Variación léxica y sintáctica del español en las redes sociales

Spanish lexical and syntactic variation in social networks

Antonio Ruiz Tinoco
Universidad Sofía

RESUMEN

En este estudio analizamos datos de un corpus extraído de *Twitter*. Gracias a su función de *geocoding*, el API nos permite la extracción automática de *microblogs* limitando la búsqueda alrededor de cualquier punto geográfico por sus coordenadas. Hemos recogido en una base de datos en un radio de 25 km alrededor de varias ciudades de España e Hispanoamérica.

Hemos analizado la alternancia de los modos indicativo y subjuntivo en estructuras como *es una lástima que* + IND/SUBJ; orden de palabras como *más nada* vs. *nada más*; así como el uso de *delante de mí* vs. *delante mío/a*.

Como resultado, podemos observar que la alternancia del modo varía no solamente según el lugar sino también de manera consistente y muy diferente en estructuras muy similares. Asimismo, podemos observar muchas estructuras estigmatizadas, mostrar los resultados fácilmente en mapas y desarrollar nuevas metodologías de investigación en geolingüística.

Palabras clave: Variación del español, geolingüística, modo verbal, orden de palabras, *Twitter*

ABSTRACT

In this study, we analyzed corpus data extracted from *Twitter*. Due to its geocoding function, its API allows automatic extraction of microblogs by limiting the search around any geographical location with its coordinates. Our data was collected within a 25 km radius of Spanish and Latin American cities and stored in a relational database for further analysis.

We analyzed the alternation of the Spanish indicative/subjunctive mood with structures such as *es una lástima que* + IND/SUB; word order such as *más nada* vs. *nada más*; and the use of *delante de mí* vs. *delante mío/a*.

As a result, we found that that mood selection not only varies depending on the location, but also quite differently and consistently among similar structures. Additionally, we can quantitatively observe many stigmatized structures, easily map the results and develop new research methodologies in geolinguistics.

Keywords: Spanish variation, geolinguistics, verb mood, word order, *Twitter*

— o —

0. INTRODUCCIÓN

El uso cada vez más extendido de las redes sociales y la facilidad de recolección de *microblogs*[213] por medio de una programación no demasiado compleja, hace posible la construcción de enormes corpus sincrónicos o bases de datos de prácticamente cualquier lengua en un breve período de tiempo. En este artículo[214] trataremos algunas investigaciones sobre el español usado en *Twitter*[215]. Utilizando el API[216] hemos aprovechado la función de geocodificación[217], que nos permite determinar el punto geográfico desde donde surgen los mensajes. Resulta especialmente útil para el estudio de la variación geográfica tanto desde el punto de vista del léxico como de la sintaxis.

En primer lugar, hemos analizado la alternancia de los modos subjuntivo e indicativo en estructuras que tradicionalmente usan el subjuntivo, como *es una lástima que*…, etc. Para ello hemos reunido *microblogs* procedentes de 22 ciudades hispanohablantes. Hemos observado que la sustitución del modo subjuntivo por el indicativo es mayor en *Twitter* que lo que se había considerado hasta ahora en otros estudios (Takagaki, Ueda, Miyamoto, Fukushima, Ruiz Tinoco, 2004) y que el grado de alternancia no se produce de un modo general para el subjuntivo en conjunto, sino que varía con cada estructura particular y las zonas geográficas.

---

213 Mensajes breves.

214 Este artículo se basa en varias ponencias presentadas parcialmente en el LVII Congreso de la Asociación de Hispanistas de Japón el 8 de octubre de 2011 en la Universidad Komazawa de Tokio, titulada «Variación geográfica del uso del modo subjuntivo en *Twitter*»; la ponencia presentada en el IV Congreso Internacional de Lingüística del Corpus, celebrado en la Universidad de Jaén el 22 de marzo de 2012, titulada «*Twitter* como corpus de variación geográfica –alternancia modal del español»; la ponencia presentada en Language Contact, Conflict, and Confluence at the Edge of Nation, 24th Conference on Spanish in the United States and 9th Conference on Spanish in Contact with Other Languages, celebrada el día 8 de marzo de 2013 en la Universidad Texas Pan American, titulada «Variación sintáctica en *Twitter* en el español fronterizo de Estados Unidos» y la ponencia presentada en el V Congreso Internacional de Lingüística del Corpus, celebrado en la Universidad de Alicante el 15 de marzo de 2013, titulada «Variación de la anteposición de *más* con adverbios de negación».

215 http://twitter.com/

216 API (Application Programming Interface, Interfaz de Programación de Aplicaciones, IPA, en español). Hemos utilizado Search API (ver. 1) para los datos de este estudio. En la actualidad, está en desuso. Para la recogida de léxico variable (carro, coche, etc.) hemos ya usado la nueva versión Streaming API (ver.1.1), más compleja técnicamente, pero con menos limitaciones que la anterior.

217 *Geocoding* en inglés. Es un proceso de asignar coordenadas a un punto de un mapa.

En segundo lugar, hemos observado la variación de *más nada* vs. *nada más*, cuyos resultados sugieren que la frecuencia de uso no parece ser la primera causa de la alternancia.

En un tercer caso, hemos observado la variación de *delante de mío*/a vs. *delante de mí* a ambos lados de la frontera entre Estados Unidos y México, 13 ciudades en total. En este caso hemos podido comprobar que la forma *delante de mío*/a es más frecuente en las ciudades norteamericanas que en las mexicanas. Al mismo tiempo, ya que delante mío/a implica la elección de género del adjetivo posesivo, hemos observado diferencias notables a ambos lados de la frontera que pueden deberse a diferentes causas.

Finalmente, mostraremos varios mapas mostrando los puntos geográficos de donde proceden algunos ejemplos de léxico variable.

1. ALTERNANCIA DE LOS MODOS INDICATIVO Y SUBJUNTIVO

El uso del subjuntivo se ha explicado tradicionalmente desde las perspectivas sintáctica, semántica y pragmática. Gili Gaya (1969) mantiene el concepto de *realis/irrealis* para explicar el uso del subjuntivo en las cláusulas subordinadas. Posteriormente, Terrell and Cooper (1974) proponen características semánticas y pragmáticas del subjuntivo, discutidas también en Fukushima, Noritaka (1978). Por otra parte, Lantolf (1978) y Blake (1981) señalan las frecuencias del indicativo en *Me alegro de que…* explicando el fenómeno como una evolución del subjuntivo hacia una reducción en una situación de inestabilidad. Otros estudios posteriores, como los de Lope Blanch, Juan M. (1990), Ridruejo, Emilio (1990), Silva-Corvalán, Carmen (1994, 1996-1997), Aleza Izquierdo, Milagros (2010), etc. también señalan la misma tendencia, especialmente en el español de Hispanoamérica. Sin embargo, son escasos los estudios cuantitativos que muestren hasta qué punto el fenómeno está extendido en la amplia geografía del español. En este estudio analizamos datos obtenidos de un corpus formado por *microblogs* de la red social *Twitter*, procedentes de 22 ciudades hispanohablantes y las zonas circundantes en un radio de 25 km. Las ciudades son: Alcalá de Henares, Barcelona, Huelva, Las Palmas, Madrid, Oviedo, Pamplona, Sevilla y Santa Cruz de Tenerife en España; Asunción, Bogotá, Buenos Aires, Caracas, La Paz, Lima, Ciudad de México, Monterrey, Montevideo, Quito, San José de Costa Rica y Santiago de Chile, en Hispanoamérica, además de la ciudad de Nueva York. Se reunieron algo más de ocho millones de mensajes, aproximadamente unos 120 millones de palabras, durante los meses de agosto y septiembre del año 2011 y se almacenaron en una base de datos MySQL[218].

