

## Feminine morphology in possessive complements of adverbial constructions in Andalusian varieties

Matti Marttinen Larsson<sup>1</sup>; Miriam Bouzouita<sup>2</sup>

Recibido: 27 de octubre de 2021 / Aceptado: 17 de febrero de 2022

**Abstract.** The present study scrutinizes the Spanish morphological variation between masculine and feminine possessive complements used with locative adverbials (e.g., *cerca mío / mía* ‘close to me’). The feminine *-a* suffix (e.g., *mía*) is the innovative variant and has been described by normative grammars as a low-frequency and stigmatized trait. This paper focuses on Andalusian Spanish, where the innovative *-a*-suffix has one of the highest frequencies of use. Twitter data from the eight provinces of Andalusia are compiled in order to determine the innovation’s diatopic distribution and linguistic diffusion. The quantitative results show that the *-a* variant predominates in all Andalusia and that the change has spread through all the adverbial locative contexts. We propose that a series of analogical extensions have taken place whereby the feminine possessive complement spreads from one adverbial context to another, creating an analogical snowballing effect.

**Keywords:** possessive pronouns, Andalusian Spanish, analogy, locative adverbial constructions

### [es] La morfología femenina en los complementos posesivos de las construcciones adverbiales en las variedades andaluzas

**Resumen.** El presente estudio indaga la variación morfológica entre los complementos posesivos masculinos y femeninos que se usan con adverbios locativos (p.ej., *cerca mío / mía*). El sufijo femenino con *-a* (p.ej., *mía*) constituye una variante innovadora que las gramáticas normativas califican de una variante estigmatizada de baja frecuencia. Este artículo se centra en el español de Andalucía donde se han documentado los usos más elevados de la variante innovadora. Se recopilan datos de Twitter de las ocho capitales andaluzas para precisar la distribución diatópica de la innovación así como su difusión en diferentes contextos lingüísticos. Los resultados cuantitativos demuestran que la variante femenina con *-a* predomina en toda Andalucía y que se ha completado el cambio en todos los contextos adverbiales locativos. Proponemos que se ha producido una serie de extensiones analógicas a través de la que el complemento posesivo femenino se va difundiendo de un contexto adverbial a otro, engendrando así un efecto analógico de bola de nieve.

**Palabras clave:** pronombres posesivos, el español andaluz, analogía, construcciones adverbiales locativas

**Cómo citar:** Marttinen Larsson, M. & Bouzouita, M. (2022). Feminine morphology in possessive complements of adverbial constructions in Andalusian varieties. *Círculo de Lingüística Aplicada a la Comunicación* 91, 157-171.

**Índice.** 1. Introduction. 2. Method. 3. Analysis. 4. Conclusions. References.

## 1. Introduction

In essentially all varieties of Spanish, there is morphosyntactic variation in the different person-referential complements used with the locative adverbials (e.g., *in front of him*). These complements present two variants: one prepositional (LOCATIVE ADVERB + <sup>complement</sup>[preposition *de* + PERSONAL PRONOUN]) and one possessive (LOCATIVE ADVERB + <sup>complement</sup>[POSSESSIVE PRONOUN]). Following grammars such as RAE & ASALE (2009), the correct construal would be with the prepositional complement, e.g., *encima de mí* (‘on top of me’), *debajo de ella* (‘underneath her’), *detrás de ti* (‘behind you’), and so on. The possessive construction yields variants such as *encima mío* (lit. ‘on top mine’), *debajo suyo* (lit. ‘underneath hers/his/theirs/yours’), *detrás tuyo* (lit.

<sup>1</sup> Stockholm University.

Correo electrónico: [matti.marttinen@su.se](mailto:matti.marttinen@su.se)

ORCID: <https://orcid.org/0000-0002-6224-7872>

<sup>2</sup> Humboldt-Universität zu Berlin.

Correo electrónico: [miriam.bouzouita@hu-berlin.de](mailto:miriam.bouzouita@hu-berlin.de)

ORCID: <https://orcid.org/0000-0001-8196-620X>

‘behind yours’), etc. This is deemed grammatically incorrect since possessive pronouns are said to modify noun phrases only and adverbs are traditionally considered to be invariable grammatical categories (In recent years, however, the latter view is being challenged by various empirical studies that show that certain speakers do indeed agree adverbs with other elements in the clause (e.g., Felú Arquiola & Pato, 2019, 2020)). As such, possessive pronouns should not modify locative adverbials. The only exception to this are the denominal adverbial locatives *al lado* (‘next to’) and *alrededor* (‘around’) which, due to their nominal status (*lado* ‘side’ and *alrededor* ‘surrounding area’), are allowed to combine with possessives, as in (2b). However, nowadays the possessive complement is also found with non-denominal adverbial locatives, such as *cerca* ‘near’, *debajo* ‘under’ or *detrás* ‘behind’, as in (2c), which do not have a nominal status in Spanish. Consequently, it appears that a reanalysis of the possessive pronoun took place due to the existence of similar constructions found in the nominal domain. To illustrate this syntactic parallelism between the nominal and adverbial domain in more detail, consider the examples in (1) and (2).

- (1) a. *el libro de Juan / María*  
 the book of Juan / María  
 ‘Juan’s/María’s book’
- b. *su libro – el libro de él / ella – el libro suyo*  
 his/her book the book of he she the book his/her.MASC  
 ‘his/her book’
- (2) a. *al lado de Juan / María – delante de Juan / María*  
 to-the side of Juan María in-front of Juan María  
 ‘next to Juan/María – in front of/before Juan/María’
- b. *a su lado – al lado de él / ella – al lado suyo*  
 to his/her side to-the side of he /she to-the side his/her.MASC  
 ‘next to him/her’
- c. *por/en su delante – delante de él / ella – delante suyo*  
 for/in his/her in-front in-front of he / she in-front his/her.MASC  
 ‘in front of/before him/her’
- d. *al lado suya – delante suya*  
 to-the side his/her.FEM in-front his/her.FEM  
 ‘next to him/her – in front of him/her’

As can be seen in (1b), Spanish exhibits so-called *Differential Possessor Expression* (cf. O’Connor, 2003; DPE henceforth) as there are various ways to express possession within the nominal domain. Concretely, Spanish has a tripartite DPE system, in which (i) prenominal possessives can be used, as in *su libro*, as well as (ii) their tonic postnominal counterparts, as in *el libro suyo*, or (iii) the prepositional construction, *el libro de él*, in which the preposition *de* is followed by a tonic personal pronoun. Observe, however, that certain varieties use the prepositional construction more frequently than others and that it is hard (but not impossible) to find it for certain grammatical persons, such as the 1<sup>st</sup> and 2<sup>nd</sup> person singular (see Bouzouita, 2022 for an overview of DPE in Spanish varieties). Diachronic results from other studies indicate that locative adverbials with nominal bases, such as *al lado* ‘next’, *alrededor* ‘around’ and *en medio* ‘in the middle/centre’, were the earliest in adopting the innovative tonic possessive variant, as illustrated with *al lado suyo* in (2b) (for Spanish, see Octavio de Toledo y Huerta, 2016, p. 114, pp. 217-219; Marttinen Larsson, in press; Marttinen Larsson & Álvarez López, 2022; for Galician, see Silva Domínguez, 2020). Note how there is masculine agreement between *lado* and the possessive *suyo* (marked with the coindexing in underscript), as also found in the nominal domain, as exemplified with *libro* and *suyo* in (1b). From the denominal adverbials, the tonic possessives extend their use to contexts with locative adverbials that do not contain a nominal base, as in *delante suyo* in (2c). Although the variation between the masculine tonic possessive and the prepositional complement is found in virtually all varieties of Spanish (Santana Marrero, 2014; Marttinen Larsson & Bouzouita, 2018), the highest frequency of use of the possessive variant can be found in River Plate Spanish (RAE & ASALE, 2009; Santana Marrero, 2014; Marttinen Larsson & Álvarez López, 2017; Marttinen Larsson & Bouzouita, 2018; Hoff, 2020). Although the examples in (1) and (2) contain possessive constructions for the 3<sup>rd</sup> person, we are not implying that the change started with this grammatical person. In fact, diachronic data indicates that the spread began with the 1<sup>st</sup> person, spreading subsequently to the 2<sup>nd</sup> person and lastly to the 3<sup>rd</sup> (Marttinen Larsson, in press). Synchronically, the possessive construction is favored especially for the 1<sup>st</sup> person (Marttinen Larsson & Bouzouita, 2018; Hoff, 2020; Marttinen Larsson, in press, 2022).

