

# *Encima de mí vs. encima mío*: un análisis variacionista de las construcciones adverbiales locativas con complementos preposicionales y posesivos en Twitter

MATTI MARTTINEN LARSSON  
MIRIAM BOUZOUITA  
Ghent University

## Abstract

El presente trabajo abarca el estudio de la variación en las construcciones adverbiales locativas, que se manifiesta en la alternancia entre la variante ADVERBIO + preposición *de* + PRONOMBRE PERSONAL (por ejemplo, *encima de mí*), y la variante ADVERBIO + PRONOMBRE POSESIVO (por ejemplo, *encima mío/a*). El objetivo consiste en estudiar el posible efecto de varios factores intra- y extralingüísticos sobre la variación bajo estudio. Con esta finalidad, se han analizado datos provenientes de Twitter. La recolección de datos se llevó a cabo en febrero del 2017. Se recogieron 1473 observaciones de construcciones adverbiales locativas de 20 países, de las cuales 537 son de la variante posesiva. Se clasificaron los datos según las siguientes variables independientes: el género y el país del tuitero, la persona y el número gramaticales del referente, y el adverbio. Además, se examinó la interacción entre la persona y el número gramaticales. Se realizó un análisis de bosque aleatorio y un árbol de inferencia condicional. También se construyó un modelo mixto de regresión logística. En su conjunto, estos análisis revelan que el país, el adverbio, la persona y el número gramaticales, y la interacción entre estas últimas variables son los factores estadísticamente significativos para predecir el uso de la variante posesiva.

**Palabras clave:** variación morfosintáctica, variación lingüística, construcciones adverbiales locativas, pronombres posesivos, Twitter, regresión logística

## 1. Introducción

La presente investigación se centra en el estudio de la alternancia entre la variante preposicional ADVERBIO + preposición *de* + PRONOMBRE PERSONAL (por ejemplo, *delante de mí*), y la variante posesiva ADVERBIO + PRONOMBRE POSESIVO (por ejemplo, *detrás mío/a*) en las construcciones adverbiales locativas. Hasta hace poco, este caso de variación morfosintáctica no ha recibido debida atención en comparación con otros fenómenos de variación gramatical. Se ilustra el empleo de las dos variantes en el ejemplo (1):

- (1) Vi que iba a atravesar un umbral y que Becerril hablaba **detrás mío**. Luego, Barbeitos me dijo una palabra larguísima, sentado **delante de mí** y yo perdía la noción de tiempo (ápuđ Marttinen Larsson y Álvarez López 2017, CREA, *La media distancia*, Gándara, Alejandro, 1984, España).

El uso del pronombre posesivo en las construcciones adverbiales locativas, como *detrás mío* en el ejemplo (1), ha sido principalmente estudiado a la luz de normas prescriptivas (cf. Llorente 1980, Martínez de Sousa 1996, RAE/ASALE 2009), aunque en los últimos años se ha empezado a analizar empíricamente esta variación

(p.ej. Almela Pérez 1991, Santana Marrero 2014, Marttinen Larsson 2015, Salgado y Bouzouita 2017, Marttinen Larsson y Álvarez López 2017). Estos estudios nos han proporcionado conocimientos más profundos sobre la evolución diacrónica y diatópica de la variante posesiva. A pesar de eso, se sabe relativamente poco acerca de la probable correlación entre factores intra- y extralingüísticos y el uso de la variante posesiva en las diferentes variedades del español. Es más, pocos estudios anteriores han logrado proporcionar suficientes datos cuantitativos, incluso para caracterizar su uso en las comunidades con las frecuencias de uso más altas. Por tanto, existe la necesidad de llevar a cabo investigaciones variacionistas sobre estas construcciones en las diversas variedades de español.

El fenómeno bajo estudio se ha investigado tanto con base en corpus escritos (Santana Marrero 2014, Marttinen Larsson 2015, Marttinen Larsson y Álvarez López 2017), como en corpus orales (Salgado y Bouzouita 2017). Sin embargo, dada la baja frecuencia del uso de estas construcciones adverbiales locativas en el habla espontánea, ha resultado difícil recoger suficientes datos cuantitativos de esta variación en corpus orales, como se observa en Salgado y Bouzouita (2017), quienes encontraron solo 204 casos relevantes en la pesquisa de 21 corpus que comprenden en total más de 7 250 000 palabras. Aunque existen corpus largos (p. ej. CORDE: 250 000 000 palabras; CREA: 160 000 000 palabras; CORPES XXI: 225 000 000 palabras, entre otros, principalmente de fuentes escritas), tienen la desventaja de que diferentes factores lingüísticos puedan operar en estos comparado con en el lenguaje oral (cf. la variación diafásica; p.ej. Moreno Fernández 2009: 99). Aun así, existe una gran variedad de géneros textuales escritos, algunos de los cuales se acercan más a los parámetros usados en el habla espontánea (cf. Koch y Oesterreicher 2007, Loureda Lamas 2009). En este sentido, el microblog Twitter es un buen ejemplo de una fuente escrita con una inmediatez comunicativa, concepto ligado a lo diafásicamente informal (Koch y Oesterreicher 2007, Estrada Arráez y de Benito Moreno 2016). Debido al carácter comunicativo inmediato de Twitter, se ha optado por analizar en este estudio casos provenientes de este medio de comunicación. Asimismo, el corpus de Twitter tiene la ventaja de poder resolver el problema de la escasez de datos disponibles en los corpus orales (cf. Salgado y Bouzouita 2017). A modo de comparación, Marttinen Larsson (2015) y Marttinen Larsson y Álvarez López (2017) recogen 541 ejemplos de construcciones adverbiales locativas con posesivos en corpus proporcionados por la RAE/ASALE (cf. Santana Marrero 2014, quien recopila 181 casos de la variante posesiva en un corpus escrito), mientras que el corpus de Salgado y Bouzouita (2017) solamente contiene 87 ocurrencias orales de la variante posesiva (a pesar de haber consultado 21 diferentes corpus). Para la presente investigación, al contrario, se ha logrado recopilar y clasificar 537 ejemplos de la variante posesiva frente a 936 casos preposicionales durante un periodo de dos días (19-20 de febrero de 2017). Se ilustra esta variación adverbial locativa en los siguientes ejemplos de Twitter<sup>1</sup>:

---

<sup>1</sup> En los ejemplos expuestos en el presente trabajo, se mantiene la ortografía original de los tuits (mayúsculas incl.). Además del ejemplo mismo, se presentan algunos metadatos, como el género

- (2) como odio q estén **detrás mía** tol día (H, España, 20/02/2017)
- (3) #DobleTentacion no veo el reality hace tiempo, quien es Julia y porque el nacho anda **detrás de ella?** (M, Chile, 20/02/2017)
- (4) Pocas canciones generan **dentro mío** una sensación como lo hace “hablando de la libertad” de la Renga (H, Argentina, 20/02/2017)
- (5) Buscando las ganas **dentro de mí** para hacer el intro de mi research paper (M, Puerto Rico, 20/02/2017)

En cuanto a la estructura de este artículo, se empieza con un repaso del estado de la cuestión (en el apartado 2), que discute los factores extra- e intralingüísticos que se han considerado influyentes en el uso de la variante posesiva. Después, en el apartado 3, se introducen las preguntas de investigación y las hipótesis que orientan el presente estudio. Una vez expuesta la metodología en la sección 4, se introduce el estudio empírico y los análisis cuantitativos en el apartado 5. Se concluye este trabajo con un sexto apartado donde se ofrecen las principales conclusiones.

## 2. Estado de la cuestión

A continuación, se resumen los distintos factores identificados en estudios anteriores que parecen condicionar la alternancia entre la variante preposicional y la posesiva. En lo que concierne a las adscripciones normativas que consideran esta variación, se acepta la variante posesiva para algunos adverbios (p. ej. *alrededor* y *al lado*: RAE/ASALE s.a.). En cuanto a los demás, se considera su combinación con los pronombres posesivos incorrecta (RAE/ASALE 2009: 18.4n), independientemente de si es un adverbio con base nominal (p. ej. *encima*) o con base adverbial (p. ej. *cerca*).

### 2.1 Factores extralingüísticos

Pese a las recomendaciones normativas sobre el uso de la variante posesiva en las construcciones adverbiales locativas, la gramática de la RAE/ASALE (2009: 18.4ñ) afirma que esta variante se ha difundido a registros más elevados que el coloquial en varios países de habla hispana, pero en “diferente medida según las zonas hispanohablantes”. En la descripción diatópica de la variante posesiva, la bibliografía revela adscripciones cuantiosas sobre su uso y se registra su presencia en, esencialmente, todo el mundo hispánico. En la Tabla 1 se sistematizan estas atestaciones diatópicas de manera geográfica, en trazos generales del norte al sur por continente.

---

del tuitero, la región geográfica dada por el programa utilizado (TAGS, Hawksey 2014) y la fecha de publicación del tuit.

Tabla 1. La distribución diatópica de la variante adverbial locativa posesiva según estudios anteriores

País	Estudios anteriores
España	Alcina Franch y Blecua (1975), Kany (1969), Llorente (1980), Almela Pérez (1991), Hernández Alonso (1996), Gómez Torrego (2009), Aleza Izquierdo (2010), Salgado y Bouzouita (2017), Marttinen Larsson y Álvarez López (2017)
Estados Unidos	
México	
Guatemala	Santana Marrero (2014), Marttinen Larsson y Álvarez López (2017)
Nicaragua	
El Salvador	Marttinen Larsson y Álvarez López (2017)
Costa Rica	Santana Marrero (2014), Marttinen Larsson y Álvarez López (2017)
Panamá	Marttinen Larsson y Álvarez López (2017)
Cuba	Santana Marrero (2014), Marttinen Larsson y Álvarez López (2017)
República Dominicana	Kany (1969)
Puerto Rico	Santana Marrero (2014), Marttinen Larsson y Álvarez López (2017)
Venezuela	Kany (1969), Marttinen Larsson y Álvarez López (2017)
Colombia	Santana Marrero (2014), Marttinen Larsson y Álvarez López (2017)
Ecuador	
Perú	Kany (1969), Santana Marrero (2014), Marttinen Larsson y Álvarez López (2017)
Bolivia	
Chile	
Paraguay	Meilán García (1998), Santana Marrero (2014), Marttinen Larsson y Álvarez López (2017)
Uruguay	
Argentina	Kany (1969), Meilán García (1998), RAE/ASALE (2009), Bertolotti (2014), Santana Marrero (2014), Di Tullio (2015), Marttinen Larsson y Álvarez López (2017)
Guinea Ecuatorial	Marttinen Larsson y Álvarez López (2017)

Como demuestra la Tabla 1, se trata de un fenómeno ampliamente extendido por la mayor parte del mundo hispanohablante y que ha gozado de varias descripciones diatópicas para los continentes de América Latina, Europa, y, en menor medida, África.

En lo referente a Hispanoamérica, es imprescindible apuntar aquí la atención que ha recibido Argentina y el área rioplatense en lo que concierne al uso de la variante posesiva en las construcciones adverbiales locativas (véase la Tabla 1): ya desde la publicación de Kany (1969), se ha atribuido a las variedades rioplatenses un uso muy elevado del posesivo (cf. Meilán García 1998: 23, RAE/ASALE 2009: 18.4ñ, Bertolotti 2014, Santana Marrero 2014: 29, Di Tullio 2015, Marttinen Larsson y Álvarez López 2017). Igualmente afirma la RAE/ASALE (2009: 18.4ñ) que, en esa área, la variante posesiva “aparece en proporción mayor en todos los niveles de lengua”. Santana Marrero (2014: 23) y Marttinen Larsson y Álvarez López (2017)

corroboran esta observación, visto que, en los corpus examinados por estos autores, el uso de las construcciones adverbiales locativas con un posesivo parece estar más generalizado que en las otras variedades examinadas.

Varios investigadores también han identificado un uso no desdeñable de la variante posesiva en España (p.ej. Almela Pérez 1991; véase Salgado y Bouzouita 2017 para una descripción de su difusión diatópica en el español peninsular), donde en algunas variedades compite casi de forma equilibrada con su homólogo preposicional. Es más, la variante posesiva parece ser la predominante en las flexiones de las 1ª y 2ª personas gramaticales en los corpus orales de Andalucía, como en *delante mío* o *encima tuya* (Salgado y Bouzouita 2017). Estos hallazgos divergen bastante de los resultados diacrónicos, que apuntan a una frecuencia menor en el español peninsular escrito y, ciertamente, lejos de ser predominante (Marttinen Larsson 2015, Octavio de Toledo y Huerta 2016: n. 196, Marttinen Larsson y Álvarez López 2017). Se podría considerar que esta diferencia se debe a factores diafásicos (corpus orales vs. corpus escritos) o diastráticos, pues se supone que hay mayor representación de varios estratos socioculturales en los corpus orales que en los escritos, aunque esto depende del tipo de corpus en cuestión.

A pesar de todo esto, son muy escasos y poco detallados los estudios sociolingüísticos que indagan la variación entre la variante adverbial locativa preposicional y la posesiva. En lo que concierne al uso de esta última variante en España, las afirmaciones divergen respecto a sus adscripciones socio-económicas, que la atribuyen o al habla de bajo nivel sociocultural (Hernández Alonso 1996: 204), o a hablantes de niveles de cultura media (Gómez Torrego 2009: 292), o a todos los estratos socioculturales (Llorente 1980: 42), pero sin considerar la influencia de otros factores sociales, como el de género. Estos autores tampoco presentan datos cuantitativos para poder sostener sus afirmaciones. Asimismo, nos faltan análisis multivariados para la mayoría de las regiones en cuanto a las variables extra- e intralingüísticas que puedan determinar el uso de la variante posesiva (p.ej. el uso por el tipo de adverbio, la persona gramatical, etc.).

## 2.2 Factores intralingüísticos

### 2.2.1 El tipo de adverbio

En cuanto a la gama de adverbios locativos usados con la variante posesiva, se expondrá primero una caracterización de los adverbios utilizados en las construcciones de interés, junto con una propuesta de su agrupación con base en estudios anteriores. Segundo, se discutirán los hallazgos de investigaciones anteriores sobre los adverbios más frecuentemente combinados con los pronombres posesivos.

Como propone Meilán García (1998: 12), el paradigma de los adverbios locativos del español consiste en dos grupos principales: los locativos deícticos (*aquí, allí, acá, allá*, etc.) y los “descriptivos” (*delante/detrás, dentro/fuera, encima/debajo*, etc.). Dentro del grupo de los “descriptivos”, este autor hace hincapié en la necesidad de distinguir entre dos subgrupos: los llamados adverbios

“direccionales”, que vienen encabezados por *a-* (p.ej. *adelante/atrás, adentro/afuera*, etc.) y los “no direccionales”. De modo similar, Almela Pérez (1991: 441) indica que los adverbios locativos que se usan con la variante preposicional y/o la posesiva son principalmente aquellos que poseen el rasgo de transitividad, es decir, la capacidad de ir seguido de un complemento. Como se puede ver en la Tabla 2, según este autor, los adverbios encabezados por una *a-* como *adelante* y *adentro* (salvo *alrededor*) carecen de transitividad en su origen. Además de transitividad, Almela Pérez (1991: 440) incluye tres características más para la descripción de los diferentes tipos de locativos; a saber, la ‘gradabilidad’, la ‘locación ductiva’ y la ‘locación aductiva’. La gradabilidad se refiere a la capacidad de los adverbios de ir acompañados por morfemas de comparación y la elación (p. ej. *muy cerca*); la locación ductiva es un concepto ligado a la representación gramatical de ‘en donde’; y finalmente, la locación aductiva es referente a la representación gramatical de ‘de donde’ (Almela Pérez 1991: 440, nota 11).