Para el análisis, de cada ciudad hemos escogido 100 ejemplos al azar para cada una de las secuencias *lástima que, pena que, necesario que, alegro (de) que* y *gus-*

[218] http://www.mysql.com

ta que. Hemos filtrado y desechado mensajes de procedencia geográfica dudosa, como los que contienen citas de otros mensajes o van marcados con RT (*retweet*).

1.1 ¿Qué es Twitter?

Twitter es un servicio de redes sociales que usa *microblogs*, también llamados *tuits* o *tweets*, de un máximo de 140 caracteres. Para participar es necesario inscribirse y en la actualidad cuenta con aproximadamente 500 millones de usuarios, que en total cada día envían alrededor de 500 millones de mensajes sobre todo tipo de temas en una gran variedad de idiomas. Aunque los mensajes pueden ser privados, accesibles solamente a los usuarios autorizados, los mensajes públicos se pueden leer en línea y pueden ser almacenados en una base de datos siempre que respeten las condiciones expresadas en las condiciones de servicio[219]. Lamentablemente para la comunidad académica, existen serias restricciones para la difusión de los datos en grandes cantidades debido principalmente a los derechos de autor y a la privacidad de los miembros.

Evidentemente, la restricción de 140 caracteres, así como la enorme cantidad de temas, estilos, errores ortográficos y gramaticales, el uso de símbolos, procedencia de los autores, etc. dificultan su análisis, pero al mismo tiempo proporcionan una rica fuente de material lingüístico de gran espontaneidad.

1.2 Obtención de datos

Hay varios métodos para obtener datos de *Twitter*. El más simple es a través de la interfaz[220] existente en línea mostrada en la Fig. 1 o la de búsqueda avanzada[221] como se muestra en la Fig. 2. Por medio de esta interfaz podemos determinar entre otros parámetros la palabra o secuencia de palabras que nos interesa, el idioma del mensaje y el lugar de procedencia del mensaje.

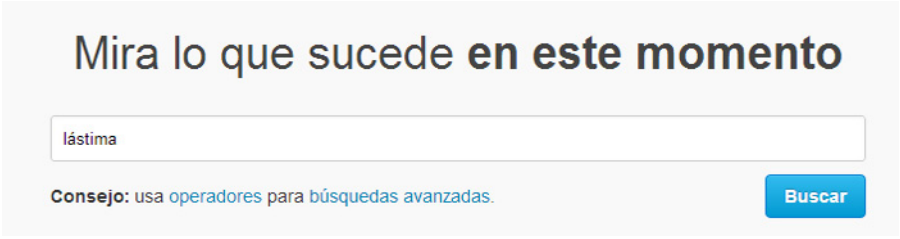


Fig. 1 Interfaz de búsqueda simple (https://twitter.com/search-home)

219 http://twitter.com/tos

220 https://twitter.com/search-home

221 https://twitter.com/search-advanced

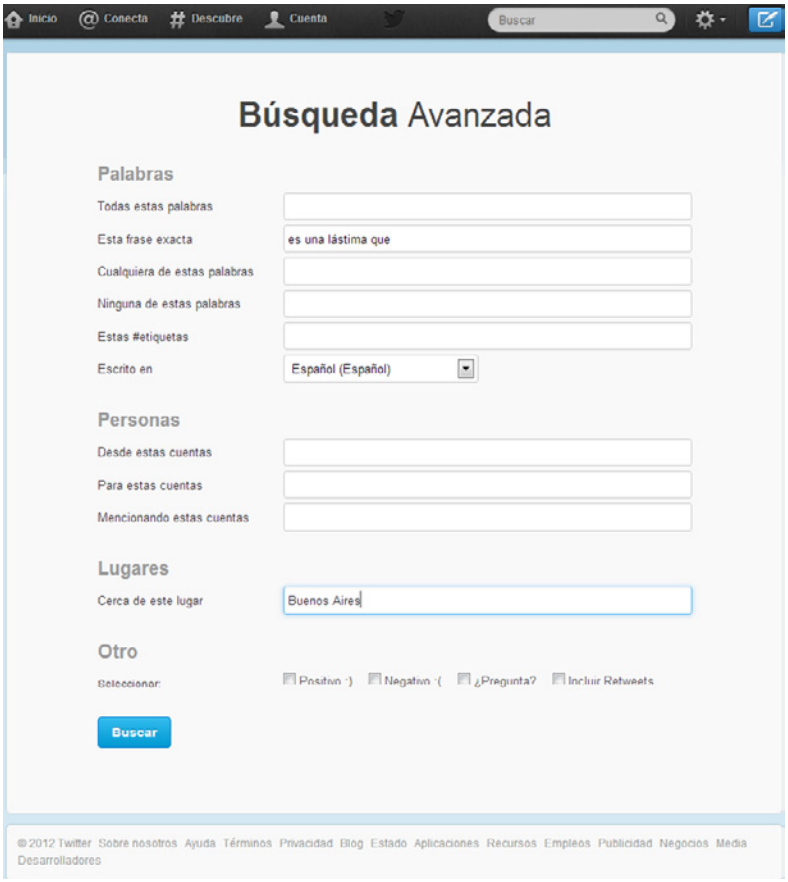


Fig. 2 Interfaz de búsqueda avanzada (https://twitter.com/search-advanced)

Estas interfaces, aun siendo útiles para búsquedas esporádicas, no nos resultan adecuadas para nuestro objetivo por varias razones. En primer lugar, solamente hay acceso a los mensajes con una antigüedad de menos de una semana o unos diez días. En segundo lugar, la determinación del lugar se basa en la declaración del usuario, por lo que no siempre es un dato fiable. Finalmente, los datos obtenidos tendrían que ser procesados manualmente para trasladarlos a una hoja de cálculo o a una base de datos, por lo que no resultaría práctico.

Como hemos mencionado anteriormente, para obtener eficientemente los datos podemos usar los APIs publicados[222] para este fin. Los detalles técnicos se pueden consultar en línea así como en varias obras publicadas recientemente como Ho, Christopher and Peri, Bess (2011). El resultado de las búsquedas se obtiene en formato JSON o ATOM, similares a XML, que tienen la ventaja de que se pueden procesar con bastante facilidad.

222 https://dev.twitter.com/docs/api

Los datos obtenidos fueron procesados mediante un sistema diseñado en lenguaje PHP y una base de datos MySQL. En la Fig. 3 se muestra la interfaz creada para hacer búsquedas detalladas de los datos reunidos. Así en la Fig. 3 se puede observar que hay 2.121 ejemplos de uso de la secuencia *lástima que* en un radio de 25 km alrededor de la ciudad de Monterrey en México. Asimismo, el resultado de cada búsqueda se puede exportar a formato EXCEL para cualquier procesamiento posterior.