The analogical extension of the use of tonic possessives to non-denominal adverbial contexts, such as *delante* (2c), and the subsequent variation between the tonic possessive and the prepositional complement with non-denominal adverbials, is brought on by parallel patrons in other syntactic environments: to wit, (i) the alternation between the postnominal possessive and the prepositional constructions in the nominal domain, as in (1b), and (ii) the subsequent variation between the tonic possessive and prepositional complements in the

denominal adverbials, illustrated in (1b) and (2b) respectively. In other words, due to various similarity-based generalizations between the nominal and adverbial systems, i.e., analogy (De Smet, 2012; Bertolotti, 2017, p. 343), the tonic possessives spread to a new syntactic environment, to wit, the non-denominal adverbial ones. These two analogical extensions of the tonic possessive to novel contexts are visualized in Table 1, in which possessives are firstly used in the nominal domain, then in denominal locative adverbial constructions before spreading to non-denominal ones. (For reasons of completeness, we also added in (2c) *por/en su delante*, which illustrates a similar change for the non-tonic possessives, although these structures fall outside of the scope of this paper. It should be noted though that the geographic distribution of this change is much more limited than the one for their tonic counterparts: it has been documented for varieties spoken in Peru and Bolivia (RAE & ASALE 2009: 1360)).

Syntactic context	Prepositional complement	Tonic possessive complement
NOMINAL DOMAIN	ARTICLE + NOUN + <i>de</i> + PERS. PRON. <i>El libro de él/ella</i>	ARTICLE + NOUN + POSS. PRON. <i>El libro suyo</i>
ADVERBIAL DOMAIN	DENOM. LOC. ADV. + <i>de</i> + PERS. PRON. <i>Al lado de él/ella</i>	DENOM. LOC. ADV. + POSS. PRON. <i>Al lado suyo</i>
	NON-DENOM. LOC. ADV. + <i>de</i> + PERS. PRON. <i>Delante de él/ella</i>	NON-DENOM. LOC. ADV. + POSS. PRON. <i>Delante suyo</i>

Table 1. Analogical extensions of tonic possessives from NPs (or DPs) to AdvPs

It should not be forgotten that the description of the change given in Table 1 only deals with the syntactic side of the change and, as such, abstracts away from the synchronic morphological variation that exists in the tonic possessive forms. Indeed, in certain varieties, such as Andalusian Spanish, feminine possessive complements – illustrated in (2d) with *al lado suya* and *delante suya* – are also attested and even predominate overwhelmingly (see section 3; for the northern varieties of Spain, the feminine variant prevails in the Spanish spoken in Galicia, see Bouzouita et al., 2021, in press). It is this morphological variation in the possessive adverbial complements that is the object of study of this article. Interestingly, this microvariation is essentially only found in European Spanish (Marttinen Larsson & Álvarez López, 2017; Salgado & Bouzouita, 2017; Hoff, 2020). Note that, in the present paper, we refer to the possessives with *-o* suffixes as masculine, whereas those with *-a* as feminine possessives. Notwithstanding the previous, the use of this terminology does not entail any claims regarding gender agreement between the grammatical gender marking of the possessive and the gender of the alluded referent, as is done in the agreement hypothesis posited by Kany (1969, p. 66) and Zamora Vicente (1967, p. 433).

The masculine *-o* suffix of the possessive is perceived to constitute the unmarked and most common variant (Alcina Franch & Blecua, 1975, p. 620; Almela Pérez, 1991, p. 438; González Calvo, 2006, p. 67; RAE & ASALE, 2009, § 18.4o; Santana Marrero, 2014, p. 17). As the following paragraphs outline, scholars have proposed a number of hypotheses concerning the driving forces behind this microvariation, the most recurrently cited ones being analogy with the final vowel of the preceding adverb (Alcina Franch & Blecua, 1975, pp. 619-620; Fernández Ramírez, 1987, p. 94) and pseudo-agreement with the gender of the alluded referent (Zamora Vicente, 1967, p. 433; Kany, 1969, p. 66).

Concerning the analogy hypothesis, it has been suggested that the motivation behind the variation in possessive suffix morphology lies in the replication of the final vowel of the preceding locative adverbial (Alcina Franch & Blecua, 1975, pp. 619-620; Fernández Ramírez, 1987, p. 94). The vast majority of Spanish locative adverbs end in */-o/* or */-a/*, which coincide with the possible suffixes of the possessives; in that sense, adverbials that end in */-o/* would yield a corresponding *-o* suffixed possessive (3a), whereas those with */-a/* would generate the analogical *-a*-suffixed possessives (3b). There also exists a third possibility as the final vowel of locatives can also be */-e/*, as shown in (3c). However, what this adverbial ending would yield as a result in the possessive morphology remains unaddressed by this hypothesis, represented in the variation in (3c).

- (3) a. *debajo*            *mío*  
underneath        mine.MASC  
'beneath me'
- b. *encima mía*  
above mine.FEM  
'above me'
- c. *enfrente mío*    –    *enfrente*        *mía*  
in-front mine.MASC    in-front        mine.FEM  
'in front of me'

Another hypothesis concerning the motivation behind the suffix variation of the possessive posits that it conveys agreement with the gender of the alluded referent (Zamora Vicente, 1967, p. 433; Kany, 1969, p. 66). Given Spanish gender marking rules, this would in practice mean that *-o*-suffixed possessives correspond to masculine referents, as in illustrated in (4a), whereas *-a*-suffixed ones are exhibiting agreement with feminine referents, as in (4b).

- (4) a. *El hijueputa que está al frente mío se cagó [...]*  
 The son-of-whore that is at-the front mine.MASC CL shat  
 ‘The son of a bitch in front of me shat himself.’ (Twitter, male, San Sebastián, 23/04/2018)
- b. *Martita se ha traído una botella igual que la que tien*  
 Martita.FEM CL has brought a bottle same that it that has  
*el chaval que está sentando en frente suya [...]*  
 the guy that is sitting in front his/her.FEM  
 ‘Martita brought the same kind of bottle [of drinks] as the guy has that is sitting in front of her.’  
 (Twitter, female, Madrid, 24/04/2018)

Further, the *-o* possessive variant started to spread significantly from the early 1900s onwards (Marttinen Larsson, in press; Marttinen Larsson & Álvarez López, 2017, 2022), whereas the *-a* variant started to diffuse in (written) European Spanish by around 1950 (Marttinen Larsson & Álvarez López, 2017). Synchronically, it also has a noticeable presence; to illustrate this claim, Hoff’s (2020) examination of Twitter data from Madrid shows that, among the documented possessive cases, 62% (1611/2591) are of the feminine *-a* variant. Salgado & Bouzouita (2017) show that this microvariation exists in oral corpus data from Valencia, Madrid and Andalusia and find that, excluding the nominal masculine locatives *al lado* ‘next’ and *alrededor* ‘around’, the *-a* variant constitutes approximately 85% (44/52) of the gathered possessive forms. What is more, focusing solely on the Andalusian data, the authors find that close to 97% (28/29) of the possessives exhibit the feminine morphology. These findings strongly contradict RAE & ASALE’s (2009, § 18.4o) postulation of the *-a* variant constitutes a rarer variant in comparison to its masculine *-o* counterpart. The fact that politicians, such as the president of the Autonomous Community of Andalusia when addressing fellow politician Pedro Sánchez in a televised debate, use the feminine possessive complement, as in *cerca tuya* ‘near you’ in (5), serves as anecdotal evidence that this feminine variant appears not to be rare at all but rather widely used (cf. also the *El País* language column in which this matter is discussed; Grijelmo, 2017, May 22). As we will see in Section 3, the current study will corroborate this observation with sound quantitative results.

- (5) *Cuando la gente que ha trabajado cerca tuya resulta que*  
 when the people that has worked near yours.FEM results that  
*no se fia de ti, deberías hacértelo ver*  
 not CL trusts of you you-have do-CL-CL see  
 ‘When it turns out that people near you don’t trust you, you should realize it.’ (Susana Díaz, 15/05/2017)

With respect to the proposed pseudo-agreement with the gender of the referent, this hypothesis is not borne out neither by the data examined by Salgado & Bouzouita (2017, p. 784) nor by Hoff (2020, p. 71). Rather, it appears that the *-a* variant is used both with feminine referents and masculine ones, as well as those of unspecified gender (e.g., collective entities and impersonal or generic subject *tú* ‘you/one’). As regards the analogy hypothesis, neither Salgado & Bouzouita (2017, p. 785) nor Hoff (2020, p. 73) find any synchronic patterns that support this hypothesis. Instead, the *-a* variant is used with all types of locatives and is thus not restricted to locatives ending in *-a*. In order to account for this, Hoff (2020, p. 73) proposes that the variation began as an analogical process (*/a/-*ending locatives yielding *-a* suffixed possessives) and has subsequently spread to other locatives not ending in *-a*. This, nonetheless, needs adequate diachronic corroboration.

