant. It was seen that compared to individuals of higher SES, people of lower and middle SES seem less likely to use this form.

Regarding the involvement of gender in the situation described earlier, a possible explanation resides in the relationship between *tú* and the lack of perceived masculinity documented here and in previous studies. We argue that given the heteronormative nature of Colombian society (Serrano Amaya 2011), there is a fairly strong social pressure for men to appear masculine. In terms of address selection, this is achieved by showing restraint in the use of *tú*, especially in the context of same-gender dyads. As to what role the SES of an individual plays in all this, we argue that, given that individuals of lower and middle social status often have no choice but to conform to rigid conceptions of masculinity and femininity due to social pressures (Lawler 2000), some may accordingly be more hesitant to fully embrace the use of a form with possible connotations of effeminacy.

Interestingly, the answers that participants provided in the survey with regard to the meanings they associate with varying address forms – which concerns our third research question – seem to confirm the interpretations we offer earlier. Regarding *tú*, for instance, it was seen that though this form is mainly thought of as a productive way of conveying familiarity, it is also overwhelmingly associated with delicate treatment, which, as one informant pointed out, could be seen as a sign of unmanliness in male, same-gender dyads. Moreover, and in agreement with Pomerantz (2005), the answers provided by participants allow for enhancing our analytic claims. Indeed, the fact that *usted* is thought of as a polysemic form that can convey closeness as much as it can signal formality and/or deference could well explain why it remains used in [–power/+close] relationships despite the dominance of *tú* in those contexts. This is most likely because it allows men to avoid possible connotations of delicateness and effeminacy without impinging upon the expression of closeness.

In considering the *sumercé* and *vos* forms, it was seen that our data support neither a projected disappearance of the former in Bogotá's address system (Bayona 2006) nor the establishment of the latter due to internal migration (Mestre de Caro 2011). Rather, they support the idea that *sumercé* is an integral part of the city's address system (Guerrero Rivera 2010) – albeit not very frequent – and that *vos* does not truly belong to it (Cepeda Ruiz 2014). Regarding *sumercé*, the study of the factors affecting its selection suggest that its very low rates are better explained by the fact that only a few types of interlocutors seem to elicit its use, namely individuals from the countryside and older relatives like parents, grandparents, and in-laws, regardless of their gender. As for the participant's comments, upon analysis, they provide insight into why this form persists despite its low frequency. It was seen that besides being thought of as a form that allows for the expression of both affection and deference, it is seen as a local feature. This being the case, we could see how its retention in the city's address system allows for distinguishing the local variety from that of other important metropolitan areas like Medellín, Cali, and Barranquilla. And though the fact that its selection was primarily elicited by individuals from the countryside remains largely unexplained by the triangulation of the current data, we could also see how this specific type of interlocutor may communicate a sense of local identity, even if they are not necessarily thought of as city dwellers.

As for *vos*, participants' comments are quite unequivocal as to where this form belongs: anywhere but in Bogotá, where its presence could only be expected in the speech of migrants. This largely explains why its selection did not even reach 1% in our data.

The complex situation described thus far invites the question as to what, in the context of a well-documented shift toward greater use of *tú* in [-power/+close] relationships, could best explain the persistence of *usted*, the retention of *sumercé*, and the rejection of *vos* in these contexts. We suggest that Bogotá's address system is best seen through the proposal that address forms may gain specific meanings within their particular context of use, despite having more conventional meanings attached to them (Blas Arroyo 2005, Fernández-Mallat 2020). As seen in our data and in the comments of the participants, the conventional meaning of *tú* would be that of expressing familiarity. However, in male, same-gender dyads, *tú* may become an index of unmanliness. To avoid this indexing, men often resort to *usted*, a form that has well-established connotations of familiarity despite being very much associated with the expression of distance and respect as well. For its part, *sumercé* adds further information highlighting the positioning of speaker to interlocutor, affection, and a sense of local identity. Regarding *vos*, while it certainly is a productive form to convey proximity in other Colombian regions (e.g., Medellín and Cali), in the context of Bogotá it is first and foremost thought of as a form that does not belong to the local address system and, if used, would most likely signal foreign status.