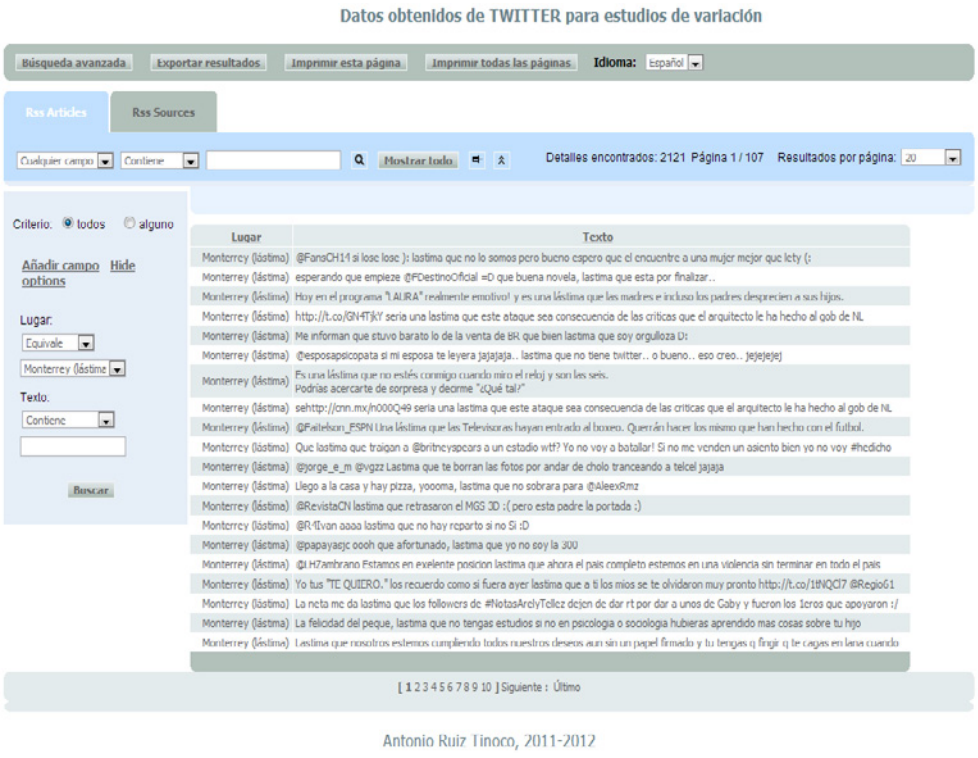


Fig. 3 Interfaz de nuestra base de datos de *Twitter*

Una de las características de *Twitter* es que con gran frecuencia contiene mensajes en forma de cita. Ya que no podemos tener la seguridad de que el mensaje original coincida con la localidad desde la que se emitió en último lugar, eliminamos tales mensajes que van señalados con las letras RT (*retweet*), como el siguiente:

(1) RT @ritrejo: @jmesa77 @raxxie_ @azul_silencio Eso pues, me gusta que mis amigos se quieran =)

También hemos eliminado de nuestro análisis los mensajes que contienen citas evidentes de otras personas como el siguiente:

(2) «Da tu primer paso con fe, no es necesario que veas la escalera completa, sólo da tu primer paso» - Martin Luther King

En las búsquedas de *Twitter* se puede determinar la lengua objeto, en nuestro caso el español. Sin embargo, con cierta frecuencia encontramos mensajes escritos en portugués o catalán que contienen casualmente las mismas secuencias de palabras por tener idéntica ortografía. Todos estos mensajes de procedencia dudosa, ininteligibles, de otras lenguas se evitan automáticamente en lo posible, aunque la decisión final la tomamos nosotros. No obstante, incluimos los mensajes que sean legibles aunque no sean correctos ortográficamente, como el siguiente, donde en lugar de *que* se usa *q*, faltan tildes, etc.

(3) Aa ya extrañaba este rollo y desvelarmeeeeee !!!!! Lastima q la mayoria ya sta dormidoooo

1.3 Estructuras analizadas

Hemos analizado varias estructuras que según la gramática prescriptiva tradicional deben ir seguidas de un verbo en modo subjuntivo. Para ello, hemos tomado 100 ejemplos al azar de la base de datos. Para la recolección de mensajes hemos utilizado como secuencias de palabras las partes subrayadas.

(a) Me alegro (de) que… + ind/sub
(b) Me gusta que… + ind/sub
(c) Es necesario que… + ind/sub
(d) Es una lástima que… + ind/sub
(e) Es una pena que… + ind/sub

1.3.1 Me alegro (de) que…

En la Fig. 4 se muestran las frecuencias relativas de uso del subjuntivo respecto al indicativo. Las ciudades españolas de Alcalá de Henares, Madrid y Oviedo, así como Asunción de Paraguay muestran el 100% de uso del subjuntivo mientras que en el resto de las ciudades el porcentaje baja hasta un mínimo de 92 en Caracas y Monterrey, y 83 en Nueva York. A continuación se muestran algunos de los ejemplos de uso de selección del modo indicativo en varias ciudades. Hemos considerado las variantes *me alegro que…* y *me alegro de que…*

(4) @JC_Varela me alegro que les fue bien en colon (Tenerife)
(5) @MassiCastillo Gracias Massiel! Me alegro que te *gusto*... jejeje:-) (Nueva York)
(6) @andyta_angel que bueno, me alegro que la *pasaron* bien, por las fotos estuvieron al lado (Quito)
(7) @S4nTePull2206 cuenta!!me alegro que *fuiste* valiente y seas feliz!ese sentimiento es bello y al final tendrs rosas en tu jardin:-) (Barcelona)

Curiosamente, el ejemplo (7) contiene en la misma cláusula un verbo en indicativo (fuiste) y otro en subjuntivo (seas).

Se advierte que, en general, hay una mayor tendencia al uso del subjuntivo en las ciudades españolas, mientras que el indicativo, aunque en poca medida, se observa algo más en las ciudades latinoamericanas, siendo la ciudad bilingüe de Nueva York la que se aleja algo más de las demás.

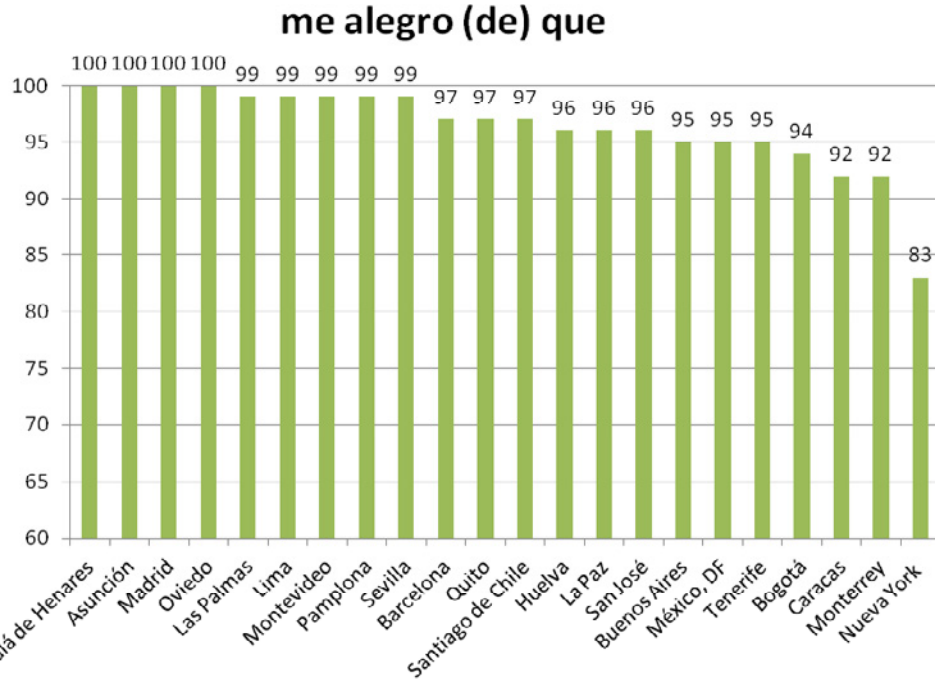


Fig. 4 Frecuencia relativa de uso del subjuntivo *[me alegro (de) que]*

1.3.2 Me gusta que…

En la Fig. 5 se muestran las frecuencias de uso del subjuntivo con el verbo *gustar*. De nuevo observamos que varias ciudades españolas e hispanoamericanas usan el subjuntivo en los 100 ejemplos de la muestra. La menor proporción se da en Asunción, con 93 casos en subjuntivo y 7 en indicativo, incluso por debajo de Nueva York, el uso del subjuntivo es casi total.