Summarizing, the *-a* variant is frequently employed in various communities of Spain and appears to have the highest frequency of use in Andalusia. So far, however, no clear patterns have emerged that help explain the motivation behind the generalization of the *-a* variant. An additional gap lies in the geography of the microvariation within Andalusia, since so far only diatopically aggregated and scarce datasets have been analyzed (cf. Salgado & Bouzouita, 2017). Considering the above, the present paper will thus try to answer the following research questions:

1. Is the *-a* variant the preferred option in Andalusian Spanish (as found in Salgado & Bouzouita, 2017)?
  2. Can diatopic variation between the *-o* and *-a* suffixes of the possessive be found within Andalusia?
  3. What mechanisms, if any, account synchronically for the use of the *-a* variant in Andalusian Spanish?
- In order to address these questions, the present study compiles 1,921 tweets stemming from the capitals of the eight Andalusian provinces. A series of predictors are coded for in an effort to determine to what extent the studied variation is subject to constraints and the variation’s geographical distribution. In doing this, this



research contributes to a greater understanding of the functioning of this phenomenon and the driving forces that have led to a greater grammaticalization of possessive pronouns in Spanish.

The paper is organized as follows. Section 2 outlines the employed method for the data collection and specifies the operationalization of the above-presented research questions. Section 3 presents the diatopic distributional analysis, as well as an examination of the influences of the coded intralinguistic predictors. Section 4 presents the overall conclusions and offers a completely new hypothesis concerning the mechanisms underlying the generalization process of the feminine *-a* possessive variant.

## 2. Method

Following the general trend within Humanities, social media are increasingly being used as corpora within the field of linguistics. This is in part due to the user metadata they offer, such as geolocation, gender and social networks. Data from social media also facilitate access to low-frequency tokens and, at the same time, it is a cheap and time-saving way of data gathering. For the present study, Twitter was used as a source for data collection. The main advantage of using Twitter for this type of diatopically fine-grained variationist study is that it allows for quantitative corpora construction of a variable that is rather infrequently found in traditional oral or written corpora, as is the case for the morphological possessive variation that this study deals with (cf. Marttinen Larsson & Álvarez López, 2017; Salgado & Bouzouita, 2017). Tapping into the Twitter API, we are able to compile original observations of colloquial synchronic language use in the specific areas that the present paper is concerned with. Ethical vetting has been applied for by the first author and approved by the Swedish Ethical Review Authority (record number 2019-01271).

A geospecified script was written using the R package *rtweet* (Kearney, 2016). The geocoordinates corresponding to the respective capitals of the eight Andalusian provinces were entered into the code with an additional geographical reach of a radius of 25 kilometers. The API searches were centered on period of 10 days and were gathered iteratively between February and September of 2018.

The search strings of the linguistic variable that the script searched for using the Twitter API is (near) exhaustive and includes the following locative adverbials with diminutive and colloquial variants, including possible orthographic errors and common phonetic elisions that are at times transferred into written language, as such casting out a large data net:

- *al lado* ‘next to’ (+ *al lao, alado, alao, a lado, a lao, a ladito, al ladito*)
- *alrededor* ‘around’ (+ *arededor, aredor, alrebedo, alrededor, alderredor, alderredo, alredo, derredor, derredo, rededor, rededo, redor, redo, arededo, aredo, alredecorito, alderredorito, alredecorito, derredorito, rededorito, redorito, aredorito, aredorito, alreredito, alderredito, alreredito, derrito, redito, aredito, alededol, aledol, deledol*)
- *arriba* ‘above’ (+ *riba, ariba, arriba, ariva, riva, arribita, aribita, ribita, arivita, arrivita, rivita*)
- *encima* ‘above’ (+ *ensima, en cima, cima, sima, en sima, encimita, cimita, ensimita, simita, en simita*)
- *debajo* ‘below’ (+ *abajo, bajo, dbajo, abaho, baho, debaho, dbhaho, abajito, bajito, debajito, dbajito, abahito, bahito, debahito, dbahito*)
- *cerca* ‘close to’ (+ *cerka, acerca, acerka, cerqua, acerqua, cerquita, cerkita, acerquita, acerkita, cerquito, cerkito, acerquito, acerkito*)
- *delante* ‘in front of’ (+ *adelante, ante, dlante, lante, delant, dlant, adlante, adelant, adlant, alante, alant, ant, lant, delantito, delantita, dlantito, dlantita, adelantito, adelantita, adlantito, adlantita, alantito, alantita, antita, antito, lantito, lantita*)
- *dentro* ‘inside’ (*adentro, dentrito, adentrito*)
- *detrás* ‘behind’ (*atrás, detras, detra, detrás, tras, tra, atrás, atra, detrasito, detrasita, detracito, detracita, atrasito, atrasita, atracito, atracita, trasito, trasita, tracito, tracita, patras, patra, p’atrás, p’atra, patracita, patrasita, p’atracita, p’atrasita*)
- *enfrente* ‘in front of’ (*al frente, frente, frent, enfrent, enfrentito, enfrentita, frentito, frentita*)
- *en lo alto* ‘above’ (*por lo alto, en lo altito, por lo altito, en alto, en altito, en altita, por alto, por altito, en lo altito, en alta, en la alta, por alta, por la altita, en altita, por altita*)
- *en torno* ‘around’ (*en tornno, en tornito*)
- *en medio* ‘in the middle’
- *fuera* ‘outside’ (*afuera, fuerita, afuerita*)
- *junto* ‘next to’ (*hunto, juntito, juntita, huntito, huntita*)
- *lejos* ‘far from’ (*lejo, lehos, leho, lejito, lejitos, lehitos, lehitos*)
- *a la derecha* ‘on the right’ (*a la derechita*)
- *a la izquierda* ‘on the left’ (*a la izquierdita*)

The variants of the dependent variable, i.e., the possessive complement following the locative adverbial, were:

- *-o* suffixes: *mío, míos, tuyo, tuyos, suyo, suyos, nuestro, nuestros, vuestro, vuestros*
- *-a* suffixes: *mía, mías, tuya, tuyas, suya, suyas, nuestra, nuestras, vuestra, vuestras*

We included the plural forms of the possessive in spite of them not being documented by earlier studies as participating in the variation. This was done so that possibly misspelled occurrences would also be included. Additionally, the following predictors were coded for:

- **Province:** includes the eight Andalusian capitals, to wit Almería, Cádiz, Córdoba, Granada, Huelva, Jaén, Málaga, Seville.
- **User:** in order to account for idiosyncratic variation in the regression model, the screen name of the user was included as an identifier (*N* of users included in sample = 1,444).
- **Gender of referent:** each tweet, whenever possible, was examined manually in order to determine the gender of the referent. This information could be derived either from the tweet itself, such as in example (4b), or by user mentions to specific profiles, which were each inspected. The gender of the referent was interpreted to have three levels: male, female and generic (non-identifiable accounts were N/A coded). ‘Generic’ refers to collective entities or impersonal/generic *tú*-constructions, such as in examples (6a)-(6b):

- (6) a. *Lo peor es la inseguridad... el sentir a alguien muy cerca de tu espalda en el bus y dar gracias a dios que es una abuela. Asegurarte de que la que camina **detrás tuya** es otra chica... #cuéntalo*  
 ‘The worst part is the insecurity... the feeling of someone close to your back on the bus and thanking God that it is an old lady. Making sure that the person walking behind you is another girl... #cuéntalo.’ (Seville, 27/04/2018)
- b. *Cuando estás tomándote el café, tranquila en una terraza. Y un camarero le grita al otro **cerca tuya** “QUE ES PA ‘2, NO PA ‘4” [...]*  
 ‘When you’re drinking a coffee on a terrace and a waiter yells to the other waiter that’s close to you “IT’S FOR N.2, NOT FOR N.4” [...].’ (Huelva, 25/04/2018)

