

La expresión gramaticalizada ‘lo que pasa es que’ en español. Estudio contrastivo de PRESEEA-Palma y PRESEEA-Alcalá

The Grammatical Expression ‘lo que pasa es que’ in Spanish. Contrastive Study of PRESEEA-Palma and PRESEEA-Alcalá

Beatriz Méndez Guerrero

UNIVERSIDAD AUTÓNOMA DE MADRID
ESPAÑA

beatriz.mendez@uam.es

Recibido: 24-VIII-2020 / **Aceptado:** 01-04-2022

DOI: 10.4067/S0718-09342022000300844

Resumen

La forma ‘lo que pasa es que’ ha sufrido un proceso de gramaticalización en español que le ha llevado a adquirir un significado procedimental muy habitual en el habla cotidiana. A pesar de ello, es poco lo que se sabe sobre las funciones discursivas de esta construcción en la interacción, las variantes que presenta y los factores sociales que influyen en su producción. En esta investigación, se presenta un estudio que trata de responder a estas cuestiones en dos comunidades de habla españolas: Palma y Alcalá de Henares. Para ello, se analizan dos corpus PRESEEA pertenecientes a dichos lugares que han permitido establecer ciertos patrones de uso de las distintas variantes de ‘lo que pasa es que’. Entre las principales conclusiones del estudio, destacan que ‘lo que pasa es que’ presenta al menos cuatro funciones discursivas distintas en la interacción: contrastiva, aclarativa/justificativa, concesiva y de objeción/negación; que algunas de las variantes de esta forma gramaticalizada representan mejor a algunas funciones que a otras; que los factores extralingüísticos, como el origen o el grupo etnolingüístico, inciden especialmente en la producción y elección de las formas; y que, normalmente, es la conjunción de las variables independientes del estudio la que determina el uso que se hace de la expresión.

Palabras Clave: ‘Lo que pasa es que’, gramaticalización, significado procesual, funciones discursivas, PRESEEA.

Abstract

The form *lo que pasa es que* has been grammaticalized in Spanish and now also has a procedural meaning that is quite frequent in everyday speech. Nonetheless, there is still little knowledge about the discursive functions of this expression in the interaction, the variants it presents, and the social factors that influence its usage. In this research, we present a study that attempts to answer these questions in two Spanish-speaking communities: Palma and Alcalá de Henares. To do this, we analyzed two PRESEEA corpus of these cities to establish the patterns of use of the different variants of *lo que pasa es que*. The main conclusions of the study suggest that *lo que pasa es que* presents four different discursive functions in the interaction: contrastive, clarifying/justificatory, concessive and objection/denial; that some of the variants of this grammaticalized form represent some functions better than others; that extralinguistic factors, such as origin or ethnolinguistic group, have a special influence on the production and selection of forms; and that, normally, the conjunction of the independent variables of the study is what determines the use of the expression.

Keywords: Lo que pasa es que, grammaticalization, procedural meaning, discursive functions, PRESEEA.

INTRODUCCIÓN

La construcción ‘lo que pasa es que’ en español se ha considerado tradicionalmente una forma pseudo-hendida con significado referencial (Moreno Cabrera, 1983, 1999; Johanson, 2001; Dufter, 2008; Feldhausen & Vanrell, 2015; Krassenberg & Lahousse, 2018; Gutiérrez Bravo, 2021). Sin embargo, esta estructura convive hoy con un uso fijo o gramaticalizado, cuyo significado es procesual y que todavía ha sido poco estudiado en situaciones de habla real (Romera, 2008; Reig, 2011; Camargo & Enrique-Arias, 2021). En este trabajo nos interesa conocer el funcionamiento de la expresión ‘lo que pasa es que’ y de sus principales variantes en la interacción cuando se utiliza con un significado pragmático o procedimental. Para ello, se estudiarán los intercambios comunicativos de distintos grupos de hablantes de dos ciudades españolas (Palma y Alcalá de Henares) a través de los corpus PRESEEA existentes en dichos lugares. En este estudio tenemos varios objetivos, por una parte, pretendemos esclarecer las funciones discursivas que presentan estas construcciones y las frecuencias de aparición en el habla cotidiana; también queremos determinar la incidencia de ciertos factores lingüísticos y extralingüísticos en la elección de las diferentes variantes utilizadas por los hablantes; y, por último, analizar las similitudes y diferencias que presentan los individuos en las zonas geográficas ya mencionadas.

Las hipótesis de las que se parten son que las formas gramaticalizadas de la expresión ‘lo que pasa es que’ presentan diferencias que están marcadas por algunos factores extralingüísticos como la edad, el sexo, el nivel educativo, el grupo etnolingüístico o el origen. Y que estas formas, asimismo, guían las inferencias de los hablantes en diferentes direcciones, dependiendo de la función que desempeñen o de la intención comunicativa que haya detrás de su uso. Concretamente, prevemos que, entre las distintas variantes de ‘lo que pasa es que’, podrán distinguirse unas más

innovadoras, que incluyen pérdida de cópula y reducción fónica, como ya ha demostrado Romera (2008), que muestran gran vitalidad en las ciudades analizadas y pueden vincularse principalmente al grupo de jóvenes y a los bilingües o catalanohablantes. En las siguientes secciones del artículo se presenta una explicación teórica en la que se concentran los estudios previos más relevantes sobre el fenómeno comentado, una explicación del marco metodológico y la muestra utilizada en el estudio, los resultados de la investigación y las principales conclusiones a las que se llegaron. Las preguntas de investigación del estudio tratan de determinar si existen distintas funciones discursivas para la expresión gramaticalizada de 'lo que pasa es que', como sugieren distintos autores. Si las variaciones que aparecen en la forma gramatical guardan relación con dichas funciones, hasta qué punto intervienen los factores extralingüísticos en su producción y qué diferencias, si las hay, se observan en los dos grupos de hablantes de ciudades analizadas.

1. La forma gramaticalizada 'lo que pasa es que' en español

La expresión 'lo que pasa es que' se estudia hace mucho tiempo desde dos perspectivas: (a) como una construcción pseudo-hendida o perífrasis relativa con significado léxico pleno¹ (Prince, 1978; Moreno Cabrera, 1983, 1999; Declerck, 1984, 1994; Collins, 1991; Johanson, 2001; Lambrecht, 2001; Romera, 2008; Reig, 2011); y (b) como una estructura en proceso de gramaticalización² con un significado procesual (y sin valor léxico), que guía la interpretación del segmento del discurso que introduce (Gallego, 2006; Garrido, 2006; Romera, 2008; Reig, 2011). Desde un punto de vista gramatical, ambas construcciones comparten estructura y, desde un punto de vista semántico-pragmático, las dos contienen nueva información contrastiva y/o limitativa (Sedano, 1990; Romera, 2008).

De acuerdo con Moreno Cabrera (1999) y Morales (2005), la perífrasis relativa o construcción pseudo-hendida está formada por tres elementos: una relativa libre introducida por un artículo o pronombre neutro 'lo', una cópula y un constituyente hendido, como, por ejemplo, 'lo que pasa es que se ha olvidado las llaves dentro y ahora no puede entrar'. Desde la perspectiva semántica, presenta un significado referencial a partir del cual 'lo que pasa es que' se refiere literalmente a algo que sucede o tiene lugar: 'lo que sucede es eso' (Reig, 2011). Esta construcción convive actualmente con un uso fijo de 'lo que [pasa es que]', que se ha gramaticalizado, al menos parcialmente, y cuyo significado es pragmático (Reig, 2011)³ La forma gramaticalizada presenta, como explica Romera (2008), pérdida de autonomía sintagmática y paradigmática, desemantización, expansión funcional y, en algunos casos, reducción fonética⁴.

- (1a) I: la gente que había antes *lo que pasa es que* se ha comprado segunda vivienda entonces qué ha hecho
E: aah aja

I: ha vendido los pisos
[PRESEEA-Palma 3M1 castellano]

(1b) E: ¿hay mucha delincuencia en Alcalá?