In closing, the results of the present study add to the existing research on address forms in Bogotá Spanish. Using data that combine reported usage and participants' metalinguistic comments, our findings contribute to a better understanding of the general shift toward a greater use of *tú* documented for the first time nearly four decades ago. By further paying attention to the factors affecting the selection of *usted* and *sumercé*, as well as to the place of *vos* within the city's address system, our findings also provide insight into why the first continues to be a relatively productive form in [-power/+close] settings despite the dominance of *tú* in these contexts, why the second has not been ousted from the system despite other scholars' suggestions that this would happen, and why the third does not seem to mix into the system despite large numbers of internal migrants from *voseante* regions.

Our study also offers opportunities for future research in several areas. This includes exploring how gender identity and presentation affect the use of *tú*, and whether the differences and associations observed in the present study regarding this form are reflected in everyday language practice or are mostly the product of ideological positions toward it. It also includes exploring the ways in which *sumercé* allows Bogotanos to construct a local identity in the same manner that, for example, *vos* functions as an identity marker for many Medellínenses and Caleños. Finally, it includes exploring the processes through which migrants and their offspring gradually adapt the address patterns found in their regions of origin to those found in Bogotá, especially among those with origins in *voseante* regions. Answering these and other questions would certainly shed further light on address forms in Bogotá in particular, and in Colombia in general.

Abbreviations

DANE	<i>Departamento Administrativo Nacional de Estadística</i> [National Administrative Department of Statistics]
SES	Socioeconomic status
2PS	Second person singular

Acknowledgments: First and foremost, we would like to express our deep gratitude to the individuals who participated in this study. We would also like to thank Ji Son Jang, Michael Sciacca, Austin Vander Wel, and the two anonymous reviewers who provided valuable feedback on previous versions of this study. Any remaining errors are ours alone.

Author contributions: Both authors have accepted responsibility for the entire content of this manuscript and approved its submission. The authors jointly designed the study, carried out data collection and analysis, and authored this article.

Conflict of interest: The authors state no conflict of interest.

Data availability statement: The datasets generated and analyzed during the current study are available from the corresponding author upon reasonable request.