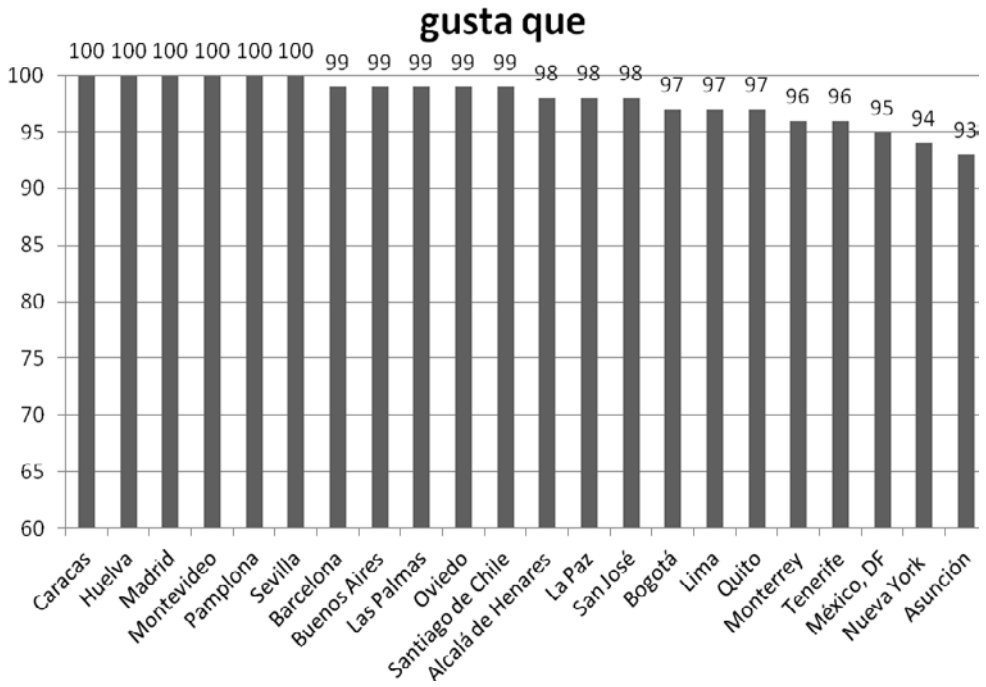


Fig. 5 Frecuencia relativa de uso del subjuntivo *[gusta que…]*

A continuación se muestran algunos de los ejemplos de uso del indicativo:

(8) @VRCaldevilla amigos eso ganamos... sino le echamos la culpa a correos y llamamos a algún S. de agrupación porque no le gusta que *tuiteamos* (Oviedo)

(9) me gusta que De Narvaez se *deja* un lado del cuello de la camisa mas alto para que no se vea el tatuaje =) (Buenos Aires)

(10) @Multiplaza_py me gusta que siempre *encuentro* todo lo que necesito y muy cerca de mi casa #multiplaza (Asunción)

1.1.3 Es necesario que…

De manera análoga a las anteriores estructuras, en el caso de *es necesario que…* observamos que en la mayoría de las ciudades el uso del subjuntivo es casi total, con tan solo algunas excepciones, como se muestra en la Fig. 6.

A continuación mostramos varios ejemplos con indicativo.

(1) Es necesario que en tres canales *están dando* el mismo partido de fútbol? A eso llaman FPT? Para todos los canales! (Buenos Aires)

(2) @elmono0311 No necesitams los permisos, pero es necesario que tu *tienes* el permiso de los autores para distrribuirlo. (Nueva York)

(3) De investigación y nos vayamos a la internet, es necesario que *tengo* el hábito de leer, para ser la diferencia;) (México, DF)

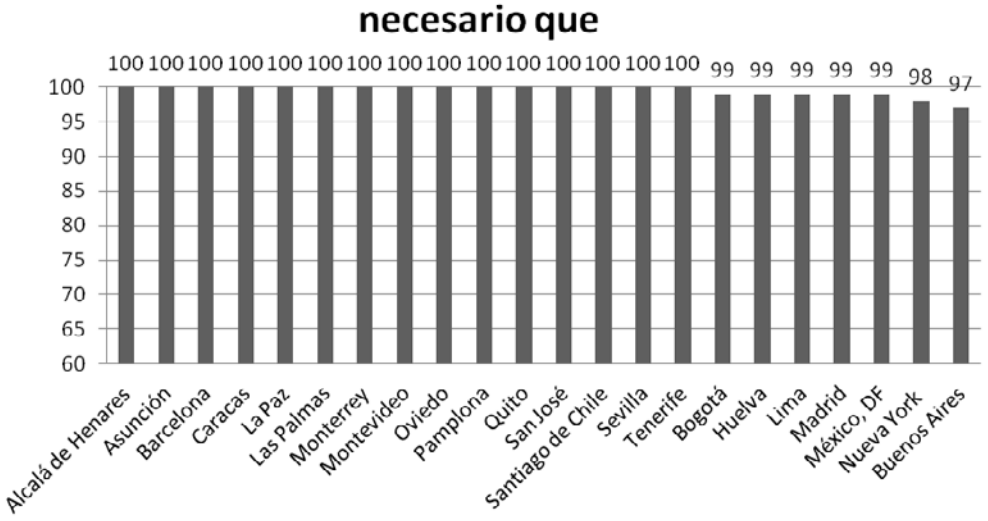


Fig. 6 Frecuencia relativa de uso del subjuntivo *[necesario que…]*

1.3.4 Es una pena que…

Para la secuencia *pena que…,* a excepción de Madrid, encontramos un menor uso del subjuntivo, que baja hasta llegar al 62% en Montevideo, como se muestra en la Fig. 7.

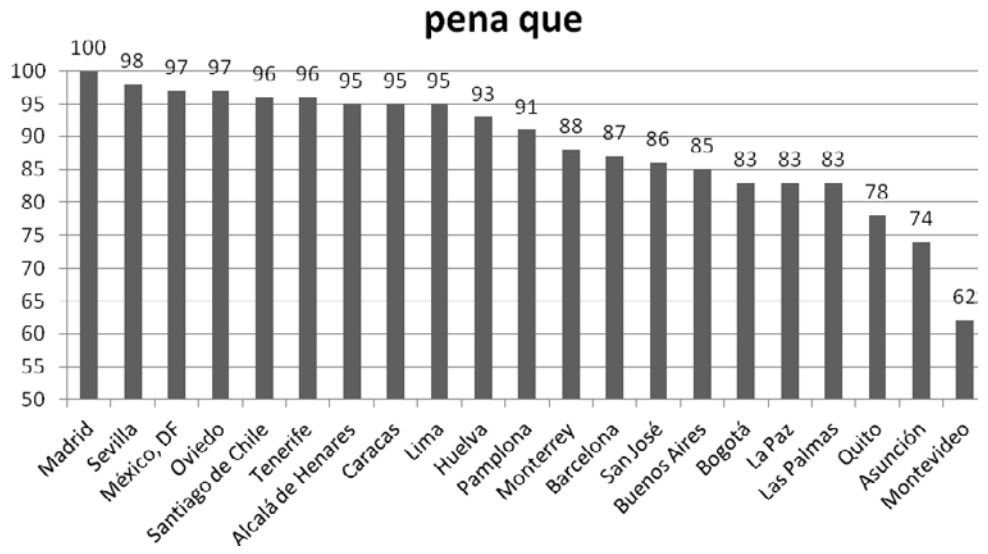


Fig. 7 Frecuencia relativa de uso del subjuntivo *[pena que…]*

A continuación mostramos algunos de los ejemplos con indicativo.

(4) Es una pena que se *fue* jaime. Se va un excelente persona. (Santiago de Chile)
(5) Todos tuvimos nuestra época paloma Sui Generis. Es una pena que los pendejos de ahora nunca la *van* a tener. ¿Qué escuchan ahora? Pobres. (Montevideo)
(6) @iMundoBizarro y es una pena que cada dia *son* mas los que la olvidan... (Caracas)

1.3.5 Es una lástima que…

Entre las estructuras estudiadas, *lástima que…* es la que muestra el menor uso del subjuntivo, como se muestra en la Fig. 8, variando desde un máximo de 87% en Caracas a los mínimos de San José, Asunción y Montevideo con 28%, 27% y 27% respectivamente.