I: la verdad es que sí la hay/ *lo que pasa es que* no:/ o sea no nos llegamos a enterar casi ni de la mitad de las cosas pero yo creo que sí que la hay
[PRESEEA-Alcalá 1M2 castellano]

(1c) I: y / y muy bien / yo / el piso es muy grande / el piso es grande porque tiene / eeh 4 habitaciones y tiene come / sala comedor y bueno que es grande el piso y estoy muy bien *lo que* la finca ya es vieja
[PRESEEA-Palma 3M2 ambas]

En el ejemplo (1a) puede verse un caso de construcción pseudo-hendida formada por una relativa libre ('lo que pasa'), la cópula ('es') y el constituyente hendido ('que se ha comprado una segunda vivienda'). En este caso, 'lo que pasa es que' tiene un significado referencial con el que literalmente se refiere a algo que sucede o tiene lugar ('lo que sucede es que se han comprado segunda vivienda y han vendido los pisos'), de ahí que muchas veces aparezcan las construcciones 'lo que pasa es que' y 'lo que sucede es que' indistintamente con este uso. En cambio, en los ejemplos (1b) y (1c) nos encontramos ante una construcción gramaticalizada que transmite información proposicional y que, en estos casos, como en muchos otros de este tipo puede sustituirse por 'pero'. En (1b) 'lo que pasa es que' y en (1c) 'lo que' presentan un significado procedimental, que guía la interpretación del segmento del discurso que le sigue 'no nos llegamos a enterar casi ni de la mitad de las cosas' en (1b) y 'la finca ya es vieja' en (1c). Adicionalmente, en (1c) observamos un caso de estructura gramaticalizada con significado pragmático que se ha visto reducida fónicamente, pero que no por ello ha perdido su significado y su función. Esta variante, como se explicará, comparte la mayoría de sus propiedades con la forma 'lo que pasa es que'²⁵.

Las construcciones pseudo-hendidas y gramaticalizadas mostradas en los ejemplos, aunque pueden presentar límites borrosos en algunos contextos, suelen diferenciarse entre sí en algunas de sus propiedades. Una de las principales diferencias tiene que ver con la rigidez de la construcción gramaticalizada, que pasa de ser una forma flexible a ser una expresión fija. Eso se aprecia en, al menos, cuatro aspectos: (1) mientras que en la forma pseudo-hendida se permiten modificaciones tanto de tiempo como de número y persona ('lo que pasaba era que', 'lo que nos pasó fue que', etc.), en la construcción gramaticalizada la forma verbal ('pasa') no puede sustituirse en ningún caso; (2) imposibilidad de insertar en la construcción gramaticalizada pronombres, adverbios, modalizaciones o perífrasis verbales, mientras que eso sí es posible en las formas pseudo-hendidas ('lo que esta mañana me ha pasado es que', 'lo que ha debido de pasar es que', etc.); (3) resulta posible revertir el orden de los elementos de las estructuras pseudo-hendidas si se dan en el discurso las circunstancias apropiadas para

ello ('que hemos llegado tarde, eso es lo que pasa'), pero esto no es posible en las construcciones gramaticalizadas; y (4) los elementos que forman la expresión pseudo-hendida pueden sustituirse paradigmáticamente por otros de la misma categoría que admitan focalización (como 'decir' o 'saber': "lo que dice es que") y eso no es posible en las formas gramaticalizadas (Romera, 2008; Reig, 2011).

Adicionalmente, Reig (2011) ha establecido dos diferencias más entre las construcciones "lo que pasa es que" gramaticalizadas y pseudo-hendidas. Por una parte, las gramaticalizadas son siempre afirmativas, mientras que las pseudo-hendidas pueden ser también negativas en ciertos momentos ('lo que pasa no es que haga calor, sino que la casa está mal ventilada'). Por último, la construcción pseudo-hendida mantiene su significado completo y léxico 'suceder' y este significado se ha perdido en la construcción gramaticalizada, la cual como se ha mencionado ya ha adquirido un significado procedimental de carácter pragmático que guía las inferencias para una correcta interpretación del mensaje y que puede cumplir con distintas funciones discursivas.

La pérdida del significado léxico de 'lo que pasa/ocurre es que' en determinadas situaciones comunicativas y la adquisición del valor procesual han llevado a considerar esta forma una posible unidad funcional del discurso (Romera, 2008)⁶. Hoy esta forma parece estar ampliamente extendida en distintas variedades del español y es incluso más frecuente que el uso original como construcción pseudo-hendida⁷. Las referencias que existen hasta la fecha asignan a esta construcción un significado pragmático adversativo (de contraste u oposición), como en (2a), y/o explicativo, como se aprecia en (2b) (Narbona, 1986; Garrido, 2006; Arumí & Esteve, 2006; Romera, 2008; Reig, 2011)⁸.

(2a) I: sí: por la:- yo es que antes venía por la tarde *lo que pasa que* me cambié por la mañana porque como los críos van al colegio por la mañana: .../

E: claro

[PRESEEA-Alcalá 1M2 castellano]

(2b) I: allí ya éramos del mismo grupillo luego perdimos el contacto porque él estudio en otro colegio yo estudié en otro y *lo que pasa es que* al vivir en una zona muy próxima pues su su madre y mi madre siempre mantenían el contacto y por tanto siempre sabíamos el uno del otro y luego hubo la casualidad que que coincidimos en la carrera

[PRESEEA-Palma 1H3 catalán]

Según Romera (2008), 'lo que [pasa/ocurre es que]' adquiere normalmente este significado pragmático adversativo en las formas gramaticalizadas que hace posible que la expresión pueda sustituirse por 'pero' con una función similar. No es extraño que, en algunos momentos, además se encuentren casos en los que 'pero'" y 'lo que pasa/ocurre es que' aparezcan juntos. Estas situaciones se deben a las pequeñas

diferencias pragmáticas encontradas entre ellos y otorgan a la construcción un matiz atenuador o más indirecto al que tiene la forma 'lo que pasa es que' (Wesch, 2002; Guerrero, Gajardo, Ibarra, González & Reyes, 2020a, 2020b). En palabras de Reig (2011), el significado de 'pero' y 'lo que pasa es que' no es exactamente equivalente porque no siempre son intercambiables ('lo que pasa es que' puede ser sustituido por 'pero' manteniendo un significado muy cercano, pero no al contrario). Esto indica que las restricciones requeridas por 'lo que pasa es que' son mayores que las de 'pero' y, por tanto, que su significado es diferente al menos en algunos matices. Estas expresiones pueden tener también un significado explicativo, que puede aparecer junto al valor adversativo o por separado, a partir del cual 'lo que pasa es que' introduce alguna información que orienta al destinatario sobre alguna información transmitida en el discurso anterior y que puede actuar a modo de justificación (Romera, 2008; Reig, 2011). Estos significados, como explican las autoras, se representan en la enunciación en al menos cuatro funciones discursivas o conversacionales: contrastiva, aclarativa/justificativa, concesiva o de objeción/negación.

(3a) I: mmm él es de Valencia *lo que pasa es que* también ha vivido desde pequeño aquí

[PRESEEA-Palma 3M1 ambas]

(3b) I: nada / descansar (risa = "E") no *lo que pasa es que* como yo / a mí mmm yo estoy solo eeh a mí me cuesta yo creo que uno de los problemas de separarme de mi segunda mujer fue que yo / no asimilo en vivir con la familia / yo no soy muy familiar ni hogareño la verdad / yo prefiero más estar solo en mi casa que irme por ejemplo / me llevo muy bien con mi ex / con la segunda eeh me invitan a comer a su casa cada año por Navidad voy pero me encuentro un poco incómodo / de hecho últimamente ya no voy y y ella ya lo nota / de hecho la semana pasada fue el cumpleaños de mi exsuegra / yo fui a llevarle el regalo eeh pero ya como vi que no me encontré bien a las 8 y media le dije que me tenía que ir porque me habían llamado del trabajo / es decir que mmm / no me adapto y prefiero más yo estar solo que que que en familia / en Navidades pues igual / al fin y al cabo hay momentos en que tengo que ir pues mira porque tengo un hijo que está allá y bueno pues voy pero que no que no / me gusta más estar solo / y la Navidad la procuro pasar solo / quitando las fechas clave

[PRESEEA-Palma 3H2 ambas]

(3c) I: bueno pues cuando me sobra tiempo (risa = "E") de trabajar estudiar y estar con mi novia / que es lo que más me gusta (risa = "todos") / hay que decirlo / no a aparte pues me gusta hacer deporte // y si y si en el futuro se pudiera me gustaría volver a competir / concre concretamente jugar a fútbol // y bueno / es básicamente en lo que paso el tiempo / si igual igual que cualquier

joven de mi edad más o menos / tiene (carraspeo) una perioricidad para ver a sus amigos y salir y demás y divertirse / pero *lo que* la tónica de mi vida lo que más veo / es que me gustaría recuperar ese hábito de hacer deporte de manera constante // y no perderlo
[PRESEEA-Palma 3H1 castellano]

(3d) I: sí *lo que pasa es que* no se lo consintió al final
[PRESEEA-Alcalá 3H1 castellano]

En (3a) encontramos un contraste de los contenidos de las distintas proposiciones, sin que este contraste implique una contradicción entre las partes, sino más bien una adición (es de Valencia y ha vivido mucho aquí). En el caso de (3b), la función es otra, pues la intención del emisor es aclarar o justificar la información presentada en la primera parte de la enunciación para evitar que el interlocutor infiera determinadas conclusiones sobre lo dicho, como, por ejemplo, que va a dedicarse únicamente a descansar y pasar tiempo con la familia durante las vacaciones cuando en realidad eso es algo que no suele hacer porque no encaja con su forma de ser. En el ejemplo (3c) observamos una función concesiva en la que se presentan argumentos de una misma tesis en una dirección y en otra. El emisor explica que le gusta mucho pasar tiempo con su novia cuando no tiene que estudiar o trabajar, pero que al mismo tiempo le gustaría también retomar la actividad física y el deporte habitual en cuanto le sea posible. Por último, en (3d) se observa una relación de objeción, pues la segunda parte de la enunciación presenta un desacuerdo hacia lo expresado anteriormente (su hermano no consintió a su jefe que lo explotara, aunque, efectivamente, el trabajo era duro).

