

## Article

# A Historical-Variationist Analysis of Subject Pronoun Expression in 19th and Early 20th Century Arizonan Spanish

Álvaro Cerrón-Palomino <sup>1,\*</sup>, Sergio Loza <sup>2</sup>  and Rosti Vana <sup>3</sup> <sup>1</sup> School of International Letters and Cultures, Arizona State University, Tempe, AZ 85281, USA<sup>2</sup> Department of Romance Languages, University of Oregon, Eugene, OR 97403, USA<sup>3</sup> Department of World Languages and Cultures, Sam Houston State University, Huntsville, TX 77341, USA

\* Correspondence: acerropa@asu.edu

**Abstract:** This diachronic study is a variationist analysis of subject pronoun expression (SPE) in two key Spanish-language Arizonan newspapers, *El Fronterizo* (1878–1914) and *El Tucsonense* (1915–1957), following Tucson’s annexation to the United States through the Gadsden Purchase, a period of great social change during which the Spanish-speaking population in the city underwent a gradual process of anglicization. Since some research on SPE in Spanish in the United States suggests that English-Spanish bilingualism increases the use of overt subject personal pronouns (SPPs) because of their almost categorical use in English, this study’s main aim is to track the initial stages of such progression in a period when social bilingualism was steadily extending in Tucson. In this respect, our results show that the presence of overt SPPs does increase over time in the data analyzed; however, lower rates of overt SPPs in contemporary Tucson and Phoenix spoken Spanish raise the possibility that the percentage surge in the aforementioned period is rather due to the offline written nature of the newspapers, which, for instance, weakens the effect of online constraints, such as switch reference, ambiguous TAM endings, and non-reflexive verbs. Even so, regression analyses with the mixed-effects statistical software Rbrul reveal that the linguistic factor groups shaping SPE in the diachronic data are essentially the same ones found operating in contemporary varieties of Spanish.

**Keywords:** language variation; change; bilingualism; linguistic convergence; U.S. Southwest Spanish



**Citation:** Cerrón-Palomino, Álvaro, Sergio Loza, and Rosti Vana. 2023. A Historical-Variationist Analysis of Subject Pronoun Expression in 19th and Early 20th Century Arizonan Spanish. *Languages* 8: 25. <https://doi.org/10.3390/languages8010025>

Academic Editor: Ian MacKenzie

Received: 26 March 2022

Revised: 2 November 2022

Accepted: 13 December 2022

Published: 13 January 2023



**Copyright:** © 2023 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

Situations of language contact inevitably bring forth the question of linguistic permeability: what features can be transferred directly or indirectly from one language to another? In the case of Spanish in contact with English in the United States, one locus communis has been to investigate whether the almost categorical use of overt subject personal pronouns (SPPs) in the latter<sup>1</sup> prompts an increase of pronominal subject rates in the former.

Some studies, mostly conducted in the U.S. East Coast, suggest that there is convergence in subject pronoun expression (SPE) between the two languages because of the influence of the English patterns (Lapidus and Otheguy 2005; Otheguy et al. 2007; Otheguy and Zentella 2012), whilst most research conducted in the U.S. Southwest has found a lack of such effect (Travis 2005, 2007; Torres Cacoullos and Travis 2010; Bayley et al. 2012; Cerrón-Palomino 2016). This disparity in the studies’ results, then, points to the existence of dialectal differences in U.S. Spanish with respect to SPE.

While the aforementioned studies examine the linguistic production of different generations of Hispanics in the U.S., the present historical sociolinguistics investigation examines SPE across time in the Spanish of Tucson, Arizona, throughout a period when contact with English was becoming more extensive.

Accordingly, the data analyzed come from two Spanish newspapers published in the decades following the annexation of Southern Arizona to the United States in 1854 through the Gadsden Purchase: *El Fronterizo* (1878–1914) and *El Tucsonense* (1915–1957). The period

covered by the aforementioned newspapers is of great importance to test the convergence account that U.S. Spanish displays more overt SPPs due to contact with the non pro-drop English patterns: it corresponds to a time of transition when a number of social changes brought about an unprecedented migration of English-speaking populations to Southern Arizona, which gradually outnumbered the preceding Spanish-speaking communities. Thus, Tucson experienced a steady process of anglicization, which changed the Hispanic population status from majority in 1880 (63.8%) to minority in 1940 (29.9%). As this fluid language contact between Spanish and English consolidated, flourishing bilingualism gave birth to heritage varieties of Spanish in a territory where Spanish monolingualism once enjoyed hegemony.

In order to investigate the validity of the convergence account at the earliest stages of substantial language contact between Spanish and English, this paper is guided by the following research questions:

- (1) How does the overt SPP rate in the 19th and 20th Century written Spanish of Tucson compare to those in contemporary varieties of spoken Arizonan Spanish?
- (2) Do linguistic factors conditioning variable SPE in the 19th and 20th Century written Spanish of Tucson differ from those in contemporary varieties?
- (3) What social factors favor overt SPPs in the 19th and 20th Century written Spanish of Tucson?
- (4) Does the rate of overt SPPs increase over time as bilingualism increases in Tucson?

With the aim of answering the aforementioned questions, we have organized this paper as follows: Section 2 offers a succinct historical overview of Tucson; Section 3 reviews the major variationist findings on Spanish SPE; Section 4 is devoted to the methodology and data employed in this study; Section 5 displays the envelope of variation of the phenomenon; in Section 6, we introduce the factors and hypotheses analyzed in this study; Section 7 is dedicated to the discussion of the results; in the last section, the concluding remarks to this study are presented.

## 2. Historical Background

Arizona first came under the gaze of the Spanish crown in the 16th Century beginning with the Vázquez de Coronado expedition in 1540, which journeyed to the lower Colorado River, the Grand Canyon, and the eastern part of the territory (Officer 1987). Although Vázquez de Coronado failed to locate mineral deposits, subsequent treks into the Sonoran Desert by Antonio Espejo (1582) and Captain Marcos Farfán (1598) resulted in promising prospects for the Spanish crown. As soldiers, profiteers, and priests began to settle the area under orders of Felipe II, religious missions promoted sparse settlement into New Spain's northern frontier, and they brought with them Spaniards to the surrounding Native American communities.

The San Xavier de Bac Mission was founded by the end of the 17th Century by Jesuit missionary Padre Kino. The violent and ruthless European colonization of the area was not met without resistance as Native Americans repelled attempts to evangelize the surrounding populations (Dobyns 1976). Nonetheless, the first permanent Spanish settlement in Arizona was established in 1752: the Tubac Presidio, forty-three miles south of Tucson (Sheridan 1995). The status of Tucson had been a *visita* of the San Xavier de Bac Mission since 1737 (Greenleaf and Wallace 1962), before the arrival of the founding father of the city, Irishman Hugo O'Connor. On 20 August 1775, after constant Apache resistance to Spanish colonization, O'Connor relocated the Tubac presidio forty miles to the north, which became the Presidio San Agustín de Tucson. Afterward, the town's name was shortened to "Tucson", which became a part of the state of Sonora, after Mexico's independence in 1821.

Undoubtedly, the process of colonization established Spanish as the language of prestige and power in Arizona for more than a hundred years until two major historical events reshaped the national boundaries as well as Arizona's sociopolitical status within the area once again. Not too long after Mexico's emancipation from the Spanish crown, in 1846, the Mexican-American War erupted, fueled by the ideology of Manifest Destiny.

The United States provoked a war with Mexico that ended in 1848 with the Guadalupe–Hidalgo Treaty, by which Mexico ceded most of the territory of present-day New Mexico, Utah, Nevada, Arizona (North of the Gila River), California, Texas, and western Colorado. Tucson, however, became part of the United States only after the Gadsden Purchase in 1854, despite Mexican troops remaining there until March 1856 (Sheridan 1995). As part of the New Mexico Territory until 1863, Tucson did not experience an accelerated influx of Anglo settlers, but instead endured a steady populating process, starting in the 1860s.