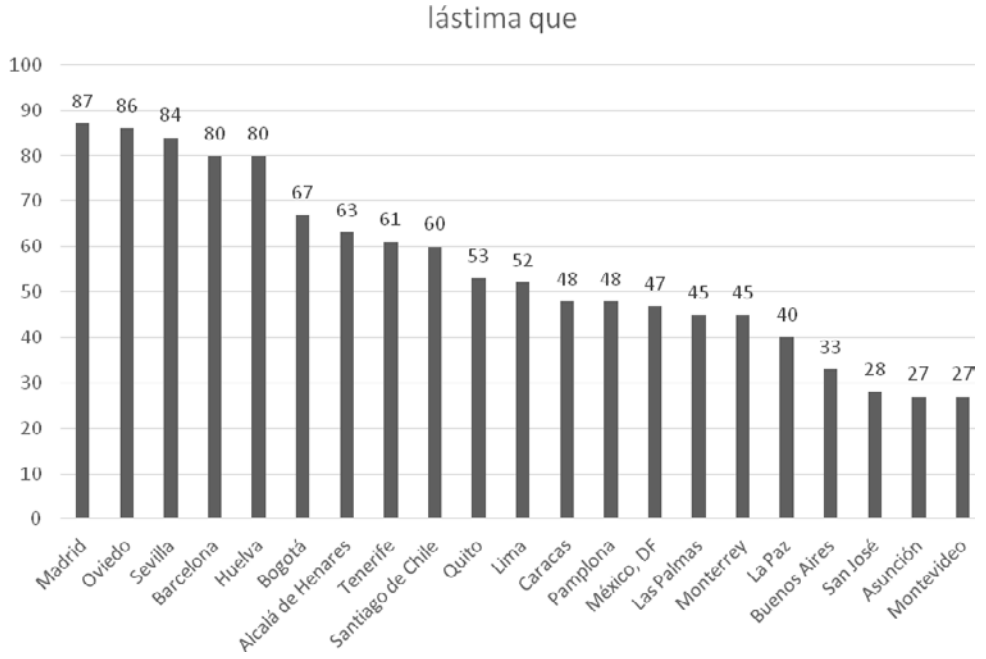


Fig. 8 Frecuencia relativa de uso del subjuntivo *[lástima que…]*

A continuación mostramos algunos de los ejemplos con indicativo.

(7) Ayer nuevamente gran salida de #YovaniGallardo. Sólo 1 carrera permitida y 11 ponches en 7in. Una lástima que se *fue* sin decisión (Tenerife)
(8) Me hubiese gustado ir, una lastima que *estuve* encerrada estudiando. La PUCP es de sus alumnos! (Pamplona)

(9) @NachoReigada Gracias seguidor. Una lastima que el partido se *vio* enturbiado por la actitud violenta de los rivales. Un saludo compañero (Oviedo)

1.3.6 Es una lástima que… vs. Es una pena que…

Entre las estructuras estudiadas, *pena que…* y *lástima que…* son las que más se alejan del uso del subjuntivo a pesar de que sintácticamente son todas análogas. En la Fig. 9 se muestran las frecuencias relativas del subjuntivo en las dos estructuras.

Comparando directamente las frecuencias de estas dos estructuras podemos apreciar:

1) La pérdida del subjuntivo en *lástima que…* es mucho más pronunciada que en *pena que…* en todas las ciudades investigadas.
2) La tendencia a la pérdida del subjuntivo es más pronunciada en las ciudades latinoamericanas que en España.
3) Cuanto mayor es la pérdida del subjuntivo en lástima que… la tendencia a perder el subjuntivo en pena que… es mayor también.

Los datos obtenidos a través de *Twitter* son completamente actuales, es decir, sincrónicos, por lo que podemos conocer con exactitud el lugar geográfico de procedencia y también podemos obtener un número suficientemente elevado de ejemplos para dar una mayor validez estadística a los resultados de los análisis.

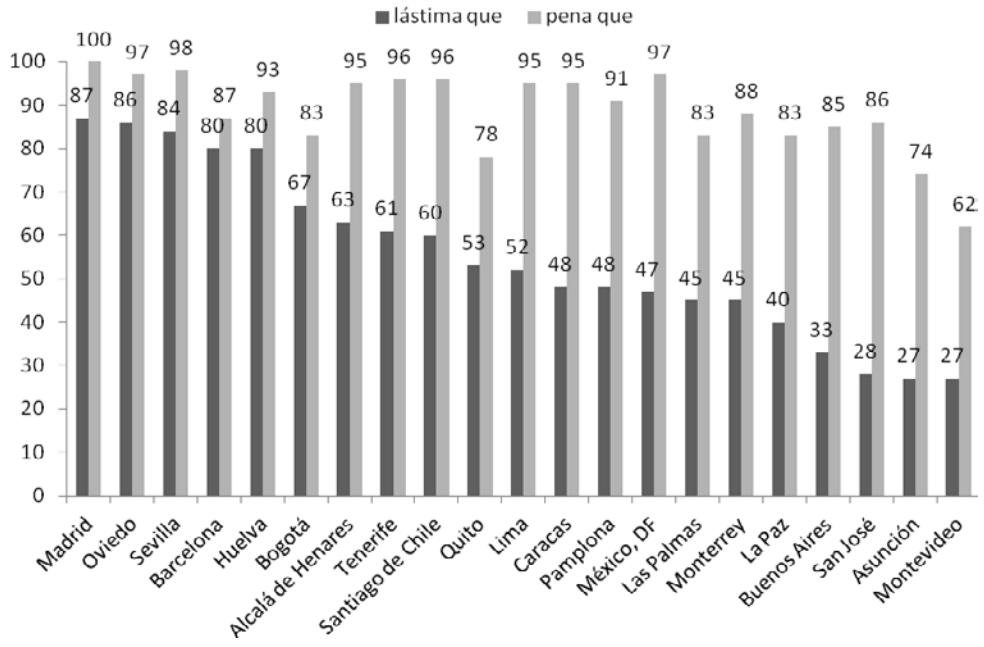


Fig. 9 Frecuencia relativa del subjuntivo en *lástima que…* vs. *pena que…*

2. VARIACIÓN DE *MÁS NADA* VS. *NADA MÁS*

La anteposición de *más* con adverbios de negación en secuencias como *más nada, más nunca, más nadie, más ninguno*, etc. se ha señalado frecuentemente por distintos autores y por la RAE. El *Diccionario Panhispánico de Dudas* indica (RAE, 2005) que cuando *más* acompaña a estos adverbios se pospone en el español normal, pero que en el habla coloquial americana, especialmente en el área caribeña, Andalucía y Canarias es frecuente la anteposición. Luis A. Ortiz López lo menciona como característica de Puerto Rico y del habla caribeña en general, Pilar García Mouton (2003), Vaquero de Ramírez (1996), Gutiérrez-Rexach y González-Rivera (2011, 2012), García Cornejo (2008) y Pato (2011) también lo señalan como característico de las Islas Canarias y del español americano. En general se describe como frecuente en el español oral y no recomendable.

Para cuantificar en lo posible la extensión del fenómeno y de manera análoga al caso anterior de alternancia del modo verbal, hemos preparado un corpus de *microblogs* extraídos de la red social *Twitter*, seleccionando con un radio de 25 km las ciudades españolas de Madrid, Barcelona, Sevilla, Bilbao, A Coruña y Las Palmas; las caribeñas de San Juan de Puerto Rico, Santo Domingo y La Habana, así como México DF, Panamá, Caracas, Bogotá, Lima, Santiago de Chile, Buenos Aires y la ciudad de Nueva York, ésta última por sus características especiales de diversas procedencias de los hispanohablantes y situación de contacto con el inglés.