A pesar de los conocimientos teóricos que se tienen para explicar la estructura gramaticalizada de ‘lo que pasa es que’ en español, todavía es escasa la evidencia empírica sobre el uso de esta forma en los intercambios comunicativos de las distintas comunidades hispanohablantes. Como se ha mencionado en la introducción, esta investigación aspira a explorar la posible influencia de ciertas variables internas y externas en la expresión ‘lo que [pasa es que]’ y sus variantes en dos comunidades de habla concretas: Palma (ciudad española de 400.000 habitantes, situada en una zona bilingüe (castellano-catalán) en la isla de Mallorca) y Alcalá de Henares (ciudad española monolingüe de 200.000 habitantes, al noreste de la Comunidad de Madrid). Para que la muestra resulte válida en términos de comparabilidad, como se explica en la siguiente sección, se recurre al uso de dos corpus PRESEEA recogidos en estas ciudades, en los que se ha aplicado la metodología del proyecto y se han utilizado las mismas variables independientes.

El potencial interés de esta investigación radica en conocer por primera vez, desde una perspectiva empírica, la evolución de la forma ‘lo que pasa es que’ en una ciudad castellana monolingüe y en otra bilingüe; en esta última, en la que la forma ‘lo que’

reducida parece estar más avanzada según sugiere Romera (2008). En trabajos anteriores, ya se ha demostrado que las ciudades de Palma y Alcalá muestran diferencias lingüísticas en castellano, relacionadas con el contacto de lenguas, como, por ejemplo, una menor progresión en el cambio lingüístico (ya documentado en el s. XVII) de la expresión de futuro (Enrique-Arias & Méndez Guerrero, 2020). El futuro morfológico ‘cantaré’, de acuerdo con lo indicado en los trabajos de Sedano (2007), Blas-Arroyo (2008), Osborne (2008), Aaron (2010) u Orozco (2015), está siendo progresivamente sustituido por el futuro perifrástico ‘voy a cantar’ en la mayor parte de los territorios hispanohablantes. Este cambio en marcha en el español actual, que incluso se ha visto acelerado en algunas zonas de contacto de lenguas español-inglés, como, por ejemplo, Estados Unidos (Gutiérrez, 1995, 2002; Orozco, 2015), muestra, en cambio, en otras zonas de contacto castellano-catalán que el declive del futuro morfológico está menos avanzado (Wesch, 1997; Blas-Arroyo, 2004, 2007, 2008; Enrique-Arias, 2008, 2014; Enrique-Arias & Méndez Guerrero, 2020), y que la forma sintética es, incluso, la preferente entre los hablantes bilingües o catalanohablantes por la inexistencia de la forma perifrástica en catalán (Enrique-Arias & Méndez Guerrero, 2020). Estos estudios, por tanto, muestran la pertinencia de realizar investigaciones sobre otros fenómenos como el cambio lingüístico que se está produciendo con ‘lo que pasa es que’, que pongan en foco en las diferencias entre las zonas monolingües castellanas y las bilingües castellano-catalán.

2. Metodología

Para realizar el estudio se han analizado las 436 situaciones en las que aparece la forma gramaticalizada ‘lo que [pasa es que]’ en alguna de sus variantes en las 108 entrevistas semidirigidas que componen los corpus de PRESEEA-Palma y PRESEEA-Alcalá⁹. En ambos casos se siguió el mismo tipo de muestreo por cuotas de sexo, edad y nivel de instrucción, dicha información puede consultarse en la página web del proyecto: <http://preseea.linguas.net/> (Moreno, 2021). En cada comunidad contamos con 54 informantes, que están igualmente divididos en hombres y mujeres de tres grupos de edad y tres niveles de formación educativa, como puede verse en la Tabla 1.

Tabla 1. Distribución de los informantes a partir de las variables sociales en Palma y Alcalá.

| | 1º Grupo de Edad (18-34 años) | | 2º Grupo de Edad (34-55 años) | | 3º Grupo de Edad (+ de 55 años) | |
|----------------------------|----------------------------------|---------|----------------------------------|---------|------------------------------------|---------|
| | Hombres | Mujeres | Hombres | Mujeres | Hombres | Mujeres |
| Nivel de instrucción bajo | 3 | 3 | 3 | 3 | 3 | 3 |
| Nivel de instrucción medio | 3 | 3 | 3 | 3 | 3 | 3 |
| Nivel de instrucción alto | 3 | 3 | 3 | 3 | 3 | 3 |

Los participantes de Palma fueron seleccionados, adicionalmente, distinguiendo su grupo etnolingüístico (Azurmendi, Bourhis, Ros García & García, 1998), esto es,

atendiendo a su pertenencia al grupo de dominantes en castellano, catalán o ambas lenguas, a partir de la información extraída de una encuesta de datos biográficos y del uso de las lenguas de la zona que completaron. Por ende, en el caso de Palma, los tres participantes en cada celda correspondían a un hablante catalán dominante, uno español dominante y un bilingüe equilibrado (incluimos esta última categoría ya que observamos que un gran número de hablantes bilingües en la ciudad de Palma hacen uso de ambos idiomas a diario)¹⁰. Cada entrevista duró alrededor de 45 minutos. Siguiendo la metodología diseñada para el proyecto PRESEEA, se introdujo un número fijo de módulos temáticos en todas las entrevistas, algunos de los cuales (bloques 3, 4 y 5: lugar donde vive, familia y amistad, y costumbres) favorecen la aparición de las formas gramaticalizadas de ‘lo que pasa es que’.

Se han descartado del estudio todas aquellas formas introducidas por ‘lo que pasa es que’ con valor léxico que, por otra parte, eran muy escasas. El resto de las formas gramaticalizadas en alguna de sus variantes (‘lo que’ y ‘lo (único) que pasa/ocurre (es que)’) constituyen la variable dependiente de la investigación. La codificación de las formas resultantes se realizó atendiendo distintas variables independientes: por una parte, un factor lingüístico, la función comunicativa que realiza; y por otra, varios criterios extralingüísticos como el sexo, la edad, el nivel de instrucción, la localidad geográfica (considerando Palma vs. Alcalá), y en el caso de Palma, el grupo etnolingüístico. Para el análisis de datos utilizamos el paquete estadístico SPSS v. 26; generamos tablas con casos (frecuencias absolutas) y porcentajes (frecuencias relativas) y realizamos test de significación estadística (chi-cuadrado y valor de p). En las siguientes secciones justificamos la elección de variables y analizamos los resultados.

Como grupo de control se ha seleccionado el *Corpus oral juvenil del español de Mallorca* (COJEM), con aproximadamente 20 horas de conversaciones coloquiales entre 10 hablantes jóvenes universitarios palmesanos (5 mujeres y 5 hombres) muy amigos entre sí. Concretamente, el corpus se compone de 7 conversaciones espontáneas de casi 3 horas cada una, recogidas en lugares muy frecuentados por los informantes (cafeterías, domicilios particulares, vehículos, playa, etc.) de la ciudad de Palma mediante las técnicas de grabación secreta y observación participante (Méndez Guerrero, 2015). De esta manera, se pretende contar con un corpus oral adicional, ajeno al estudio, que sirva para validar los resultados.

3. Análisis de los resultados

3.1. El uso de la expresión ‘lo que pasa es que’ en las ciudades de Palma y Alcalá

Uno de los objetivos de este trabajo es explicar el funcionamiento de la construcción gramaticalizada ‘lo que pasa es que’ y sus variantes en dos comunidades de hablantes españoles (una monolingüe y otra bilingüe). La Tabla 2 ofrece las frecuencias de aparición en los dos corpus PRESEEA manejados en la investigación.