Tabla 2. Rasgos sintácticos de los adverbios locativos (ápud Almela Pérez 1991: 440)

Adverbios de lugar	Transitividad	Gradabilidad	Locación ductiva	Locación aductiva
<i>encima</i>	+	+	+	-
<i>debajo</i>	+	+	+	-
<i>delante</i>	+	+	+	-
<i>detrás</i>	+	+	+	-
<i>dentro</i>	+	+	+	-
<i>fuera</i>	+	+	+	-
<i>enfrente</i>	+	-	+	+
<i>alrededor</i>	+	-	+	+
<i>cerca</i>	+	+	+	+
<i>lejos</i>	+	+	+	+
<i>arriba</i>	-	+	+	+
<i>abajo</i>	-	+	+	+
<i>adelante</i>	-	+	-	+
<i>atrás</i>	-	+	+	+
<i>adentro</i>	-	+	-	+
<i>afuera</i>	-	+	-	+
<i>aquí</i>	-	-	+	-
<i>allí</i>	-	-	+	+
<i>acá</i>	-	-	+	-
<i>allá</i>	-	-	+	+
<i>ahí</i>	-	-	+	+

Con base en las características señaladas en la Tabla 2, Almela Pérez (1991: 441) establece la existencia de los siguientes subgrupos de adverbios:

- El primer subgrupo que incluye los adverbios *encima, debajo, delante, detrás, dentro* y *fuera* posee los tres primeros rasgos, esto es, transitividad, gradabilidad y locación ductiva.

- El segundo subgrupo con los adverbios *enfrente* y *alrededor* posee el rasgo de transitividad y las locaciones ductiva y aductiva, pero carece de gradabilidad.
- El tercer subgrupo incluye los adverbios *cerca* y *lejos* que ostentan los cuatro rasgos.
- Los adverbios *arriba*, *abajo*, *adelante*, *atrás*, *adentro* y *afuera* forman el cuarto subgrupo. Estos tienen todos gradabilidad y locación aductiva, aunque carecen de transitividad, y, con respecto a la locación ductiva, son heterogéneos.
- El quinto subgrupo contiene los adverbios *aquí*, *allí*, *acá*, *allá* y *ahí* que carecen de transitividad y gradabilidad. Sin embargo, tienen locación ductiva, a pesar de no ser homogéneos con respecto a la locación aductiva.

Almela Pérez propone que los adverbios del primer y del segundo subgrupo forman un “microsistema bien trabado de relaciones mutuas de emparejamientos y oposiciones, lo cual hace que las formas del [primer] subgrupo ‘atraigan’ hacia estos usos a las formas del [cuarto] subgrupo, venciendo la resistencia producida por el doble hecho de que ninguna de estas posee la transitividad (\**arriba del árbol*, \*¡*afuera de casa!*) y de que no todas poseen la locación ductiva” (Almela Pérez 1991: 441).<sup>2</sup> En otras palabras, los adverbios que vienen introducidos por *a-* también pueden tener rasgos transitivos. De ahí que se produzcan las construcciones del tipo *adelante mío*, como en (6) y (7), que proceden de Twitter:

- (6) Viene una loca **adelante mío** en el bondi que es insoportable, todo el viaje gritando (M, Uruguay, 19/02/2017)
- (7) Hay un poli **adelante mío** le voy a gritar aguante la federal como un turro gritaría guante el bajo :\$ macri gato (M, Argentina, 20/02/2017)

De lo anterior se puede concluir que tanto los adverbios sin el prefijo *a-* como los con *a-* pueden entrar en esta variación. Con respecto a los adverbios que aparecen sin este prefijo, Almela Pérez (1991) sostiene que se combinan más a menudo con pronombres posesivos que los prefijales, aunque esta afirmación no se basa en datos de corpus orales, sino en pruebas de aceptación gramatical. Por ello, propone la siguiente categorización de tipos de adverbios locativos (Almela Pérez 1991: 435):<sup>3</sup>

- Sector A: *encima*, *debajo*, *delante*, *detrás*, *dentro*, *fuera*, *alrededor*, *enfrente*, *cerca* y *lejos*
- Sector B: *arriba*, *abajo*, *adelante*, *atrás*, *adentro* y *afuera*

---

<sup>2</sup> Según Almela Pérez (1991: 441), la frase *arriba del árbol* es agramatical. Sin embargo, como demuestra Octavio de Toledo y Huerta (2016: 51-52) se documenta este tipo de construcción desde los primeros textos cuando expresa localización en la zona más alta de la entidad del término, como lo demuestra el siguiente ejemplo que data de principios del siglo XIII:

(1) Arriba del mançanar / otro uaso ui estar (*Siesta*, 27-28, ápuð Octavio de Toledo 2016: 52).

<sup>3</sup> Existen también otros adverbios que pueden aparecer con un posesivo, como *al lado*, que no están incluidos en estos grupos (cf. Salgado y Bouzouita 2017).

Como se ha mencionado antes, otra distinción que se ha introducido en estudios anteriores se basa en la direccionalidad de los adverbios (Meilán García 1998: 12). Esta categorización también recoge a la diferenciación binaria entre los adverbios con vs. sin el prefijo *a-*. Consideramos que esta categorización resulta algo inadecuada si solo se basa en la presencia/ausencia del prefijo *a-* de los adverbios, dado que el prefijo en sí no determina la posible direccionalidad del adverbio, como ilustran los ejemplos (6) y (7): *adelante mío* aparece con el verbo de movimiento *venir* en (6), mientras que coaparece en (7) con *haber*, que es un verbo estático. Asimismo, se encuentran frases con adverbios sin el prefijo *a-* en contextos direccionales, como ilustra el ejemplo (8) con el verbo de movimiento *ir*:

- (8) Cada silencio es una nube que va **detrás de mí** sin parar de llorar. (H, México, 20 de febrero del 2017).

Por lo tanto, esta terminología debería interpretarse con cierta precaución. En efecto, hoy en día la distinción entre los llamados adverbios no direccionales y los direccionales es, para ciertas variedades de Hispanoamérica (cf. Meilán García 1998: 18, RAE/ASALE 2009: 2307), meramente ortográfica, aunque en su origen había una justificación semántico-sintáctica para esta división.

Pese a la posible deficiencia de estas categorizaciones de los adverbios, sobre todo el trabajo de Almela Pérez (1991) es pionero en el estudio de la variación morfosintáctica en los complementos de los adverbiales locativos. Siendo así, se comprobará estadísticamente la validez de su propuesta y se investigará su afirmación sobre la diferencia en comportamiento entre los adverbios del Sector A y B en cuanto a la facilidad de tomar un complemento posesivo (Almela Pérez 1991: 435). Por lo que concierne al adverbio que más frecuentemente se combina con los pronombres posesivos, la Tabla 3 ofrece una visión de conjunto de las observaciones hechas por estudios previos.

Tabla 3. Los adverbios más frecuentemente combinados con pronombres posesivos según estudios anteriores

Estudios	Adverbios			
Almela Pérez (1991)	<i>cerca</i>	<i>detrás / enfrente</i>	<i>encima</i>	<i>delante</i>
RAE/ASALE (2009)	<i>encima</i>	<i>cerca</i>	<i>detrás</i>	<i>delante</i>
Santana Marrero (2014)	<i>detrás</i> 27% (53/193)	<i>delante</i> 35% (47/135)	<i>enfrente</i> 53% (18/34)	<i>alrededor</i> 64% (27/42)
Salgado y Bouzouita (2017)	<i>delante</i> 30% (17/57)	<i>al lado</i> 55% (21/38)	<i>alrededor</i> 77% (10/13)	<i>enfrente</i> 78% (7/9)
<b>Frecuencia</b>	-	←	→	+

Santana Marrero (2014) y Salgado y Bouzouita (2017) coinciden a la hora de concluir que *alrededor* y *enfrente* son los adverbios que muestran la mayor frecuencia de uso con la variante posesiva. No es de sorprender que *alrededor*, así como *al lado*, cuente con un uso frecuente de posesivo, ya que se admite por la

gramática debido a su base nominal (RAE/ASALE 2005: s.v.). Como se puede observar, *enfrente*, *delante* y *detrás* son los adverbios más frecuentemente mencionados por los estudios previos.

### 2.2.2 La persona y el número gramaticales del referente

Estudios anteriores han identificado comportamientos fuertemente distintos en cuanto a los referentes gramaticales que (des)favorecen la variante posesiva. Por un lado, los resultados de Santana Marrero (2014: 14) indican que existe una preferencia por el uso de la variante posesiva con referentes de la 3ª persona gramatical en los medios de comunicación digitales, ejemplificado en (9):

- (9) Porque casi no pierde con nadie que esté **debajo suyo** en el ranking, así como rara vez lo hace contra varios que lo superan en esa clasificación, como el mencionado Tsonga, Berdych o Tipsarevic (ápuđ Santana Marrero 2014; Perfil.com, 30/6/2012)

Por otro lado, el estudio de Salgado y Bouzouita (2017), centrado en datos peninsulares, ha revelado tendencias opuestas. Pese a que la variante preposicional sigue siendo predominante, la 3ª persona casi no se usa con el posesivo, y se presenta más variación entre las variantes con referentes gramaticales de las 1ª y 2ª personas gramaticales, como queda ilustrado en los ejemplos (10)-(19):

- (10) Pareja que está **delante de mí** en la biblioteca, podéis escribir más de una palabra seguida sin besaros constantemente. Qué grima, chico. (M, España, 20/02/2017)
- (11) Quiero dormir pero el prof está **delante mío** (M, Perú, 20/02/2017)
- (12) Chocaron dos camioneta que iban como a 3 coches **adelante de nosotros**, no se pudieron esperar que pasáramos?? (H, México, 19/02/2017)
- (13) Hay un pibe que va al estrada ,recién pasó por **adelante nuestro** y mi viejo dice “sintió tu olor y se hace el lindo” (M, Argentina, 19/02/2017)
- (14) A veces el amor de tu vida está **frente de ti** y no lo vez por estar llenando estereotipos, yo por ejemplo tengo enfrente está pizza. (H, México, 20/02/2017)
- (15) No hay mejor anticonceptivo que tener un niño jodiendo **frente tuyo** en una fila (H, Puerto Rico, 20/02/2017)
- (16) [...] como se nota a los buitres para saltar **encima de vosotros** (H, España, 20/02/2017)
- (17) @xxx Tienen pinta de estar **encima vuestro** to el día, sí. (H, España, 20/02/2017)
- (18) @xxx jajaja salí anoche pero voy cuatro horas **delante de ustedes** jaja (H, Chile, 19/02/2017)
- (19) @xxx el Ecuador entero con ojos **encima suyo**. ¡Los resultados! Si nosotros no dormimos usted menos. ¡Los resultados! (M, Ecuador, 20/02/2017)

Cabe mencionar que, aunque Santana Marrero (2014: 14-15, Tablas 3 y 4) no analiza explícitamente el número gramatical, los datos de los posesivos que presenta (sin la comparación con el homólogo preposicional) permiten identificar que los posesivos recopilados son principalmente casos del singular. Aunque es difícil comparar por falta de coherencia metodológica, el análisis presentado por Salgado y Bouzouita (2017) apunta a que el número no parece ser un factor decisivo en determinar la alternancia en complemento, ya que entre los referentes que se combinan más frecuentemente con el posesivo se encuentran tanto las

observaciones singulares como plurales. Sin embargo, los autores no analizan separadamente la persona y el número, y tampoco llevan a cabo pruebas estadísticas para verificar la distribución.

En resumen, estudios anteriores han encontrado tendencias opuestas en lo que concierne a la persona y el número gramaticales que (des)favorecen la variante posesiva. Dichas diferencias en los resultados podrían posiblemente verse condicionadas por factores diatópicos, visto que Santana Marrero (2014), quien abarca tanto el español europeo como el hispanoamericano, no detalla la distribución geográfica de los referentes gramaticales, una observación ya hecha por Salgado y Bouzouita (2017). También consideramos adecuado contemplar la posibilidad de que estas diferencias, así como las que se pueden observar en el lenguaje escrito frente al habla oral, se deban a factores diafásicos (en este caso, conversaciones informales, como las que investigan Salgado y Bouzouita (2017), frente a medios de comunicación digitales, que pueden variar en su grado de formalidad; cf. Biber (1995).

### **3. Las preguntas de investigación**

Respecto a la variación entre las dos variantes bajo investigación, en el presente estudio se propone identificar qué predictores favorecen o desfavorecen el uso de la variante posesiva y parte de cuatro hipótesis. Primero, se asume la idea que la variante posesiva se registra en todo el mundo hispánico y cuenta con una alta frecuencia de uso en el español rioplatense (Kany 1969; Meilán García 1998: 23; RAE/ASALE 2009; 18.4ñ; Marttinen Larsson y Álvarez López 2017). Segundo, en cuanto al género del tuitero, no contamos con ningún estudio anterior que haya prestado suficiente atención a esta variable. Debido a esto, la hipótesis sobre el posible efecto del género se basa en los hallazgos registrados en estudios variacionistas, esto es, una supuesta predilección por parte de las mujeres por las variantes innovadoras (cf. Labov 2001) – en este caso, la variante posesiva. Respecto a las variables intralingüísticas, se trabajará sobre la hipótesis de que los adverbios pertenecientes al Sector A son los que aparecen significativamente más frecuentemente con la variante posesiva, de acuerdo con los resultados obtenidos por Almela Pérez (1991). Finalmente, se propone que se exhibe más variación con referentes de las 1<sup>as</sup> y 2<sup>as</sup> personas gramaticales, como observado por Salgado y Bouzouita (2017) en un conjunto de corpus orales peninsulares. En el apartado 4.3 se especifica la metodología empleada para evaluar estas hipótesis.

### **4. El corpus y la metodología**

A continuación, se presentarán el corpus y el método usado para la extracción los datos. Asimismo, se entablará la discusión sobre el tipo de expresión lingüística que se registra en Twitter. También se presentan los métodos cuantitativos usados para intentar verificar las hipótesis.

#### 4.1 El corpus: Twitter

Para la presente investigación, se han recogido datos del microblog Twitter, una red social en la cual se permite la publicación de tuits que consisten en un máximo de 140 caracteres (Brown 2016: 49). El uso de Twitter como corpus conlleva la ventaja de que facilita la extracción cuantitativa de datos que muchas veces resultan difíciles de documentar (Bouzouita y Casanova 2017, Brown 2016: 49, Estrada Arráez y de Benito Moreno 2016), tal como es el caso del fenómeno del que se ocupa el presente estudio (cf. Salgado y Bouzouita 2017). La utilidad de datos provenientes de Twitter se debe a tres aspectos principales: 1) la posibilidad de buscar material lingüístico, 2) la existencia de datos con geolocalización, y 3) el carácter coloquial o informal de los tuits (Estrada Arráez y de Benito Moreno 2016: 78). Efectivamente, pese a que se trata de una fuente escrita, el lenguaje documentado en las redes sociales arroja luz sobre las similitudes que tiene a la modalidad oral (Berlanga y Martínez 2010: 57; Mancera Rueda y Pano Alamán 2013: 7). Mancera Rueda y Pano Alamán (2013: 81), por ejemplo, apuntan a que el carácter del español en las redes sociales se explica por lo “inmediato, espontáneo, personalizado y dialógico de la situación comunicativa y, en parte, por las vacilaciones de los hablantes a la hora de escribir.” Varios estudios corroboran la caracterización coloquial del lenguaje en las redes sociales (p.ej. Berlanga y Martínez 2010: 51, Betancourt Arango 2010: 7-8, Estrada Arráez y de Benito Moreno 2016: 84, entre otros), que se debe a la inmediatez comunicativa que expone (Koch y Oesterreicher 2000). Visto que el habla espontánea también se caracteriza por la inmediatez comunicativa, las observaciones que se harán en este estudio sobre el lenguaje contenido en los tuits tienen el potencial de ser válidas para otras formas espontáneas de comunicación.