As shown in Table 1 (Sheridan 1986), Tucson experienced a gradual process of ethnic and linguistic reconfiguration. This social and demographic shift resulted in the Hispanic population’s decline from 70.6% in 1866 to 29.9% in 1940. Furthermore, such changes in Tucson’s milieu also signaled the beginning of continuous Spanish-English language contact, which brought about the minoritized status of Spanish-speaking people and the emergence of heritage varieties of Spanish. As the Anglo population outnumbered the Spanish-speaking population by the 1900s, Dubord (2013) notes that they imposed their ideologies, which included embedding the English language as an American ideal within public schools. As public institutions began to emerge in the new frontier, Tucson’s Mexican elites contested such indoctrination and sought to maintain Spanish as a cultural symbol and as an economic resource by supporting private Catholic schools in lieu of sending their children to Protestant public schools (Dubord 2013). Over time, the Americanization of Mexican children into the American value systems, which positioned Spanish as unpatriotic and non-modern or non-progressive, established English as the dominant language of Tucson. This period of anglicization coincides largely with the time when *El Fronterizo* and *El Tucsonense* were in print, so their linguistic content is an ideal display of the Spanish utilized in U.S.A.-annexed Tucson, and they constitute the consummate resource to study changes of Spanish in its earliest contact with English in the Southwest. The 1882 excerpt from *El Fronterizo* in (1) soundly summarizes the situation the Spanish language was facing in Tucson after the continuous arrival of English-speaking Americans:

- (1) Nuestra lengua, que hace dos siglos era la más extendida de todas, no ha cesado de perder terreno y está amenazada de ser ahogada por sus vecinos, y sobre todo por los ingleses, sus más terribles rivales. (*El Fronterizo*, News, 20 January 1882)  
 ‘Our language, which two centuries ago was the most extended of all, hasn’t stopped losing ground and is threatened to be drowned by its neighbors, and mainly by the English, its most terrible rivals’.

**Table 1.** Population in Tucson, 1860–1940.

Population	1866	1880	1900	1920	1940
Total	925	7007	7531	20,337	36,818
Hispanic	653	4469	4122	7489	11,000
Hispanic %	70.6	63.8	54.7	36.8	29.9

Our historical corpora thus offer a glimpse of the development of subject expression patterns at the outset of social bilingualism in Tucson.

### 3. SPE in Spanish

Subject pronoun expression (SPE) has become widely known as the showcase variable in the field of variationist sociolinguistics (Bayley et al. 2012). SPE has been extensively studied in both monolingual and bilingual Spanish varieties (Silva-Corvalán 1982; Bentivoglio 1987; Ranson 1991; Cameron 1993; Cameron and Flores-Ferrán 2004; Flores-Ferrán 2010; Samper Padilla et al. 2006; Orozco and Guy 2008; Carvalho and Child 2011; Holmquist 2012; Padilla 2020, 2021, forthcoming). To exemplify, studies on monolingual varieties have demonstrated two patterns. The first is that Hispanic-American Spanish varieties have been found to have more overt subject personal pronouns (SPPs) than peninsular varieties.

The second is that Caribbean Spanish varieties account for much higher overt SPPs than those of the mainland (Orozco and Guy 2008, p. 70).

In bilingual Spanish-English communities, due to the nearly obligatory usage of overt SPPs in English, some studies suggest that bilinguals tend to implement more overt SPPs in their Spanish in comparison to their monolingual counterparts (Bayley and Pease-Alvarez 1997; Lapidus and Otheguy 2005; Otheguy et al. 2007; Otheguy and Zentella 2012; Abreu 2012). However, other scholars suggest that there is no effect of English on overt SPP production when bilinguals speak Spanish (Silva-Corvalán 1994; Flores-Ferrán 2010; Travis 2005, 2007; Torres Cacoullós and Travis 2010; Bayley et al. 2012; Cerrón-Palomino 2016). Nevertheless, despite the amount of research relating to SPE in bilingual U.S. Spanish, there are very few studies on this variable in Arizona (Anderson 2013; Cerrón-Palomino 2016). To date, Anderson (2013) is the only study focusing on Tucson. In this light, the present study takes upon a historical variationist lens by investigating SPE in two historical Tucson newspapers.

#### 4. Data and Methodology

The present study follows the variationist framework (Labov 2001). However, as we look at a Spanish-language variety utilized during the second half of the 19th century up until the first half of the 20th century, the data obtained fit within the historical linguistics tradition (Schneider 2017) given that they come from written records. Using written records from the past to analyze the early stages of a language exhibits a normative style of writing, rather than a written vernacular form. Labov discusses how this is “the art of making the best use of bad data” (Labov 1994, p. 11). Despite the diaphasic difference between written and oral data, Mas I Miralles (2003, p. 3) comments “it is perfectly possible to distinguish greater or lesser levels of formality depending on the document type, such that lesser formality will tend to bring us close to the spoken language”. Nonetheless, this morpho-syntactic variable analyzed is not confined solely to a vernacular language variety: overt SPPs and null subjects can be used in many different Spanish styles and can be authenticated in historical Spanish written records.

As mentioned previously in Section 2, the data for this study were collected from two different digitized Tucson newspapers. Specifically, we gathered the data from various issues of each of the newspapers *El Fronterizo* (1878–1908) and *El Tucsonense* (1915–1957). As seen, the publication dates span from the end of the 19th century up until the first half of the 20th century.

Founded by Carlos Ygnacio Velasco, a Sonoran Mexican, *El Fronterizo* was a bi-weekly newspaper for U.S.–Mexico border regionalism and Mexican-American cultural pride. According to Kanellos and Martell (2000), Velasco was an anti-assimilationist given his active involvement in pushing for the repatriation of the Mexican-American population in Arizona to Mexico. After *El Fronterizo* ceased printing, another conservative Sonoran Mexican businessman, Francisco Moreno, established a new newspaper titled *El Tucsonense*. Through bi-weekly publication, *El Tucsonense* focused on the Mexican-American middle class and their own personal interests and strived to bring to light a sense of Americanism among the Mexican-Tucsonan population (Kanellos and Martell 2000), while also supporting refugees who were fleeing the Mexican revolution.

Although both *El Fronterizo* and *El Tucsonense* were highly conservative and therefore linked to the Republican Party, they supported Hispanic candidates running for local, state, and national government positions even if holding different ideals and running for a different party. This showcases the newspapers’ push for Hispanic American identity and unity, as can be seen in (2).

- (2) Así mismo les hemos hecho ver la conveniencia de no dividir su voto cuando se trate de elegir a uno de nuestra raza, sino por el contrario procurar favorecerlo todos; para que él nos represente en cualquiera parte que sea en el gobierno de la cosa pública. (*El Fronterizo*, Editorial, 13 September 1890)  
 ‘Likewise, we have pointed out to them the convenience of not dividing their votes when it comes to electing one of our race, but instead for all of us to favor him, so that he represents us anywhere in the Government or public administration’.

For the present study, we analyzed issues available online corresponding to the years 1882–1908 for *El Fronterizo* (32 issues) and 1924–1956 for *El Tucsonense* (32 issues). As we can see, the period available for *El Fronterizo* was shorter (26 years) than that of *El Tucsonense* (32 years). Our guiding principle was to use data from one issue per year, selecting issues that exhibited both variants across the majority of genres: news, editorials, advertisements, short stories, letters, essays, and miscellaneous. However, doing so would have left *El Fronterizo* with a shortage of six issues. In addition, there were no digitalized issues for *El Fronterizo* for the years 1885–1886, 1898–1903, and 1906–1907. In order to balance out the data distribution, we added *El Fronterizo* issues corresponding to years where the occurrences of both variants of the variable were scant. The distribution of the articles analyzed in this study for each newspaper is shown in Table 2.

**Table 2.** Articles analyzed for each newspaper by journalistic genre.

	News	Editorials	Short Stories	Advertisement	Miscellaneous	Essays	Letters	Total
<i>El Fronterizo</i>	29	20	31	17	23	12	28	160
<i>El Tucsonense</i>	41	23	25	35	27	30	19	200
Total	70	43	56	52	50	42	47	360

Following a strict course of selection and exclusions detailed in the section below, the data studied amounted to 2010 tokens, 830 (41.3%) corresponding to *El Fronterizo* and 1180 (59.7%) corresponding to *El Tucsonense*. The tokens were analyzed using the statistical mixed-effects software Rbrul (Johnson 2012).

### 5. Envelope of Variation and Exclusions

The present study, following the vast literature on Spanish variable SPE, includes only tensed verbal forms with a human subject referent, expressed as an SPP or null. The two variants are shown in (3): an overt pronominal subject, *yo*, subject of *estaré*, and five null subjects, corresponding to the verbs *crees*, *preguntará*, *dices*, *estoy*, and *comprenderá*.

- (3) –Yo estaré en mi cuarto.  
 – ¿Ø Crees que Ø preguntará por ti?  
 – Le Ø dices que Ø estoy enferma y con eso ya Ø comprenderá.

(*El Tucsonense*, Short Story, 9 March 1945)

- ‘– I will be in my room.  
 – Do (you) think (she) will ask for you?  
 – (You) tell her that (I) am sick and with that (she) will understand’.