Los datos obtenidos reflejan frecuencias similares en ambos corpus, con una pequeña predominancia de casos en el corpus de la ciudad de Alcalá (229 casos, 52,5%) frente al corpus de la ciudad de Palma (207 casos, 47,5%).

Tabla 2. Frecuencia de formas gramaticalizadas ‘lo que pasa es que’ en Palma y Alcalá.

| | Casos Palma | | Casos Alcalá | | Total | |
|-----------------------|-------------|-------------|--------------|-------------|------------|--------------|
| | F | % | F | % | F | % |
| lo que | 32 | 15,5 | 0 | 0 | 32 | 7,3 |
| lo que ocurre es que | 9 | 4,3 | 1 | 0,4 | 10 | 2,3 |
| lo que pasa | 11 | 5,3 | 10 | 4,4 | 21 | 4,8 |
| lo que pasa es | 1 | 0,5 | 2 | 0,9 | 3 | 0,7 |
| lo que pasa es que | 77 | 37,2 | 89 | 38,9 | 166 | 38,1 |
| lo que pasa que | 61 | 29,5 | 109 | 47,6 | 170 | 39,0 |
| lo único | 8 | 3,9 | 8 | 3,5 | 16 | 3,7 |
| lo único que | 6 | 2,9 | 7 | 3,1 | 13 | 3,0 |
| lo único que pasa es | 2 | 1,0 | 0 | 0 | 2 | 0,5 |
| lo único que pasa que | 0 | 0 | 3 | 1,3 | 3 | 0,7 |
| Total | 207 | 47,5 | 229 | 52,5 | 436 | 100,0 |

$X^2= 57,314$, $p\text{-value}= 0,000$

Las frecuencias observadas en la tabla también nos permiten establecer diferencias entre las formas específicas o variantes más extendidas en la muestra. Las expresiones más habituales en ambas comunidades de hablantes son ‘lo que pasa es que’, utilizada en mayor medida en Palma (37,2%) y en menor medida en Alcalá (38,9%), y ‘lo que pasa que’, la más habitual en Alcalá (47,6%) y la segunda más utilizada en Palma (29,5%). En el caso de Palma, también se aprecia una frecuencia considerable, en tercera posición, de la forma más reducida ‘lo que’ (15,5%). Estos datos son similares a los presentados por Romera (2008) en su estudio contrastivo en las ciudades de Palma y Soria (ciudad castellana a 200 km al norte de Madrid), en el cual la expresión gramaticalizada más extendida en el global de los datos era ‘lo que pasa es que’ (49,5%), seguida por ‘lo que pasa que’ (31,2%). El tercer lugar lo ocupaba la forma ‘lo que’ (4,3%), cuya presencia se limitaba también a la ciudad de Palma¹¹.

Estos datos sugieren además que se produce una alternancia de las formas plenas (lo que pasa/ocurre es que) en el 41,5% (181 casos) con la pérdida de la cópula, (lo que pasa/ocurre que) en el 44,5% (194 casos) y con la pérdida además de la perífrasis de relativo (lo único que, lo que) en el 14% (61 casos) en las ciudades españolas estudiadas. Así pues, se observa aquí una tendencia algo mayor que en el trabajo de Romera hacia la reducción de las expresiones, principalmente por pérdida de la cópula y pérdida de la perífrasis de relativo (alrededor del 40% en total en el trabajo de Romera¹²). Sin embargo, los resultados por ciudades muestran una pequeña diferencia en las formas predominantes de cada grupo de hablantes que resulta estadísticamente significativa: en Palma la frecuencia es mayor en las formas con cópula (89 casos, 43%) frente a las que han perdido la cópula (72 casos, 34,8%) y las que también han perdido la perífrasis de relativo (46 casos, 22,2%); mientras que en la zona peninsular de Alcalá las formas sin cópula son las predominantes (122 casos, 53,3%) frente a las

expresiones con cópula (92 casos, 40,2%) y las formas sin perífrasis de relativo (15 casos, 6,5%). Por ende, en realidad, el proceso de gramaticalización que experimentan estas estructuras parece llevar ‘distintas velocidades o tendencias’ según la ciudad.

Al comparar estos resultados con los aportados por el grupo de control, conformado por 10 informantes palmesanos del corpus COJEM, observamos una tendencia similar a la del corpus PRESEEA-Palma con un predominio de la expresión plena (80%) frente a las formas reducidas (20%). Si bien, en este caso, la tendencia de la forma plena ‘lo que pasa/ocurre es que’ es incluso mayor que en los corpus manejados en el estudio. En futuras investigaciones, habrá que evaluar si las diferencias en los corpus están determinadas por la lengua predominante de los hablantes, por el contacto de lenguas o por cualquier otra circunstancia. En esta línea, Camargo y Enrique-Arias (2021) avanzan que la forma reducida ‘lo que’ se documenta también en otros corpus recogidos en Mallorca como el *Corpus Oral y Sonoro del Español Rural* (COSER) y el *Corpus de Wasaps – Mallorca* (CORWAS), así como en otros corpus PRESEEA no bilingües castellano-catalán como el de La Habana o el de Ciudad de Guatemala. Gallego (2006), por su parte, también documenta la forma ‘lo que’ en el castellano peninsular.

Resulta igualmente interesante analizar las preferencias de los distintos grupos de hablantes en relación con el verbo o adverbio que acompaña al relativo y precede a la cópula y la perífrasis.

Tabla 3. Variantes agrupadas de las formas gramaticalizadas ‘lo que pasa es que’ en Palma y Alcalá.

| | Casos Palma | | Casos Alcalá | | Total | |
|----------------------------|-------------|-------------|--------------|-------------|------------|--------------|
| | F | % | F | % | F | % |
| lo que | 32 | 15,5 | 0 | 0 | 32 | 7,3 |
| lo que pasa (es) (que) | 152 | 73,4 | 211 | 92,1 | 363 | 83,3 |
| lo que ocurre es que | 7 | 3,4 | 0 | 0 | 7 | 1,6 |
| lo único (que pasa es/que) | 16 | 7,7 | 18 | 7,9 | 34 | 7,8 |
| Total | 207 | 47,5 | 229 | 52,5 | 436 | 100,0 |

$X^2= 47,719$, $p\text{-value}= 0,000$

Los resultados presentados en la Tabla 3 no dejan lugar a duda sobre la preferencia de los hablantes respecto a la forma más extendida en sus ciudades. Este hecho queda especialmente claro en la ciudad de Alcalá, donde hasta el 92,1% de los casos corresponden a la variante ‘lo que pasa (es) (que)’ frente al resto de formas gramaticalizadas recogidas en los corpus: ‘lo que ocurre es que’, ‘lo único (que pasa es/que)’ o ‘lo que’. Por su parte, la ciudad de Palma presenta mayor variación en la elección de las expresiones, aunque los resultados son igualmente significativos y muestran un predominio evidente de la forma ‘lo que pasa (es) (que)’ (73,4%). Al comparar estos resultados globales con los obtenidos por Romera (2008), se aprecia una tendencia similar con más de un 80% de ‘lo que pasa (es) (que)’ y frecuencias entre el 4% y el 6% del resto de variantes. Diferentes autores apuntan que la elección

de los hablantes de una forma frente a las demás puede responder también a las necesidades discursivas (Morales, 2005; Romera, 2008). En el siguiente apartado vamos a determinar qué funciones comunicativas asignan los hablantes a dichas construcciones y qué relación existe entre algunas de las formas más habituales y sus funciones.

3.2. Incidencia de las variables lingüísticas en la expresión ‘lo que pasa es que’

Como se explicó en el apartado 1 de este trabajo, son varias las funciones discursivas asignadas a la forma gramaticalizada ‘lo que pasa es que’ y a todas sus variantes (Romera, 2008; Reig, 2011). En lo que sigue, tanto para el análisis de las variables internas (funciones discursivas) como para el de las variables externas (sexo, edad, nivel formativo y grupo etnolingüístico, que se presentarán en el punto 3.3.), continuaremos la explicación del análisis agrupando, en primer lugar, los casos por las variantes presentadas en la Tabla 3: ‘lo que’, ‘lo que pasa/ocurre (es) (que)’, ‘lo que ocurre es que’ y ‘lo único (que pasa es/que)’. De ese modo evitaremos las frecuencias bajas que se han obtenido en algunas expresiones, que limitan el análisis en cuanto a una muestra suficiente. Así veremos más claramente cuáles son las formas preferentes de los hablantes frente a la función discursiva y al perfil sociocultural del informante. No obstante, en ciertos momentos se recurrirá también a los datos no agrupados para presentar los resultados. La Tabla 4 muestra las funciones discursivas observadas en el corpus PRESEEA-Palma en relación con las variantes agrupadas.