Para la búsqueda de los ejemplos en todas las ciudades mencionadas hemos usado tanto la anteposición como la posposición (*más nada / nada más, más nunca / nunca más, más nadie / nadie más y más ninguno / ninguno más*) para conocer también las frecuencias relativas. Como hemos hecho en estudios anteriores, (Ruiz Tinoco, A. 2011, 2012a, 2012b, 2013a, 2013b, 2013c), mediante un programa informático se han reunido los ejemplos en una base de datos para mayor facilidad de análisis.

En total, entre diciembre de 2012 y febrero de 2013 se reunieron 171.661 *microblogs* de uso de *más* tanto en anteposición como en posposición, con distintos porcentajes entre las distintas ciudades y, como era de esperar, el fenómeno se da en mayor o menor grado en todas las ciudades estudiadas como veremos más adelante.

2.1 Anteposición y posposición de *más*

Análogamente al estudio de la alternancia del modo verbal, escogimos una muestra de 100 ejemplos al azar por cada una de las secuencias de anteposición y posposición de *más* con *nada, nadie, ninguno*, y nunca eliminando manualmente los *microblogs* que no eran adecuados. Los resultados de las proporciones de anteposición y posposición fueron cuantitativamente bastante similares. En la Fig. 10 mostramos el caso de *más nadie* vs. *nadie más*.

Como ya se había señalado en estudios anteriores, la anteposición de *más* se da frecuentemente en las ciudades caribeñas, no solamente en San Juan de Puerto Rico y Santo Domingo, sino también en Caracas o Panamá. Curiosamente, en La Habana la proporción de anteposición es mucho más baja que en otras ciudades caribeñas, debido tal vez a que la mayoría de los *tuits* los escriben viajeros que no son originalmente cubanos, ya que el uso de Internet en Cuba no está extendido por su alto coste. Nueva York también muestra un porcentaje bastante alto, que creemos se debe a que la mayoría de la población hispana es de origen puertorriqueño y dominicano. En España, la proporción más alta de anteposición se da en Sevilla (Andalucía) y Las Palmas (Islas Canarias). En A Coruña, donde esperábamos un alto grado de convergencia debido a que en gallego es bastante usual la secuencia *mais nunca* mostró una baja proporción de anteposición. En el resto de las ciudades estudiadas se nota cierto grado de anteposición, especialmente en Buenos Aires, comprobando así que el fenómeno está bastante extendido en todas las ciudades hispanohablantes.

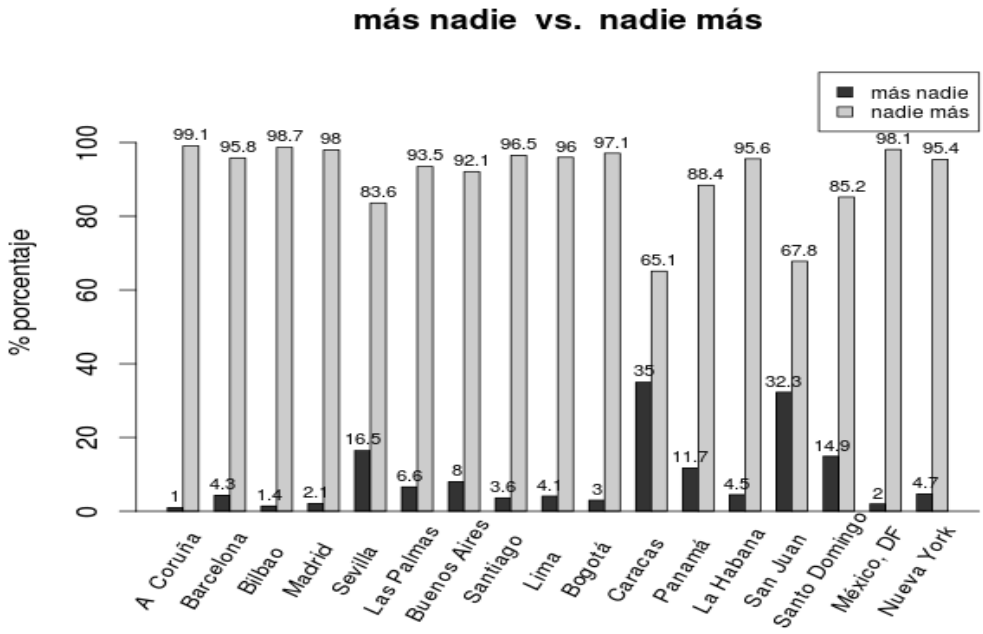


Fig. 10 *más nadie* vs. *nadie más*

2.2 Frecuencia de uso vs. variación

Al comparar las frecuencias absolutas de uso (expresadas en 1/10.000) y la proporción de variación, podemos observar en la Fig. 11 que la frecuencia de las secuencias con *nada* es la más alta y la secuencia con *ninguno* la más baja. Sin embargo, el grado de variación es similar entre 10,56% y 14,69% aproximadamente. Estas proporciones sugieren que la alternancia de la anteposición y posposición ha

avanzado casi por igual con cada uno de los adverbios negativos estudiados como si se tratara de una especie de subida de *más*.

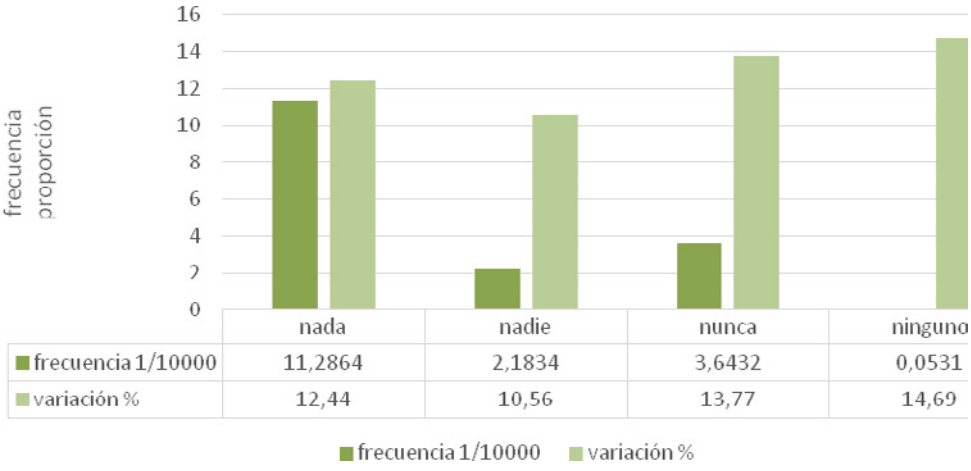


Fig. 11 Frecuencia absoluta vs. Grado de variación

3. VARIACIÓN DE *DELANTE DE MÍO/A* VS. *DELANTE DE MÍ*

El uso de los adverbios locativos con determinantes posesivos, como *cerca mío(a), lejos mío(a), delante mío(a), detrás mío(a), dentro mío(a), encima mío(a), debajo mío(a),* etc. respecto a las locuciones preposicionales como *cerca de, lejos de, delante de, detrás de, dentro de, encima de, debajo de,* etc., se ha señalado frecuentemente en cuestionarios de variación sintáctica en el *Diccionario Panhispánico de Dudas* (RAE, 2005) o en la *Nueva Gramática de la Lengua Española* (RAE, 2010). Aunque la RAE no aconseja su uso, algunos autores argumentan que se puede justificar gramaticalmente (Gómez Torrego, 2009), que se trata de un americanismo (Černý, 2002) o incluso de un catalanismo (Martínez de Sousa, 2008). Esta alternancia de uso está bastante extendida en todos los países hispanohablantes, sin exceptuar Estados Unidos, cuya situación es más compleja debido a una serie de fenómenos de contacto, bilingüismo, etc.