As post-verbal pronominal subjects cannot be considered part of an analysis on SPE, but rather on subject position (Cerrón-Palomino and Vana 2022), all instances of post-verbal subjects, as *él* in (4), were omitted from the analysis.

- (4) En compañía de su hijo James, volvía él de un viaje al otro lado del río.  
 ‘In Company of his son James, he was returning from a trip from one side of the river to the other’. (*El Fronterizo*, News, 17 October 1896)

Sentences whose subjects were categorically null, like subject relative clauses, as in (5) were also excluded from our analysis.

- (5) Y son ladrones aquellos que Ø tesoran [sic] a costa del sudor del campesino a quien pagan un miserable jornal, mientras ellos guardan ricas cosechas en sus amplios graneros. (*El Fronterizo*, Short Story, 6 January 1882)  
 ‘And (they) are thieves who get rich at the expense of farmers, to whom (they) pay miserable daily wages, while they stack abundant harvests in their spacious barns’.

In addition, newspaper sections that displayed only one of the variants but lacked the presence of the other were excluded from our study. Finally, texts that appeared more than once in the issues examined, such as short stories and advertisements, were only considered once in the analysis.

## 6. Factor Groups Analyzed and Their Predictions

### 6.1. Linguistic/Internal Factor Groups

Following previous studies on SPE in Spanish (Silva-Corvalán 1994; Travis 2005, 2007; Otheguy and Zentella 2012; Lastra and Martín Butragueño 2015), a total of six internal (linguistic) factor groups were analyzed: grammatical person and number, verb class, verb lexeme, referential connection, ambiguity of the TAM ending, and reflexive use of the verb. In the sections that follow, they are described in more detail.

#### 6.1.1. Grammatical Person/Number

Throughout the SPE literature, there has been a variety of results with respect to the conditioning of overt SPPs by the grammatical person/number of the verb’s subject. Thus, the single generalization that can be inferred from this factor group is that singular verbal forms favor the overt variant more than plural verbal forms (Carvalho et al. 2015, p. xiv). In consequence, we expect that verbs conjugated in their singular forms like *conocía* in (6) will favor overt subjects more than their plural counterparts, like *hubieron caminado*, *necesitaban*, and *serían molestados*.

- (6) Cuando hubieron caminado corta distancia, el Señor Miller, según referencias, ordenó a la escolta se devolviera, que no necesitaban de su compañía; porque él conocía bien a los indios y estos a él, razón pues, porque no serían molestados en ninguna manera. (*El Fronterizo*, Short Stories, 5 May 1905)  
 ‘When (they) had walked the short distance, Mr. Miller, according to references, ordered the escort to turn around, that (they) didn’t need their company; because he knew the Indians very well and they knew him, [which is the] reason why (they) would not be disturbed at all’.

#### 6.1.2. Verb Class

In this study, we follow Travis’s (2007) verb semantic classification. We divided verbs into five categories: psychological, like *saber* (know), *pensar* (think), *entender* (understand); speech acts, like *decir* (say), *pedir* (ask); copulas, like *ser*, *estar* (be); motion, like *ir* (go), *venir* (come); and other, such as *tener* (have); and *querer* (want), which encompasses verbs not suitable in the previous groups. Following previous SPE studies (Carvalho et al. 2015, p. xv), we expect psychological verbs to favor the presence of overt SPPs in contrast to the rest.

#### 6.1.3. Verb Lexeme

Researchers like Posio (2011) and Orozco (2016) had cast doubts on the accuracy of using categories based on semantic features to investigate verbal effects on SPE. As Orozco and Hurtado (2020) point out, some recent theoretical proposals, mainly within the usage-based and exemplar models (Bybee 2001; Pierrehumbert 2001; Bybee and Torres Cacoullos 2008), suggest that speakers store detailed and extensive information about the words and expressions that they use more frequently. In this sense, researchers like Erker and Guy (2012) were among the first to propose analyzing verb frequency within the corpus. This type of analysis has been only possible with the advent of mixed-effects models statistical software like R and Rbrul, and has shown that verb lexemes operate differently on SPE in

speech communities regardless of the categories they have been traditionally assigned to in previous studies (Orozco 2016; Orozco and Hurtado 2020; Del Carpio forthcoming).

In consequence, in this study, we adhere to the newer proposals and analyze verbs as a random factor in order to determine if verb frequency plays a role in promoting the use of overt SPPs in offline corpora.

#### 6.1.4. Referential Connection (Same/Switch Reference)

The literature has attested without exception that overt SPPs tend to be favored when a referential switch of the second of two consecutive subjects occurs (Travis 2005, p. 329). This can be seen in (7), where the subjects of *compramos* and *vendió* occur with an overt SPP given the switch in reference, while *compró* does not, given the same reference.

- (7) ¿Necesita usted dinero? **Nosotros** compramos su derecho o sea su contrato, hipoteca o nota, si es que **usted** vendió o  $\emptyset$  compró su casa, lote, terreno o negocio en abonos! (*El Tucsonense*, Advertisement, 6 April 1951)  
 ‘Do you need money? We buy your property right or your contract, mortgage, or bill, if you sold or bought your house, lot, property, or business in installments.’

Hence, according to these previous studies on Spanish SPE (Abreu 2012; Alfaraz 2015; Otheguy and Zentella 2012; Padilla 2021; Silva-Corvalán 1994), we hypothesize that a change in subject reference will favor overt SPPs, while the same reference will favor null subjects.

#### 6.1.5. Ambiguity of the Tense/Aspect/Mood (TAM) Form

We present a binary division between ambiguous endings of TAM (conditional, imperfect, pluperfect, and subjunctive) and not ambiguous (present indicative, preterit, future perfect, and preterit indefinite). The logic underlying this division is due to reasons of clarification or disambiguation. Scholars have noted that speakers are inclined to use more expressed subjects with ambiguous TAM, and fewer pronouns with verbs of unambiguous TAM (Cameron 1993; Cerrón-Palomino 2014; Lastra and Martín Butragueño 2015; De Prada Pérez 2015; Padilla 2021). As such, it is hypothesized that ambiguous TAMs will favor overt SPPs. The factors alluded to are illustrated in (8): the verb *dije* with an unambiguous TAM and a null subject, and the verb *estaba* with an ambiguous TAM and overt SPP *usted*. On the other hand, *diría* is ambiguous but lacks an overt SPP.

- (8) **Usted** estaba en la capital de México, y  $\emptyset$  dije a ud. por carta que en otra  $\emptyset$  diría el motivo de mi vuelta, y fue este: las terribles incursiones de los indios. (*El Fronterizo*, News, 2 May 1884)  
 ‘You were in Mexico’s capital and (I) told you in a letter that in another one (I) would say the cause of my return, and (it) was this one: the terrible incursions of the Indians’.

#### 6.1.6. Reflexive Use of the Verb

Previous studies have reported that reflexive verbs tend to co-occur less frequently with overt pronouns, whereas verbs lacking the reflexive pronoun usually favor overt subjects (Abreu 2012; Carvalho and Child 2011; Michnowicz 2015; Otheguy et al. 2007; Otheguy and Zentella 2012). To this end, reflexive pronouns present additional referential information that disambiguates without the need for an overt SPP. In contrast, non-reflexive verbs can cause speakers to use overt personal pronouns to make the subject’s referent more apparent. We hypothesize that the Tucson newspapers will follow the aforementioned pattern, as is shown in (9): *amó* has an overt SPP with a non-reflexive construction, whereas *quería vengarse* has a null subject.

- (9) **Ella** no amó nunca por cálculo, sino por capricho, hay que decirlo así. E inculpaba de todo a Nana, y quería a toda costa vengarse de ella. (*El Tucsonense*, Short Stories, 28 September 1943)  
 ‘She never loved by calculation, but by whim, we have to say it like it is. And (she) blamed Nana for everything, and (she) wanted to take revenge on her at any cost’.

## 6.2. Social/External Factor Groups

Since the data for this study came from historical documents, it was clear that the potential social factor groups were scant. In the section that follows, they are described in detail.

### 6.2.1. Journalistic Genre

As already discussed, SPP rates can be affected by genre (Travis 2007) or conversation topic (Flores-Ferrán 2010). It is only logical to consider such effects on SPE given the textual genres in written language. For the analysis, we follow the classification of genres used by Gubitosi (2010) in her study on passive expression in 19th and early 20th Century Californian and New Mexican Spanish newspapers. The reason for utilizing the same classifications is that the time span coincides with the newspapers analyzed in this present study and because the layouts of Southwest Mexican-American newspapers are comparable.