- **Locative adverb:** in order to account for residual variation attributed to the individual locative adverbials, the locative was included as a random effect in the regression model (cf. Gries & Hilpert, 2010, p. 305).
- **Grouped locative adverbials:** this predictor is also item-concerned, but collapses variants of the same adverbial into a grouped adverb variable. In that sense, *debajo* contains occurrences of both *debajo* as well as *bajo, abajo*, etc., which are considered variants of *debajo*. This predictor was constructed so that it could be compared with ‘locative adverb’ in terms of fit during the model selection process.
- **Vowel of adverbial’s final syllable:** in an attempt to operationalize the hypothesized analogy between the final vowel of the locative adverb and the suffix of the possessive, different options were available as to how to code this. Since the hypothesis proposed by Alcina Franch & Bleca (1975, pp. 619-620) suggests that adverbs ending in /-o/ generate the *-o* variant of the possessive, while those ending in /-a/ yield the feminine *-a* variant, it remains unaddressed what the hypothesis predicts for other possible endings of the locatives, viz. blocked-syllable endings (such as in *alrededor, detrás/atrás/tras* and *lejos*). Similarly, it remains unclear what /-e/-ending locatives (such as *enfrente* and *delante*) would yield as an analogical result. In the present study, we interpret Alcina Franch & Bleca’s (1975, pp. 619-620) hypothesis as limited to locatives with open-syllable endings (e.g., *cerca mía, debajo suyo*, etc.) but as extendible to include closed-syllable endings (e.g., *detrás mía, alrededor suyo*). We implement this interpretation for two reasons: firstly, we interpret these linguists’ theory (1975, pp. 619-620) as one in which the phonological vowel-value present in the last syllable plays a pivotal role and not the final phoneme of the preceding adverbial. Secondly, in Andalusian Spanish, word-final consonants, especially /s/ but also /r/, are frequently aspirated, debilitated or completely elided (Llorente Maldonado de Guevara, 1962, p. 228; Narbona Jiménez et al., 2011, p. 23; Fernández-Ordóñez, 2016), which means that, in practice, all these adverbials could potentially be open-syllable endings. Obviously, this is more so for oral than written language; yet, it is plausible to assume that oftentimes changes that take place in spoken language and the mechanisms underlying them are transferred into written modes (see Marttinen Larsson, in press, for an illustrative case study). In that sense, the distinction between for example *cerca* and *detrás* might be interpreted as purely orthographic since, in the spoken word, these can be both pronounced as open-syllable ending adverbials, to wit when the final /-s/ is elided. Considering all of this, the factor concerning the vowel of adverbial’s final syllable contains three levels: /-o/ (*al lado, dentro, alrededor, lejos, debajo, abajo, en torno, en/por lo alto*); /-a/ (*encima, arriba, detrás, atrás, tras, cerca*); and /-e/ (*enfrente, delante, adelante, ante*).

- **Grammatical number:** the number of the possessive is included since earlier studies have found that this factor influences the macrovariation between prepositional and possessive complements (examples (2) and (3); Salgado & Bouzouita, 2017; Marttinen Larsson & Bouzouita, 2018; Hoff, 2020; Marttinen Larsson, in press; Marttinen Larsson & Álvarez López, 2022). It contains two levels, these being singular (*mío/a*, *tuyo/a* and *suyo/a*) and plural (*nuestro/a*, *vuestro/a*, *suyo/a*). Since *suyo* is ambiguous in terms of number and can thus refer both to singular and plural referents, each observation was inspected manually. As indicated above, the search strings originally also included morphological plural variants of the possessives (e.g., *míos/as*, etc.), but the very few tokens that were compiled of such variants (very likely orthographical errors) have been grouped with their respective singular counterparts.
- **Grammatical person:** this factor is also known to influence the microvariation between prepositional and possessive complements significantly (cf. Salgado & Bouzouita, 2017; Marttinen Larsson & Bouzouita, 2018; Hoff, 2020; Marttinen Larsson, in press; Marttinen Larsson & Álvarez López, 2022, among others) and includes the following levels: 1<sup>st</sup> person (*mío*, *nuestro*), 2<sup>nd</sup> person (*tuyo*, *vuestro*) and 3<sup>rd</sup> person (*suyo*). Deferential uses are coded as 3<sup>rd</sup> person.

During the codification process, the following types of occurrences were removed from the analysis: duplicates, quotes (identified by the use of quotation marks or “name dropping” the source of a given tweet), retweets, lyrics, false instances of the variable (e.g., *bajo nuestro control* ‘under our control’, *tras vuestro evento* ‘after your event’, etc.), tweets with obvious non-European Spanish dialectal traits, such as the use of *voseo* (typical of River Plate Spanish and some other Latin American varieties of Spanish, which might have appeared due to modified VPN tunnel where overseas users “change” their device’s location in order to access shows available on other countries’ streaming services, e.g. Spain’s Netflix), and metalinguistic comments on the use of the studied linguistic variants, such as in examples (7a) and (7b):

- (7) a. @javier\_mrquez @valentingarcia2 @RadioSevilla Ahora es un buen momento para dejar de decir “al lado mía” “al lado suya”, etc....saludos  
 ‘@javier\_mrquez @valentingarcia2 @RadioSevilla This would be a good moment to stop saying *al lado mía*, *al lado suya*, etc....greetings.’ (Seville, 13/08/2018)
- b. @Renfestal Por aquí es muuy común decir “delante mio/suyo/suya”, seguramente se me escape xDD  
 ‘@Renfestal Here it is suuper common to say *delante mio/suyo/suya*, I’m sure I have used it xDD.’ (Seville, 13/08/2018)

While tweets such as these are interesting in themselves due to the sociolinguistic perceptions of speakers on the use of the possessive variants, we cannot retain them in our analysis due to their metalinguistic status, i.e., they do represent original linguistic uses.

For the descriptive parts of the analysis, a final dataset containing 1,921 observations is used. However, since this dataset has empty cells for some variables (e.g., ‘gender of referent’, where it was not always possible to determine), a reduced dataset containing no empty cells was used for the inferential statistical analysis in order to aid the model fit process. This dataset with only complete observations consists of 1,675 observations. In what follows, the different analyses conducted on these datasets will be presented.

### 3. Analysis

As concerns the analysis of the variation, it comprises two parts: firstly, a distributional diatopic analysis will be presented to shed light on the geographic frequency variability in the use of the feminine *-a* variant. Secondly, we turn to inferential statistical measurements in order to assess the effects of the included predictor variables described in Section 2.

#### 3.1. Diatopic distribution

For the diatopic part of the analysis, the full dataset is used. Out of these 1,921 observations, 91.3% (1,754) are of the *-a* variant of the possessive. In Figure 1, the proportions of the *-a* variant are mapped out for the respective Andalusian provinces. As indicated by the chi-square test (Figure 1), there is no statistically significant difference in terms of variant preference between the eight provinces.

The large proportion of the *-a* variant is rather remarkable given the fact that the variant is considered innovative, marked, less common and more stigmatized than the *-o* variant (RAE & ASALE, 2009, § 18.4o; Marttinen Larsson & Álvarez López, 2017). It coincides, nonetheless, largely with the dataset from oral corpora analyzed by Salgado & Bouzouita (2017, pp. 782-783). Their oral Andalusian data demonstrate a proportion of 82.2% (37/45) of the *-a* variant of the possessive with all adverbials; excluding the nominal masculine locatives *al lado* and *alrededor*, this proportion increases to 96.6% (28/29). Despite Salgado & Bouzouita

(2017) consulting more than twenty oral corpora and only obtaining a limited dataset for Andalusian Spanish, the similarities with the results obtained in the present analysis are striking and confirm the high degree of generalization of the *-a* variant in this Spanish variety.