Para estudiar este fenómeno de variación en el español fronterizo de Estados Unidos con México, se ha elaborado un corpus de tales estructuras a partir de *microblogs* de *Twitter* de forma análoga a los casos anteriormente expuestos. Las zonas estudiadas son East Los Ángeles, San Diego, Tucson, Albuquerque, El Paso, San Antonio y Houston en la parte estadounidense para contrastar resultados, Tijuana, Ciudad Juárez, Chihuahua y Monterrey de la parte mexicana. Para cada ciudad se ha considerado un círculo de 25 km de radio. Los resultados globales se muestran en la Fig. 12. Podemos comprobar que la forma con adjetivo posesivo es bastante más frecuente del lado estadounidense que del lado mexicano. También se puede

observar que las ciudades mexicanas más cercanas a la frontera, como El Paso o Tijuana muestran proporciones cercanas a las ciudades estadounidenses.

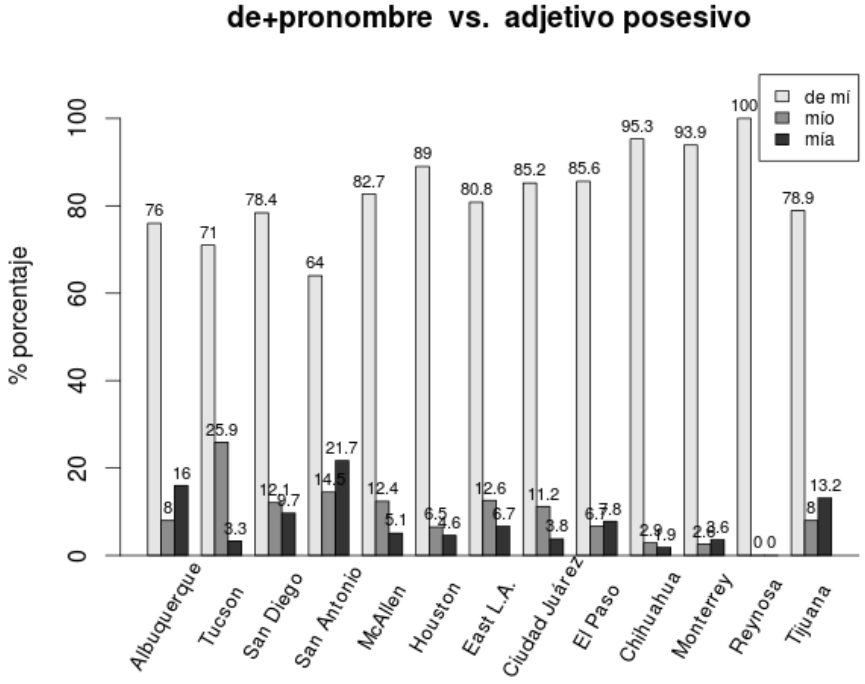


Fig. 12 adverbio de lugar + (de + pronombre vs. adjetivo posesivo masculino/femenino)

A ambos lados de la frontera se usan los posesivos tanto en femenino como en masculino con algunas diferencias claras que resumimos en la Tabla 1. Por ejemplo, con *cerca* y *dentro* se usa más el género masculino para el adjetivo posesivo a ambos lados de la frontera. No hay tendencia clara para *debajo* ni para *lejos*. Sin embargo, mientras del lado mexicano delante se inclina más al femenino, en la parte estadounidense no hay diferencia. También detrás se inclina más al masculino en la parte mexicana mientras que en la parte estadounidense no se aprecia tampoco ninguna preferencia. Finalmente, en el caso de encima resulta mayoritariamente masculino en el lado mexicano pero femenino en el estadounidense.

	México	*EE. UU.*
cerca mío/a	*M*	*M*
debajo mío/a	-	-
delante mío/a	*F*	-
dentro mío/a	*M*	*M*
detrás mío/a	*M*	-
encima mío/a	*M*	*F*
lejos mío/a	-	-

Tabla 1. Género del adjetivo posesivo con adverbios de lugar

4. DISTRIBUCIÓN GEOGRÁFICA DE GEOSINÓNIMOS

Finalmente, ya que hemos obtenido las coordenadas de los lugares desde donde se emitieron los mensajes, podemos preparar mapas que indiquen su posición. Aunque hay muchas opciones para esta tarea con programas de Sistemas de Información Geográfica (SIG)[223] como ArcGIS[224], QGIS[225], etc., en esta ocasión utilizaremos el sistema gratuito de Google Fusion Tables[226]. Como ejemplo, mostraremos algunos mapas parciales de la distribución de los geosinónimos de *automóvil*.

En las Fig. 13, 14, 15 y 16 se muestran los puntos donde se encontraron *tuits* en España con los geosinónimos *carro, coche, automóvil*, y *auto* respectivamente.



Fig. 14 carro



Fig. 15 coche



Fig. 16 automóvil



Fig. 17 auto

Comprobamos que la aparición de *coche* en España es mucho más frecuente que los demás geosinónimos. Utilizando el mismo método podemos preparar un

223 GIS en inglés, Geographic Information System

224 http://www.esrij.com/

225 http://www.qgis.org/es/site/

226 http://www.google.com/drive/apps.html#fusiontables

mapa de calor (heat map) para la distribución de *coche* y *carro*, como se muestra en las Fig. 18 y 19.

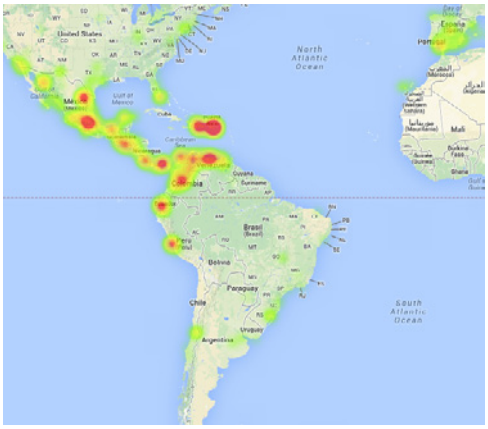


Fig. 18 carro

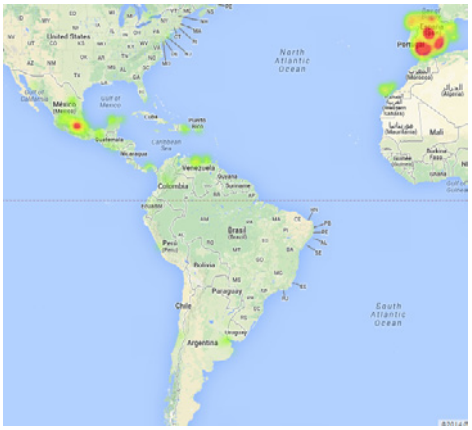


Fig. 19 coche

En la Fig. 18 podemos comprobar que *carro* es más frecuente en México, Centroamérica, el Caribe, Ecuador y Perú, mientras que *coche* destaca en España y aparte de la Ciudad de México donde se puede comprobar su uso, en menor escala que en España, hay una gran diferencia con las demás zonas hispanohablantes.

5. CONCLUSIONES

Aunque solamente hemos observado muy pocas estructuras, en todas se da cierto grado de alternancia de los modos indicativo y subjuntivo, no solamente en el español atlántico, sino también en el peninsular y en las zonas bilingües en contacto con otras lenguas, como son Nueva York o Asunción. La sustitución del subjuntivo por el indicativo parece ser más frecuente en general en el español atlántico.

La pérdida del subjuntivo en favor del indicativo no se produce con la misma frecuencia en las estructuras estudiadas incluso entre las que son sintácticamente similares. Asimismo, en oraciones con un significado prácticamente sinónimo, como el caso de *lástima que* vs. *pena* que, no se dan las mismas proporciones de elección de modo, lo que sugiere que la pérdida del subjuntivo no se debe solamente a causas sintácticas ni semánticas.