As Jakobson (1960) mentions, the poetic function of language plays an important role in short stories. We, for this reason, hypothesize that short stories should favor overt SPPs given the stylistic freedom they convey and the poetic function they abound in. We also hypothesize that advertisements and letters, where the addressee-appealing conative function takes precedence, will favor overt SPPs. Conversely, as editorials are more formal and cater toward a more emotive function (Cerrón-Palomino and Vana 2022), they might promote more null SPPs. Finally, following Gubitosi (2010), we consider the news genre to be formal and primarily referential, thus promoting more null SPPs than the other genres under analysis.

### 6.2.2. Individual Journal

We divided the data according to their source: whether they were extracted from *El Fronterizo* or *El Tucsonense*. This was done to account for possible idiosyncrasies of each data set such as the overall Spanish proficiency of the staff, their bilingual status, (or lack thereof), Spanish instruction, the percentage of Mexican writers compared to that of U.S.-born writers, etc. We have no hypothesis regarding this factor group, as it is rather exploratory.

### 6.2.3. Year of Publication

As previously mentioned, *El Fronterizo* was actively publishing between 1878 and 1914, whereas *El Tucsonense* was operative between 1915 and 1957. Also, as stated above, we included in our study digitally available issues of each newspaper: for *El Fronterizo*, we analyzed the years 1882–1908 and for *El Tucsonense*, the period between 1924–1956. The Rbrul software allowed us to analyze this factor group as a continuous variable rather than as a discrete variable.

Using the year of publication as a continuous variable offers us the advantage of avoiding a somewhat arbitrary chronological division of the particular periods analyzed. Instead, if the year of publication is found to be statistically significant, we will be able to see if there is a positive or negative correlation with the variant labeled as the application value. In other words, a regression with a continuous factor group can provide us evidence of whether the likelihood of the publication year increases or decreases as time progresses. If the structural convergence hypothesis is correct, then there should be a positive correlation between overt SPPs and the year of publication: they would increase over time.

## 7. Results and Discussion

Table 3 shows the distribution of the variants under scrutiny in this study: 24.6% of overt SPPs and 75.4% of null subjects.



**Table 3.** Distribution of overt SPPs and null subjects in the Spanish of Tucson 19th and 20th century newspapers.

Variant	N	%
SPP	494	24.6
Null	1516	75.4
Total	2010	

In answering our first research question, the overall Tucson newspapers’ overt SPP rate is very close to the 20.2% attested for contemporary Tucson Spanish (Anderson 2013) and slightly higher than the 17.8% reported for contemporary Phoenix Spanish (Cerrón-Palomino 2016). In other words, there has only been a small fluctuation of the overt variant distribution across time.

In order to answer our second research question, we conducted a mixed methods multiple regression analysis of the linguistic factors constraining the presence of overt SPPs in our data with Rbrul, and the results are presented in Table 4.

**Table 4.** Rbrul analysis of internal factors contributing to overt SPPs in Tucson 19th and 20th century newspapers.

Factor	Total N		2010		
	Factor Weight	% Overt SPP	Tokens	p-Value	Range (fw)
<b>Grammatical Person &amp; Number</b>					
<i>tú + usted</i>	0.75	47.1	170		
<i>yo</i>	0.59	33.3	547		
<i>ustedes + vosotros</i>	0.51	26.7	120	0.000	44
<i>nosotros,-as</i>	0.43	20.8	255		
<i>él/ella</i>	0.41	17.5	681		
<i>ellos/ellas</i>	0.31	11.8	237		
<b>Lexeme (random)</b>		Standard deviation 0.494		N.A.	34
<b>Reference</b>					
switch	0.57	29.3	1191	0.000	14
same	0.43	17.7	819		
<b>Verb Class</b>					
psychological	[0.61]	36.5	178		
movement	[0.55]	22.2	108	N.A.	-
copula	[0.49]	32.4	210		
other	[0.49]	23	1319		
speech act	[0.36]	16.9	195		
<b>Reflexivity</b>					
non-reflexive	[0.51]	24.9	1852	N.A.	-
reflexive	[0.49]	20.3	158		
<b>Ambiguity of TAM</b>					
non-ambiguous	[0.5]	28.5	911	N.A.	-
ambiguous	[0.5]	21.3	1099		

As we can see, of the six factor groups mentioned in Section 4, only three obtained statistical significance: grammatical person and number, lexeme, and reference, whereas verb class, reflexivity, and ambiguity of the TAM ending were disregarded in the Rbrul runs.

First, we will refer to the factor groups that were shown to have a statistical effect in predicting the occurrence of overt SPPs. As with almost every SPE study, grammatical person and number, as well as reference, turned out to be factor groups decisively affecting the preference for overt SPPs in Spanish. Lexeme, on the other hand, has not been studied as much because of the lack of use of mixed-model statistical packages in previous decades.

With a robust range of 44, grammatical person and number was the strongest factor group favoring overt SPP use in our data. In order to preserve orthogonality in the results and to avoid possible interactions spotted performing cross tabulations, *tú + usted* were collapsed, as were *ustedes + vosotros*. Within the factor group, *tú + usted* was the strongest predictor of the variant at issue with a probability of 0.75, followed by *yo* with 0.59, and *ustedes + vosotros* exerting little to no influence with a factor weight of 0.51. As usual, the plural forms *nosotros/as* and *ellos/ellas* disfavored overt SPPs, and *él/ella*, which in very rare cases promotes pronominal subjects, patterned with the plural forms. In general, the factors behave similarly to what has been found in the literature. The random factor group of lexeme is ranked second with a range of 34, reflecting great lexical variation involved in the choice of overt SPPs and null subjects, which we will address in more detail later. The third factor group that achieved statistical significance, reference, has a rather small range compared to the strongest one: 14. As expected, switching the reference of a subject favors the occurrence of a pronominal subject (0.57), as opposed to maintaining the same referential subject across verbs, which promotes null subjects (0.43). The low probability, despite the statistical significance achieved, is likely due to the offline nature of journalistic genres, which is less sensitive to pragmatic features than the online character of speech.

Now we will discuss the linguistic factor groups rendered not significant in the regressions. Verb class, a factor group that has been relatively stable across studies as affecting SPE choice, was discarded in the analysis. This result is directly related to analyzing verb lexeme in the same run: in regressions not including the random factor group, verb class was selected as significant, with copulas as the strongest predictors of overt SPPs.

Table 5 offers a more detailed look at the results of the lexeme factor group. Out of the 531 different verb lexemes attested in the corpus, this table only includes a rather short list of them, i.e., those with 20 occurrences or more. As can be seen, the strongest predictor of the pronominal variant is *ser*, the verb with the highest number of occurrences (137). We then attest a decrease of significance until none is achieved, at around 30 occurrences, with the exception of verbs *estar* (75) and *ver* (49).

**Table 5.** Lexical frequency effect of the verb on overt SPPs.

Factor	Prob.	Token N	% Overt	% Data
<i>ser</i> 'to be'	0.7	137	43	6.81
<i>hacer</i> 'to do'	0.6	83	28.9	4.1
<i>decir</i> 'to say'	0.6	104	22	5.2
<i>tener</i> 'to have'	0.55	96	28.1	4.8
<i>saber</i> 'to know'	0.52	62	37	3
<i>creer</i> 'to believe'	0.52	42	38	2
<i>dar</i> 'to give'	0.47	30	20	1.5
<i>estar</i> 'to be'	0.45	75	1	3.7
<i>conocer</i> 'to know'	0.44	22	18.2	1
<i>dejar</i> 'to leave'	0.43	23	13	1.1
<i>esperar</i> 'to wait'	0.42	21	1	1
<i>ver</i> 'to see'	0.38	49	14.3	2.4
<i>poner</i> 'to put'	0.37	21	4.8	1

These results give us a clearer pattern of the verb constraint: the more frequent verbs favor more occurrences of pronominal subjects, and the least frequent ones promote null subjects, corroborating the findings of Orozco (2016), Orozco and Orozco (2022), and Del Carpio (forthcoming). For instance, *estar*, another copulative verb, does not favor overt SPPs in the present study, despite being relatively frequent. It seems evident, then, that the first regressions we run without lexeme as a random factor rendered copulas as the strongest predictor of the factor group mainly because of *ser*, which accounts for 60.5% of the tokens of the category.