#### 4.2 La extracción de datos

Se ha realizado la recolección de datos mediante TAGS (Hawksey 2014), un programa de Search API que permite, a través de una conexión con Twitter, la extracción de una muestra de datos de un máximo de 3000 tuits publicados en un período de 7 días. Es posible definir el lema o la secuencia de lemas, el período temporal del que se ocupa (1-7 días), el número máximo de seguidores deseado, el número máximo de tuits y, finalmente, el tipo de búsqueda deseado (tuits, nombres de usuarios, etc.).<sup>4</sup> Cabe subrayar que el objetivo principal de las búsquedas del Search API consiste en encontrar datos relevantes, aunque no siempre son exhaustivos, como indica Twitter (Twitter Inc. 2017): es decir, los resultados devueltos constituyen un índice de tuits recientes y/o populares.<sup>5</sup>

---

<sup>4</sup> Los adverbios y las variantes buscados son los siguientes: *abajo y bajo; adelante, alante y ante; adentro; afuera; alrededor, rededor y redor; arriba; atrás; cerca; debajo; delante; dentro; detrás; encima; enfrente y frente; fuera; y lejos.*

<sup>5</sup> Ya que en el momento en el que se realizaron las búsquedas TAGS se imponía un límite de 3000 resultados, se han restringido las búsquedas a las últimas 24 horas para evitar búsquedas con más de 3000 resultados, puesto que, una vez alcanzado el máximo de 3000 casos, las frecuencias se vuelven estáticas e incomparables.

Al realizar las búsquedas, el primer paso ha consistido en la eliminación de *retuits*, puesto que estos no son pruebas del propio uso lingüístico del tuitero en cuestión. Asimismo, se han eliminado todos los tuits que son citas, identificables por ejemplo por ir introducidas por comillas o, en el caso de letras de canciones, por símbolos de notas musicales o auriculares. Luego, se han examinado los tuits que no se repiten palabra por palabra, pero que comparten una secuencia de su contenido para verificar si se trata de un refrán/un dicho/una cita, un llamado meme de internet, o si se trata de un tuit original. Los casos no originales se han descartado.

Tras estos pasos de depuración, lo esencial ha sido filtrar todas las ocurrencias para eliminar las no geolocalizadas, visto que una de las variables independientes en este estudio es el país y la zona geográfica de residencia del tuitero. Después de haber clasificado las observaciones geolocalizadas, han sido filtradas para solamente analizar ocurrencias de países hispanohablantes.<sup>6</sup>

Visto que otra variable de este estudio es el género del usuario, otro paso ha sido su clasificación. A diferencia de la geolocalización, esta información no viene indicada por TAGS, sino que ha sido necesario hacer este trabajo manualmente. Para cada ocurrencia se ha entrado en los perfiles de los usuarios que corresponden a cada tuit y se han clasificado los usuarios según una categorización binaria entre mujeres y hombres, la cual ha sido posible a partir de la fotografía del perfil y/o la información personal que viene publicada debajo de esta. Los casos dudosos de género han sido descartados, así como las ocurrencias dobles.<sup>7</sup>

También es necesario destacar la existencia de casos ambiguos, como los recogidos en (20) y (21). Visto que el pronombre posesivo *suyo* puede tener 6 diferentes referentes y, por tanto, puede ser el equivalente semántico de *de él*, *de ella*, *de ellos*, *de ellas*, *de usted* y *de ustedes*, no siempre resulta clara la interpretación del referente de este posesivo por falta de suficiente co(n)texto. Por ello se descartan este tipo de ejemplos ambiguos, todos ilustrados abajo:

(20) Estar **cerca suyo** nunca es bueno! (M, Argentina, 20/02/2017)

(21) **Fuera suyo** ahí si le creo jajaja (M, Colombia, 20/02/2017)

Las ocurrencias de *suyo* en los ejemplos (20) y (21) se contrastan con el uso de este posesivo en el ejemplo (22), donde la identificación de su referencia no causa problemas por el co(n)texto.

---

<sup>6</sup> Puesto que el presente estudio es el primer estudio sociolingüístico con un enfoque cuantitativo sobre el fenómeno bajo discusión, optamos por no incluir en el análisis países en que el español no ocupa un estatus oficial, excepción hecha de los Estados Unidos debido al número muy elevado de su población hispanohablante. Concretamente, hemos descartado datos de Alemania, Brasil, Canadá, Emiratos Árabes Unidos, Escocia, Francia, Gales, Grecia, Inglaterra, Irlanda, Islandia, Israel, Italia, Jamaica, Japón, Madagascar, Mali, Nueva Zelanda, Portugal, Rusia, Suiza y Zimbabue.

<sup>7</sup> Descartamos los casos en que la fotografía de perfil del tuitero no representa al usuario (p.ej. fotos de una celebridad, un animal, etc.), por lo cual desconocemos su género. También se han apartado del corpus tuits procedentes de cuentas empresariales u organizacionales.

- (22) y el pato cree que me voy a dejar de él, no todo gira **alrededor suyo**, so, no sea tan patético, plis honey. (M, Costa Rica, 19/02/2017)

Como se ilustra en el ejemplo (22), la mención de *el pato* y *dejar de él* posibilita la interpretación del referente gramatical como la 3ª persona del singular.

Tras todos estos pasos de depuración de datos, el corpus final incluye un total de 1473 observaciones válidas para el análisis.

### 4.3 Metodología

La variable dependiente consiste en dos variantes, siendo estas i) la variante preposicional y ii) la variante posesiva. En la presente investigación, se han realizado análisis distribucionales con pruebas de chi cuadrado para identificar si existe una relación estadísticamente significativa entre la elección de variante y los diferentes factores considerados. Junto con esa prueba, se presenta también el coeficiente *V* de Cramér que arroja luz sobre la fuerza de la relación entre las variables. Asimismo, se ha construido un bosque aleatorio (*random forest*; Breiman 2001) para examinar la importancia de los factores investigados sobre la variación en la variable dependiente. También se ha elaborado un árbol de inferencia condicional (Hothorn, Hornik y Zeileis 2006) para proporcionar una profundización y visualización de la importancia de los factores. Adicionalmente, se ha llevado a cabo un análisis de regresión logística, una técnica usada para medir la relación entre dos o más variables cuando la variable dependiente es categórica (Paolillo 2002: 153, Tagliamonte 2012: 121). Se considera legítimo el uso de este tipo de análisis cuando la variable dependiente consiste en la elección entre dos o más formas y el objetivo es predecir en qué medida las variables independientes afectan a la aparición de la forma de interés. El modelo de regresión logística genera una indicación de la precisión del modelo estadístico, calcula la importancia relativa de cada variable independiente y ofrece un sumario de la exactitud del modelo en cuanto a su capacidad de clasificar las observaciones (Levshina 2015: 254).

Visto que uno de los objetivos del presente estudio es investigar la utilidad de los sectores de adverbios propuestos por Almela Pérez (1991: 435), se evalúa esta agrupación mediante un análisis distribucional de estos dos sectores, los cuales se señalan a continuación junto con los adverbios<sup>8</sup> que recogen respectivamente:

- Sector A: *encima, debajo, delante, detrás, dentro, fuera, alrededor, enfrente, cerca y lejos*
- Sector B: *arriba, abajo, adelante, atrás, adentro y afuera*

No obstante, en el modelo mixto de regresión logística se introducen los adverbios como un factor aleatorio. La ventaja de usar modelos mixtos es la posibilidad de incluir cada ítem (en este caso, cada adverbio) como un nivel individual dentro de

---

<sup>8</sup> En la codificación, se han juntado las observaciones de la variante *rededor* con el adverbio *alrededor*, ya que se trata de una variante que no difiere en significado. Lo mismo vale para el adverbio *enfrente*, que también incluye la variante *frente*.

un factor, pese a que varios niveles contengan frecuencias muy bajas o ceros. A diferencia de un modelo simple de regresión, los modelos mixtos manejan estas discrepancias cuantitativas sin que se distorsione el análisis. Asimismo, el modelo mixto podrá determinar cuál es el peso de cada adverbio en relación con la elección de variante.

En lo que concierne al número y la persona gramatical del referente, se considera tanto la codificación empleada por Salgado y Bouzouita (2017), esto es 1ª persona del singular, 1ª persona del plural, 2ª persona del singular, etc., así como una codificación que separa el número y la persona. Así pues, también se analizará separadamente la persona gramatical (1ª persona, 2ª persona y 3ª persona) y el número gramatical (singular vs. plural).

Para los factores extralingüísticos, se ha codificado el género del tuitero según una distinción binaria: i) el género femenino vs. ii) el masculino. Con respecto a la variación diatópica, se presentarán análisis que consideran tanto el efecto de los países individuales sobre la variación en la elección de variante, como según zonas dialectales para verificar si se trata de un fenómeno homogéneo dentro de diversas regiones lingüísticas. Siendo así, se han agrupado los países codificados en ocho grupos dialectales (expuestos en el apartado 5.1, Tabla 4, basados en la clasificación de zonas lingüísticas propuestas por la RAE/ASALE (2013: 9).

## 5. El análisis distribucional e multivariado

Para el presente estudio se ha recopilado un total de 1473 casos de construcciones adverbiales locativas. De estos 1473 casos, 936 casos (63,5%) son preposicionales, mientras que 537 ejemplos contienen la variante posesiva (36,5%). En los apartados que siguen, se exponen las tablas de contingencia correspondientes a las variables independientes respectivas –esto es, el factor diatópico, el género, los diferentes tipos de adverbios, y la persona y el número gramaticales del referente– para luego presentar el análisis multivariado. Todos los análisis han sido llevados a cabo en R (R Core Development Team 2015).

### 5.1 La zona dialectal

En la Tabla 4, se exponen las frecuencias de las dos variantes según la zona dialectal y el país. Como se observa, las frecuencias relativas más altas del posesivo se observan, en orden descendente, en Argentina (66,2%, 331/500), Chile (52,8%, 28/53) y Uruguay (51,9%, 28/54). Por el contrario, examinando solo los países para los cuales se ha logrado representatividad, el uso posesivo tiene la menor vitalidad en México (2,4%, 6/245), seguido de Venezuela (4,2%, 4/95).

El uso elevado de la variante posesiva en Argentina, Chile y Uruguay corrobora las observaciones de Kany (1969), Meilán García (1998), RAE/ASALE (2009), Santana Marrero (2014) y Marttinen Larsson y Álvarez López (2017). Meilán García (1998: 23) y Santana Marrero (2014: 24) especifican además que la variante posesiva es la predominante en la zona rioplatense, afirmación que concuerda con nuestros datos salvo en el caso de Paraguay. Por el contrario, la frecuencia de uso de la variante posesiva en México y Venezuela es extremadamente baja, conforme

a los resultados obtenidos por Santana Marrero (2014: 23) quien no registra ningún caso de la variante posesiva en Venezuela y solamente 4 en México (la autora no presenta las frecuencias de la variante preposicional, lo que impide una comparación más precisa).

Tabla 4. Los complementos de adverbios locativos según la zona dialectal

Zona dialectal y país	Complemento		Total
	Preposicional	Posesivo	
España	71,9% (159)	28,1% (62)	221
Hispanoamérica	62,1% (777)	37,9% (475)	1252
Estados Unidos	71,4% (35)	28,6% (14)	49
México y Centroamérica	$p < 0,001$ 94,8% (276)	5,2% (15)	291
México	97,6% (239)	2,4% (6)	245
Guatemala	90,9% (10)	9,1% (1)	11
Honduras	100% (3)	0% (0)	3
Nicaragua	100% (6)	0% (0)	6
El Salvador	100% (6)	0% (0)	6
Costa Rica	62,5% (5)	37,5% (3)	8
Panamá	58,3% (7)	41,7% (5)	12
Antillas	$p > 0,05$ 86% (37)	14% (6)	43
Cuba	0	0	0
República Dominicana	91,7% (22)	8,3% (2)	24
Puerto Rico	78,9% (15)	21,1% (4)	19
Caribe continental	$p = 0,001$ 88% (168)	12% (23)	191
Venezuela	95,8% (91)	4,2% (4)	95
Colombia	80,2% (77)	19,8% (19)	96
Zona andina	$p > 0,05$ 57,6% (34)	42,4% (25)	59
Ecuador	58,6% (17)	41,4% (12)	29
Perú	58,6% (17)	41,4% (12)	29
Bolivia	0% (0)	100% (1)	1
Chile	47,2% (25)	52,8% (28)	53
Río de la Plata	$p < 0,05$ 35,7% (202)	64,3% (364)	566
Paraguay	58,3% (7)	41,7% (5)	12
Argentina	33,8% (169)	66,2% (331)	500
Uruguay	48,1% (26)	51,9% (28)	54
<b>Total</b>	<b>63,5% (936)</b>	<b>36,5% (537)</b>	<b>1473</b>

$\chi^2 (7, N = 1473) = 386,27: p < 0,001$  (por zona);  $V = 0,512$ .<sup>9</sup>

$\chi^2 (1, N = 1473) = 7,5: p < 0,01$  (por continente);  $V = 0,005$ .

En lo que concierne al uso con posesivo en España, si se comparan los resultados del presente estudio con los obtenidos por Martinen Larsson y Álvarez López (2017), se observa una diferencia notable: en nuestro corpus de tuits, el posesivo llega al 28,1% (62/221) frente al 3,1% (163/5277) en el estudio de Martinen Larsson y Álvarez López (2017). De manera similar, se registra para Argentina un uso mucho más elevado del posesivo en los tuits frente a los corpus examinados

<sup>9</sup> El nivel de significación usado en el presente estudio es 0,05. El resultado se presenta de la siguiente manera:  $\chi^2$  (grados de libertad, número de observaciones) = valor del chi cuadrado, valor que - junto con los grados de libertad - se asocia con un determinado nivel de significación:  $p$ .

por Marttinen Larsson y Álvarez López (2017): a saber, un 66,2% (331/500) frente a un 18,7% (147/787) del total argentino. En otras palabras, nuestras frecuencias son sustancialmente más altas, lo que posiblemente se debe al registro más informal y familiar de Twitter que los textos encontrados en el CORDE, CREA y CORPES XXI.

La diferencia entre las zonas dialectales es altamente significativa ( $p < 0,001$ ), como se indica en la Tabla 4. Sin embargo, no hay que perder de vista que la fuerza de la relación es intermedia según el coeficiente  $V$  de Cramér ( $V = 0,512$ ). Asimismo, la fuerza de relación es sumamente débil ( $V = 0,005$ ) si se comparan los datos de España y Hispanoamérica, pese a que la diferencia sea estadísticamente significativa ( $p < 0,01$ ). Así pues, para indagar las diferencias diatópicas en más detalle y averiguar si el comportamiento de los diferentes países dentro de la zona dialectal es uniforme, se ha realizado la prueba exacta de Fisher para los países incluidos en cada zona, cuyo resultado se presenta al lado de las zonas. Efectivamente, el resultado es muchas veces heterogéneo con respecto a la variante preferida: para México y Centroamérica, el Caribe continental y Río de la Plata se han obtenido resultados estadísticamente significativos. Es decir, la preferencia por las distintas variantes difiere dentro de las zonas dialectales, por lo cual no es posible llegar a conclusiones generales para las zonas dialectales. Por esa razón, se introducen los países individuales como factor aleatorio en el análisis de regresión mixta con el fin de obtener sus coeficientes respectivos.