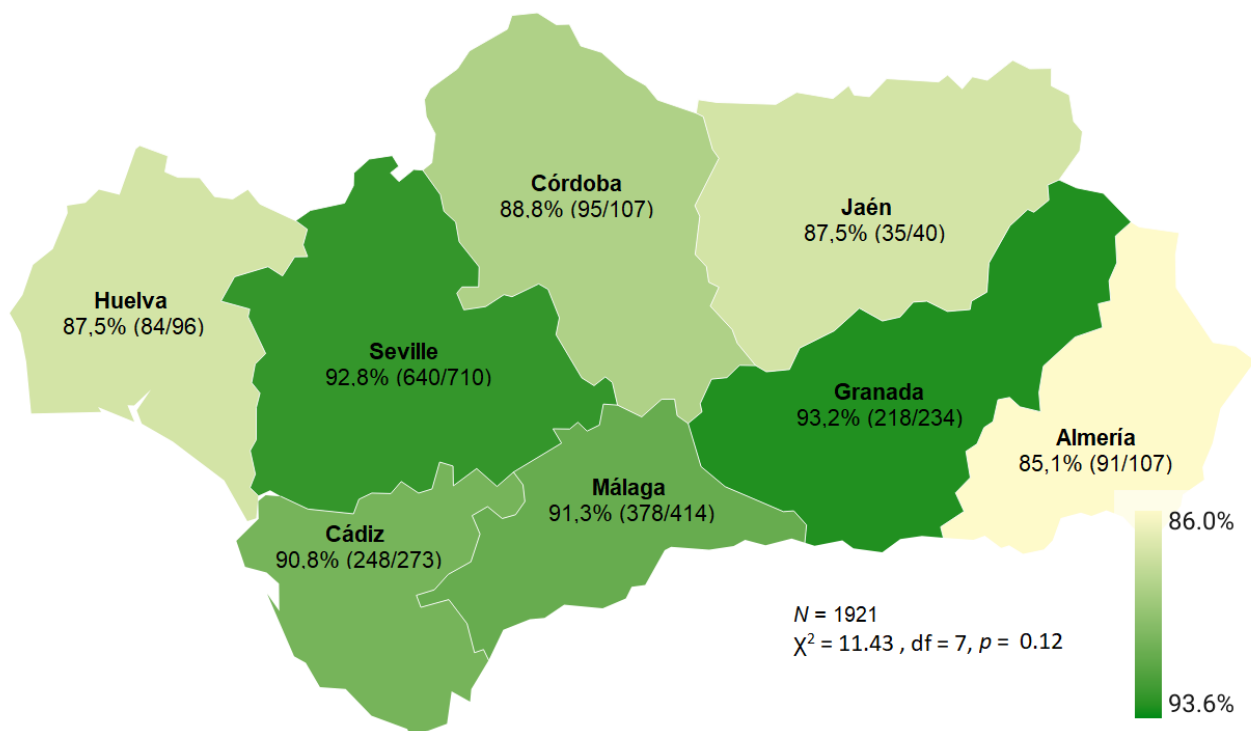


Figure 1. Percentage of *-a* variant per Andalusian province

As Figure 1 shows, the proportion of the *-a* variant ranges from 85.1% to 93.2% in the different provinces. This testifies to a highly generalized use of the feminine variant in at least colloquial language use on Twitter, where it is the preferred option over the *-o* variant. From a bird's eye view, the feminine morphology thus appears to have diffused entirely and evenly throughout the Andalusian territory. The next part of the analysis deals with whether the little variation that we do find between the two variants is synchronically conditioned by any language-internal predictors.

### 3.2. The influence of language-internal constraints

In order to determine the influence of the coded predictors on the suffix variation, we constructed a mixed-effects regression model using the *lme4* package in R (Bates, Mächler, Bolker & Walker, 2015). A stepwise model selection process was followed, fitting predictors iteratively and only maintaining significant predictors that contributed to the improvement of the model fit (as determined by between-model ANOVA comparisons). We included 'user' – in order to account for idiosyncratic variability – and 'locative adverb' as random factors (cf. Tagliamonte & Baayen, 2012). However, for the latter, a model overfit occurred which yielded a singular model. In an attempt to remedy this, we used the variable 'grouped locative adverbials' in which variants of the same adverbial were collapsed into adverbial variables and included this as a fixed effect which, nonetheless, produced convergence errors. There appears thus to be too little variation to allow for such a fine-grained and somewhat unbalanced predictor, since it produces a singular model. Instead, the influence of the different locative adverbials will be scrutinized more thoroughly through a tabulated analysis.

The final model indicates that none of the predictors have any significant effect on the variation (Table 2). The effect sizes (estimates) do signal, nonetheless, that locative adverbials ending in */-e/* have the highest probability of yielding the *-a* variant of the possessive, which will be examined in more detail in the tabulated analysis of the influence of the individual locatives. In addition, 3<sup>rd</sup> person possessives favor the *-a* variant of the possessive, as well as generic referents; yet, the observed differences are non-significant in the regression model. In all, the largest effect sizes are indeed produced by the locative adverbials whose final nucleus consists of */-e/*.

Because of the observed effect size in Table 2 ('Estimate'), the influence of analogy and the different vowel-endings of the adverbials will be the focus in this analysis in an effort to identify if any such constraints are



regulating the studied variation in Andalusian Spanish. To this end, a more detailed descriptive analysis will be provided in what follows.

Factor	Estimate	Std. Error	z value	Pr(> z )
Intercept	10.92392	1.87958	5.812	<0.00000001***
Vowel of adverbial's last syllable: /-a/	0.07391	0.79735	0.093	0.926
Vowel of adverbial's last syllable: /-e/	1.79906	1.31111	1.372	0.170
Number: Singular	0.43721	1.74370	0.251	0.802
Person: 2 <sup>nd</sup>	0.34438	1.05167	0.327	0.743
Person: 3 <sup>rd</sup>	0.85889	1.63099	0.527	0.598
Gender of referent: Generic	0.47989	1.61500	0.297	0.766
Gender of referent: Masculine	-0.30137	0.89373	-0.337	0.736

Table 2. Mixed-effects logistic regression model

Locative adverbial by final syllable's vowel	-o suffix (%)	-a suffix (%)	-o suffix (N)	-a suffix (N)	Total (N)	-a suffix % total by vowel group
/o/	<i>al lado</i>	12.6%	60	417	477	86.5%
	<i>debajo</i>	9.1%	3	30	33	
	<i>alrededor</i>	26.7%	12	33	45	
/a/	<i>cerca</i>	11.1%	18	144	162	91.7%
	<i>encima</i>	8.8%	24	248	272	
	<i>detrás</i>	6.1%	21	324	345	
	<i>atrás</i>	20%	3	12	15	
/e/	<i>delante</i>	4.2%	21	483	504	95.7%
	<i>enfrente</i>	5.3%	4	71	75	
Total	8.6%	91.4%	166	1762	1928	

Table 3. Locative adverbial per vowel group and possessive suffix variant

In Table 3, the data is tabulated according to the individual locative adverbials but grouped by the vowel of their final syllable. This analysis uses the full dataset since it is univariate and not dependent on the other variables containing complete observations. This said, adverbials and their variants that exhibit very low frequencies (below 10) have been removed from the analysis. The removed locative adverbials are the following: *abajo* (4 occurrences, all with the *-a* variant), *dentro* (8 occurrences, out of which 7 have the *-a* variant), *en/por lo alto* (4 occurrences, out of which 3 with the *-a* variant), *en torno* (2 occurrences, none with the *-a* variant), *lejos* (2 occurrences, out of which 1 with the *-a* variant), *arriba* (10 occurrences, all with the *-a* variant), *tras* (1 occurrence with the *-a* variant), *ante* (1 occurrence with the *-a* variant) and *adelante* (1 occurrence with the *-o* variant).

Another methodological choice concerns the definition of the variable context: generally, the variationist methodology recommends removing contexts or items that exhibit zero variation in terms of variant selection (cf. Tagliamonte, 2012, pp. 10-11). However, since what the following analysis sets out to test is the hypothesis that there could be a certain degree of analogy between the vowel of the preceding adverb's last syllable and the suffix of the possessive, we believe that it is pertinent to include categorical contexts as well, since they would testify to a high degree of entrenchment in a particular analogical construction (such as in the case of *debajo*, *encima*, *detrás*, *delante* and *enfrente*; see Table 3). Furthermore, the inclusion of such categorical contexts does not pose any challenge to the following analysis, since adverbials are not compared on individual levels, but on aggregate group levels using the final syllables of the respective locative adverbials.

As Table 3 shows, there are virtually no locative adverbials that strongly disfavor the *-a* variant. Among the locative adverbials included in Table 3, the proportion of the *-a* variant ranges from 73.3% for *alrededor* to 95.8% for *delante*.

As stated in Section 1, various locative adverbials have nominal bases. This is of particular interest to us when examining the effect of individual adverbials since there might be a correlation between the grammatical gender of the nominal base and the suffix of the possessive. The nominal adverbials with a masculine base included in the dataset are the following, with their respective percentages of the *-a* variant between parentheses: *al lado* (87.4%) and *alrededor* (73.3%). As can be evidenced by the frequency of use of the *-a* variant of these locative adverbials, even the masculine noun-based ones strongly attract the *-a* variant of the possessive and

their overall frequency of use of the *-a* variant reaches 86.2% (450/522). These findings are very much in line with the data from Salgado & Bouzouita's study (2017, pp. 782-783) which demonstrates that, when including *al lado* and *alrededor*, the frequency of use of the *-a* variant constitutes 82.2% (37/45) in Andalusia. This proportion increases to 96.6% (28/29) once these denominal adverbs are excluded. Tellingly, these denominal adverbials occur with the *-o* possessive variant in 85.7% (18/21) of the instances from Andalusia, Madrid and Valencia (Salgado & Bouzouita, 2017, p. 782). In light of this, the authors argue that the preference for this *-o* pattern might be explained by the original masculine nominal character of these adverbials. While this appears to be the case for their limited dataset, the oral corpora included in their study range from approximately 1987 (data from Seville) to 2012. Consequently, the diachrony of the spread of the *-a* suffix as well as the trajectory of grammaticalization of the *-a* variant of the possessive remains to be explored in detail. In an attempt to contribute to this issue, we turn to the comparison between the distributions presented by Salgado & Bouzouita (2017, pp. 782-783) and the data presented in Table 3. Undoubtedly, in the Twitter dataset from 2018 analyzed here the use of the *-a* possessive variant predominates substantially and is completely generalized, even for the masculine denominal adverbials *al lado* and *alrededor*. Notwithstanding this, they are among the ones that appear least with the feminine possessive. It might thus be inferred that these denominal adverbials have resisted change the most due to their original nominal masculine traits. Synchronically, though, this effect is not evident, since all adverbials in the group of */-o/-*-ending final syllables – nominal or not – appear to manifest similar tendencies in terms of variant preferences.