Tabla 4. Funciones discursivas de las formas gramaticalizadas ‘lo que pasa es que’ en Palma (datos agrupados).

| Casos Palma | lo que | lo que pasa (es) (que) | lo que ocurre es que | lo único (que pasa es/que) | Total |
|---------------------------|--------------------|------------------------|----------------------|----------------------------|--------------------|
| Contrastiva | 25 (78,1%) | 132 (86,8%) | 7 (100,0%) | 11 (68,8%) | 175 (84,5%) |
| Aclarativa/ justificativa | 2 (6,3%) | 18 (11,8%) | 0 (0,0%) | 3 (18,8%) | 23 (11,1%) |
| Concesiva | 4 (12,5%) | 2 (1,3%) | 0 (0,0%) | 2 (12,5%) | 8 (3,9%) |
| Objeción/ negación | 1 (3,1%) | 0 (0,0%) | 0 (0,0%) | 0 (0,0%) | 1 (0,5%) |
| Total | 32 (100,0%) | 152 (100%) | 7 (100%) | 16 (100%) | 207 (100%) |

$X^2= 20,848$, $p\text{-value}= 0,013$

Los resultados obtenidos en la muestra de los informantes isleños sugieren una consistencia estadística y reflejan la existencia de una función discursiva preferente: la contrastiva (84,5%), para la expresión gramaticalizada introducida por ‘lo que’. Dicha función se representa especialmente a través de la forma ‘lo que pasa (es) (que)’ (86,8%). La segunda función discursiva que presenta una mayor presencia es la aclarativa/justificativa, si bien la frecuencia es mucho menor que en el caso de la función contrastiva (11,1%). Esta función se representa mayormente con la forma ‘lo único (que pasa es/que)’ (18,8%). Los casos registrados en el PRESEEA-Palma con funciones discursivas diferentes a las mencionadas hasta el momento son muy limitados: el 3,9% para la función concesiva, con una presencia mayoritaria de la

forma 'lo único (que pasa es/que)' y 'lo que' (12,5%), y tan solo el 0,5% para la función de objeción/negación, que se representa exclusivamente con la forma 'lo que' (3,1%).

Tabla 5. Funciones discursivas de las formas gramaticalizadas 'lo que pasa es que' en Alcalá (datos agrupados).

| Casos Alcalá | lo que pasa (es) (que) | lo único (que pasa es/que) | Total |
|---------------------------|------------------------|----------------------------|--------------------|
| Contrastiva | 175 (82,9%) | 13 (72,2%) | 188 (82,1%) |
| Aclarativa/ justificativa | 32 (15,2%) | 4 (22,2%) | 36 (15,7%) |
| Concesiva | 3 (1,4%) | 1 (5,6%) | 4 (1,7%) |
| Objeción/ negación | 1 (0,5%) | 0 (0,0%) | 1 (0,4%) |
| Total | 211 (100,0%) | 18 (100%) | 229 (100%) |

$X^2= 2,465$; $p\text{-value}= 0,482$

Los datos obtenidos en la ciudad de Alcalá (Tabla 5) muestran diferencias respecto a la ciudad de Palma. La mayoría de ellas están relacionadas principalmente con la forma de representar las funciones y no tanto con la proporción o frecuencia de aparición de las distintas funciones discursivas. La función discursiva predominante en el PRESEEA-Alcalá es igualmente la contrastiva, si bien el porcentaje de aparición es algo menor que en Palma (82,1%). A esta le sigue la función aclarativa/justificativa con un porcentaje ligeramente superior al registrado en el PRESEEA-Palma (15,7%). Una vez más, son muy pequeñas las frecuencias de la función concesiva (1,7%) y la de objeción/negación (0,4%). Estos datos coinciden con los presentados por Cuadros (2015), donde la contrastiva o contraargumentativa es la función principal de 'lo único (que)'. En las formas de representar las funciones, como decíamos, se aprecian más diferencias. El primer hecho notorio es que hay menos variación en Alcalá en las formas gramaticalizadas introducidas por 'lo que' que utilizan los hablantes (no hay casos de 'lo que ocurre es que' y 'lo que'), cuestión que ha repercutido en esa menor variación al representar las funciones. Eso no impide, no obstante, que predomine la forma 'lo único que pasa es que' para las funciones aclarativa/justificativa (22,2%) y concesiva (5,6%), como ya ocurría con los informantes palmesanos, y la forma 'lo que pasa es que' para las funciones contrastiva (82,9%) y la de negación/objeción (2,5%).

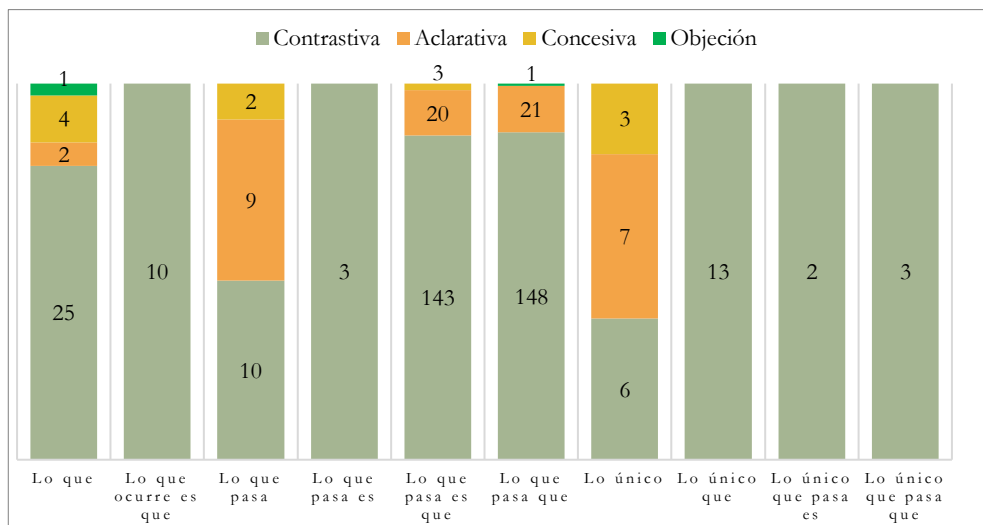
Tabla 6. Funciones discursivas de las formas gramaticalizadas 'lo que pasa es que' (datos globales).

| | Casos Palma | | Casos Alcalá | | Total | |
|---------------------------|-------------|--------------|--------------|--------------|------------|--------------|
| | F | % | F | % | F | % |
| Contrastiva | 175 | 84,5 | 188 | 82,1 | 363 | 83,3 |
| Aclarativa/ justificativa | 23 | 11,1 | 36 | 15,7 | 59 | 13,5 |
| Concesiva | 8 | 3,9 | 4 | 1,7 | 12 | 2,8 |
| Objeción/ negación | 1 | 0,5 | 1 | 0,5 | 2 | 0,5 |
| Total | 207 | 100,0 | 229 | 100,0 | 436 | 100,0 |

$X^2= 3,562$; $p\text{-value}= 0,313$

Los datos globales (Tabla 6) permiten ver con más claridad la frecuencia de las funciones discursivas asociadas a las formas gramaticalizadas de 'lo que pasa es que': contrastiva (83,3%) con una tendencia algo mayor en la ciudad de Palma; aclarativa/justificativa (13,5%) con mayor incidencia en Alcalá; concesiva (2,8%) con una frecuencia de más del doble en Palma; y de objeción/negación (0,5%) con tan solo un caso registrado en cada corpus. Estos resultados coinciden con los presentados por Romera (2008), donde se apreciaba igualmente una mayor presencia de la función contrastiva en torno al 50% del total de casos registrados, frente al 25% aproximadamente de la función aclarativa y el 15% y 10% de las funciones de objeción y concesión respectivamente. El grupo de control analizado refleja igualmente esta tendencia al predominar claramente la función contrastiva (en torno al 83%) y registrar pocos casos del resto de funciones.

Los resultados globales de las distintas funciones discursivas son igualmente interesantes si se revisan atendiendo cada una de las variantes registradas en los corpus.



$$X^2 = 80,459; p\text{-value} = 0,000$$

Gráfico 1. Funciones discursivas de las formas gramaticalizadas 'lo que pasa es que' (datos sin agrupar).