## 5.2 El género del tuitero

En cuanto a la posible correlación entre el género del tuitero y la elección de la variante adverbial locativa, los datos de la Tabla 5 muestran que la variante preposicional predomina tanto para los tuiteros femeninos como los masculinos. Las diferencias porcentuales entre el uso de la variante preposicional y la posesiva son ligeramente más pronunciadas para los hombres, quienes prefieren la variante preposicional. Además, la diferencia de la forma adverbial locativa entre tuiteros masculinos y femeninos es estadísticamente significativa, aunque la fuerza de la asociación es muy débil ( $V = 0,097$ ).

Tabla 5. Los complementos de adverbios locativos según el género del tuitero

Género	Complemento		Total
	Preposicional	Posesivo	
Hombre	69,1% (422)	30,0% (189)	611
Mujer	59,6% (514)	40,4% (348)	862
<b>Total</b>	<b>63,5% (936)</b>	<b>36,5% (537)</b>	<b>1473</b>

$\chi^2(1, N = 1473) = 13,345$ ;  $p < 0,001$ ;  $V = 0,097$ .

## 5.3 La persona y el número gramaticales del referente

En lo que concierne a la influencia de la persona y el número gramaticales del referente de la forma, Salgado y Bouzouita (2017) han registrado más variación entre la variante preposicional y la posesiva para los casos con referentes de 1ª y 2ª personas gramaticales, tanto del singular como del plural, mientras que para los

casos de 3ª persona de ambos números observan una clara preferencia por la variante preposicional. En la Tabla 6 se expone la distribución del complemento preposicional y el posesivo según persona y número. Como indica el valor del chi cuadrado, la diferencia observada entre personas y números gramaticales de la forma adverbial en tuits es altamente significativa. Cabe señalar que el valor *V* de Cramér especifica que la fuerza de la asociación entre los factores bajo discusión es intermedia.

Como detalla la Tabla 6, en este estudio de tuits se registra más frecuentemente el pronombre posesivo en las construcciones adverbiales con referentes singulares que con plurales: 39,7% (494/1245) vs. 18,9% (43/228), una diferencia altamente significativa, aunque el valor *V* indica una asociación débil. Dicho de otra manera, es más probable que se encuentren tuits con construcciones del tipo *delante mío* o *delante tuyo* que, por ejemplo, *delante nuestro*. La 1ª persona del singular es la que predomina para el uso posesivo (64,7%, 279/431). La distribución de su equivalente plural indica una preferencia por la variante preposicional, pues solo el 29,7% (30/101) de estos casos plurales aparece con un pronombre posesivo. Si comparamos los casos de 2ª persona del singular y del plural con los de la 1ª persona del plural, observamos que presentan porcentajes muy similares: 32,6% (196/601), 22,2% (6/27) y 29,7% (30/101) respectivamente.<sup>10</sup> El uso adverbial posesivo más bajo se registra para la 3ª persona: un 8,9% (19/213) para el singular y un 7% (7/100) para el plural.<sup>11</sup> Como es de esperar, la diferencia en el comportamiento de las diferentes personas en el singular, así como la para el plural son altamente significativas.<sup>12</sup>

Tabla 6. Los complementos de adverbios locativos según la persona y el número gramaticales del referente

Número y persona gramaticales	Complemento		Total
	Preposicional	Posesivo	
Singular	60,3% (751)	39,7% (494)	1245
1ª persona	35,3% (152)	64,7% (279)	431
2ª persona	67,4% (405)	32,6% (196)	601
3ª persona	91,1% (194)	8,9% (19)	213
Plural	81,1% (185)	18,9% (43)	228
1ª persona	70,3% (71)	29,7% (30)	101
2ª persona	77,8% (21)	22,2% (6)	27
3ª persona	93% (93)	7% (7)	100
<b>Total</b>	<b>63,5% (936)</b>	<b>36,5% (537)</b>	<b>1473</b>

$\chi^2 (5, N = 1473) = 264,12; p < 0,001$  (por número gramatical);  $V = 0,423$ .

$\chi^2 (1, N = 1473) = 35,163; p < 0,001$  (por persona gramatical);  $V = 0,156$

Resumiendo, la 1ª persona del singular es el único contexto que favorece claramente la variante posesiva. Dicho esto, conviene recordar que no contamos con muchas

<sup>10</sup> La diferencia entre estos resultados no es significativa:  $\chi^2 (2, N = 729) = 1,5289; p > 0,05$ .

<sup>11</sup> La diferencia entre estos resultados no es significativa:  $\chi^2 (1, N = 313) = 0,3294; p > 0,05$ .

<sup>12</sup> Para el singular:  $\chi^2 (2, N = 1245) = 209,7693; p < 0,001$ ; para el plural:  $\chi^2 (2, N = 228) = 17,151; p < 0,001$ .

ocurrencias para la 2ª persona del plural, por lo cual habría que interpretar este resultado con cierta precaución.

Estos resultados para Twitter concuerdan en grandes líneas con las observaciones de Salgado y Bouzouita (2017) sobre el uso de los complementos posesivos en el español peninsular, puesto que en su estudio también se registran los porcentajes más altos del uso posesivo para la 1ª persona singular y plural: 47,9%, (56/117) y 44,8% (13/29) respectivamente (vs. 64,7% (279/431), y 29,7% (30/101) en nuestro corpus de tuits). No obstante, es de notar que las diferencias entre la 1ª persona del singular y del plural son mucho más pronunciadas en nuestro corpus de tuits, ya que la frecuencia de uso de la variante posesiva con el singular es más del doble que su homólogo plural.

Salgado y Bouzouita (2017) observan una frecuencia más alta en su corpus oral peninsular para la variante posesiva con la 2ª persona singular que en este estudio de tuits, a saber 57,1% (20/35) vs. 32,6% (196/601) (no registran ninguna observación para la 2ª persona del plural, ni de la variante preposicional ni de la posesiva). Sin embargo, ambos estudios coinciden en que para los casos de 3ª persona gramatical, tanto del singular como del plural, la variante preposicional es palmariamente más prominente, ya que se documentan frecuencias por arriba del 90%. En concreto, el 96,3% (77/80) y el 100% (30/30) de los casos de 3ª persona singular y plural aparecen con la variante preposicional en el corpus oral peninsular analizado por Salgado y Bouzouita (2017) frente al 91,1% (194/213) y el 93% (93/100) de los tuits examinados en el presente estudio.

Resulta difícil comparar nuestros resultados de la variante adverbial en tuits con los de Santana Marrero (2014) para los medios de comunicación digitales, puesto que esta autora se limita a presentar tasas porcentuales de las diferentes personas gramaticales dentro del paradigma de los posesivos, sin compararlas con sus homólogos preposicionales. No obstante, según nuestros cálculos, el 52% (97/187) de sus datos de uso posesivo corresponde a la 3ª persona singular. Este resultado es muy llamativo, ya que se opone diametralmente a las observaciones de la presente investigación. Contrariamente a esta frecuencia alta para la 3ª persona del singular, solo un 2,13% (4/187) de su corpus concierne a la 3ª persona del plural, un porcentaje bajo que está en consonancia con el del presente análisis de datos de Twitter (7%, 7/100). En cuanto a la 1ª persona del singular, Santana Marrero (2014) también identifica un uso considerable de la variante posesiva (31,02%, 58/187). Sin embargo, se encuentran frecuencias bajas de la variante posesiva tanto para la 1ª persona del plural como en la 2ª persona del singular y plural, lo cual no está en sintonía con los resultados presentados aquí.

Aún queda por resolver si la persona gramatical contribuye significativamente a la variación estudiada. Los resultados presentados en la Tabla 7 confirman que, efectivamente, la diferencia entre las tres personas gramaticales es altamente significativa ( $p < 0,001$ ), y el valor  $V$  indica que hay una fuerza intermedia en la asociación.<sup>13</sup> Más específicamente, se usa la variante posesiva primariamente con

---

<sup>13</sup> También hay significación estadística entre cada persona:

los referentes de la 1ª persona: en el 58,1% (309/532) de los casos de 1ª persona aparece la construcción adverbial locativa con posesivo. En comparación con la 1ª persona, la tasa porcentual del uso posesivo disminuye considerablemente para la 2ª persona, que prefiere la variante preposicional (67,8%, 426/628), así como la 3ª persona, para la cual se registra el porcentaje más bajo del uso adverbial posesivo, a saber, un 8,3% (26/313).

Tabla 7. Los complementos de adverbios locativos según la persona gramatical del referente

Persona gramatical	Complemento		Total
	Preposicional	Posesivo	
1ª persona	41,9% (223)	58,1% (309)	532
2ª persona	67,8% (426)	32,2% (202)	628
3ª persona	91,7% (287)	8,3% (26)	313
<b>Total</b>	63,5% (936)	36,5% (537)	1473

$\chi^2 (2, N = 1473) = 219,46: p < 0,001; V = 0,386.$

En conclusión, se puede constatar que tanto el número como la persona gramatical influyen en la elección de variante, aunque, a juzgar por el coeficiente *V*, la persona influye más (*V* = 0,386 para la persona vs. *V* = 0,156 para el número). Se investigará este hallazgo más profundamente en el análisis multivariado, así como la posible interacción entre estos factores.

#### 5.4 El tipo de adverbio

Con respecto de la posible influencia del tipo de adverbio, como se ha señalado en el apartado 2.2.1, verificaremos cuantitativamente la validez de la categorización propuesta por Almela Pérez (1991). Según este autor, los adverbios pertenecientes al denominado Sector A se usan más frecuentemente en contextos posesivos que los del Sector B. En lo que sigue, se investiga esta hipótesis mediante un análisis distribucional de los sectores en relación con la variante adverbial locativa (Tabla 8).

El análisis de los diferentes usos por adverbio, presentado en la Tabla 8, revela lo siguiente. Primero, la escala de frecuencia de los adverbios que más frecuentemente escogen el pronombre posesivo consiste en orden decreciente *adelante* (80,3%, 53/66), *arriba* (73,7%, 42/57), *enfrente* (62,5%, 80/128), *atrás* (60,6%, 77/127), *adentro* (57,9%, 11/19), *alrededor* (53,2%, 25/47) y *encima* (47,2%, 68/144). Estos resultados van en la dirección casi totalmente contraria de los presentados por Almela Pérez (1991), visto que 4 de los 6 adverbios que sobrepasan el 50% de combinación con un pronombre posesivo pertenecen al denominado Sector B, los adverbios que, según este autor, aparecen menos frecuentemente con posesivos. En lo que concierne a los adverbios sin el prefijo *a-* (del denominado Sector A), nuestros resultados sí concuerdan en gran parte con los

1ª persona vs. 2ª persona:  $\chi^2 (1, N = 1160) = 78,494: p < 0,001;$

2ª persona vs. 3ª persona:  $\chi^2 (1, N = 941) = 64,7694: p < 0,001;$

1ª persona vs. 3ª persona:  $\chi^2 (1, N = 845) = 204,0505: p < 0,001.$

presentados por Santana Marrero (2014) y Salgado y Bouzouita (2017), dado que en su estudio los adverbios *enfrente* y *alrededor* son los más frecuentemente combinados con pronombres posesivos (cf. la Tabla 3 en la sección 2.2.1).

Tabla 8. Los complementos de adverbios locativos según el sector del adverbio

Sectores	Adverbio	Complemento		Total
		Preposicional	Posesivo	
A <i>p</i> < 0,001	<i>alrededor</i>	46,8% (22)	53,2% (25)	47
	<i>bajo</i>	50% (1)	50% (1)	2
	<i>cerca</i>	72,4% (160)	27,6% (61)	221
	<i>debajo</i>	100% (10)	0% (0)	10
	<i>delante</i>	59% (62)	41% (43)	105
	<i>dentro</i>	83,3% (189)	16,7% (38)	227
	<i>detrás</i>	84,7% (124)	17,3% (26)	150
	<i>encima</i>	52,8% (76)	47,2% (68)	144
	<i>enfrente</i>	37,5% (48)	62,5% (80)	128
	<i>fuera</i>	100% (36)	0% (0)	36
	<i>lejos</i>	92,2% (119)	7,8% (10)	129
	<b>Total</b>	<b>70,6% (847)</b>	<b>29,4% (352)</b>	<b>1199</b>
B <i>p</i> < 0,01	<i>abajo</i>	0% (0)	100% (2)	2
	<i>adelante</i>	19,7% (13)	80,3% (53)	66
	<i>adentro</i>	42,1% (8)	57,9% (11)	19
	<i>afuera</i>	100% (3)	0% (0)	3
	<i>arriba</i>	26,3% (15)	73,7% (42)	57
	<i>atrás</i>	39,4% (50)	60,6% (77)	127
	<b>Total</b>	<b>32,5% (89)</b>	<b>67,5% (185)</b>	<b>274</b>
<b>Total</b>		<b>63,5% (936)</b>	<b>36,5% (537)</b>	<b>1473</b>

$\chi^2(1, N = 1473) = 136,87$ ; *p* < 0,001 (por sector); *V* = 0,307.

Aunque, en general, los miembros del Sector A favorecen el complemento preposicional (70,6%, 847/1199 vs. 29,4%, 352/1199 para la variante posesiva), observamos una preferencia por el posesivo en el caso de *enfrente* (62,5%, 80/128) y una distribución bastante equilibrada para *alrededor*. Otros, al contrario, exhiben una clara preferencia por el complemento preposicional, como por ejemplo *fuera* (100%, 36/36), *lejos* (92,2%, 119/129), *dentro* (83,3%, 189/227), *detrás* (84,7%, 124/150) y *cerca* (72,4%, 160/221), siguiendo por tanto la tendencia general de este sector. Los adverbios del Sector B, contrariamente a la tendencia general observada para los del Sector A, tienden a preferir el complemento posesivo (67,5%, 185/274 vs. 32,5% 89/274 para el uso preposicional). Además, los adverbios del Sector B parecen vacilar menos que los de A, ya que para el uso de la variante posesiva casi todos muestran porcentajes entre el 57,9% (11/19 para *adentro*) y el 80,3% (53/66 para *adelante*), conforme a la tendencia general del Sector B. La única posible excepción podría ser *afuera*, adverbio para el cual no se ha recopilado ninguna observación con la variante posesiva; no obstante, no consideramos que se haya logrado representatividad, visto que solamente contamos con 3 ocurrencias de este. Lo mismo vale para el adverbio *abajo*, para el cual se registran apenas dos observaciones.

Puesto que el resultado del chi cuadrado es estadísticamente significativo ( $p < 0,001$ ), se confirma que el comportamiento de los sectores de adverbios difiere considerablemente el uno del otro. No obstante, como indica el valor  $V$ , la fuerza de la asociación no es muy fuerte ( $V = 0,307$ ). Asimismo, la prueba exacta de Fisher revela que el comportamiento de los diferentes adverbios con respecto a la elección del complemento es muy heterogéneo dentro de los sectores (cf. los valores  $p$  debajo de los títulos de los sectores). En otras palabras, este resultado indica que la categorización de Almela Pérez (1991) no encaja adecuadamente la distribución observada.