However, Travis and Torres Cacoullós (2021) have argued that frequent verb-form combinations interact with semantic verb classes. They claim, in particular, that highly frequent cognition verbs (‘psychological’, in this paper) like *creer* affect SPE in Spanish in the first person singular, e.g., *(yo) creo*, but not across different grammatical person/number combinations. In this respect, Travis and Torres Cacoullós claim that frequent sequences like *yo creo* must be interpreted as chunks rather than as selected combinations. They substantiate their claim by showing that *(yo) creo* represents 84% of the occurrences of its lexical paradigm, 27% of all the instances of overt *yo + verb*, and 14% of all instances of verbs inflected for the first person singular, both overt and null.

In order to test the validity of this claim in our data, we followed the methodology used by Travis and Torres Cacoullós (2021) to obtain the overall token frequency and the relative frequency of all the instances of the *(yo) + verb* combinations of the verbs displayed in Table 5, whose recurrence favored overt SPPs in the data analyzed. The distribution of said combinations across the data studied is shown in Table 6.

**Table 6.** Proportion of *(yo) + frequent verb* in the diachronic newspapers’ data.

Verb	% Within Lexical Type	N	% Within <i>yo + Verb</i> (N = 182)	% Within <i>(yo) + Verb</i> (N = 547)
<i>ser</i>	37.2%	137	17%	5.7%
<i>hacer</i>	15.7%	83	3.8%	1.3%
<i>decir</i>	25%	104	4.4%	1.5%
<i>tener</i>	33.3%	96	6%	2%
<i>saber</i>	37.1%	62	5.5%	1.8%
<i>creer</i>	50%	42	5.5%	1.8%

As we can see in the second column, *(yo) creo* represents 50% of the occurrences of its lexical type, and *(yo) sé* comprises 37.1% of all *saber* inflected forms. Unlike the findings of Travis and Torres Cacoullós, where *(yo) creo* constituted 84% of the occurrences, it does not seem that either cognition verb had achieved the status of a particular unit at the time, although *(yo) creo* shows the highest proportions of any first person singular verb within its respective paradigm. The other frequent non-psychological verbs, whose rates go from 15.7% to 37.2%, do not seem to be experiencing chunking either.

The fourth column shows that the strings *yo creo* and *yo sé* each account for 5.5% of all the instances of *yo + verb*, strikingly lower than the 27% found by Travis and Torres Cacoullós for the former string, with *yo soy* being the most frequent *yo + verb* sequence at 17%. When all instances of verbs inflected in first person singular are taken into account (fifth column), *yo creo* and *yo sé* each represent only 1.8% of all occurrences, considerably lower than the 14% attested by the aforementioned authors for the former sequence, with *yo soy* being once more the most recurrent string at 5.7%. The aforementioned numbers thus suggest that *yo creo* and *yo sé* were not treated as chunks in the Spanish of the Tucson newspapers, either because the process of chunking was still incipient or because it was not a part of formal written styles.

Another factor group that was discarded in the regressions was reflexivity, which traditionally was not a robust factor group across studies. This factor group was perhaps even neutered further given the offline setting of the data, where the extra referential subject clue conveyed by the reflexive pronoun was not deemed redundant enough to favor null subjects over overt SPPs in a decisive way. Despite 24.9% of non-reflexive verbs attesting overt SPPs in contrast with 20.3% of reflexive verbs, following previous studies’ trends, the difference in the distribution of the variants was too meager to achieve statistical significance.

By the same token, we hypothesize that the lack of an online setting, where phonetic features can be crucial in aiding or hampering sentence processing, prevented the journals’

staff from resorting to the use of overt SPPs to clarify the referent of the subjects of verbs with ambiguous TAM endings.

In sum, and in answering our second research question, the linguistic factor groups that have proven to be most robust across the literature are also significant in these diachronic data: grammatical person and number, and reference. On the other hand, the factor groups that were not constantly significant in the literature, reflexivity, and ambiguity of the TAM, were discarded in our regression analysis. In addition, a factor group somewhat stable in constraining SPE, the semantic class of the verb, was deemed non-significant only when included in the same regression with verb lexeme as a random factor.

Now, to answer our third and fourth research questions, we move on to the social factors analyzed in this study. Table 7 displays the results of the regression analysis of three external factor groups, two of which were statistically significant: newspaper genre and the continuous variable of the year of publication, whereas specific journal was rendered not significant.

**Table 7.** Rbrul analysis of external factors contributing to overt SPPs in Tucson 19th and 20th century newspapers.

Factor	Total N		2010		
	Factor Weight	% Overt SPP	Tokens	p-Value	Range (fw)
<b>Genre</b>					
advertising	0.69	39.4	165	0.000	34
letters	0.57	37.5	40		
short stories	0.56	25.7	908		
editorials	0.55	28.1	114		
news	0.42	20.2	514		
miscellaneous	0.37	16.7	216		
essays	0.35	17	53		
<b>Year of publication</b>					
Continuous +1				0.000	——
Log odds −0.015					
<b>Newspaper</b>					
<i>El Tucsonense</i>	[0.55]	20.4	1180	N.A.	——
<i>El Fronterizo</i>	[0.46]	30.5	830		

The predictors of overt SPPs in the genre factor group are advertising (0.69), letters (0.57), short stories (0.56), and editorials (0.55) with the rest of the factors favoring null subjects. Regarding the year of publication, more overt SPPs tend to show up as time progresses, which supports prima facie the hypothesis that language contact favors English patterns to show up in Tucson Spanish as the territory is gradually anglicized and answers in a positive way our fourth research question.

As for the factor group that turned out to be non-significant, the specific journal from where the tokens were extracted, our results rule out possible hidden factors that could cause the overt SPP rate to spike in one way or another, such as the Spanish proficiency of the staff, their bilingual status, (or lack thereof), Spanish instruction, and so on. In general, the results of the social factors regressions go along with the hypotheses presented in Section 6.

The answer to our fourth research question takes us to the matter of whether there is an ongoing change in overt SPP rate in the Spanish of Tucson that started with the anglicization of the area after the Gadsden Purchase. If this were true, the overt SPP rate in contemporary varieties of Tucsonan Spanish should be higher than the 24.6% found in the early newspapers’ data, in a continuing linguistic change. However, as mentioned earlier, the rate found by Anderson (2013) was lower: 20.2%, which goes against the trend found in the present study. How can we account, then, for this seeming contradiction?

A plausible explanation of the chronological increase of overt SPPs in the data analyzed is that there is a surge of switch reference contexts across time favoring the rise of overt pronominal subjects. In order to explore this hypothesis, we performed independent crosstabulations of overt SPP/null subjects by reference and time period for the three grammatical person + number combinations showing probabilities of 0.5 or higher in Table 4: second person singular, first person singular, and second person plural. Tables 8 and 9 display the results for first person singular subjects across time in switch reference and same reference contexts, respectively.

**Table 8.** Crosstabulation of first person singular verbs in switch reference contexts across time.

		1882–1908	1924–1956	Total
SPP	N	61	76	137
	%	46.2	31.8	36.9
Null	N	71	163	234
	%	53.8	68.2	63.1
Total	N	132	239	371
	%	100	100	100

$p \leq 0.301$ .

**Table 9.** Crosstabulation of first person singular verbs in same reference contexts across time.

		1882–1908	1924–1956	Total
SPP	N	28	17	45
	%	38.4	16.5	25.6
Null	N	45	86	131
	%	61.6	83.5	74.4
Total	N	73	103	176
	%	100	100	100

$p \leq 0.017$ .

Table 8 shows that switch reference contexts increase over time (from 132 to 239), but the overt SPP rate decreases in said contexts from 46.2% to 31.8%, albeit not in a statistically significant way ( $p \leq 0.301$ ). However, the number of overt SPPs shows an increase from 61 to 76, which goes along the lines of the hypothesis.

Similarly, in Table 9 we see that same reference contexts also increase over time, from 73 to 103. Analogous to the switch reference contexts, the overt SPP rate decreases in coreferential contexts from 33.8% to 16.5%, also lacking statistical significance ( $p \leq 0.017$ ). However, in contrast to the number increase of overt SPPs experienced across time in switch reference contexts, same reference contexts show a decrease from 22 to 14 cases. Omitted here for reasons of space, the results for second person singular also exhibit a switch reference increase across time, albeit with a percentage and number decrease of overt SPPs, whereas the switch reference contexts decrease as well as the rates and number of overt SPPs for second person plural.

In sum, we do not attest in our data a surge of switch reference contexts across time that promotes the rise of overt pronominal subject rates. However, an increase of *yo* pronoun occurrences across time does take place, although this increase is not statistically significant and is not reflected in a rate increase.