References

- Avenidaño de Barón, Gloria S. 2014. “Formas pronominales de tratamiento y cortesía en el habla de Tunja, Colombia.” *Folios* 39, 31–49. doi: 10.17227/01234870.39folios31.49.
- Bates, Douglas, Martin Mächler, Ben Bolker and Steve Walker. 2015. “Fitting linear mixed-effects models using lme4.” *Journal of Statistical Software* 67(1), 1–48. doi: 10.18637/jss.v067.i01.
- Bayona, Patricia. 2006. “Sociolinguistic competence in the use of Colombian pronouns of address.” In *Proceedings of the 2006 Canadian Linguistics Association Annual Conference (CLA’06), 27-30 May 2006, Toronto, Canada*, edited by P. C. Gurski and M. Radisic.
- Benavides, Carlos. 2003. “La distribución del voseo en Hispanoamérica.” *Hispania* 86(3), 612–23. doi: 10.2307/20062914.
- Blas Arroyo, José Luis. 2005. *Sociolingüística del español: Desarrollos y perspectivas en el estudio de la lengua española en contexto social*. Madrid: Cátedra.
- Brown, Roger and Albert Gilman. 1960. “The pronouns of power and solidarity.” In *Style in language*, edited by Thomas A. Sebeok, p. 253–76. Cambridge: MIT Press.
- Calderón Campos, Miguel and María Teresa García-Godoy. 2020. “The European roots of the present-day Americanism *su merced*.” In *Address in Portuguese and Spanish: Studies in diachrony and diachronic reconstruction*, edited by Martin Hummel and Célia dos Santos Lopes, p. 413–41. Berlin: De Gruyter.
- Canfield, D. Lincoln. 1988. *El español de América: Fonética*. Barcelona: Editorial Crítica.
- Carricaburo, Norma. 2015. *Las fórmulas de tratamiento en el español actual* (2nd edn.). Madrid: Arco Libros.
- Cepeda Ruiz, Cristal Yeseidy. 2014. “Usted, tú, sumercé y vos: Formas pronominales de tratamiento en el español de Bogotá (Colombia).” Master’s thesis. México D. F.: Universidad Nacional Autónoma de México, México.
- Cepeda Ruiz, Cristal Yeseidy. 2017. “¿Tú o usted? Estigmatización del tuteo en Bogotá.” *Anuario de Letras. Lingüística y Filología* 5(2), 33–65.
- Chappell, Whitney. 2018. “The importance of motivated comparisons in variationist studies.” In *Contemporary trends in Hispanic and Lusophone linguistics: Selected papers from the Hispanic Linguistic Symposium 2015*, edited by Jonathan E. MacDonald, p. 143–68. Amsterdam: John Benjamins.
- Claes, Jeroen. 2011. “¿Constituyen las Antillas y el Caribe continental una sola zona dialectal?: Datos de la variable expresión del sujeto pronominal en San Juan de Puerto Rico y Barranquilla, Colombia.” *Spanish in Context* 8(2), 191–212. doi: 10.1075/sic.8.2.01cla.
- Clyne, Michael, Catrin Norrby, and Jane Warren. 2009. *Language and human relations: Styles of address in contemporary language*. Cambridge: Cambridge University Press.
- Cortés López, Miguel Alfonso. 2017. “Análisis del crecimiento en Bogotá y su impacto a la calidad de vida en la población de la región.” Undergraduate’s thesis. Bogotá, D. F.: Universidad de La Salle, Colombia.
- Crawley, Michael J. 2007. *The R book*. Chichester: John Wiley.
- Díaz-Campos, Manuel. 2014. *Introducción a la sociolingüística hispánica*. Malden: Wiley-Blackwell.
- Dos Santos Caliri, Izabel. 2019. “Los usos de la forma de tratamiento de segunda persona singular ‘vos’ entre los estudiantes en la Universidad de Nariño – Colombia.” Undergraduate’s thesis. São Carlos: Federal University of São Carlos, Brazil.
- Drager, Katie. 2018. *Experimental research methods in sociolinguistics*. London: Bloomsbury.
- Ferguson, Christopher J. 2009. “An effect size primer: A guide for clinicians and researchers.” *Professional Psychology: Research and Practice* 40(5), 532–38. doi: 10.1037/a0015808.
- Fernández-Mallat, Víctor. 2020. “Forms of address in interaction: Evidence from Chilean Spanish.” *Journal of Pragmatics* 161, 95–106. doi: 10.1016/j.pragma.2020.03.006.
- File-Muriel, Richard J. 2012. “A laboratory approach to s-lenition in the Spanish of Barranquilla, Colombia.” In *Colombian varieties of Spanish*, edited by Richard J. File-Muriel and Rafael Orozco, p. 127–40. Madrid: Iberoamericana.
- Fox, John and Sanford Weisberg. 2019. *An R companion to applied regression* (3rd edn.). Thousand Oaks: Sage.
- Guerrero Rivera, Javier. 2010. “Una reflexión sobre las fórmulas de tratamiento: ‘sumercé venga le digo.’” *Revista Interacción* 10, 55–61.
- Hawkins, Jennifer Morey. 2018. “Thematic analysis.” In *The Sage encyclopedia of communication research methods*, edited by Mike Allen, p. 1757–60. Thousand Oaks: Sage.
- Jang, Ji Son. 2015. “Matiz feminizante del tuteo y el futuro del voseo en el departamento de Antioquia (Colombia).” *Estudios Filológicos* 56, 85–99.
- Kany, Charles E. 1945. *American-Spanish syntax*. Chicago: University of Chicago Press.
- Labov, William. 1972. *Sociolinguistic patterns*. Philadelphia: University of Pennsylvania Press.
- Landis, J. Richard and Gary G. Koch. 1977. “The measurement of observer agreement for categorical data.” *Biometrics* 33(1), 159–74. doi: 10.2307/2529310.
- Lawler, Steph. 2000. “Escape and escapism: Representing working-class women.” In *Cultural studies and the working class: Subject to change*, edited by Sally R. Munt, p. 113–28. London: Cassell.