Así pues, a primera vista no parece que la categorización propuesta por Almela Pérez (1991: 436) sea muy plausible para los datos aquí examinados. No obstante, como se ha señalado en el apartado 2.2.1, la preferencia por los diferentes tipos de adverbios también podría ser condicionada por factores diatópicos visto que en las variedades hispanoamericanas se está perdiendo la distinción semántica entre los adverbios “direccionales” y los “no direccionales” (cf. Meilán García 1998: 18, RAE/ASALE 2009: 2307). Además, no hay que olvidar que la propuesta de Almela Pérez (1991) se formuló con base en datos peninsulares. Considerando tales circunstancias, conviene examinar los sectores en función de la distinción diatópica entre España y los demás países hispanohablantes (Tabla 9).

Tabla 9. Los complementos de adverbios locativos según el sector del adverbio y la distribución diatópica

Continentes y sector del adverbio	Complemento		Total
	Preposicional	Posesivo	
Hispanoamérica	62,1% (777)	37,9% (475)	85% (1252)
Sector A	68,2% (688)	29,8% (291)	78,2% (979)
Sector B	32,7% (89)	67,3% (184)	21,8% (273)
España	71,9% (159)	28,1% (62)	15% (221)
Sector A	72,3% (159)	27,7% (61)	99,5% (220)
Sector B	0% (0)	100% (1)	0,5% (1)
<b>Total</b>	<b>63,5% (936)</b>	<b>36,5% (537)</b>	<b>1473</b>

Como se señala en la Tabla 9, se ha encontrado un solo ejemplo peninsular del Sector B (0,5%, 1/221), el cual aparece con un complemento posesivo. A consecuencia de esta falta de casos para el Sector B, la afirmación de Almela Pérez (1991: 436) de que los adverbios del Sector A se combinan más frecuentemente con un posesivo no se puede confirmar para las variedades peninsulares. Para las hispanoamericanas, al contrario, los adverbios del Sector B se combinan con complementos posesivos en el 67,3% (184/273) de los casos, en comparación con los del Sector A cuya combinatoria con el posesivo llega a tan solo 29,8% (291/979). En consecuencia, estas observaciones para los datos hispanoamericanos no validan la conclusión de Almela Pérez (1991: 436).

En lo que concierne a la distribución de los dos complementos en el Sector A, una prueba de chi cuadrado revela que no hay diferencia estadísticamente significativa entre España y Hispanoamérica ( $\chi^2 (1, N = 1199) = 0,26; p > 0,05; V$

= 0,0171). En otras palabras, a pesar de la diferencia significativa entre los continentes en su totalidad (cf. la Tabla 4), no parece haber diferencia diatópica en cuanto a los adverbios del Sector A.

Así, considerando el claro desequilibrio en la distribución entre los datos de España y los de América provenientes del Sector B, junto con el hecho de que la prueba de Fisher indicara heterogeneidad estadísticamente significativa dentro de los sectores respectivos (Tabla 8), a continuación se examinará el efecto de los adverbios individualmente en vez de partir de una categorización *a priori*.

### 5.5 Análisis de bosque aleatorio

Ahora que se ha introducido el análisis distribucional, se presenta el análisis del bosque aleatorio (*random forest*), seguido de un árbol de inferencia condicional, para explorar la variación entre los dos tipos de complemento de los adverbios locativos y los factores que determinan esta variación.<sup>14</sup>

Como explican Tagliamonte y Baayen (2012: 159), la metodología del bosque aleatorio constituye una innovación bastante reciente en el ámbito estadístico. Se trata de un tipo de prueba no paramétrica desarrollada por Breiman (2001) que ha ido ganando terreno en varias disciplinas, entre ellas la (socio)lingüística. Los análisis de bosque aleatorio y de regresión logística comparten el mismo objetivo, esto es, predecir cuál de las variantes investigadas es la más probable, dado un conjunto de factores. Sin embargo, los pasos realizados por estos dos métodos de análisis para cumplir con su objetivo no son idénticos. El método del bosque aleatorio se basa en un algoritmo que construye varios árboles de inferencia condicional a diferencia de los modelos logísticos, cuyo resultado se basa en fórmulas matemáticas (Tagliamonte y Baayen 2012: 159). Estos árboles de inferencia condicional analizan individualmente un subgrupo de los datos, a partir del cual proporcionan estimaciones de la probabilidad del valor de la variante mediante una serie de preguntas binarias sobre los valores de los factores. Luego se comparan estas predicciones con los datos reales del resto de la muestra, una comparación a partir de la cual se constata la precisión de la clasificación del árbol. Así pues, cada árbol constituye un voto para el evento más probable. El bosque aleatorio final es el resultado de la alternativa que ha recibido más votos (Tagliamonte y Baayen 2012: 161). Como indican Tagliamonte y Baayen (2012: 161), existen varias ventajas del método del bosque aleatorio. Una de ellas es la opción de poder trabajar con pocos datos pero con muchos factores, visto que las predicciones de los árboles de inferencia condicional solamente se basan en un subgrupo de la totalidad de los datos. Asimismo, a diferencia de los modelos de regresión, los bosques aleatorios no se ven afectados por la posible colinealidad de factores. Por último, también es posible analizar datos con casillas vacías. El producto final de este método indica la importancia condicional de las variables,

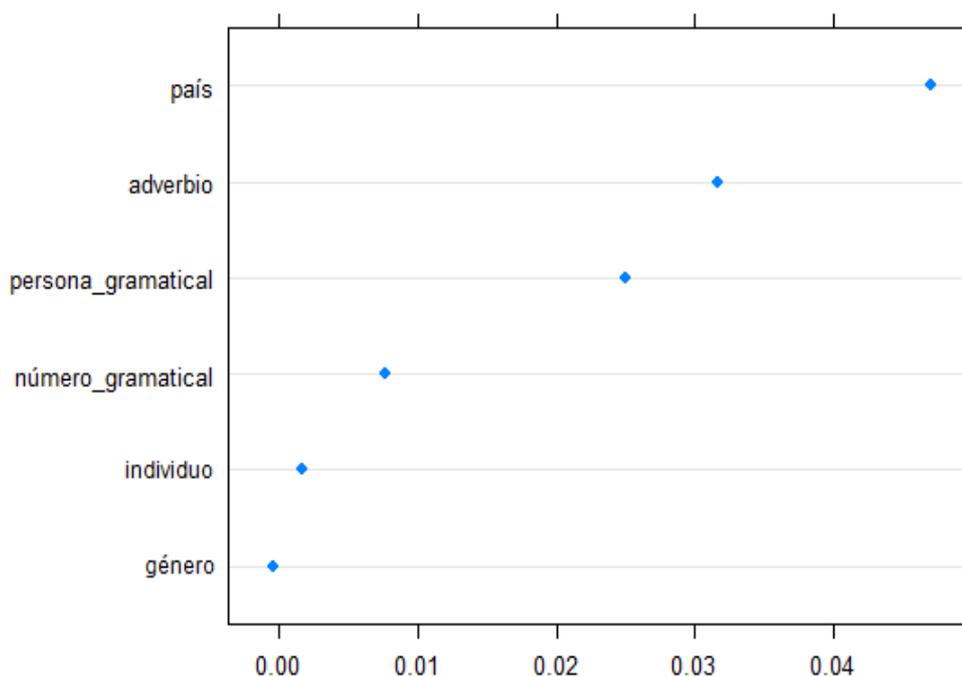
---

<sup>14</sup> Para estos análisis se ha usado el paquete *party* en R, usando las funciones *cforest* (Hothorn et al. 2006, Strobl et al. 2007, Strobl et al. 2008) y *cree* (Hothorn, Hornik y Zeileis 2006).

siendo los factores más influyentes posicionados hacia la derecha, mientras que los menos influyentes se hallan cerca del cero en la margen izquierda.

En la Figura 1 se presentan los resultados de la aplicación del método del bosque aleatorio a los datos relacionados al uso de la construcción adverbial locativa y sus complementos en tuits. El índice  $C$  de concordancia presenta un valor considerablemente alto ( $C = 0,95$ ), de ahí que el bosque aleatorio elaborado constituye un modelo muy fuerte en predecir la variación.

Figura 1. La importancia de las variables basada en la permutación condicional de todos los factores.  $C = 0,95$



De manera interesante, los resultados apuntan a que el país constituye el factor predominante en condicionar la alternancia, aunque el bosque aleatorio no permite determinar qué países favorecen o desfavorecen la variante posesiva. El segundo factor más importante para la explicación de la variación estudiada es el adverbio que precede al complemento, seguido por dos factores intralingüísticos adicionales, la persona y el número gramaticales, aunque este último en menor medida. Como se puede observar, se ha incluido el individuo como un factor de interés, ya que algunos tuiteros en nuestro corpus han producido varios tuits. Sin embargo, el bosque aleatorio indica que la variación individual entre los tuiteros no es un factor determinante para condicionar el tipo de complemento de los adverbios locativos. Lo mismo vale para el género del tuitero, que constituye el factor que menos influye en la elección de variante en nuestro corpus.

Aunque el país constituye el factor que mejor predice la variación investigada, es evidente que son sobre todo los factores intralingüísticos los que la condicionan. En el árbol de inferencia condicional, que se presenta en la Figura 2 del apartado 5.6 que sigue, se ilustra más en detalle cómo estos factores predicen la variación en su conjunto, así como valores  $p$  para cada nivel del árbol.<sup>15</sup>

### 5.6 Análisis de árbol de inferencia condicional

Como ya explicamos, la funcionalidad de los árboles de inferencia condicional consiste en que calculan la probabilidad de obtener el valor de la variable dependiente (en este caso, el complemento preposicional o el complemento posesivo) (cf. Tagliamonte y Baayen 2012: 159). Esta probabilidad se estima a partir de clasificaciones binarias sobre los valores de los distintos factores introducidos en el modelo. Los árboles clasifican las observaciones y las dividen en subgrupos cuando esta división es estadísticamente significativa. La ventaja de este tipo de análisis es que deja ver claramente cómo los factores considerados funcionan en su conjunto y, además, modela interacciones que muchas veces son difíciles de identificar y especificar en modelos lineales generalizados (Tagliamonte y Baayen 2012: 163).

Como se puede observar en la Figura 2<sup>16</sup>, el árbol de inferencia condicional indica un índice  $C$  de concordancia algo menor ( $C = 0,92$ ) que el del bosque aleatorio ( $C = 0,95$ ), aunque, con un valor de esa magnitud, sigue siendo un modelo muy adecuado para predecir la variación estudiada. Antes de todo, este árbol de inferencia condicional indica, al igual que el bosque aleatorio (cf. Figura 1), que el país es el factor más significativo en el modelo ( $p < 0,001$ ). En efecto, la primera división del árbol se hace entre los países que usan la variante posesiva más frecuentemente, los cuales aparecen a la izquierda, a saber, Argentina, Bolivia, Chile, Ecuador, Panamá, Paraguay, Perú y Uruguay. Los que presentan las frecuencias de uso posesivo más bajas figuran en la parte derecha del árbol: estos son Costa Rica, Colombia, EE.UU., España, Guatemala, Honduras, México, Nicaragua, Puerto Rico, República Dominicana, El Salvador y Venezuela.

En cuanto al segundo factor más influyente, el árbol de inferencia condicional identifica varias tendencias. Para empezar, el segundo nodo de la parte izquierda indica que la persona gramatical es el segundo factor más influyente para determinar la variación observada en Argentina, Bolivia, Chile, Ecuador, Panamá,

---

<sup>15</sup> Los significados de las abreviaturas usadas en el árbol de inferencia condicional (Figura 2) son los siguientes: AB = abajo; ADEL = adelante; ADEN = adentro; AFU = afuera; ALR = alrededor; ARR = arriba; ATR = atrás; BA = bajo; CE = cerca; DEB = debajo; DEL = delante; DEN = dentro; DET = detrás; ENC = encima; ENF = enfrente; FU = fuera; LEJ = lejos; ARG = Argentina; BOL = Bolivia; C.RI. = Costa Rica; CHI = Chile; COL = Colombia; ECU = Ecuador; EE.UU. = Estados Unidos; ES = España; GUAT = Guatemala; HON = Honduras; MX = México; NIC = Nicaragua; P.RI. = Puerto Rico; PAN = Panamá; PAR = Paraguay; R.DOM. = República Dominicana; PER = Perú; SALV = El Salvador; URU = Uruguay; VEN = Venezuela.

<sup>16</sup> En los nodos, el gris claro representa la proporción preposicional, mientras que el gris oscuro representa la proporción posesiva.

Paraguay, Perú y Uruguay ( $p < 0,001$ ). Más específicamente, se presenta una subdivisión en la que, por un lado, se encuentra la 3ª persona, y, por otro, las 1ª y 2ª personas. Como también ha indicado el análisis distribucional (cf. la Tabla 7), la 3ª persona se usa menos frecuentemente con el complemento posesivo. Aun así, el árbol de inferencia condicional logra arrojar luz sobre el hecho de que ciertos adverbios evidencian variación estadísticamente significativa incluso en el ámbito de la 3ª persona. En concreto, el nodo 4 comprende a *adentro*, *delante*, *encima* y *enfrente*, los adverbios que hasta con la 3ª persona se encuentran en un estado de variación ( $p < 0,01$ ), puesto que se combinan en casi un 40% de los casos de la 3ª persona con la variante posesiva. Los adverbios del nodo 5, es decir *adelante*, *afuera*, *alrededor*, *atrás*, *cerca*, *debajo*, *dentro*, *detrás*, *fuera* y *lejos*, se usan sumamente poco con la variante posesiva en casos con referentes de 3ª persona. Estos resultados constituyen un hallazgo novedoso que no ha sido descrito previamente.