Turning now to the potential influence of analogy, Table 3 demonstrates the proportional differences between the *-a* and *-o* variants in the respective vowel groups (*/-o/*, */-a/* and */-e/*). The group that least frequently adopts the *-a* possessive variant is the */-o/-*-ending adverbial group. Interestingly, the group of adverbials ending in */-a/* is only intermediate among these three, presenting an *-a* variant frequency of 91.7%. Surprisingly, the adverbial group favoring the *-a* variant the most is the */-e/-*-ending group (95.7%). A chi-square test indicates that there is a statistically significant difference between the three vowel groups ( $\chi^2 = 30.605$ ,  $df = 2$ ,  $p < 0.0001$ ). In addition, there is a significant difference between the two latter groups (*/-a/* and */-e/*;  $\chi^2 = 8.633$ ,  $df = 1$ ,  $p = 0.003$ ). This */-e/* preference is palpable and constitutes a novel finding. The question thus arises as to why we evidence this pattern.

If analogy were the triggering factor of the spread of the *-a* possessive variant, we would expect to find the following pattern: (1) */-a/* adverbials favor the *-a* possessive variant; (2) */-o/* adverbials do not favor the *-a* variant; and (3) those with */-e/* might exhibit free variation. However, */-e/* adverbials prefer the *-a* variant – how do we explain this? In what follows, we will propose a new hypothesis regarding the development of this linguistic change.

As found by Marttinen Larsson & Álvarez López (2017), the innovative *-a* variant's spread is first documented in written European Spanish from the 1950s. At the same time, in Andalusian Spanish 1950s' spoken language, a separate but parallel process was taking place consisting of a recurrent phonetic fronting of */-a/* in VC-syllables in word-final position (Alonso, 1956; Llorente Maldonado de Guevara, 1962, p. 229; Narbona Jiménez et al., 2011, pp. 192-194). This */-a/-*-fronting in final position yielded */-e/*, a change led mainly by children and women, particularly those belonging to lower socio-economic groups and elderly women. Furthermore, consonants in syllable-final positions are often debilitated. These phonological alternations lead to words such as *berenjenas* 'eggplants' being pronounced during a period of time as */be.ren.'je.ne/* with consonant debilitation of word-final */-s/* and phonetic fronting of word-final */-a/*. Meanwhile, we hypothesize that adverbials whose final nucleus consists of */-a/* trigger the use of the *-a* variant of the possessive as an analogical process. This process becomes entrenched and obligatory for adverbs with final */-a/* nucleus (e.g., *cerca mía* lit. 'close mine'; *detrá[s] tuya* lit. 'behind yours'). As the parallel phonetic fronting change is advancing, constructions such as *detrá[s] tuya* may be pronounced as */de.'tre tu.ya/* (cf. Alonso, 1956, p. 18). We posit that, through similarity-based generalization, this use is possibly transferred to adverbs ending in */-e/*, thus yielding *enfrente mía*, *delante tuya*, and so on. In other words, we might conceptualize the changes at hand as a series of analogical extensions (e.g., De Smet, 2012), which can be detailed in the following way:

Stage 1: At this stage, two processes take place, which might have coincided synchronically. Firstly, due to the formal similarity between the non-denominal adverbials with */-a/* nucleus and the feminine suffix, the *-a* possessive variant starts being used in this adverbial locative context (e.g., *cerca mía*, *detrás tuya*). Through frequency of use the collocational pattern [adverbial */a/* + possessive *-a*] becomes increasingly obligatory and entrenched, replacing progressively the prepositional complement construction. Secondly, the fronted [a] of the non-denominal adverbial becomes */-e/* in final nucleus, resulting thus in the use of examples such as */de.'tre 'mi.a/*. The analogical extension of Stage 1 provides the necessary stepping stone for the next analogical extension to take place.

Stage 2: Due to the formal similarity between the denominal adverbials with */-a/* nucleus, which are pronounced as */-e/* as in */de.'tre/*, and those ending in */-e/*, such as *delante*, another analogical extension takes place as the use of the feminine possessive is transposed to this new adverbial context, giving rise to cases such as *delante mía*.

Stage 3: As the use of the *-a* variant is becoming increasingly obligatory for */-a/* and */-e/-*-ending locative adverbials, another analogical extension takes place: the */-o/-*-ending adverbials, which initially admitted the masculine possessive complement as a result of an analogical extension due to a similarity-based generalization with the masculine denominal adverbials, now also starts to admit the feminine variant, as in *debajo mía*.

Stage 4: As a final step in the analogical spread of the feminine possessive morphology, the *-a* possessive variant extends its scope to denominal masculine locatives, replacing thus their masculine counterpart (*alrededor mía*, *al lado suya*).

Summarizing, the spreading of the feminine morphology in the adverbial locative contexts operates through analogy: more concretely, a series of analogical extensions spreads the feminine possessive complement from one adverbial context to another, creating a so-called analogical chain or analogical snowballing effect, whereby one extension becomes possible due to the previous analogical change (De Smet, 2012, p. 8; Bouzouita & Marttinen Larsson, 2020, pp. 19-24 for more details on other analogical changes affecting the tonic possessives in non-nominal contexts). Additionally, the diffusional changes in this chain are replaceive in nature as either the original prepositional complement is replaced (in Stage 1 and 2) or the masculine possessive variant (in Stage 3 and 4). The proposed analogical chain is a change from below, probably led by women and lower socio-economic groups, below the level of consciousness (cf. Narbona Jiménez et al., 2011, pp. 192-194). Normatively, the use of the feminine possessive variant continues to be highly stigmatized and perceived as a low-frequency trait (cf. example (5)).

Interestingly, similar analogical extensions of the feminine possessive complement (Stages 1-4) appear to have been observed diachronically for Galician. More specifically, the first documentations of the masculine possessive complement date from the end of the 19<sup>th</sup> century and appear in the masculine denominal adverbial context. From the 1930s onwards, the feminine variant is attested for non-denominal adverbials ending with *-a* and *-e*, to wit with *arriba* ‘above’, *diante* ‘in front’ and *detrás* ‘behind’. Only later, from the 1970s and 1980s onwards, do the non-denominal adverbials also start admitting the masculine possessive complement. Finally, the denominal adverbials start accepting the feminine possessive variant from 1950s onward, especially since the 1990s (see Silva Domínguez, 2020, for more details). This said, further research on Galician is needed to examine the extent to which the same diffusional analogical changes are observed synchronically and the motivations behind them. Moreover, unlike for Andalusian Spanish where a phonetic change occurred in Stage 1 and could explain the analogical extension to Stage 2, the authors are not aware of a similar development for Galician. However, this is not to exclude that other co-occurring processes of phonological variation might have played a role in the case of Galician. In all, this issue remains open, and more research is needed to address this gap in knowledge.

Although we do not present here diachronic data for Andalusian Spanish, the synchronic results do buttress the proposed analogical chain of extensions as we have shown that, synchronically, (i) adverbials that end in */-a/* and */-e/-* exhibit the largest probabilities of being used with the feminine *-a* suffix of the possessive in Andalusian Spanish, a reflection of Stage 1 and 2 of the chain respectively, and (ii) non-denominal */-o/-*-ending adverbials and masculine denominal adverbials follow behind, exhibiting less high probabilities of presenting a feminine possessive complement, mirroring thus Stage 3 and 4 of the proposed analogical chain. However, one final issue that requires further discussion is the very high frequency with which */-e/-*-ending locatives appear with the feminine possessive. How come these */-e/-*-ending locatives appear more frequently with the *-a* suffix than the source context of the feminine possessive, viz. */-a/-*-ending locative adverbs? Why has the */-e/-* ... *-a* pattern outrun the analogous */-a/* ... *-a* combination in terms of productivity? This finding will be briefly dwelled upon in what follows.