Another possible explanation to explore for the chronological overt SPP increment displayed in Table 7 is that the different journalistic genres sway the overt SPP rate, on the one hand, and disengage the variable from oral speech patterns, on the other. In order to investigate whether the overt SPP rate spike across time is due to the array of data across the newspaper genres, Table 10 displays the number of tokens distributed by journalistic genre in each newspaper’s period and the rate of overt SPPs for each genre. For instance, if there were an increase of tokens from the genres that favor overt SPPs the most, i.e.,

advertising, letters, short stories, and editorials, it could explain the overt variant's surge over time.

**Table 10.** Distribution of overt SPP tokens across journalistic genres for each period studied.

	News	Editorials	Short Stories	Advertising	Miscellaneous	Essays	Letters	Total
1882–1908	24.8% 286	19% 42	43.8% 281	47.5% 48	11% 109	12.1% 33	38.5% 31	830
1924–1956	14.1% 228	33.3% 72	25% 627	36.8% 117	22.2% 107	25% 20	33.3% 9	1180
Total	514	114	908	165	216	53	40	2010

As we can see, three of the four favoring genres show a token increase, with only letters experiencing a decrease, from 31 to 9. In contrast, advertising contexts increase from 48 to 117, short stories increase from 281 to 627, while editorials increase from 42 to 72. However, the overall increase of tokens in these genres does not translate into an overt SPP rate increase: the latter increases only in editorials, from 19% to 33.3%, but decreases in the other three favoring genres. Conversely, the overt SPP rate also increases over time in genres not favorable for pronominal subjects, such as miscellaneous (11% to 22.2%) and essays (12.1% to 25%).

In summary, the data distribution by genre in the two periods does not seem to have skewed the results in favor of the overt variant.

However, there is the possibility that it is not the journalistic genres, but rather the more general offline nature of written texts that causes overt SPP patterns to differ—even if slightly—from those of spoken Spanish. In fact, in written texts, authors have the choice of correcting and amending their own production for different purposes before its publication, such as the avoidance of redundancies or the enhancement of emphasis. Along these lines of reasoning, the low range of the referential connection factor group shown in Table 6 seems to suggest that written Spanish differs from oral Spanish with regards to pragmatic features such as switch reference, as illustrated in (10), where overt SPP *ella*, preceded by two coreferential null subjects, would be more likely to be interpreted as referring to someone other than *Mrs. Diamond* in most varieties of spoken Spanish.

- (10) Mrs. Diamond, quien lloraba, con el pelo descompuesto y fumando cigarrillo tras cigarrillo, gritaba que ella era inocente del crimen. (*El Tucsonense*, News, 19 December 1931)  
'Mrs. Diamond, who was crying, with messy hair and smoking one cigarette after another, was yelling that she was innocent of the crime'.

By the same token, the impact of an ambiguous TAM ending and non-reflexive verbs, closely related to online production, seems to be neutralized in written output, where authors can edit an utterance at will, and the subject's referential retrieval does not hamper the structure's parsing, as seen in (11).

- (11) No descansan ni pueden dormir tranquilos **estos bellos ejemplares de indiscreción**, hasta no saber, con buenos datos, cuáles son los recursos con que cuenta Fulano para pasar la vida; averiguan cuántos matrimonios están en paz y cuántos riñen como perros y gatos; pueden decir cuáles comerciantes andan mal en sus negocios y si están próximos a declararse en quiebra; de las demás pueden revelar ocultos secretos tocantes a la legítima propiedad del color del pelo y rostro y si son naturales los lunares que tanto le agracian; son capaces, en fin, de investigar minuciosamente la vida y milagros de todo bicho viviente que cae bajo la jurisdicción de su impertinente curiosidad. (*El Fronterizo*, Essay, 4 October 1891)  
 ‘These pretty exemplars of indiscretion can neither rest nor sleep calmly until it is found out, from good sources, what resources John Doe has at his disposal to live his life; (they) find out how many couples live peacefully and how many fight like cats and dogs; (they) can point out what businesses are going through a rough time and whether or not they are on the verge of bankruptcy; regarding the rest, (they) can reveal hidden secrets related to the legitimacy of their hair and face color, and whether the beauty marks that embellish them are natural or not; (they) are capable of investigating meticulously the life of every living creature that falls under the jurisdiction of their impertinent curiosity’.

In (11), the four main verbs following the postverbal subject *estos bellos ejemplares* have an ambiguous TAM and are also non-reflexive, and despite having numerous clauses with different subject referents intervening between them, they all have null subjects. It is possible, however, that the last three verbs are instances of null subject priming, which was not studied in this paper due to the relative shortness of many texts, particularly advertisements, the most favorable journalistic genre for overt SPP use.

Conversely, in (12), we attest verbs with non-ambiguous TAMs like *discutimos*, *amamos*, and *planteamos* which, in addition, are inflected in the plural, a non-favorable context for overt SPPs, following each an instance of the overt pronoun *nosotros*.

- (12) ¿Habrà alguno de los amigos y camaradas que me escuchan que crean, después de la forma democrata que **nosotros discutimos** nuestros problemas, que diga que **nosotros amamos** ser dictadores? ¿Habrà alguien de los presentes que crea que **nosotros planteamos** conspiraciones o violencias? Mil veces “no”. (*El Tucsonense*, News, 23 June 1936)  
 ‘Is there any one among the friends and comrades listening to me who believes, after the democratic way in which we discuss our problems, who says that we love being dictators? Is there any one among the attendants who believes that we advocate for conspiracies or violence?’

One could argue that it is a case of overt SPP priming, although there is only one clause intervening between the first and second verb and one clause between the second and third verb, as opposed to the multiple intervening clauses that favor priming (Travis 2007; Ramos 2016). Instead, it appears to be a case of stylistic emphasis not commonly found, for instance, in oral sociolinguistic interviews.

Another example of these seeming stylistic licenses can be seen in (13), where the general spoken Spanish patterns are not being followed.

- (13) Antes les dirigió un discurso conmovedor, diciéndoles que los perdonaba, y que moría por Dios y por la patria. Varios se conmovieron, uno de los soldados, vencido por la emoción, arrojó el rifle, diciendo: “Yo no tiro, joven, yo pienso como usted, yo soy católico”. (*El Tucsonense*, Short Story, 22 March 1928)

In (13) we could assume there is a priming effect on the overt variant. However, the probability of overt SPPs priming when there are non-coreferential verbs between the trigger and the target is higher than when there are zero intervening non-coreferential verbs between them (Travis 2007; Ramos 2016), which would make this one an infrequent case of priming.

More to the point, almost the totality of SPE studies revolves around oral data collected through sociolinguistic interviews, with Ramos (2016) being the only exception that we are aware of. Ramos analyzed SPE for the first person singular in three Spanish books covering the Old and Middle Spanish stages: *El Conde Lucanor* (1335), *La Celestina* (1499), and *El Lazarillo de Tormes* (1554). Interestingly, the *yo* rates found in the three texts are not constant:

28%, 15%, and 19%, respectively. Ramos offers no explanation as to why this difference is attested, but he does state that no changing pattern is evident over time: overt SPP use does not increase or decrease consistently as time progresses. It seems, then, that overt SPP rates are in a league of their own in written Spanish. When we compare the *yo* rates of said texts and the one found in the present study, 33.3%, we corroborate that no discernable pattern can be found.

In sum, the results of this section show that overt SPP rates in historical newspapers do not differ substantially from those of contemporary spoken varieties of Arizonan Spanish. In addition, SPE in our data was conditioned essentially by the same factor groups shaping the variable in contemporary spoken varieties of Spanish. However, the offline nature of written data seems to weaken and even neutralize constraints that operate more vigorously in oral, online data. In particular, the effect of the switch reference constraint was considerably weaker than in most studies. Additionally, the constraints of ambiguous TAM and non-reflexive verbs showed no impact on SPE in historical newspapers.

## 8. Conclusions

This paper was motivated by four research questions: (1) How does the overt SPP rate in the 19th and 20th Century written Spanish of Tucson compare to those in contemporary spoken Arizonan Spanish? (2) Do linguistic factors conditioning variable SPE in the 19th and 20th Century written Spanish of Tucson differ from those in contemporary varieties? (3) What social factors favor overt SPPs in the 19th and 20th Century written Spanish of Tucson? and (4) Does the rate of overt SPPs increase over time as bilingualism increases in Tucson?

In answering our first research question, at 24.6% of overt SPPs, the Tucson newspapers' data show a percentage similar to spoken contemporary Tucson Spanish (20.2%) and Phoenix Spanish (17.8%), albeit slightly higher.