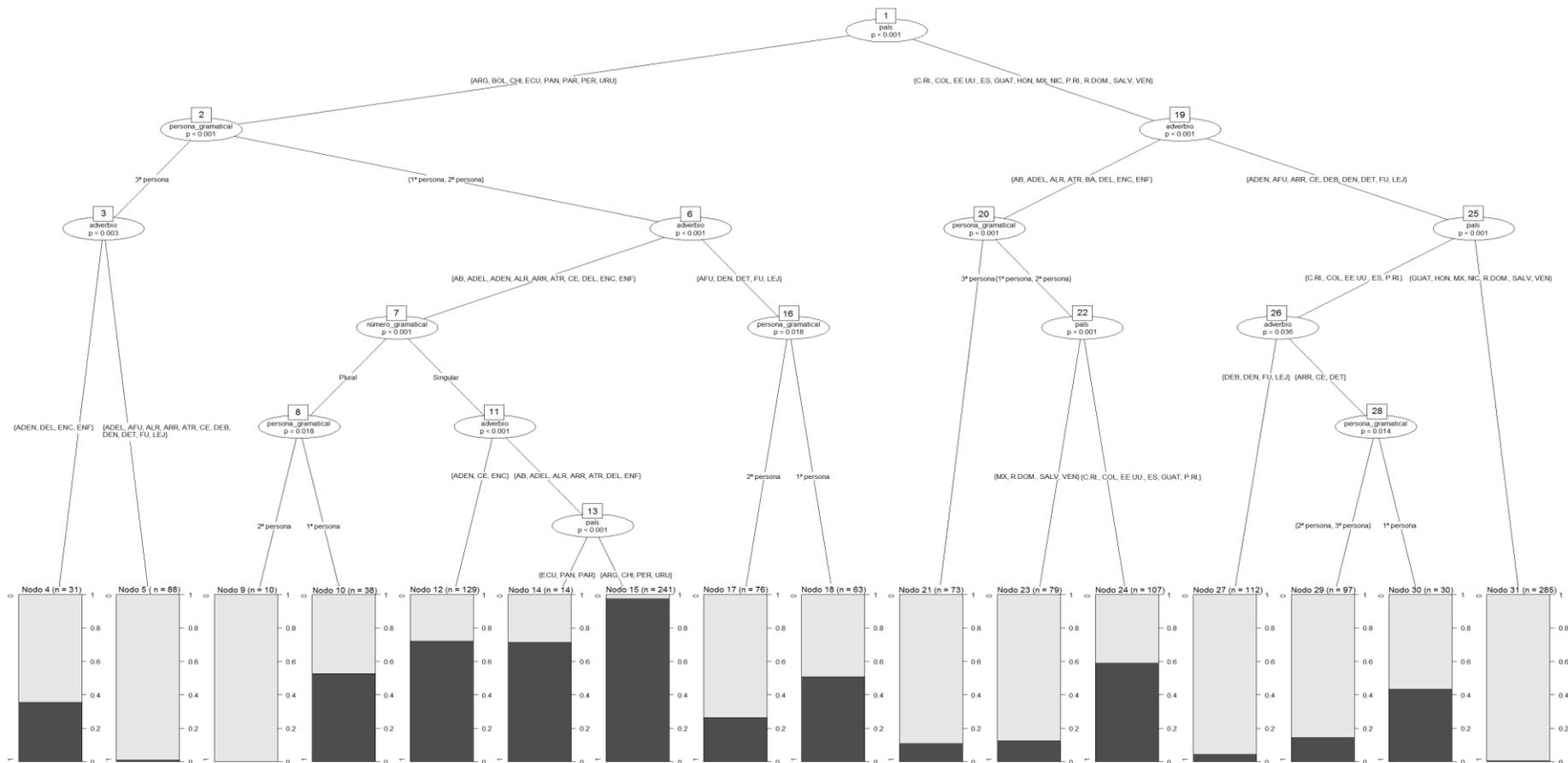
En lo que concierne a las 1ªs y 2ªs personas en los mismos países, estas también vienen condicionadas por el adverbio (nodo 6;  $p < 0,001$ ). Los adverbios *abajo*, *adelante*, *adentro*, *alrededor*, *arriba*, *atrás*, *cerca*, *delante*, *encima* y *enfrente* vienen condicionados por el número gramatical (nodo 7;  $p < 0,001$ ). Los casos plurales de las observaciones de estos adverbios vienen, a su vez, determinados por la persona gramatical (nodo 8;  $p = 0,018$ ). Efectivamente, se observa que, en los casos plurales de la 2ª persona con los adverbios mencionados anteriormente, no se registran observaciones del uso de la variante posesiva (nodo 9); contrariamente, cuando aparece la 1ª persona del plural con estos adverbios, se usa la variante posesiva en más del 50% de los casos (nodo 10).

Con respecto a los casos del singular, los nodos 11 y 12 ( $p < 0,001$ ) indican que, en Argentina, Bolivia, Chile, Ecuador, Panamá, Paraguay, Perú y Uruguay, los adverbios *adentro*, *cerca* y *encima* se usan en aproximadamente un 70% de los casos con la variante posesiva. Así también es el caso para los adverbios *abajo*, *adelante*, *alrededor*, *arriba*, *atrás*, *delante* y *enfrente* en Ecuador, Panamá y Paraguay (nodos 13 y 14;  $p < 0,001$ ). Sorprendentemente, el árbol de inferencia condicional muestra asimismo que estos adverbios se usan casi en un 100% de los casos singulares con el complemento posesivo en Argentina, Chile, Perú y Uruguay (nodo 15).

La variación en los complementos de los adverbios *afuera*, *dentro*, *detrás*, *fuera* y *lejos* viene condicionada por la persona gramatical (nodo 16;  $p = 0,018$ ): la variante posesiva aparece con la 2ª persona en casi un 30% de los casos (nodo 17) y con la 1ª persona en aproximadamente un 50% de los casos (nodo 18). Los resultados de los nodos 10 y 18, que demuestran la prevalencia posesiva para los casos de 1ª persona, están en sintonía con los del análisis distribucional (cf. la Tabla 7), así como con los de Salgado y Bouzouita (2017).

Los tuiteros de Costa Rica, Colombia, los EE.UU., España, Guatemala, Honduras, México, Nicaragua, Puerto Rico, la República Dominicana, El Salvador

Figura 2. El árbol de inferencia condicional de la variación de complemento del adverbio locativo.  $C = 0,92$ .



y Venezuela usan la variante posesiva menos que los de los demás países hispanohablantes incluidos en este estudio. A pesar de la predominancia de la variante preposicional, hay también en estos países cierta variación. Recuérdese que, entre los países favorecedores de la variante posesiva, la variación en los complementos viene principalmente determinada por la persona gramatical. En los países propensos al uso del complemento preposicional, al contrario, el adverbio condiciona primordialmente la variación (nodo 19;  $p < 0,001$ ).

Para los adverbios *abajo*, *adelante*, *alrededor*, *atrás*, *bajo*, *delante*, *encima* y *enfrente* la persona gramatical es el factor más influyente (nodo 20;  $p < 0,001$ ). Además, se encuentra la misma tendencia que antes, puesto que para la 3ª persona aparece con muy poca frecuencia el pronombre posesivo con estos adverbios (nodo 21). Sin embargo, en construcciones con estos mismos adverbios cuyo referente es una 1ª o 2ª persona, el país determina la variación (nodo 22;  $p < 0,001$ ): en México, la República Dominicana, El Salvador y Venezuela, la frecuencia del posesivo es sumamente baja, llegando tan solo a aproximadamente un 10% (cf. el nodo 23). En Costa Rica, Colombia, los EE.UU., España, Guatemala y Puerto Rico, el porcentaje del uso con posesivo llega a constituir casi un 60% de las observaciones (nodo 24).

En lo que concierne a los adverbios *adentro*, *afuera*, *arriba*, *cerca*, *debajo*, *dentro*, *detrás*, *fuera* y *lejos*, la variación en el complemento también está condicionada por el país (nodo 25;  $p < 0,001$ ). Por un lado, los tuiteros de Guatemala, Honduras, México, Nicaragua, la República Dominicana, El Salvador y Venezuela evidencian un rechazo casi total de la variante posesiva (nodo 31). Por otro lado, existe más variación en Costa Rica, Colombia, los EE.UU., España y Puerto Rico, que también se ve afectada por el adverbio (nodo 26;  $p = 0,036$ ). Más específicamente, los adverbios *debajo*, *dentro*, *fuera* y *lejos* casi no se usan con la variante posesiva en dichas comunidades (nodo 27), pues tal variante constituye aproximadamente un 5% de los casos. Por el contrario, los adverbios *arriba*, *cerca* y *detrás* muestran más variabilidad en Costa Rica, Colombia, los EE.UU., España y Puerto Rico, pero la variación en el complemento viene determinada en gran medida por la persona gramatical (nodo 28;  $p = 0,014$ ). Las 2ª y 3ª personas gramaticales llegan a aproximadamente un 15% (nodo 29), mientras que la 1ª persona sobrepasa el 40% de la variante posesiva (nodo 30).

En resumen, el árbol de inferencia condicional señala que el factor más influyente en determinar la variación es el país. Asimismo, apunta a una red enormemente compleja de interacciones virtualmente imposibles de detectar con análisis meramente descriptivos. Así pues, este método de análisis logra identificar sistemas y reglas de variación que no han sido descritos por la bibliografía anteriormente; a saber, la prominencia de ciertos adverbios (p.ej. *adentro*, *delante*, *encima* y *enfrente*) con la variante posesiva con referentes de la 3ª persona en algunas variedades, pese a que la 3ª persona ha sido indicada por el análisis distribucional y el estudio de Salgado y Bouzouita (2017) como un factor rechazador del posesivo. Asimismo, el árbol de inferencia condicional también apunta a que existen frecuencias extremadamente altas de la variante posesiva, pues llega esencialmente al 100%, cuando esta se halla bajo ciertas circunstancias, más específicamente, en

casos de referentes de 1ª y 2ª personas de singular en combinación con los adverbios *abajo*, *adelante*, *alrededor*, *arriba*, *atrás*, *delante* y *enfrente* en Argentina, Chile, Perú y Uruguay.

### 5.7 Análisis de regresión logística de efectos mixtos

En este apartado se llevará a cabo un análisis mixto de regresión logística para evaluar el impacto de los factores sobre la probabilidad de que los tuits exhiban la variante posesiva.<sup>17</sup> Los factores que se han considerado como efectos fijos en el modelo son los discutidos anteriormente: género, persona y número. También se considera la interacción entre persona y número (persona \* número en la Tabla 10). A continuación, se expone el tratamiento estadístico de los factores adverbio y país.

Como ya se ha señalado en el apartado 5.4, la categorización de los adverbios propuesta por Almela Pérez (1991) no parece ser válida para los datos examinados en el presente trabajo. Recuérdese que una prueba de Fisher ha señalado una heterogeneidad significativa dentro de los sectores propuestos por este autor; esto es, los adverbios dentro de los sectores muestran variabilidad en cuanto al complemento preferido (cf. la Tabla 8). Por eso, se ha ignorado todo tipo de clasificaciones *a priori* e introducido los ítems como un factor aleatorio en un modelo mixto de regresión con la finalidad de determinar los efectos individuales de los adverbios sobre el uso de la variante posesiva. Asimismo, como se ha indicado en el apartado 5.1, la elección del complemento del adverbio locativo tampoco es uniforme dentro de las zonas dialectales investigadas. Por esta razón, también se incluirán los países como un factor aleatorio para así poder determinar los coeficientes de los países respectivos y, por tanto, qué países favorecen vs. desfavorecen la variante posesiva (cf. el razonamiento técnico detrás esta decisión en el apartado 4.3). Por último, como se ha mencionado en el apartado 5.5, el individuo también constituye un factor aleatorio porque en algunos casos se han registrado varios tuits del mismo individuo. Según estudios anteriores (p.ej. Tagliamonte y Baayen 2012), incluso un número muy bajo de medidas repetidas puede afectar el resultado del análisis, si no se incluye el individuo como un factor aleatorio. Pasemos ahora a los resultados del modelo mixto de regresión logística, proporcionados en la Tabla 10.

Como se señala al final de la Tabla 10, el índice *C* de concordancia del modelo de regresión logística es 0,92, que es el mismo valor que el del árbol de inferencia condicional (Figura 2), pero inferior al del bosque aleatorio (Figura 1; 0,95). Como muestra la Tabla 10, el factor género no es significativo en el modelo para determinar la variación en el complemento del adverbio locativo. Este resultado concuerda con el de los análisis anteriores, pues el bosque aleatorio coloca este factor hacia la margen izquierda (Figura 1), mientras que el árbol de inferencia condicional ni siquiera lo incluye en el modelo (Figura 2). Por lo que concierne al factor persona, se puede identificar que no hay ninguna diferencia estadísticamente significativa entre la 1ª y la 2ª persona ( $p > 0,05$ ).

---

<sup>17</sup> Se ha usado el paquete *lme4* en R.

Tabla 10. Modelo mixto de regresión logística<sup>18</sup>

Variables	Nivel	Beta	E.E.	Z	P
(Constante)		-1,82079	0,57303	-3,177	< 0,01 **
<b>Género</b>	Hombre	Nivel de referencia			
	Mujer	-0,05939	0,16511	-0,360	> 0,05
<b>Persona</b>	1 <sup>a</sup>	Nivel de referencia			
	2 <sup>a</sup>	-0,63729	0,60508	-1,053	> 0,05
	3 <sup>a</sup>	-1,90717	0,52057	-3,664	< 0,001 ***
<b>Número</b>	Plural	Nivel de referencia			
	Singular	1,97398	0,32362	6,100	< 0,001 ***
<b>Persona * Número</b>	1 <sup>a</sup> * Plural	Nivel de referencia			
	2 <sup>a</sup> * Singular	-0,36583	0,63104	-0,580	> 0,05
	3 <sup>a</sup> * Singular	-1,74804	0,60392	-2,984	< 0,01 **
<b>Evaluación del modelo</b>	AIC		1121,9		
	índice C de concordancia		0,92		
	dxy de Somer		0,84		

No obstante, por el valor negativo del coeficiente beta, es posible notar que la 2<sup>a</sup> persona desfavorece la variante posesiva más que la 1<sup>a</sup> persona (beta = -0,63729). En cuanto a la 1<sup>a</sup> y la 3<sup>a</sup> persona, la diferencia es altamente significativa ( $p < 0,001$ ). El coeficiente beta es -1,90717, lo que indica que la 3<sup>a</sup> persona desfavorece fuertemente la variante posesiva en comparación con la 1<sup>a</sup> persona. Este resultado está en sintonía con los presentados por Salgado y Bouzouita (2017) quienes identifican que se presenta más variación con la 1<sup>a</sup> y 2<sup>a</sup> personas, mientras que la 3<sup>a</sup> persona se usa casi únicamente con el complemento preposicional.

El valor  $p$  del factor número apunta a que la diferencia entre el plural y el singular es altamente significativa ( $p < 0,001$ ). Además, el coeficiente beta positivo para el singular nos lleva a concluir que favorece palmariamente la variante posesiva en comparación con el plural (beta = 1,97398). Aunque este factor no se examina explícitamente y/o separadamente de la persona gramatical en los trabajos de Santana Marrero (2014) y Salgado y Bouzouita (2017), nuestros resultados concuerdan parcialmente con los datos expuestos por Santana Marrero (2014), quien registra más posesivos singulares que plurales. Sin embargo, dicha autora no presenta ninguna comparación con el complemento preposicional, por lo cual se presentan problemas metodológicos a la hora de relacionar sus resultados con los aquí presentados. No obstante, en lo que concierne a la influencia del número, como el valor  $V$  de Cramér calculado a partir de los datos de la Tabla 6 indica una fuerza

<sup>18</sup> Beta: el coeficiente de la constante del modelo nulo, también llamado el *intercepto*.

E.E.: el error estándar del coeficiente.

Z: El número de desviaciones estándares del promedio.

P: el valor  $p$  calculado a partir de la estadística Z.

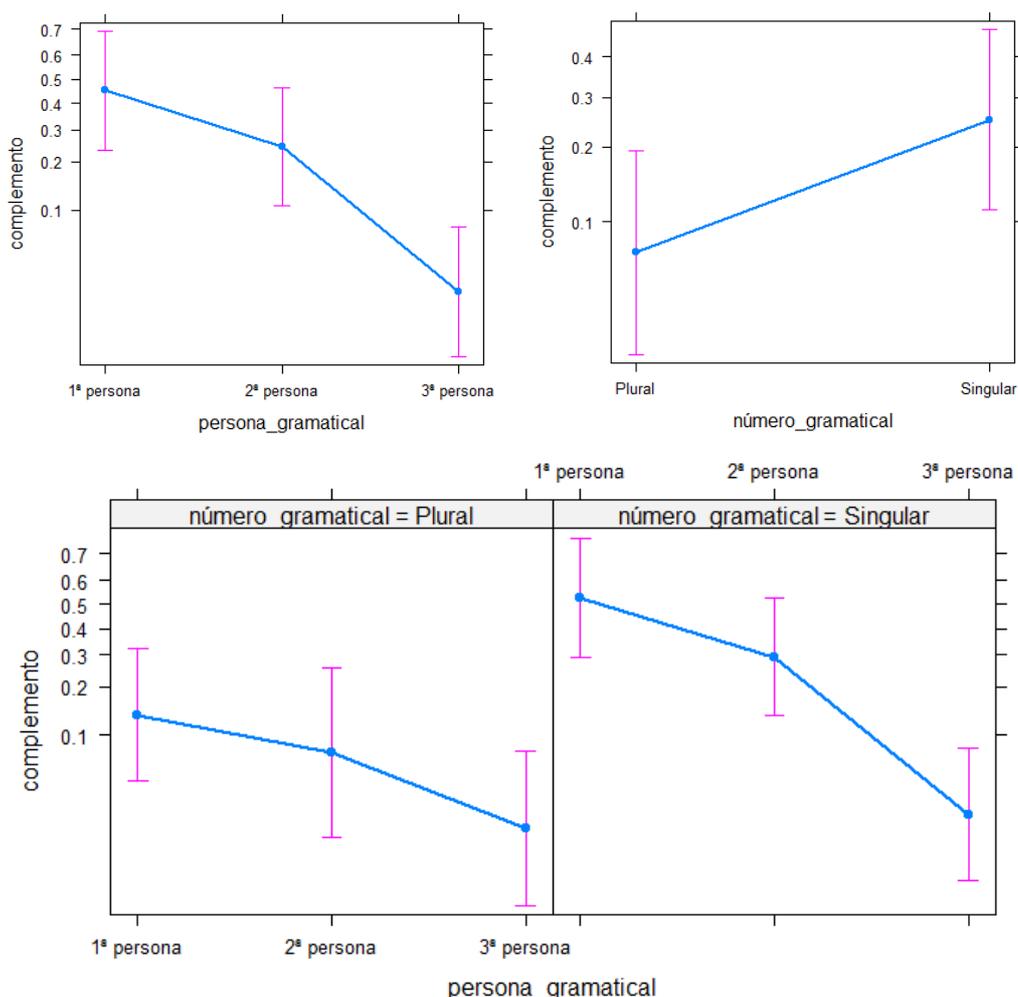
AIC: *Akaike information criterion*. Una estimación de la calidad del modelo estadístico.