- Lenth, Russell. 2020. *emmeans: Estimated marginal means, aka least-squares means*. R package version 1.5.0. <https://CRAN.R-project.org/package=emmeans>.
- Lingård Arja. 2016. “El ustedeo de familiaridad en Colombia.” Master’s thesis. Tromsø: Norges Arktiske Universitet, Norway.
- Lipski, John M. 2004. *El español de América* (3rd edn.). Madrid: Cátedra.
- Lüdecke, Daniel, Mattan S. Ben-Shachar, Indrajeet Patil, Philip Waggoner and Dominique Makowski. 2021. “Performance: An R Package for assessment, comparison and testing of statistical models.” *The Journal of Open Source Software* 6(60), 3139. doi: 10.21105/joss.03139.
- Mestre de Caro, Pilar. 2011. “Alternancia de pronombres en el habla de Bogotá.” *Enunciación* 16(2), 17–30.
- Michnowicz, Jim and Vanessa Quintana Sarria. 2020. “A new look at forms of address in the Spanish of Cali, Colombia.” *Hispanic Studies Review* 4(2), 121–39.
- Millán, Mónica. 2014. “‘Vos sos paisa’: A study of address forms in Medellín, Colombia.” In *New directions in Hispanic linguistics*, edited by Rafael Orozco, p. 92–111. Newcastle upon Tyne: Cambridge Scholars.
- Montes Giraldo, José Joaquín. 1967. “Sobre el voseo en Colombia.” *Thesaurus* 22(1), 21–44.
- Montes Giraldo, José Joaquín. 1982. “El español de Colombia: Propuesta de clasificación dialectal.” *Thesaurus* 37(1), 23–92.
- Montes Giraldo, José Joaquín. 2000. *Otros estudios sobre el español de Colombia*. Bogotá: Instituto Caro y Cuervo.
- Morgan, Terrell A. and Scott A. Schwenter. 2016. “Vosotros, ustedes, and the myth of the symmetrical Castilian pronoun system.” In *Inquiries in Hispanic linguistics: From theory to empirical evidence* edited by Alejandro Cuza, Lori Czerwionka and Daniel Olson, p. 263–80. Amsterdam: John Benjamins.
- Newall, Gregory M. 2012. “*Second person singular address forms in Caleño Spanish: Applying a theory of language regard*.” PhD’s thesis. Bloomington, IN: Indiana University, USA.
- Newall, Gregory. 2016. “Second person singular forms in Cali Colombian Spanish: Enhancing the envelope of variation.” In *Forms of address in the Spanish of the Americas*, edited by María Irene Moyna and Susana Rivera-Mills, p. 149–70. Amsterdam: John Benjamins.
- Orozco, Rafael. 2015. “Pronominal variation in Colombian Costeño Spanish.” In *Subject pronoun expression in Spanish: A cross-dialectal perspective*, edited by Ana M. Carvalho, Rafael Orozco, and Naomi Lapidus Shin, p. 17–37. Washington, D. C.: Georgetown University Press.
- Orozco, Rafael and Gregory Guy. 2008. “El uso variable de los pronombres sujetos: ¿Qué pasa en la costa Caribe colombiana?” In *Selected Proceedings of the Fourth Workshop on Spanish Sociolinguistics*, edited by Maurice Westmoreland and Juan Antonio Thomas, p. 70–80. Somerville: Cascadilla Proceedings Project.
- Otheguy, Ricardo and Ana Celia Zentella. 2012. *Spanish in New York: Language contact, dialectal leveling, and structural continuity*. Oxford: Oxford University Press.
- Pomerantz, Anita. 2005. “Using participants’ video-stimulated comments to complement analyses of interactional practices.” In *Conversation and cognition*, edited by Hedwig te Molder and Jonathan Potter, p. 93–113. Cambridge: Cambridge University Press.
- R Core Team. 2021. *R: A language and environment for statistical computing*. R Foundation for Statistical Computing, Vienna, Austria. <http://www.R-project.org/>.
- Ringaila, Bárbara and María Teresa Cristina. 1966. “La familia, el ciclo de vida y algunas observaciones sobre el habla en Bogotá.” *Thesaurus* 21(3), 551–649.
- Rojas Cárdenas, Julián David. 2018. “Formas y fórmulas de tratamiento en el español hablado en Sogamoso (Colombia).” Master’s thesis. Sogamoso: Universidad Pedagógica y Tecnológica de Colombia, Colombia.
- Serrano Amaya, José Fernando. 2011. *Challenging or reshaping heteronormativity with public policies? A case study from Bogotá, Colombia (IDS Working Paper 361)*. Brighton: Institute of Development Studies at the University of Sussex.
- Uber, Diane Ringer. 1985. “The dual function of usted: Forms of address in Bogotá, Colombia.” *Hispania* 68(2), 388–92. doi: 10.2307/342216.
- Weyers, Joseph R. 2016. “Medellín cuenta con vos: The changing role of voseo in written communication.” *Comunicación* 35, 67–81.
- Zamora Munné, Juan C. and Jorge M. Guitart. 1982. *Dialectología hispanoamericana: Teoría, descripción, historia*. Salamanca: Ediciones Almar.
- Zwislser, Joshua James. 2017. “Tú, usted and the construction of male heterosexuality in young, working class men in Tolima.” *Cuadernos de Lingüística Hispánica* 29, 39–62.