As research on frequency effects in morphosyntax has shown repeatedly (Boas, 2008; Bybee, 1985, 2006, 2007, 2010; Schmid, 2010, among others), an item’s or construction’s frequency of use brings about a progressive entrenchment in speakers’ mental and cognitive storage of grammar. In other words, a high frequency of use and the recurring repetition of forms are correlated with a higher degree of establishment of the structures’ mental representation in grammar. Conversely, forms that are less frequent are not as deep-rooted in their grammatical representation and are, as such, in greater flux in the speakers’ grammar. This thus means that high-frequency structures tend to resist change longer than low-frequency structures since the former are more anchored in their grammatical representation and have a so-called “conserving effect” (Langacker, 1987, p. 59; Bybee, 2006, 2007, 2010). In view of this, the reason for which */-e/-*-ending locatives could favor the *-a* suffix more than the */-a/-*-ending locatives might potentially be that they appear less frequently in discourse and their combination with the original *-o* possessive suffix is less entrenched. Due to space limitations, we will not enter into an extensive discussion on this subject. This said, the following data appear at first sight to corroborate the posited hypothesis. Searching the different locatives that appear in Table 3 (see above), the *Corpus de Referencia del Español Actual* (CREA) reveals that the */-a/-*-ending adverbials found in Spain have a group frequency of 41,664 (mean = 10,416), whereas the */-e/-*-ending ones, such as *delante* and *enfrente*, have a group frequency of 9,497 and a mean of 4,748.5. Consequently, our prediction

appears to be born out: comparing the respective means, /-a/-ending locatives are much more frequent than the /-e/-ending ones. The same pattern is observed when consulting the frequencies of use of the locative adverbials in other corpora: in the European Spanish subcorpus of the *Corpus del Español del Siglo XXI* (CORPES XXI), the group frequency of the /-a/-ending locatives is 70,272 with a mean of 17,568 whereas the ones ending in /-e/ have a group frequency of 15,929 and a mean of 7,964.5. Lastly, the Andalusian data obtained from uneducated, rural informants of the *Corpus Oral y Sonoro del Español Rural* (COSER), we find that the /-a/-ending group frequency is 299 and the mean 74.8, while the /-e/-ending locatives are found in 93 instances and have a group mean of 46.5. All the previous supports our hypothesis that more frequent items – in this case, /-a/-ending locatives – have a conserving effect on the observed process of variation. Therefore, they appear to resist the implementation of the innovation to a larger extent than less frequent structures, such as the adverbials with /-e/. Importantly, this hypothesis needs further systematic empirical testing before it can be assigned any factual explicatory power. We thus invite future studies to scrutinize this matter.

#### 4. Conclusions

The present paper deals with the morphological variation found in Spanish adverbial locative constructions with possessive pronouns in one of the communities where the innovative variant is used most frequently, to wit Andalusia. Almost 2,000 tweets were compiled in order to determine the incoming variant's geography and functioning. The following detail the examined research questions and their answers.

1. Is the *-a* possessive variant the preferred option in Andalusian Spanish, as found in Salgado & Bouzouita (2017)?

The findings presented in this paper show that the *-a* variant is indeed the predominant variant in all provinces of Andalusia. It is by far the preferred option with frequencies ranging between 85 and 93 percent. These indications strongly contradict earlier postulations of the *-a* variant being a rare variant (RAE & ASALE, 2009, § 18.4o), but confirms what empirically-oriented investigations have found for this region, albeit with a limited dataset (Salgado & Bouzouita, 2017).

2. Can diatopic variation between the *-o* and *-a* suffixes of the possessive be found within Andalusia?

There appears to be no significant diatopic variation between the two possessive variants within Andalusia. The *-a* variant is the most frequently found form in all eight provinces. However, considering that the geospecified searches centered on the capitals of the provinces plus an additional radius of 25 kilometers, we are focusing largely on urban varieties. In other words, there is potentially more diatopic variation than the present dataset allows us to discern. Generally speaking though, the *-a* variant appears to be highly conventionalized in all eight provinces of Andalusia, with the caveat that there might be additional sub-regional variation within these provinces in the rural varieties. The same goes for the frontier zones of the provinces of Andalusia that are in dialect contact with other varieties. Future studies should also examine possible diastratic differences in the use of the *-a* possessive variant with locative adverbials.

3. What mechanisms, if any, account synchronically for the use of the *-a* possessive variant in Andalusian Spanish?

From a synchronic point of view, the only factor that appears to have a slight effect on the studied variation is that of analogy. Inferring the diachronic process from synchronic data, we have hypothesized a series of analogical extensions with four different stages that spread the *-a* possessive variant from one adverbial locative context to another. These replacive changes are driven by similarity-based generalizations, whether these be syntactic similarities or phonetic ones. The diffusional spread of the feminine morphology in the locative adverbial possessive structures appear to be close to completion since it now affects all possible adverbial contexts and its frequency of use is very high.

In all, the present study provides a clear case study for the functioning of analogical snowballing effects and analogical chains (cf. De Smet, 2012, p. 8, p. 255) as the identified stages reflect how one analogical extension paves the way for a subsequent one. As argued by De Smet (2012, p. 6), diffusional change reflects speakers' regularization of grammar with its phasedness being determined by the degree of resistance in the target context. The innovation firstly finds its grip in the least resistant environment and provokes an increasing pressure onto a neighboring context, allowing the innovation to diffuse through analogy (De Smet, 2012, p. 6, p. 8).

Lastly, the present study has some limitations that need to be articulated. It should be pointed out that, although the tweets that have been analyzed here can give us a good idea about the diatopic diffusion of certain phenomena, they are nonetheless inadequate for the study of diastratic and diaphasic conditioning since they do not contain the necessary metadata. Additionally, it should not be forgotten that due to the methodology used for this study, whereby we only focus on the various urban centers of Andalusia, the rural areas are de



facto excluded. Consequently, if there were to be diatopic variation in those rural zones, the analyses presented here would not be able to render this. Further, the typical restrictions related to Twitter data apply: to wit, Twitter users tend to belong to the younger generations, as is well known (Longley, Adnan & Lansley, 2015; Statista, 2021). Given that the sociolinguistic status of the innovative *-a* variant of the possessive has been frequently characterized as diastatically marked, an interesting potential venue for further empirical scrutiny thus lies in the collection and analysis of elicitation data, grammaticality judgement tests or perceptual data that can satisfactorily operationalize diastatic and diaphasic variables that might influence the variation at hand. Moreover, future studies are urged to replicate the analysis brought forward here for other dialects of European Spanish that are not as advanced in their diffusion of the innovative variant in order to infer the diachrony from synchronic patterns in different areas. In doing this, more light can be shed onto the evolutionary paths along which the studied phenomenon has unfolded.

## Acknowledgements

Different versions of this paper have been commented on by participants at the conference *Investigando las hablas andaluzas: New approaches to Andalusian Spanish* (Innsbruck, May 24-25 of 2018), at the conference *Possessive Constructions in Romance* (Ghent, June 27-29 of 2018) and at *The International Conference on Language Variation in Europe (ICLaVE) 10* (Leeuwarden, June 26–28 of 2019). We are thankful for the feedback that we received. In addition, we would like to thank Carlota de Benito Moreno, Johannes Kabatek, Lars Fant, Laura Álvarez López, José Luis Blas Arroyo, Martin Hummel and Antonio Fábregas who have all provided valuable feedback during the elaboration of this paper. Naturally, the usual disclaimers apply.

## References

- Alcina Franch, J. & Manuel Blecua, J. (1975). *Gramática española*. Barcelona: Ariel.
- Almela Pérez, R. (1991). Las construcciones del tipo ‘delante suyo’. In C. Hernández (Ed.), *El español de América. Actas del III Congreso Internacional del Español de América, Valladolid, 3 a 9 de julio de 1989*, vol. 1 (pp. 435-444). Valladolid: Universidad de Valladolid.
- Alonso, D. (1956). *En la Andalucía de la E: Dialectología pintoresca*. Madrid: Clavileño.
- Bates, D, Mächler, M., Bolker, B & Walker, S. (2015). “Fitting Linear Mixed-Effects Models using lme4.” *Journal of Statistical Software* 67(1). DOI: <http://doi.org/10.18637/jss.v067.i01>
- Bertolotti, V. (2017). Pronombres posesivos y cambios gramaticales en español. Análisis de la variedad rioplatense. In C. Company Company & N. Huerta Flores (Eds.), *La posesión en la lengua española* (pp. 325-350). Madrid: Consejo Superior de Investigaciones Científicas.
- Boas, H. C. (2008). Determining the structure of lexical entries and grammatical constructions in Construction Grammar. *Annual Review of Cognitive Linguistics* 6, 113-144. DOI: <http://doi.org/10.1075/arcl.6.06boa>
- Bouzouita, M. (2022). Las estructuras posesivas nominales en español: ¿un universal vernáculo? In Á. Di Tullio & E. Pato (Eds.), *Universales vernáculos en la gramática del español* (pp. 71-96). Madrid / Frankfurt: Iberoamericana / Vervuert.
- Bouzouita, M., Castillo Lluch, M. & Pato, E. (2021). Dialectos del español : une application pour l'étude de la variation linguistique dans le monde hispanophone. In M. Avanzi, N. LoVecchio, A. Millour & A. Thibault (Eds.), *Nouveaux regards sur la variation dialectale – New ways of analyzing dialectal variation* (pp. 291-304). Strasbourg: Éditions de Linguistique et de Philologie.
- Bouzouita, M., Castillo Lluch, M. & Pato, E. (in press). Dialectos del español: presentación de la *app* y primeros resultados. *Revista Internacional de Lingüística Iberoamericana*.
- Bouzouita, M. & M. Marttinen Larsson. (2020). Variation and change in the Romance possessive constructions: An overview of nominal, adverbial and verbal uses. *Moderna språk* 114(3), 1-44.
- Bybee, J. (1985). *Morphology: An Inquiry into the Relationship between Meaning and Form*. Amsterdam/Philadelphia: John Benjamins.
- Bybee, J. (2006). From usage to grammar: The mind’s response to repetition. *Language* 82(4), 711-733. DOI: <http://doi.org/10.1353/lan.2006.0186>
- Bybee, J. (2007). *Frequency of use and the organization of language*. Oxford: Oxford University Press. DOI: <http://doi.org/10.1093/acprof:oso/9780195301571.001.0001>
- Bybee, J. (2010). *Language, Usage, and Cognition*. Cambridge: Cambridge University Press. DOI: <http://doi.org/10.1017/CBO9780511750526>
- CORPES XXI = Real Academia Española: Banco de datos (CORPES) [online]. *Corpus del español del siglo XXI*. <http://www.rae.es> [Last consulted: 2021-10-05].
- COSER = Fernández-Ordóñez, I. (2005-). *Corpus Oral y Sonoro del Español Rural*. <http://www.corpusrural.es> [Last consulted: 2021-10-05].