Índice C de concordancia: una medida para la bondad de ajuste usada en modelos de regresión logística.

dxy de Somer: una medida de asociación ordinal entre dos variables dependientes.

bastante débil entre estas dos variables (nota 9:  $V = 0,156$ ), se puede considerar que este resultado se debe a cierta interacción con la persona gramatical. Efectivamente, en comparación con la 1ª persona de plural, la probabilidad de usar el posesivo con un adverbio locativo disminuye si el referente es de la 3ª persona de singular ( $p < 0,01$ ). Este hallazgo corrobora las observaciones hechas por Salgado y Bouzouita (2017). Por el contrario, no hay diferencia significativa ( $p > 0,05$ ) entre la 1ª persona de plural y la de la 2ª persona de singular.

Figura 3. Gráficos del efecto de los factores significativos del modelo mixto de regresión logística



En la Figura 3 se presentan los gráficos de los efectos de los factores significativos del modelo mixto.<sup>19</sup> Como se puede observar, la 1ª persona y el singular presentan las probabilidades más altas de usar el complemento posesivo con un adverbio

<sup>19</sup> Valores mayores del eje vertical indican que la probabilidad de usar la variante posesiva es mayor. Los bigotes indican variabilidad y certeza de la estimación.

locativo. Asimismo, hay una interacción significativa entre justamente la 1ª persona y el singular. De manera interesante, como apuntan los resultados de la Tabla 10 junto con las indicaciones de la Figura 3, las 1ª y 2ª personas nunca se diferencian significativamente entre sí, mientras que sí se contrastan claramente con la 3ª persona. Efectivamente, las probabilidades más bajas las encontramos con la 3ª persona y el plural. No obstante, como el árbol de inferencia condicional indica claramente (Figura 2), se trata de una red compleja de interacciones, por lo cual, bajo ciertas circunstancias, incluso referentes de la 3ª persona y de plural pueden producir tasas porcentuales altas de la variante posesiva. De esta manera, aunque la 3ª persona sea el referente que más desfavorece el uso del posesivo, este rechazo no es completamente sistemático y, por ende, debería investigarse más profundamente en trabajos futuros.

La Figura 4 contiene los gráficos que representan los factores aleatorios, que en nuestro modelo son el adverbio y el país.<sup>20, 21, 22</sup> Primero, los adverbios que son más propensos a combinarse con el complemento posesivo son *alrededor*, *adelante*, *enfrente*, *arriba*, *atrás*, *bajo*, *abajo*, *encima*, *delante* y *adentro*. En este sentido, los adverbios encabezados por el prefijo *a-*, los del llamado Sector B (*abajo*, *adelante*, *adentro*, *afuera*, *arriba* y *atrás* en el presente corpus) en términos de Almela Pérez (1991), son relativamente uniformes en su ocurrencia con la variante posesiva, al contrario de lo propuesto por este autor.<sup>23</sup> El único adverbio del Sector B que desfavorece el posesivo es *afuera*. Santana Marrero (2014) y Salgado y Bouzouita (2017) también observan la preferencia de *alrededor* y *enfrente* por el uso posesivo en sus corpus. Los adverbios *cerca*, *debajo*, *detrás*, *dentro*, *fuera* y *lejos*, todos pertenecientes al Sector A que supuestamente se combinan más frecuentemente con el posesivo según Almela Pérez (1991), presentan en este corpus de tuits más variabilidad en cuanto al tipo de complemento preferido. Al igual que el análisis distribucional (Tabla 8) y el árbol de inferencia condicional (Figura 2), el gráfico de efectos aleatorios del adverbio (Figura 4) demuestra que este grupo de adverbios no es homogéneo en la elección del complemento. Por consiguiente, se puede concluir que la diferenciación entre estos dos grupos de adverbios con respecto a la variante preferida no parece tener validez.

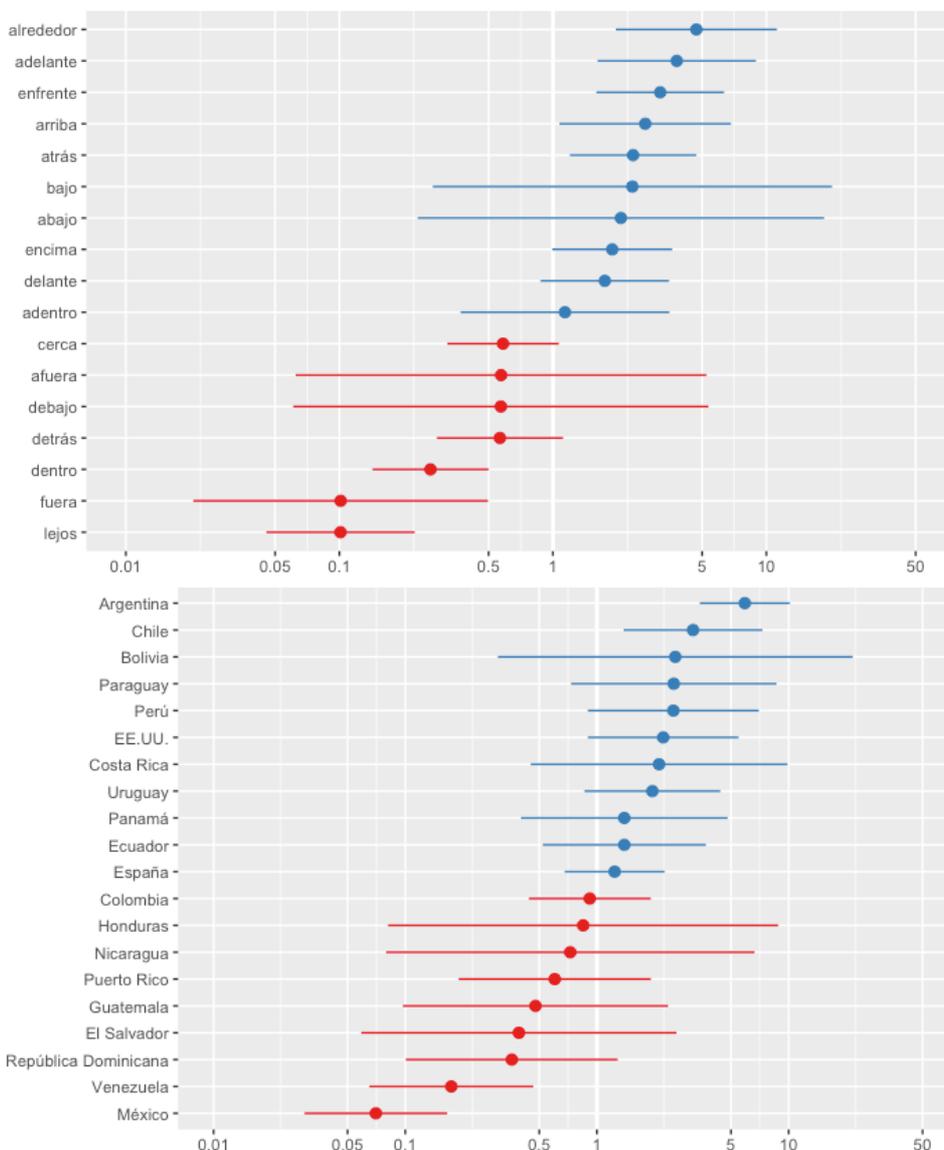
Figura 4. Gráficos de predicción de los factores aleatorios (adverbio y país) del modelo de regresión logística de efectos mixtos

<sup>20</sup> Se han construido los gráficos mediante el paquete *sjPlot* (Lüdecke 2017).

<sup>21</sup> Los valores que aparecen a la derecha del valor 1 (en azul) son aquellos que presentan más probabilidad de usar la variante posesiva, mientras que los con las probabilidades más bajas aparecen a la izquierda (en rojo). Por añadidura, los bigotes de las estimaciones indican variabilidad (cuanto mayor bigote, mayor variabilidad; un bigote grande se observa más frecuentemente en casos con pocos datos y/o datos con gran dispersión).

<sup>22</sup> Ya que el individuo no constituye un factor contribuyente al modelo (cf. p. ej. la Figura 1) y a que las medidas repetidas de los tuiteros son muy reducidas, no se presenta ningún gráfico que represente la distribución según el usuario.

<sup>23</sup> No obstante, hay que reconocer que *abajo* y *adentro* presentan mucha dispersión, por lo cual habría que interpretar este resultado con cierta precaución.



De los diez adverbios que muestran mayor preferencia por la variante posesiva, seis constituyen pares de complementariedad semántica entre los del Sector A y el Sector B: a saber, *delante* - *adelante*, *encima* - *arriba*, y *bajo* - *abajo*. En lo que concierne a los adverbios que desfavorecen el posesivo, también se encuentran pares entre los coeficientes negativos. En efecto, como revela la Figura 4, en el lado izquierdo se registra otro par complementario: *fuera* y *afuera*. Así pues, los únicos pares de complementariedad semántica que se separan por los dos lados del punto de quiebre son *detrás* - *atrás*, y *fuera* - *afuera* (también se encuentra el adverbio *debajo* entre los desfavorecedores, mientras que tanto las variantes *bajo* como *abajo* son favorecedores de la variante posesiva). Entonces, de manera general, se observa que los pares de los adverbios suelen comportarse de manera similar con respecto

a la variante preferida. Es decir, parece que no es tanto el sector de adverbio que afecta la variación en cuestión, sino que lo condicionante es justamente el adverbio mismo; si se usa el posesivo con una de las formas del adverbio, es probable que se use con la otra forma también.

En cuanto al factor diatópico, los países en que los tuiteros son más propensos a usar la variante posesiva son, en orden descendente, Argentina, Chile, Bolivia, Paraguay, Perú, los EE.UU., Costa Rica, Uruguay, Panamá, Ecuador y España. Este estudio corrobora, por tanto, la prevalencia de la variante posesiva en la zona rioplatense observada por otros estudios (Kany 1969, Meilán García 1998, Santana Marrero 2014, Marttinen Larsson y Álvarez López 2017). Asimismo, concuerda con Kany (1969), Santana Marrero (2014) y Marttinen Larsson y Álvarez López (2017) en cuanto a la alta frecuencia del uso posesivo en la zona andina. La Figura 4 resalta una tangible heterogeneidad entre los países centroamericanos: como se acaba de señalar, Costa Rica y Panamá parecen ser propensos a usar el posesivo, contrariamente a Honduras, Nicaragua, Guatemala y El Salvador, donde se desfavorece fuertemente este tipo de complemento.

A diferencia del árbol de inferencia condicional, la división generada por el modelo mixto de regresión también incluye los EE.UU. y España entre los países favorecedores. Evidentemente, aunque la variante posesiva no sea la predominante en estas regiones, sí cuenta con una frecuencia de uso considerable (cf. Santana Marrero 2014, Salgado y Bouzouita 2017 para España).

Concluyendo el análisis del presente trabajo, se considera apto discutir brevemente la utilidad metodológica de los diversos tipos de análisis presentados en este estudio. En el análisis, se han presentado resultados tanto del bosque aleatorio y el árbol de inferencia condicional, como del modelo de regresión logística. Comparando estos resultados, observamos que el árbol de inferencia condicional ha arrojado luz sobre diversos hallazgos (p. ej. el uso de la variante posesiva con referentes de la 3ª persona bajo ciertas circunstancias, un referente que hasta ahora se ha considerado rechazador del complemento posesivo [Salgado y Bouzouita 2017]) que no han sido posibles de identificar con los otros métodos de análisis. En realidad, se puede considerar el uso del árbol de inferencia condicional casi indispensable en el ámbito de la (socio)lingüística. En efecto, el modelo mixto de regresión logística -cuyo uso es ya estándar en trabajos lingüísticos con una orientación cuantitativa- solo permite estudiar el fenómeno desde una perspectiva plana sin posibilidad de identificar redes de microvariación dentro de los factores incluidos en el modelo. No obstante, no hay que perder de vista que -desde una perspectiva empírica- los modelos de regresión logística proporcionan más medidas empíricas para la evaluación del modelo estadístico (AIC, BIC,  $R^2$ , etc.), fortaleciendo así la validez empírica y el control de calidad de los modelos. Sin embargo, si en trabajos futuros la elección fuera obligatoria y binaria, consideramos necesario contemplar la posibilidad de incluir un análisis de árbol de inferencia condicional más bien que un modelo de regresión logística, siempre que se incluya el índice *C* de concordancia u otra medida de evaluación equivalente.

## 6. Conclusión

Con todo lo anterior, se puede concluir que hay varios factores que condicionan el uso de la variante posesiva. En concreto, según el bosque aleatorio, el árbol de inferencia condicional y el modelo mixto de regresión logística son significativos los factores país, adverbio, persona, número, y la interacción entre persona y número.

Respecto a las hipótesis planteadas en el apartado 3, se puede concluir lo siguiente. Primero, se corrobora la hipótesis de que se registra la variante posesiva en la mayor parte del mundo de habla hispana. De acuerdo con las conclusiones de Santana Marrero (2014), Marttinen Larsson (2015) y Marttinen Larsson y Álvarez López (2017), la variante posesiva cuenta con una alta frecuencia de uso en el español rioplatense.

Segundo, en cuanto al posible efecto de género, ni el bosque aleatorio, ni el árbol de inferencia condicional, ni el modelo de regresión logística indican su significación. Sin embargo, esto no es un resultado inesperado, ya que se aborda el fenómeno desde un panorama global, por lo que una estratificación social universal hubiera sido imprevista.

En tercer lugar, siguiendo a Almela Pérez (1991), la hipótesis planteada acerca del adverbio fue que los del Sector A se usan más frecuentemente con el complemento posesivo. No obstante, no ha sido posible estudiar este aspecto para los datos peninsulares por falta de casos para el Sector B. Aun así, tanto el análisis descriptivo como los multivariados revelan que la categorización propuesta no tiene validez para los casos hispanoamericanos, ya que el uso de la variante posesiva es predominante con adverbios del Sector B. Por añadidura, el análisis también ha concluido que los sectores propuestos no encajan adecuadamente con los datos examinados en el presente estudio, ya que hay una heterogeneidad estadísticamente significativa dentro de los sectores.