As a means to answer our second research question, we used Rbrul to conduct a mixed-methods multiple regression analysis of the internal factors constraining the presence of overt SPPs in our data. Only three out of the six factor groups analyzed turned out to be statistically significant: grammatical person and number, lexeme, and reference. Verb class, reflexivity, and ambiguity of the TAM ending were discarded as non-significant. These results are in line with those found in the vast literature on Spanish SPE: grammatical person/number and reference have been constant in conditioning SPE, whereas reflexivity and ambiguity of TAM have not. Nonetheless, verb class was discarded when included in the same regressions with the random factor of lexeme, which turned out to be significant. Following [Travis and Torres Cacoulios \(2021\)](#), we investigated whether there was a psychological/cognition verb interaction with first person singular tokens only. However, our results show no evidence of such interplay, either because of the written nature of the data or because this relationship was in its early stages and was not, therefore, echoed in written Spanish.

Answering our third and fourth research questions required a regression analysis of the external factor groups hypothesized to influence SPE in our data. Two out of the three factor groups analyzed were statistically significant: journalistic genre and the continuous variable of the year of publication, whilst specific journal was not significant.

In accordance with the language functions predominant in them, advertising, letters, short stories, and editorials were the genres favoring overt pronominal subjects, whereas news, miscellaneous, and essays favored null subjects.

With respect to the continuous factor group of the year of publication, there was a positive correlation between the year the issues were published and the use of overt SPPs: the latter increased over time. This result, in principle, supports the convergence hypothesis by which as contact with English increased, the Spanish of Tucson received indirect interference from it, manifested, in the concrete case of SPE, in an increment of overt SPP rates.



However, as mentioned before, a study of contemporary Tucson Spanish (Anderson 2013) found a lower rate of overt SPPs: 20.2%, which goes against the aforementioned hypothesis. We conjectured that this discrepancy could be due to one of three aspects of the data: an increase of switch reference contexts over time, a surge of the journalistic genres that favor overt SPPs, or simply, the offline nature of the texts analyzed. We performed crosstabulations to investigate whether there was an increase of switch reference contexts entailing an overt SPP spike in the person/number combinations favoring the overt variant, but there was no rate surge despite the context increase. With respect to a possible rise across time of the journalistic genres favoring overt SPPs triggering a surge of the overt variant, crosstabulations showed that only the genre of editorials experienced a rate increase over time. Given these results, the most feasible reason for the slight overt variant increase is the offline nature of written data, which often bypasses online constraints, as attested in the weakening of switch reference and the neutralization of ambiguous TAM and non-reflexive verbs.

In other words: our data do not present evidence that the initial stages of increasing language contact introduced English patterns in the SPE of Tucson Spanish. Instead, the factor groups that condition the variable in contemporary spoken Arizonan Spanish were already operating in the newspaper varieties in the period studied, with the aforementioned characteristics that distinguish written output from oral production.

For further research, it would be ideal to study variable SPE in Mexican newspapers of the same period, such as *El Heraldito Mexicano* (1895–1915), to compare rates, factor group rankings and constraints ranking with the ones arrived at in this paper. In addition, we would like to study contemporary Mexican newspapers to test the hypothesis that written Spanish may allow a higher rate of overt SPPs, by comparing it to its corresponding spoken Spanish rate.

**Author Contributions:** Conceptualization: Á.C.-P.; Data Curation: Á.C.-P., S.L. and R.V.; Formal Analysis: Á.C.-P.; Investigation: Á.C.-P., S.L. and R.V.; Methodology: Á.C.-P., S.L. and R.V.; Project Administration: Á.C.-P.; Resources: S.L. and R.V.; Validation: Á.C.-P., S.L. and R.V.; Visualization: Á.C.-P., S.L. and R.V.; Writing: original draft preparation, Á.C.-P.; Writing, reviewing and editing: Á.C.-P., S.L. and R.V. All authors have read and agreed to the published version of the manuscript.

**Funding:** This research received no external funding.

**Institutional Review Board Statement:** Not Applicable.

**Informed Consent Statement:** Not Applicable.

**Data Availability Statement:** The data used for this study can be found at <https://content.library.arizona.edu/digital/collection/p16127coll7> (*El Fronterizo*) (accessed on 12 December 2022) <https://content.library.arizona.edu/digital/collection/p16127coll3> (*El Tucsonense*) (accessed on 12 December 2022).

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

## Note

<sup>1</sup> Torres Cacoullós and Travis (2019) found 3% of null subjects in English for 1sg and 3sg human subjects.

## References

- Abreu, Laurel. 2012. Subject pronoun expression and priming effects among bilingual speakers of Puerto Rican Spanish. In *Selected Proceedings of the 14th Hispanic Linguistics Symposium*. Edited by Kimberly Geeslin and Manuel Díaz-Campos. Somerville: Cascadilla Proceedings Project, pp. 1–8.
- Alfaraz, Gabriela. 2015. Variation of overt and null subject pronouns in the Spanish of Santo Domingo. In *Subject Pronoun Expression in Spanish: A Cross-Dialectal Perspective*. Edited by Ana Maria Carvalho, Rafael Orozco and Naomi Shin. Washington: Georgetown University Press, pp. 3–16.
- Anderson, Hope. 2013. La influencia de la persona gramatical sobre la expresión del pronombre sujeto en el español del sur de Arizona. *Divergencias: Revista de Estudios Lingüísticos y Literarios* 11: 35–47.
- Bayley, Robert, and Lucinda Pease-Alvarez. 1997. Null pronoun variation in Mexican-descent children's narrative discourse. *Language Variation and Change* 9: 349–71. [CrossRef]