Como se puede apreciar, en todas las formas destaca la función contrastiva, al ser la predominante en la muestra. Sin embargo, estos resultados nos permiten conocer otra cuestión interesante: las formas que han representado funciones distintas a la contrastiva y las que se han limitado exclusivamente a expresar la función contrastiva. En este sentido, observamos cómo algunas de las variantes sin cópula, como 'lo que', 'lo que pasa', 'lo que pasa que' y 'lo único', presentan más variación a la hora de representar funciones discursivas distintas. Es más, en el caso de la función concesiva y la de objeción/negación, no encontramos casos en los que dicha función sea

representada por una construcción plena. La función aclarativa, al igual que la contrastiva, probablemente al presentar más casos muestra ser más versátil y admitir una representación más variada de expresiones. Resta determinar si dicha variabilidad guarda además relación con los factores sociales de los informantes que conforman los distintos grupos de hablantes y que pasamos a analizar en el apartado 3.3.

3.3. Incidencia de las variables extralingüísticas en la expresión ‘lo que pasa es que’

En la parte del análisis relativa a la influencia de los factores sociales en las formas gramaticalizadas de ‘lo que pasa es que’, se han considerado las siguientes variables independientes: edad, sexo, nivel de instrucción y grupo etnolingüístico. En las siguientes tablas, puede verse la incidencia de la primera de estas variables, la edad, en la producción de la expresión estudiada en los corpus PRESEEA-Palma y PRESEEA-Alcalá.

Tabla 7. La ‘edad’ y las formas gramaticalizadas ‘lo que pasa es que’ en Palma (datos agrupados).

| Casos Palma | lo que | lo que pasa (es) (que) | lo que ocurre es que | lo único (que pasa es/que) | Total |
|--------------|------------------|------------------------|----------------------|----------------------------|-------------------|
| 18-34 años | 14 (43,8%) | 46 (30,3%) | 0 (0,0%) | 7 (43,8%) | 67 (32,4%) |
| 35-55 años | 8 (25,0%) | 52 (34,2%) | 7 (100,0%) | 6 (37,5%) | 73 (35,3%) |
| +55 años | 10 (31,3%) | 54 (35,5%) | 0 (0,0%) | 3 (18,8%) | 67 (32,4%) |
| Total | 32 (100%) | 152 (100%) | 7 (100%) | 16 (100%) | 207 (100%) |

$X^2= 17,403$, $p\text{-value}= 0,008$

Los resultados presentados en la Tabla 7 no permiten hablar de una incidencia clara de la edad con una mayor producción de construcciones gramaticalizadas de la forma ‘lo que pasa es que’, pues los tres grupos etarios presentan frecuencias similares: grupo joven (32,4%), grupo adulto (35,3%) y grupo de mayores (32,4%). En cambio, sí resulta más significativo el uso de las distintas variantes de la expresión ‘lo que pasa es que’ que realizan los informantes según su edad. Por ejemplo, llama la atención el hecho de que sea el grupo de jóvenes de entre 18 y 34 años el que más recurra a la forma más reducida, ‘lo que’ (43,8%). Esto ocurre también con otra de las formas más innovadoras ‘lo único (que pasa es/que)’ (43,8%), muy frecuente en los jóvenes y con poca incidencia en los mayores de 55 años (18,8%). También es llamativo que el grupo de más de 55 años sea el que mayoritariamente prefiere la forma plena o, dicho de otro modo, la más tradicional y la que más tiempo ha estado entre los hablantes (35,5%) frente al grupo de jóvenes que es el que menos la utiliza (30,3%). El grupo de mediana edad, de entre 35 y 55 años, solo muestra mayor incidencia que el resto de los grupos en la forma ‘lo que ocurre es que’ y en el resto de las variantes agrupadas se encuentra en una posición intermedia entre la mayor incidencia en los mayores en un caso y la de los jóvenes en otro. En la forma ‘lo que’, la forma más innovadora y que no se ha registrado en la ciudad peninsular de Alcalá, se aprecia que el grupo medio de edad presenta una tendencia algo menor de producción que los otros dos grupos, dato

que podría explicarse tal vez por la cautela del grupo hacia esta forma más innovadora de la expresión gramaticalizada, que no se utiliza en todas las ciudades de habla castellana. Así pues, con estos resultados parece posible afirmar que hay una tendencia a la incorporación de formas más novedosas por parte del grupo de jóvenes seguido por el grupo de adultos, salvo en la forma más reducida, más restringida geográficamente, preferida por los jóvenes y después por los mayores. Este dato coincide con lo observado también en el corpus CORWAS (Camargo, 2021). Los mayores de 55 años se muestran más apegados a las formas tradicionales y extendidas en otros lugares como Alcalá (o Soria en el estudio de Romera). Así pues, la apreciación es que los adultos tienden a mostrarse más cercanos a lo prestigioso o normativo que los jóvenes (más innovadores/rompedores) y los mayores (más apegados al vernáculo, a los usos propios (idiolectales) y locales (geolectales) y menos a los prestigiosos o normativos (Zimmerman, 2002; Eckert & McConnell-Ginet, 2007).

Tabla 8. La variable ‘edad’ y las formas gramaticalizadas ‘lo que pasa es que’ en Alcalá (datos agrupados).

| Casos Alcalá | lo que pasa (es) (que) | lo único (que pasa es/que) | Total |
|--------------|------------------------|----------------------------|-------------------|
| 18-34 años | 81 (38,4%) | 4 (22,2%) | 85 (37,1%) |
| 35-55 años | 54 (25,6%) | 8 (44,4%) | 62 (27,1%) |
| +55 años | 76 (36,0%) | 6 (33,3%) | 82 (35,8%) |
| Total | 211 (100,0%) | 18 (100%) | 229 (100%) |

$X^2= 3,378$, $p\text{-value}= 0,185$

En el caso de Alcalá, la incidencia de casos registrados por edad tampoco parece ser determinante: grupo de entre 18-34 años (37,1%), entre 35-55 años (27,1%) y más de 55 años (35,8%). Y la menor variabilidad de formas para expresar la expresión ‘lo que pasa es que’ en la ciudad solo permite establecer una diferencia entre la forma más tradicional y una de las formas innovadoras. En este sentido y comparando los resultados con los obtenidos en la ciudad de Palma, llama la atención especialmente que sea el grupo de jóvenes el que menos recurre a la forma catalogada como más novedosa (‘lo único que pasa es/que’) con un 22,2% de los casos y que este mismo grupo sea el que más utiliza la forma tradicional (38,4%). Para que estos resultados se entiendan mejor, deben verse junto a los relativos a la elección de la forma gramaticalizada y al grupo etnolingüístico (Tabla 9).

Tabla 9. El ‘grupo etnolingüístico’ y las formas gramaticalizadas ‘lo que pasa es que’ en Palma (datos agrupados).

| Casos Palma | lo que | lo que pasa (es) (que) | lo que ocurre es que | lo único (que pasa es/que) | Total |
|------------------------|------------------|------------------------|----------------------|----------------------------|-------------------|
| Bilingües equilibrados | 10 (31,3%) | 60 (39,5%) | 0 (0,0%) | 3 (18,8%) | 73 (35,3%) |
| Dominantes castellano | 10 (31,3%) | 59 (38,8%) | 0 (0,0%) | 6 (37,5%) | 75 (36,2%) |
| Dominantes catalán | 12 (37,5%) | 33 (21,7%) | 7 (100,0%) | 7 (43,8%) | 59 (28,5%) |
| Total | 32 (100%) | 152 (100%) | 7 (100%) | 16 (100%) | 207 (100%) |

$X^2= 24,887$, p-value= 0,000

Los datos presentados en esta tabla sugieren que el grupo etnolingüístico de los informantes es una de las variables que más incidencia tiene en la elección de las formas gramaticalizadas de ‘lo que pasa es que’. En la ciudad de Palma, se aprecia que las mayores frecuencias en todas las formas consideradas más innovadoras son las producidas por el grupo de hablantes catalanes: ‘lo que’ (37,5%), ‘lo que ocurre es que’ (100%) y ‘lo único que pasa es que’ (43,8%); seguidos en dos de los tres casos por los hablantes que tienen un manejo equilibrado del castellano y el catalán¹³. La única forma en la que predomina el grupo castellano frente al catalán es la forma tradicional ‘lo que pasa es que’ (solo el 21,7% de las realizaciones pertenecen a catalanohablantes), como se ha visto que ocurre en Alcalá (Tabla 8). Así pues, tras ver estos resultados, sorprende menos que en la ciudad de Alcalá, en la que solo se han tenido en cuenta informantes monolingües de castellano, predomine aún más la forma tradicional que en Palma, donde es mayoritariamente representada por los castellanohablantes. Esto se ve reflejado también en el resto de las edades en la ciudad de Palma.