La simple observación de la distribución de las frecuencias de uso en distintas ciudades sugiere que la preferencia por el indicativo se inicia en algunas estructuras concretas, con mayor o menor fuerza según la ciudad. Es el caso de *lástima que*, al que siguen otras estructuras, que aunque sean sintáctica o semánticamente análogas, son de menor frecuencia, como *pena que*, de modo que la disminución del uso del subjuntivo no se da por estructuras sintácticas en conjunto, sino en

grados diferentes y progresivos con las más frecuentes en primer lugar. Evidentemente, la frecuencia de las estructuras no será el único factor en el cambio, pero podemos comprobar por la gran cantidad de ejemplos que la frecuencia es un factor clave en el cambio.

La anteposición de *más* ocurre prácticamente en las mismas proporciones sin diferencias significativas independientemente del adverbio negativo a pesar de las grandes diferencias en la frecuencia de uso. Este fenómeno nos sugiere una posible subida general de *más* ante dichos adverbios negativos en un proceso gramatical ya iniciado en determinadas zonas hispanohablantes.

En cuanto al uso de los adverbios de posición con adjetivos posesivos, no solamente encontramos que el fenómeno ocurre con más frecuencia en el lado estadounidense de la frontera con México, lo cual es tema de futuras investigaciones debido a la complejidad lingüística de las zonas de contacto de lenguas, sino que también se le asigna género al adverbio y esta asignación muestra algunas tendencias diferenciadas a ambos lados de la frontera.

Las redes sociales nos facilitan la creación de corpus absolutamente sincrónicos que se pueden utilizar para la observación y el análisis de la variación lingüística. La enorme cantidad de ejemplos no solo nos ayuda a conocer con gran exactitud la distribución geográfica de los fenómenos lingüísticos, sino que también resultan indispensables para analizar los detalles de los procesos del cambio lingüístico.

Aunque los datos obtenidos a través de *Twitter* no son totalmente representativos de la totalidad de las ciudades estudiadas, su análisis nos proporciona algunos aspectos cuantitativos importantes de la distribución de cada una de las estructuras, por lo que creemos que la red social *Twitter*, y a pesar de no tener acceso a todos los atributos de las personas que la utilizan, se puede utilizar fácilmente como corpus para el estudio global de la variación geográfica de una lengua tan extendida como es el español.

BIBLIOGRAFÍA

Aleza Izquierdo, Milagros; Enguita Utrilla, José María (eds.) (2010), La lengua española en América. Normas y usos actuales. Universitat de València.

Blake, Robert. (1981), «Some Empirically Based Observations on Adult Usage of the Subjunctive in Mexico City». En James Lantolf and Gregory Stone (eds.) (1981) Current Research in Romance Languages, pp. 13-22. Bloomington, IN: Indiana University Linguistics Club.

Černý, Jiří. (2002), «Algunas observaciones sobre el español hablado en América». En Acta Universitatis Palackianae Olomucencis, Facultas Philosophica Philologica 74, pp. 39-48.

Fukushima, Noritaka. (1978), «La aserción y el modo español». En: Lingüística Hispánica, I, pp.75-94.

García Cornejo, Rosalía. (2008), «El adverbio „más« y las palabras negativas en la historia del español». En Revista de historia de la lengua española 3, pp. 29-77.
García Mouton, Pilar (ed.) (2003), El español de América. Consejo Superior de Investigaciones Científicas. Madrid.
Gili Gaya, Samuel. (1969), Curso Superior de Sintaxis Española. Barcelona: Bibliograf, S. A.
Gómez-Torrego, L. (2009), Hablar y escribir correctamente. Tomo II. Madrid, Arco Libros, 3.ª edición.
Gutiérrez-Rexach, Javier y Melvin González-Rivera. (2011), «Negation, modification, and the syntax of PR Spanish.» En N. Faraclas et al (eds.)(2011): Anansi's Defiant Webs: Contact, continuity, convergence and complexity in the Languages, Literatures and Cultures of the Greater Caribbean. Willemstad: University of Curaçao and Fundashon pa Planifikashon di Idioma.
— (2012), «Negative Quantification and Degree Restriction: The Case of más nada in Puerto Rican Spanish». En Kimberly Geeslin and Manuel Díaz-Campos (2012): Selected Proceedings of the 14th Hispanic Linguistics Symposium, ed., pp. 285-292. Somerville, MA: Cascadilla Proceedings Project. www.lingref.com, document #2672.
Ho, Christopher and Peri, Bess. (2011), Sams teach yourself Twitter API in 24 hours. Pearson Education, Inc.
Lantolf, James. (1978), «The Variable Constraints on Mood in Puerto Rican-American Spanish.» En Margarita Suñer (ed.) (1978): Contemporary Studies in Romanc Linguistics, pp. 193-217. Washington, DC, Georgetown University Press.
Lope Blanch, Juan M. (1990), «Algunos usos de indicativo por subjuntivo en oraciones subordinadas». En I. Bosque (ed.) (1990), Indicativo y subjuntivo. Taurus, Madrid, pp. 180-182.
Martínez de Sousa, J. (2008), Diccionario de usos y dudas del español actual. Gijón, Ediciones Trea, 4.ª edición.
Pato, Enrique. (2011), «La construcción más nunca en el español moderno». Ponencia presentada en el XVI Congreso Internacional de la Asociación de Lingüística y Filología de la América Latina (ALFAL), 6-9 de julio de 2011, Alcalá de Henares, España.
Real Academia Española. (2005), Diccionario panhispánico de dudas. Madrid, Santillana.
Real Academia Española, Asociación de Academias de la Lengua Española. (2010), Nueva Gramática de la Lengua Española: morfología y sintaxis. Madrid, Espasa.
Ridruejo, Emilio. (1990), «¿Cambios iterados en el subjuntivo español?». En I. Bosque (ed.) (1990), Indicativo y subjuntivo. Taurus, Madrid, pp. 361-382.
Ruiz Tinoco, Antonio. (2011), «Variación geográfica del uso del modo subjuntivo». Ponencia presentada en LVII Congreso de la Asociación Japonesa de Hispanistas, Universidad Komazawa.
— (2012a), «Twitter como corpus de variación geográfica –alternancia modal del español». Ponencia presentada en el IV Congreso Internacional de Lingüística de Corpus, Asociación Española de Lingüística de Corpus, Universidad de Jaén.

— (2012b), «Alternancia indicativo/subjuntivo en Twitter en el español de Estados Unidos». Ponencia presentada en II ALFALito, Cuestiones lingüísticas en relación con la diáspora latinoamericana», The Graduate Center, City University of New York.
— (2013), «Twitter como corpus para estudios de geolingüística del español». En Sophia Lingüística LX, pp. 147-163, The Graduate School of Languages and Linguistics, Linguistic Institute for International Communication, Sophia University, Tokyo.
— (2013a), «Variación sintáctica en Twitter en el español fronterizo de Estados Unidos». Ponencia presentada en Language Contact, Conflict, and Confluence at the Edge of the Nation, 24th Conference on Spanish in the United States and 9th Conference on Spanish in Contact with Other Languages, The University of Texas Pan American.
— (2013b), «Variación de la anteposición de más con adverbios de negación», Ponencia presentada en V International Conference on Corpus Linguistics, Universidad de Alicante.
Silva-Corvalán, Carmen. (1994), Language contact and change: Spanish in Los Angeles. Oxford: Oxford, University Press.
— (1996-97), «Avances en el estudio de la variación sintáctica». En Cuadernos del Sur 27, pp. 35-49.
Takagaki, Toshihiro; Ueda, Hiroto; Miyamoto, Masami; Fukushima, Noritaka; Ruiz Tinoco, Antonio. (2004), Encuesta sobre problemas sintácticos de la lengua española. Ministerio de Educación, Cultura y Deporte de España.
Terrell, Tracy and Joan Hooper. (1974), «A Semantically Based Analysis of Mood in Spanish». En Hispania 57, pp. 484-494.
Vaquero de Ramírez, María. (1996), El español de América II. Morfosintaxis y léxico. Madrid. Arco/Libros.