- Bayley, Robert, Norma Cárdenas, Belinda Treviño, and Carlos Vélez. 2012. Spanish Dialect Contact in San Antonio, Texas: An Exploratory Study. In *Selected proceedings of the 14th Hispanic Linguistics Symposium*. Edited by Kimberly Geeslin and Manuel Díaz-Campos. Somersville: Cascadilla Proceedings Project, pp. 48–60.
- Bentivoglio, Paola. 1987. *Los Sujetos Pronominales de Primera Persona en el Habla de Caracas*. Caracas: Universidad Central de Venezuela, Consejo de Desarrollo Científico y Humanístico.
- Bybee, Joan. 2001. *Phonology and Language Use*. Cambridge: Cambridge University Press.
- Bybee, Joan, and Rena Torres Cacoullous. 2008. Phonological and grammatical variation in exemplar models. *Studies in Hispanic and Lusophone Linguistics* 1: 399–414. [CrossRef]
- Cameron, Richard. 1993. Ambiguous agreement, functional compensation, and nonspecific tú in the Spanish of San Juan, Puerto Rico, and Madrid, Spain. *Language Variation and Change* 5: 305–34. [CrossRef]
- Cameron, Richard, and Nydia Flores-Ferrán. 2004. Perseveration of subject expression across regional dialects of Spanish. *Spanish in Context* 1: 41–65. [CrossRef]
- Carvalho, Ana Maria, and Michael Child. 2011. Subject pronoun expression in a variety of Spanish in contact with Portuguese. In *Selected Proceedings of the 5th Workshop on Spanish Sociolinguistics*. Edited by Jim Michnowicz and Robin Dodsworth. Somersville: Cascadilla Proceedings Project, pp. 14–25.
- Carvalho, Ana Maria, Rafael Orozco, and Naomi Shin. 2015. *Subject Pronoun Expression in Spanish: A Cross-Dialectal Perspective*. Washington, DC: Georgetown University Press.
- Cerrón-Palomino, Álvaro. 2014. Ser más pro o menos pro: Variación en la expresión de sujeto pronominal en el castellano limeño. *Lingüística* 30: 61–83.
- Cerrón-Palomino, Álvaro. 2016. (No) English interference on U.S. southwest Spanish? A look at variable subject expression in Phoenix Spanish-English bilinguals. *Sociolinguistic Studies* 10: 383–408. [CrossRef]
- Cerrón-Palomino, Álvaro, and Rosti Vana. 2022. Tracking subject position change in U.S. Southwest Spanish: What 19th and 20th century Tucson newspapers tell us. *Journal of Historical Sociolinguistics* 8: 269–92. [CrossRef]
- De Prada Pérez, Ana. 2015. First person singular subject pronoun expression in Spanish in contact with Catalan. In *Subject Pronoun Expression in Spanish: A Cross-Dialectal Perspective*. Edited by Ana Maria Carvalho, Rafael Orozco and Naomi Shin. Washington, DC: Georgetown University Press, pp. 121–42.
- Del Carpio, Leslie. forthcoming. First-person singular subject pronoun expression of young Spanish speakers from Quito, Ecuador. *Spanish in Context*. in press.
- Dobyns, Henry. 1976. *Spanish Colonial Tucson: A Demographic History*. Tucson: University of Arizona Press.
- Dubord, Elise. 2013. Language, church and state in territorial Arizona. In *A Political History of Spanish: The Making of a Language*. Edited by José del Valle. Cambridge: Cambridge University Press, pp. 260–77.
- Erker, Daniel, and Gregory Guy. 2012. The role of lexical frequency in syntactic variability: Variable subject personal pronoun expression in Spanish. *Language* 88: 526–57. [CrossRef]
- Flores-Ferrán, Nydia. 2010. ¡Tú no me hables! Pronoun expression in conflict narratives. *International Journal of the Sociology of Language* 203: 61–82. [CrossRef]
- Greenleaf, Cameron, and Andrew Wallace. 1962. Tucson: Pueblo, presidio, and American city. A synopsis of its history. *Arizoniana* 3: 18–27.
- Gubitosi, Patricia. 2010. Passive expressions as isogloss between New Mexican and Californian Spanish dialect areas. In *Spanish of the U.S. Southwest: A Language in Transition*. Edited by Susana Rivera-Mills and Daniel Villa. Frankfurt: Iberoamericana/Vervuert, pp. 207–21.
- Holmquist, Jonathan. 2012. Frequency rates and constraints on subject personal pronoun expression: Findings from the Puerto Rican highlands. *Language Variation and Change* 24: 203–20. [CrossRef]
- Jakobson, Roman. 1960. Linguistics and poetics. In *Style in Language*. Edited by Thomas Sebeok. Cambridge: MIT Press, pp. 350–77.
- Johnson, Daniel. E. 2012. *Rbrul* (Version 2.05). Available online: <http://www.danielezrajohnson.com/rbrul.html> (accessed on 15 December 2021).
- Kanellos, Nicolás, and Helvetia Martell. 2000. *Hispanic Periodicals in the United States, Origins to 1960: A Brief History and Comprehensive Bibliography*. Houston: Arte Público Press.
- Labov, William. 1994. *Principles of Linguistic Change, Volume 1: Internal Factors*. Oxford: Blackwell.
- Labov, William. 2001. *Principles of linguistic change Volume 2: Social factors*. Oxford: Blackwell.
- Lapidus, Naomi, and Ricardo Otheguy. 2005. Overt nonspecific Ellos in Spanish in New York. *Spanish in Context* 2: 157–54. [CrossRef]
- Lastra, Yolanda, and Pedro Martín Butragueño. 2015. Subject pronoun expression in oral Mexican Spanish. In *Subject Pronoun Expression in Spanish: A Cross-Dialectal Perspective*. Edited by Ana Maria Carvalho, Rafael Orozco and Naomi Shin. Washington, DC: Georgetown University Press, pp. 39–58.
- Mas I Miralles, Antoni. 2003. Historical sociolinguistics: An alternative to the analysis of linguistic change. *Noves SL. Revista de Sociolingüística* 3: 1–6.
- Michnowicz, Jim. 2015. Subject pronoun expression in Yucatan Spanish. In *Subject Pronoun Expression in Spanish: A Cross-Dialectal Perspective*. Edited by Ana Maria Carvalho, Rafael Orozco and Naomi Shin. Washington, DC: Georgetown University Press, pp. 101–20.
- Officer, James. 1987. *Hispanic Arizona 1536–1856*. Tucson: University of Arizona Press.

- Orozco, Rafael. 2016. Subject pronoun expression in Mexican Spanish: ¿Qué pasa en Xalapa? *Proceedings of the Linguistic Society of America* 1: 1–15. [\[CrossRef\]](#)
- 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 4th Workshop on Spanish Sociolinguistics*. Edited by Maurice Westmoreland and Juan Antonio Thomas. Somersville: Cascadilla Proceedings Project, pp. 70–80.
- Orozco, Rafael, and Johnny Orozco. 2022. The effect of the verb on pronominal expression: A reanalysis. *Proceedings of the Linguistic Society of America* 7: 5286. [\[CrossRef\]](#)
- Orozco, Rafael, and Luz Marcela Hurtado. 2020. A variationist study of subject pronoun expression in Medellín, Colombia. *Languages* 6: 5. [\[CrossRef\]](#)
- Otheguy, Ricardo, Ana Celia Zentella, and David Livert. 2007. Language contact in Spanish in New York: Toward the formation of a speech community. *Language* 83: 770–802. [\[CrossRef\]](#)
- Otheguy, Ricardo, and Ana Celia Zentella. 2012. *Spanish in New York. Language Contact, Dialectal Leveling, and Structural Continuity*. Oxford: Oxford University Press.
- Padilla, Lille. 2020. Subject Pronoun Expression in an L2-Only Environment: The Case of Equatorial Guinea. Ph.D. dissertation, Arizona State University, Tempe, AZ, USA.
- Padilla, Lillie. 2021. First person singular subject pronoun expression in Equatoguinean Spanish. *Journal of Monolingual and Bilingual Speech* 3: 171–94. [\[CrossRef\]](#)
- Padilla, Lillie. forthcoming. Subject Pronoun Expression in Spanish in Contact with Fang in Equatorial Guinea. *Onomazein*. in press.
- Pierrehumbert, Janet. 2001. Exemplar dynamics: Word frequency, lenition, and contrast. In *Frequency and the Emergence of Linguistic Structure*. Edited by Joan Bybee and Paul Hopper. Amsterdam: John Benjamins, pp. 137–58.
- Posio, Pekka. 2011. Spanish subject pronoun usage and verb semantics revisited: First and second person singular subject pronouns and focusing of attention in spoken Peninsular Spanish. *Journal of Pragmatics* 43: 777–98. [\[CrossRef\]](#)
- Ramos, Miguel. 2016. Continuity and change. First person singular subject pronoun expression in earlier Spanish. *Spanish in Context* 13: 103–27. [\[CrossRef\]](#)
- Ranson, Diana. 1991. Person marking in the wake of /s/ deletion in Andalusian Spanish. *Language Variation and Change* 3: 133–52. [\[CrossRef\]](#)
- Samper Padilla, José Antonio, Clara Eugenia Hernández Cabrera, and Magnolia Troya Déniz. 2006. Presencia/ausencia del sujeto pronominal de primera persona singular en la norma culta de España. *El español en América. Diatopía, Diacronía e Historia*, 87–109.
- Schneider, Edgar. 2017. Written data sources. In *Data Collection in Sociolinguistics*. Edited by Christine Mallinson, Becky Childs and Gerard Van Herk. London: Routledge, pp. 169–78.
- Sheridan, Thomas. 1986. *Los Tucsonenses: The Mexican Community in Tucson, 1854–1941*. Tucson: University of Arizona Press.
- Sheridan, Thomas. 1995. *Arizona. A History*. Tucson: University of Arizona Press.
- Silva-Corvalán, Carmen. 1982. Subject expression and placement in Mexican-American Spanish. In *Spanish in the United States: Sociolinguistic Aspects*. Edited by Jon Amastae and Lucía Elías-Olivares. Cambridge: Cambridge University Press, pp. 93–120.
- Silva-Corvalán, Carmen. 1994. *Language Contact and Change: Spanish in Los Angeles*. Oxford: Oxford University Press.
- Torres Cacoullos, Rena, and Catherine Travis. 2010. Variable yo expression in New Mexico: English influence? In *Spanish of the U.S. Southwest: A Language in Transition*. Edited by Susana Rivera-Mills and Daniel Villa. Frankfurt: Iberoamericana/Vervuert, pp. 185–206.
- Torres Cacoullos, Rena, and Catherine Travis. 2019. Variationist typology: Shared probabilistic constraints across (non-) null subject languages. *Linguistics* 57: 653–92. [\[CrossRef\]](#)
- Travis, Catherine. 2005. The yo-yo effect. In *Theoretical and Experimental Approaches to Romance Linguistics: Selected Papers from the 34th Linguistic Symposium on Romance Languages*. Edited by Randall Gess and Edward J. Rubin. Amsterdam: John Benjamins, pp. 329–49.
- Travis, Catherine. 2007. Genre effects on subject expression in Spanish: Priming in narrative and conversation. *Language Variation and Change* 19: 101–36. [\[CrossRef\]](#)
- Travis, Catherine, and Rena Torres Cacoullos. 2021. Categories and frequency: Cognition verbs in Spanish subject expression. *Languages* 6: 126. [\[CrossRef\]](#)

**Disclaimer/Publisher's Note:** The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.