En cuanto a la influencia de la persona, se comprueba parcialmente la hipótesis de que el caso analizado de variación morfosintáctica ocurre principalmente con los referentes gramaticales de la 1ª y la 2ª persona: la persona gramatical que más favorece la variante posesiva es la de la 1ª persona. La 3ª persona gramatical sigue siendo la que menos se combina con posesivos en este tipo de sintagmas. De acuerdo con Salgado y Bouzouita (2017), creemos que esta ausencia de los referentes de la 3ª persona gramatical en el uso de la variante posesiva tal vez podría estar condicionado por el sincretismo de la forma *suyo/a*. Es decir, se podría interpretar la ausencia como una tendencia de evitar ambigüedad referencial, ya que estas formas expresan información gramatical sobre el número, el género y la deferencialidad del referente (se refiere a tanto *de él*, *de ella* y *de usted* como *de ellos*, *de ellas* y *de ustedes*).<sup>24</sup>

Con respecto a la cuestión del factor número, los resultados ilustrados en la Figura 3 apuntan a que la variante posesiva se usa más con referentes singulares

---

<sup>24</sup> Cf. el español canario, donde el posesivo tónico *suyo* se refiere únicamente a la persona *usted*, mientras que se expresan las terceras personas mediante las construcciones preposicionales (Catalán 1989: 155).

que plurales. Asimismo, se identifica que la variante posesiva es más frecuente con la 1ª persona de singular y menos probable con la 3ª persona de plural (Tabla 10 y Figura 3).

Al fin y al cabo, uno de los hallazgos más sustanciales de los análisis presentados en el presente trabajo ha sido la identificación de factores favorecedores de la variante posesiva anteriormente no descritos en la bibliografía (Figura 2). Más específicamente, el árbol de inferencia condicional ha podido desenredar que, en Argentina, Bolivia, Chile, Ecuador, Panamá, Paraguay, Perú y Uruguay, se emplea el complemento posesivo con algunos adverbios con la 3ª persona en un 40% de los casos. Así pues, este modelo matiza la noción del fuerte rechazo de la 3ª persona hacia el complemento posesivo. Asimismo, de manera interesante, el árbol expone que las frecuencias más altas de la variante posesiva (esencialmente en un 100% de los casos) se observan cuando se usa con la 1ª y 2ª persona gramatical en singular con los adverbios *abajo*, *adelante*, *alrededor*, *arriba*, *atrás*, *delante* y *enfrente* en Argentina, Chile, Perú y Uruguay (Figura 2).

Puesto que los resultados de este estudio coinciden en gran medida con los de Salgado y Bouzouita (2017), quienes investigan datos orales peninsulares, en cuanto a (i) el uso infrecuente del posesivo con referentes de 3ª persona, y (ii) la prevalencia del posesivo con referentes de 1ª persona, se puede interpretar la coincidencia como un argumento a favor de la observación que tanto los tuits como el habla espontánea son expresiones de lenguaje de carácter comunicativo inmediato. Asimismo, es de notar que nuestros resultados se distancian de los datos de medios de comunicación digitales analizados por Santana Marrero (2014). Teniendo en cuenta los resultados obtenidos por otros estudios que examinan datos de Twitter (cf. Berlanga y Martínez 2010, Betancourt Arango 2010, Estrada Arráez y de Benito Moreno 2016, Mancera Rueda y Pano Alamán 2013), consideramos adecuado afirmar que los tuits, por su carácter de inmediatez comunicativa, constituyen una expresión lingüística que se asocia más a la modalidad oral presente en los corpus analizados por Salgado y Bouzouita (2017) que a la escrituralidad del corpus de Santana Marrero (2014) (Koch y Oesterreicher 2007). Por falta de espacio y no por falta de interés, dejamos la indagación más profunda del carácter de los tuits a futuras investigaciones.

Asimismo, quedan pendientes estudios futuros que se enfoquen en el uso de los posesivos con el sufijo *-a*, variante frecuentemente registrada en el español peninsular (cf. Marttinen Larsson y Álvarez López 2017, Salgado y Bouzouita 2017, Bouzouita y Marttinen Larsson en prep.) y que tomen en cuenta las variables aquí investigadas, así como otros factores, como diferentes grupos etarios estratificados y el estatus socioeconómico, entre otros.

Por último, conviene contextualizar el uso del pronombre posesivo tónico adecuadamente. Es decir, cabe recordar que el uso del pronombre posesivo no se limita al ámbito adverbial locativo, pues se usa también en contextos nominales (p.ej. *el hermano suyo*) y, más sorprendentemente, en entornos verbales (p.ej. *habla mal mío*; *gusta tuyo*; *no dudo tuyo*, etc.; Bertolotti 2014, 2017) en diferentes países latinoamericanos, como en Uruguay, Argentina y Venezuela (Bouzouita y Pato en

prep.; Bouzouita y Casanova 2017). Además, esta variación también se registra en ciertas variedades brasileñas fronterizas con Uruguay (Bouzouita y Pato en prep.; Bouzouita y Casanova 2017). Aunque queda por examinar cómo estos diferentes usos posesivos se relacionan exactamente entre sí y cómo se expandieron de un ámbito a otro, no parece inverosímil la hipótesis propuesta por Bertolotti (2017), y desarrollada en más detalle por Bouzouita y Pato (en prep.) y Bouzouita y Casanova (2017), según la cual el uso del posesivo tónico se extiende por reanálisis del sintagma nominal al adverbial y al verbal formando parte de una cadena de gramaticalización. Aparte de que la gramaticalización de los posesivos parece ser un proceso hacia la abstracción en el cual predomina el rasgo de ‘persona’ y no de ‘posesión’ (cf. Huerta Flores 2003; Bertolotti 2017), existen otros argumentos a favor de esta hipótesis. El primer argumento es de índole diacrónica, dado que el uso nominal se registra ya en la Edad Media, el adverbial se generaliza a partir de principios del siglo XX (Marttinen Larsson y Álvarez López 2017: 95) y el verbal es de uso mucho más reciente (Bouzouita y Casanova 2017). El otro es de índole diatópica, puesto que el español rioplatense es la variedad en la cual no solo encontramos las frecuencias de uso más altas en el mundo hispanohablante para el uso posesivo adverbial (véase la sección 5.1), sino también las del ámbito verbal (Bouzouita y Casanova 2017).

### **Bibliografía**

- Alcina Franch, Juan & José Blecha (1975), *Gramática Española*. Barcelona: Ariel.
- Aleza Izquierdo, Milagros (2010), “Observaciones gramaticales de interés en el español de América”, en Aleza Izquierdo, Milagros & José M. Enguita Utrilla (coords.), *La Lengua Española de América: Normas y usos actuales*. Valencia: Universitat de València, 92-123.
- Almela Pérez, Ramón (1991), “Las construcciones del tipo ‘delante suyo’”, en Hernández, César (ed.), *Actas del III Congreso Internacional del Español de América*, vol. 1. Valladolid: Junta de Castilla y León, 435-444.
- Berlanga, Inmaculada & Estella Martínez (2010), “Ciberlenguaje y principios de retórica clásica. Redes sociales: el caso Facebook”, *Revista Venezolana de Información, Tecnología y Conocimiento*, 7(2): 47-61.
- Bertolotti, Virginia (2014), “Pronombres posesivos en el español rioplatense: Tres casos de reanálisis”, *Traslaciones. Revista Latinoamericana de Lectura y Escritura*, 1: 59-74.
- Bertolotti, Virginia (2017), “Pronombres posesivos y cambios gramaticales en español. Análisis de la variedad rioplatense”, en Company Company, Concepción & Norohella Huerta Flores (eds.), *La posesión en la lengua española*. Madrid: Consejo Superior de Investigaciones Científicas: 325-350.
- Betancourt Arango, Amanda (2010), “El ciberlenguaje como variedad lingüística (Artículo de reflexión no derivado de investigación)”, *Revista Q*, 4(8): sin páginas.

- Biber, Douglas (1995), *Variation across speech and writing*. Cambridge: Cambridge University Press.
- Bouzouita, Miriam & Vanessa Casanova (2017), “‘De los creadores de *al lado mío* llega *él habla suyo*’: El complemento posesivo en el ámbito verbal. Una aproximación diatópica y sociolingüística al fenómeno”. Comunicación presentada en el XVIII Congreso Internacional de la ALFAL, Bogotá, Universidad Nacional de Colombia (24-28 de julio de 2017).
- Bouzouita, Miriam & Enrique Pato (en prep.), “¿*Por qué no gustas mío, como yo gusto tuyo?* El verbo *gustar* y su complemento posesivo en las variedades americanas del español actual”.
- Bouzouita, Miriam & Matti Marttinen Larsson (en prep.), “On the morphology of possessive pronouns in Spanish adverbial constructions: some diatopic and sociolinguistic remarks”.
- Breiman, Leo (2001), “Random forests”, *Machine Learning*, 45: 5-32.
- Brown, Earl K. (2016), “On the utility of combining production data and perceptual data to investigate linguistic variation: The case of Spanish experimental *gustar* ‘to like, to please’ on Twitter and in an online survey”, *Journal of Linguistic Geography*, 3: 47-59.
- Catalán, Diego (1989), *El español: Orígenes de su diversidad*. Madrid: Paraninfo.
- Di Tulio, Ángela (2015), “Condiciones para una normativa del español lengua policéntrica”, *Signos Universitarios*, 34(51): 11-34.
- Estrada Arráez, Ana & Carlota de Benito Moreno (2016), “Variación en las redes sociales: datos twilectales”, *RILI*, 2(28): 77-111.
- Gómez Torrego, Leonardo (2009), *Hablar y escribir correctamente: Gramática normativa del español actual* (3ª ed.). Madrid: Arco Libros.
- Hernández Alonso, César (1996), “Castellano: Castilla la Vieja”, en Alvar, Manuel (ed.), *Manual de dialectología hispánica: El español de España*. Barcelona: Ariel, 197-212.
- Hothorn, Torsten, Peter Buehlmann, Sandrine Dudoit, Anette Molinaro & Mark Van Der Laan (2006), “Survival Ensembles”, *Biostatistics*, 7(3): 355-373.
- Hothorn, Torsten, Kurt Hornik & Achim Zeileis (2006), “Unbiased Recursive Partitioning: A Conditional Inference Framework”, *Journal of Computational and Graphical Statistics*, 15(3): 651-674.
- Huerta Flores, Norohella (2003), “Gramaticalización del posesivo: Un proceso hacia la abstracción”, *Medievalia*, 35: 183-197.
- Kany, Charles (1969), *Sintaxis hispanoamericana*. Madrid: Gredos.
- Koch, Peter & Wulf Oesterreicher (2007), *Lengua hablada en la Rumania: español, francés e italiano*. Madrid: Gredos.
- Labov, William (2001), *Principles of linguistic change, Volume 2: Social factors*. MA: Blackwell.
- Levshina, Natalia (2015), *How to do Linguistics with R: Data exploration and statistical analysis*. Amsterdam: John Benjamins.
- Llorente, Antonio Maldonado de Guevara (1980), “Consideraciones sobre el español actual”, *Anuario de letras*, 18:5-61.

- Loureda Lamas, Óscar (2009), *Introducción a la tipología textual*. Madrid: Arco/Libros.
- Lüdecke, Daniel (2017), *sjPlot: Data Visualization for Statistics in Social Science*. R package version 2.4.0. Disponible en <https://CRAN.R-project.org/package=sjPlot> [Fecha de consulta: 31 de enero de 2018].
- Mancera Rueda, Ana & Ana Pano Alamán (2013), *El español coloquial en las redes sociales*. Madrid: Arco/Libros.
- Martínez de Sousa, José (1996), *Diccionario de usos y dudas del español actual*. Barcelona: Bibliograf.
- Marttinen Larsson, Matti (2015), *El trabajo delante suyo: Un estudio diacrónico y diatópico de la variación en el uso de las locuciones adverbiales locativas*. Tesis de licenciatura. Estocolmo: Universidad de Estocolmo.
- Marttinen Larsson, Matti & Laura Álvarez López (2017), “‘Delante suyo’ vs ‘Delante de él’: El uso de las locuciones adverbiales locativas desde una perspectiva diacrónica y diatópica”, *Signo y Seña*, 31(1): 85-104.
- Meilán García, Antonio (1998), *Construcciones locativas y cuantitativas*. Madrid: Arco Libros.
- Moreno Fernández, Francisco (2009), *Principios de sociolingüística y sociología del lenguaje*. Barcelona: Ariel.
- Octavio de Toledo y Huerta, Álvaro S. (2016), *Los relacionantes locativos en la historia del español*. Berlin/Boston: Mouton de Gruyter.
- Paolillo, John C. (2002), *Analyzing linguistic variation: Statistical models and methods*. Stanford: CSLI Publications.
- R Core Team (2015), *R: A language and environment for statistical computing*. R Foundation for Statistical Computing, Viena: Austria. Disponible en <http://www.R-project.org> [Fecha de consulta: 29 de enero de 2018].
- Real Academia Española/Asociación de Academias de la Lengua Española (sin año), *Detrás de mí, encima de mí, al lado mío*. Disponible en <http://www.rae.es/consultas/detras-de-mi-encima-de-mi-al-lado-mio> [Fecha de consulta: 8 de octubre de 2017].
- Real Academia Española/Asociación de Academias de la Lengua Española (2005), *Diccionario panhispánico de dudas*. Madrid: Real Academia Española. Disponible en <http://www.rae.es/recursos/diccionarios/dpd> [Fecha de consulta: 9 de abril de 2017].
- Real Academia Española/Asociación de Academias de la Lengua Española (2009), *Nueva gramática de la lengua española*, vol. I-II. Madrid: Santillana. Disponible en <http://www.rae.es/recursos/gramatica/nueva-gramatica> [Fecha de consulta: 9 de abril de 2017].
- Real Academia Española/Asociación de Academias de la Lengua Española (2013), *Corpus del español del siglo XXI (CORPES): Descripción del sistema de codificación*. Libros y prensa. Madrid: Real Academia Española.
- Salgado, Hugo & Miriam Bouzouita (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, Juana (2014), “La estructura adverbio + posesivo en medios de comunicación digitales”, *Español actual*, 101:7-30.
- Strobl, Carolin, Anne-Laure Boulesteix, Achim Zeileis & Torsten Hothorn (2007), “Bias in Random Forest Variable Importance Measures: Illustrations, Sources and a Solution”, *BMC Bioinformatics* 8(25). Disponible en <https://bmcbioinformatics.biomedcentral.com/articles/10.1186/1471-2105-8-25> [Fecha de consulta: 31 de enero de 2018].
- Strobl, Carolin, Anne-Laure Boulesteix, Thomas Kneib, Thomas Augustin & Achim Zeileis (2008), “Conditional Variable Importance for Random Forests”, *BMC Bioinformatics* 9(307). Disponible en <https://bmcbioinformatics.biomedcentral.com/articles/10.1186/1471-2105-9-307> [Fecha de consulta: 31 de enero de 2018].
- Tagliamonte, Sali A. (2012), *Variationist sociolinguistics: change, observation, interpretation*. Oxford: Wiley-Blackwell.
- Tagliamonte, Sali A. & R. Harald Baayen (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.
- Twitter, Inc. (2017), *Twitter Developer Documentation: The Search API*. Disponible en <https://dev.twitter.com/rest/public/search> [Fecha de consulta: 20 de abril de 2017].