Tabla 10. El ‘sexo’ y las formas gramaticalizadas ‘lo que pasa es que’ en Palma (datos agrupados).

| Casos Palma | lo que | lo que pasa (es) (que) | lo que ocurre es que | lo único (que pasa es/que) | Total |
|--------------|------------------|------------------------|----------------------|----------------------------|--------------------|
| Hombre | 19 (59,4%) | 80 (52,6%) | 0 (0,0%) | 15 (93,8%) | 114 (55,1%) |
| Mujer | 13 (40,6%) | 72 (47,4%) | 7 (100,0%) | 1 (6,3%) | 93 (44,9%) |
| Total | 32 (100%) | 152 (100%) | 7 (100%) | 16 (100%) | 207 (100%) |

$X^2= 18,860$, p-value= 0,000

Los resultados en relación con la variable sexo tampoco muestran frecuencias muy diferentes entre la frecuencia de producción de las formas gramaticalizadas y la variación de sexo (los hombres el 55,1% y las mujeres el 44,9%), pero sí que se aprecian algunas diferencias más significativas en la incidencia de ciertas formas en hombres y mujeres. Esto se aprecia especialmente en el caso de la forma ‘lo único que pasa es que’, que es producida en el 93,8% de los casos por hombres, o en la forma ‘lo que ocurre es que’, cuya realización es exclusivamente femenina. En el caso de las formas más habituales, ‘lo que pasa es que’ y ‘lo que’, las frecuencias en cuanto al sexo están más equilibradas, si bien en ambos casos es algo mayor en los hombres. Esto se observa especialmente en el caso de la forma ‘lo que’, donde las realizaciones masculinas son casi hasta un 20% superiores. Varios estudios sociolingüísticos han determinado que las mujeres son generalmente las impulsoras de los cambios o

innovaciones lingüísticas, pero que esto únicamente ocurre cuando la forma innovadora es prestigiosa o cuando ya está normativizada (García Mouton, 2003; Molina, 2006). Este hecho puede ayudar a comprender mejor lo que quizá ocurra con las formas gramaticalizadas de 'lo que pasa es que', pues las mujeres utilizan y, en algunos casos, lideran el uso de las formas innovadoras, como 'lo que ocurre es que' y 'lo único que pasa es que', en el caso de Alcalá (Tabla 12), pero no se decantan de igual modo por las formas menos consolidadas o todavía no normalizadas en otros grupos de hablantes de su variedad o en otras zonas, como en este caso ocurre con la forma 'lo que'.

Tabla 11. El 'sexo' y las formas gramaticalizadas 'lo que pasa es que' en Alcalá (datos agrupados).

| Casos Alcalá | lo que pasa (es) (que) | lo único (que pasa es/que) | Total |
|--------------|------------------------|----------------------------|--------------------|
| Hombre | 96 (45,5%) | 8 (44,4%) | 104 (45,4%) |
| Mujer | 115 (54,5%) | 10 (55,6%) | 125 (54,6%) |
| Total | 211 (100,0%) | 18 (100%) | 229 (100%) |

$X^2= 0,007$, $p\text{-value}= 0,931$

Los resultados de Alcalá muestran en cambio que son las mujeres, aunque con una diferencia muy pequeña, las que en mayor medida recurren a las formas gramaticalizadas como objeto de estudio (las mujeres el 54,6% y los hombres el 45,4%). La elección de las variantes de la expresión en este caso no parece estar determinada por el sexo de los informantes, pues en todas las variantes las diferencias entre géneros no superan el 5%. Esta incidencia, sin embargo, es mucho mayor en el caso del nivel de instrucción (Tablas 12 y 13).

Tabla 12. El 'nivel de instrucción' y las formas gramaticalizadas 'lo que pasa es que' en Palma (datos agrupados).

| Casos Palma | lo que | lo que pasa (es) (que) | lo que ocurre es que | lo único (que pasa es/que) | Total |
|-------------------|------------------|------------------------|----------------------|----------------------------|-------------------|
| Instrucción Baja | 12 (37,5%) | 43 (28,3%) | 0 (0,0%) | 5 (31,3%) | 60 (29%) |
| Instrucción Media | 9 (28,1%) | 53 (34,9%) | 0 (0,0%) | 3 (18,8%) | 65 (31,4%) |
| Instrucción Alta | 11 (34,4%) | 56 (36,8%) | 7 (100,0%) | 8 (50%) | 82 (39,6%) |
| Total | 32 (100%) | 152 (100%) | 7 (100%) | 16 (100%) | 207 (100%) |

$X^2= 13,984$, $p\text{-value}= 0,030$

En relación con el nivel formativo, observamos que en la ciudad de Palma son las personas con formación universitaria las que utilizan en mayor proporción las formas gramaticalizadas (39,6%). Asimismo, a medida que el nivel de formación baja también disminuye la producción de este tipo de construcciones: estudios secundarios (31,4%) y estudios primarios (29%). Respecto a la variabilidad en este tipo de expresiones, solo hay una forma ('lo que') que presenta frecuencias superiores en un nivel de instrucción (el bajo) diferente al nivel de instrucción superior con un 37,5%. Por tanto, vemos una vez más cómo la forma más innovadora y restringida geográficamente presenta una incidencia menor en el grupo de personas más formadas, como ya ocurría con las

mujeres y con el grupo de adultos, más preocupados por el uso de las formas prestigiosas y/o normativas. La forma tradicional ‘lo que pasa es que’ es en la que se muestran las diferencias más pequeñas en cuanto a su producción en los distintos grupos formativos. Este hecho tiene que ver, probablemente, con ser la forma más extendida entre la población en general.

Tabla 13. El ‘nivel de instrucción’ y las formas gramaticalizadas ‘lo que pasa es que’ en Alcalá (datos agrupados).

| Casos Alcalá | lo que pasa (es) (que) | lo único (que pasa es/que) | Total |
|-------------------|------------------------|----------------------------|--------------------|
| Instrucción Baja | 73 (34,6%) | 4 (22,2%) | 77 (33,6%) |
| Instrucción Media | 98 (46,4%) | 10 (55,6%) | 108 (47,2%) |
| Instrucción Alta | 40 (19,0%) | 4 (22,2%) | 44 (19,2%) |
| Total | 211 (100,0%) | 18 (100%) | 229 (100%) |

$X^2= 1,139$, $p\text{-value}= 0,566$

En Alcalá, por su parte, los datos muestran una tendencia diferente a los descritos en Palma. En este caso es el grupo de instrucción medio el que mayores frecuencias presenta (47,2%) y el grupo con estudios superiores el que menos (19,2%). Pero no es hasta que se analizan los datos por variantes cuando puede apreciarse que el grupo de informantes con estudios superiores, como ya ocurría en Palma, recurre en mayor medida a la forma más innovadora ‘lo único que pasa es que’ (22,2%) frente a la forma tradicional ‘lo que pasa es que’ (19%). Por tanto, aunque en menor medida, también vemos aquí mayor innovación en el grupo de nivel de instrucción alto.

Los datos globales de las frecuencias en relación con los factores extralingüísticos permiten ver con más claridad las diferentes tendencias en algunos casos entre las ciudades estudiadas y las coincidencias en otros momentos que refuerzan la dirección del posible cambio lingüístico (Tablas 14, 15, 16 y 17).

Tabla 14. La ‘edad’ y las formas gramaticalizadas ‘lo que pasa es que’ (datos globales).

| | Casos Palma | | Casos Alcalá | | Total | |
|--------------|-------------|--------------|--------------|--------------|------------|--------------|
| | F | % | F | % | F | % |
| 18-34 años | 67 | 32,4 | 85 | 37,1 | 152 | 34,9 |
| 35-55 años | 73 | 35,3 | 62 | 27,1 | 135 | 31,0 |
| + 55 años | 67 | 32,4 | 82 | 35,8 | 149 | 34,2 |
| Total | 207 | 100,0 | 229 | 100,0 | 436 | 100,0 |

$X^2= 3,437$, $p\text{-value}= 0,179$

Tabla 16. El ‘nivel de instrucción’ y las formas gramaticalizadas ‘lo que pasa es que’ (datos globales).

| | Casos Palma | | Casos Alcalá | | Total | |
|-------------------|-------------|------------|--------------|------------|------------|------------|
| | F | % | F | % | F | % |
| Instrucción baja | 60 | 29 | 77 | 34 | 77 | 31 |
| Instrucción media | 65 | 31 | 108 | 47 | 108 | 40 |
| Instrucción alta | 82 | 40 | 44 | 19 | 44 | 29 |
| Total | 207 | 100 | 229 | 100 | 436 | 100 |

$X^2= 23,207$, $p\text{-value}= 0,000$

Tabla 15. El ‘sexo’ y las formas gramaticalizadas ‘lo que pasa es que’ (datos globales).