- CREA = Real Academia Española: Banco de datos (CREA) [online]. *Corpus de referencia del español actual*. <http://www.rae.es> [Last consulted: 2021-10-05].
- De Smet, H. (2012). *Spreading Patterns: Diffusional Change in the English System of Complementation*. Oxford: Oxford University Press. DOI: <http://doi.org/10.1093/acprof:oso/9780199812752.001.0001>
- Felíu Arquiola, E. & Pato, E. (2019). “¿Realmentes existen?: la ‘pluralización’ de los adverbios en *-mente* en español actual.” *Onomázein* 44: 166-190. DOI: <http://doi.org/10.7764/onomazein.44.08>
- Felíu Arquiola, E. & Pato, E. (2020). *En torno a la denominada “concordancia adverbial” en español: tres casos de variación*. Madrid: Consejo Superior de Investigaciones Científicas.
- Fernández-Ordóñez, I. (2016). Dialectos del español peninsular. In Javier Gutiérrez Rexach (Ed.), *Enciclopedia de Lingüística Hispánica*, vol. 2 (pp. 387-404). Oxford: Routledge.
- Fernández Ramírez, S. (1987). *Gramática española*, vol. 3.2: *El pronombre*. Madrid: Arco Libros.
- González Calvo, J. M. (2006). Vacilaciones, variaciones e incorrecciones sintácticas en el español actual. In J. A. Moya Corral & M. Sosinski (Eds.), *Lexicografía y enseñanza de la lengua española. Actas de las XI Jornadas sobre la enseñanza de la lengua española* (pp. 65-79). Granada: Universidad de Granada.
- Gries, S. Th. & Hilpert, M. (2010). Modeling diachronic change in the third person singular: a multifactorial, verb- and author-specific exploratory approach. *English Language and Linguistics* 14(3), 293-320. DOI: <http://doi.org/10.1017/S1360674310000092>
- Grijelmo, Álex. (2017, May 22). El uso incorrecto y “desprestigiado” del posesivo tras un adverbio se va extendiendo en la política. *El País: La Punta de la Lengua*. [https://elpais.com/elpais/2017/05/19/opinion/1495189162\\_983717.html](https://elpais.com/elpais/2017/05/19/opinion/1495189162_983717.html)
- Hoff, M. (2020). *Cerca mío/a or cerca de mí? A variationist analysis of Spanish locative + possessive on Twitter*. *Studies in Hispanic and Lusophone Linguistics* 13(1), 51-78. DOI: <http://doi.org/10.1515/shll-2019-2017>
- Kany, C. (1969). *Sintaxis hispanoamericana*. Spanish version by M. Blanco Álvarez. Madrid: Gredos.
- Kearney, M. W. (2016). rtweet: Collecting Twitter data. *CRAN*. Available at <https://cran.r-project.org/package=rtweet>.
- Langacker, R. W. (1987). *Foundations for cognitive grammar. Theoretical prerequisites*. Stanford: Stanford University Press.
- Llorente Maldonado de Guevara, A. (1962). Fonética y fonología andaluzas. *Revista de Filología Española* vol. XLV(14), 227-240.
- Longley, P. A., Adnan, M. & Lansley, G. (2015). The geotemporal demographics of Twitter usage. *Environment and Planning A: Economy and Space* 47(2), 465-484. DOI: <http://doi.org/10.1068/a130122p>
- Marttinen Larsson, M. (in press). Modelling incipient probabilistic grammar change in real time: the grammaticalisation of possessive pronouns in European Spanish locative adverbial constructions. *Corpus Linguistics and Linguistic Theory* (ahead of print). DOI: <http://doi.org/10.1515/cllt-2021-0030>
- Marttinen Larsson, M. (2022). *Las construcciones adverbiales del tipo “delante mío” y sus condicionantes: Patrones probabilísticos, reflejos de gramaticalización y encadenamiento analógico en variedades diacrónicas y diatópicas del español* [Doctoral dissertation]. Stockholm University, Department of Romance Studies and Classics.
- Marttinen Larsson, M. & Álvarez López, L. (2017). ‘Delante suyo’ vs. ‘delante de él’: el uso de las locuciones adverbiales locativas desde una perspectiva diacrónica y diatópica. *Signo y Señal* 31, 85-104.
- Marttinen Larsson, M. & Álvarez López, L. (2022). El uso de los adverbios locativos con pronombres posesivos en el español uruguayo: un análisis diacrónico y probabilístico. *Studies in Hispanic and Lusophone Linguistics* 15(1), 175-210. DOI: <https://doi.org/10.1515/shll-2022-2059>
- Marttinen Larsson, M. & Bouzouita, M. (2018). Encima de mí vs. encima mío: un análisis variacionista de las construcciones adverbiales locativas con complementos preposicionales y posesivos en Twitter. *Moderna språk* 112(1), 1-39.
- Narbona Jiménez, A., Cano Aguilar, R. & Morillo Velarde-Pérez, R. (2011). *El español hablado en Andalucía*. Sevilla: Universidad de Sevilla.
- O’Connor, M. C. (2003). Differential Possessor Expression: Are Pair-Wise Comparisons Ever Required? In M. Butt & T. Holloway (Eds.), *Proceedings of the LFG03 Conference* (pp. 347-366). Stanford: CSLI Publications.
- Octavio de Toledo y Huerta, Á. (2016). *Los relacionantes locativos en la historia del español*. Berlin: De Gruyter Mouton. DOI: <http://doi.org/10.1515/9783110458510>
- RAE & ASALE 2009 = Real Academia Española & Asociación de Academias de la Lengua Española. (2009). *Nueva gramática de la lengua española*. Madrid: Espasa Calpe.
- Salgado, H. & Bouzouita, M. (2017). El uso de las construcciones adverbiales locativas con pronombre posesivo en el español peninsular: un primer acercamiento diatópico. *Zeitschrift für romanische Philologie* 133(3), 766-794.
- Santana Marrero, J. (2014). La estructura adverbio + posesivo en medios de comunicación digitales. *Español actual* 101, 7-30.
- Schmid, H.-J. (2010). Entrenchment, Salience, and Basic Levels. In D. Geeraerts & H. Cuyckens (Eds.), *The Oxford Handbook of Cognitive Linguistics* (pp. 117-138). Oxford: Oxford University Press. DOI: <http://doi.org/10.1093/oxfordhb/9780199738632.013.0005>
- Silva Domínguez, C. (2020). El posesivo en construcciones de valor locativo de la lengua gallega: innovación lingüística e interferencia a partir del análisis de corpus contemporáneos. *Moderna språk* 114(3), 226-263.
- Statista. (2021). *Twitter in Spain*. Statista Research Department. <https://www.statista.com/study/85972/twitter-in-spain/> [Last consulted: 2021-10-10].

- Tagliamonte, S. A. (2012). *Variationist Sociolinguistics: Change, Observation, Interpretation*. Hoboken: John Wiley & Sons.
- Tagliamonte, S. A. & Baayen, R. H. 2012. Models, forests, and trees of York English: *Was/were* variation as a case study for statistical practice. *Language Variation and Change* 24(2), 135-178. DOI: <http://doi.org/10.1017/S0954394512000129>
- Zamora Vicente, A. (1967). *Dialectología española*. Madrid: Gredos.