| | Casos Palma | | Casos Alcalá | | Total | |
|--------------|-------------|------------|--------------|------------|------------|------------|
| | F | % | F | P | F | % |
| Hombres | 114 | 55 | 104 | 45 | 218 | 50 |
| Mujeres | 93 | 45 | 125 | 55 | 218 | 50 |
| Total | 207 | 100 | 229 | 100 | 436 | 100 |

$X^2= 4,056$, p-value= 0,044

Tabla 17. El ‘grupo etnolingüístico’ y las formas gramaticalizadas ‘lo que pasa es que’ (datos globales).

| | Casos Palma | | Casos Alcalá | | Total | |
|--------------|-------------|--------------|--------------|--------------|------------|--------------|
| | F | % | F | % | F | % |
| Bilingües | 73 | 35,3 | 0 | 0 | 73 | 16,7 |
| Castellano | 75 | 36,2 | 229 | 100,0 | 304 | 69,7 |
| Catalán | 59 | 28,5 | 0 | 0 | 59 | 13,5 |
| Total | 207 | 100,0 | 229 | 100,0 | 436 | 100,0 |

$X^2= 209,436$, p-value= 0,000

Mientras que en la ciudad de Palma la mayor incidencia global de casos se da en los hombres, el grupo adulto, los informantes universitarios y los castellanohablantes. En Alcalá, las frecuencias de realización más elevadas pertenecen a las mujeres, al grupo joven y al grupo con estudios medios. Estas tendencias diferentes pueden explicarse por las distintas funciones discursivas que asignan principalmente cada grupo de hablantes a las formas gramaticalizadas y por las diversas variantes que utilizan en las realizaciones registradas en los corpus. También resulta interesante revisar los resultados comentados en este apartado en relación con las formas gramaticalizadas sin agrupar.

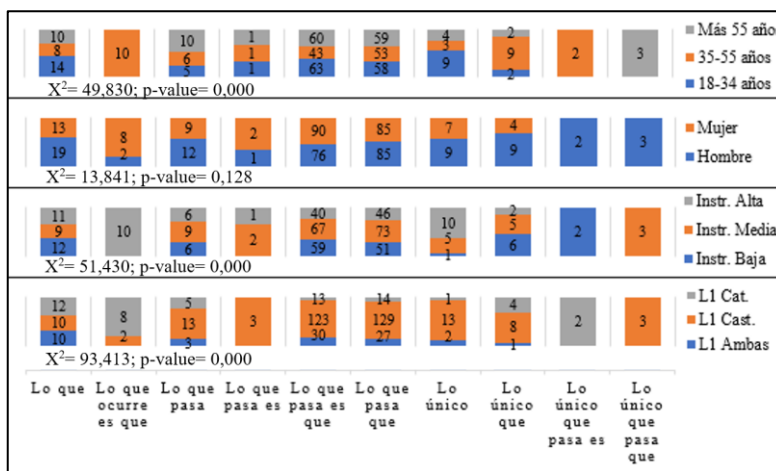


Gráfico 2. Factores extralingüísticos en relación con las formas gramaticalizadas ‘lo que’ (datos sin agrupar).

Los datos no agrupados mostrados en la Gráfica 2 permiten comprobar, en primer lugar, cómo algunas de las variantes de las formas gramaticalizadas solo aparecen en

ciertos grupos de informantes. Este es el caso, por ejemplo, de las formas 'lo único que pasa es' o 'lo único que pasa que', cuyas realizaciones se limitan al grupo de hombres adultos catalanohablantes con estudios primarios, en el caso de la primera forma; y al grupo de hombres mayores castellanohablantes con instrucción media, en la segunda forma. Algo parecido sucede con la forma 'lo que ocurre es que' y 'lo que pasa es', aunque esto no afecta a todas las variables extralingüísticas como en los casos anteriores. Aquí son el grupo adulto con un nivel formativo alto, en el primer caso, y los castellanohablantes, en el segundo, los que producen en exclusiva esta forma gramaticalizada. El resto de las frecuencias, como puede apreciarse, están más o menos repartidas entre los distintos grupos de hablantes y las incidencias varían en mayor o menor medida.

El análisis multivariable realizado para este estudio confirma que estas relaciones entre variables sociales y la realización de cualquiera de las variantes de 'lo que pasa es que' que acabamos de mencionar son estadísticamente significativas. En total, se ha determinado que resultan significativas las siguientes relaciones: forma gramaticalizada-funciones discursivas-ciudad ($X^2= 28,235$, p-value= 0,001); forma gramaticalizada-edad-ciudad ($X^2= 19,228$, p-value= 0,004); forma gramaticalizada-sexo-ciudad ($X^2= 12,694$, p-value= 0,005); forma gramaticalizada-instrucción-ciudad ($X^2= 20,897$, p-value= 0,002); forma gramaticalizada-lengua-ciudad ($X^2= 78,087$, p-value= 0,000). Así pues, estos resultados indican que en ocasiones es la conjunción del origen del informante (Palma o Alcalá) junto al resto de variables independientes (función discursiva, edad, sexo, nivel formativo y grupo etnolingüístico) lo que determina la incidencia de casos y la elección de la forma gramaticalizada.

CONCLUSIONES

El uso gramaticalizado de la expresión 'lo que pasa es que' ha ganado protagonismo entre los hablantes de español en los últimos tiempos por el uso predominante que se hace de esta forma (Romera, 2008). Aunque todavía no está claro su valor como marcador discursivo o unidad funcional (Portolés, 2004; Romera, 2008; Briz, Pons, Portolés, 2008; Reig, 2011; Cuadros, 2015), su papel en la interacción es fundamentalmente procedimental: guía las inferencias de los hablantes hacia las distintas interpretaciones dependiendo de cuál sea la intención o función que le asignen los hablantes. Uno de los objetivos fundamentales de esta investigación ha sido, precisamente, determinar las funciones discursivas que desempeñan estos elementos en la comunicación. Entre las funciones que se han asignado en distintos trabajos a estas construcciones (Romera, 2008; Reig, 2011) hemos identificado en los corpus PRESEEA de Palma y Alcalá analizados la función contrastiva, la más frecuente, y las aclarativa/justificativa, concesiva y de objeción/negación con porcentajes mucho más bajos. También se ha visto que los hablantes del estudio prefieren algunas de las variantes de esta forma gramaticalizada para representar ciertas funciones. Las formas predominantes son aquellas que han perdido la cópula,

al igual que ocurría en el trabajo de Romera (2008), como ‘lo único que’ para la función aclarativa/justificativa y concesiva o ‘lo que’ también para esta última función.

Cumpliendo con otro de los objetivos de la investigación, se ha estudiado también la incidencia de los factores extralingüísticos en la producción y elección de formas o variantes de la construcción pragmática. En este sentido se ha determinado que las variables sociales (como la edad, el nivel formativo o el grupo etnolingüístico) inciden especialmente en la producción y elección de las formas; y que, en ocasiones, es la conjunción de las variables independientes del estudio la que determina el uso que se hace de la expresión. El grupo adulto es el que se muestra más reacio a la incorporación de las formas más novedosas ‘lo que’ o ‘lo único que’, frente a los jóvenes, los más proclives e innovadores, y los mayores, también favorables y más apegados al idiolecto o estilo comunicativo propio no necesariamente normativo (Zimmerman, 2002; Eckert & McConnell-Ginet, 2007). Los catalanohablantes y bilingües equilibrados son igualmente más propicios a estas formas innovadoras. Esto mismo ocurre con los hombres frente a las mujeres, contraviniendo la idea de que las mujeres son más innovadoras y las propulsoras de los cambios lingüísticos, pues en realidad los cambios solo son difundidos por ellas cuando la forma es prestigiosa y está normativizada (García Mouton, 2003; Molina, 2006). En el futuro será necesario determinar qué otros factores incluyen sobre la propagación de las formas innovadoras en el castellano peninsular (Camargo & Enrique-Arias, 2021).

Por último, se ha comprobado en relación con el tercer objetivo del trabajo, esto es, el establecimiento de diferencias en Palma y Alcalá que ambas ciudades presentan diferencias en las formas más frecuentes y en las preferencias de los hablantes, muchas de las cuales tienen que ver con su perfil social. Este es el caso de la forma ‘lo único que’, solo empleada por hombres adultos y mayores con estudios medios y bajos o la expresión ‘lo que ocurre’ únicamente utilizada por hablantes con formación universitaria. El análisis multivariable también confirma relaciones entre varias variables independientes al mismo tiempo y la forma gramaticalizada de ‘lo que pasa es que’, como, por ejemplo, que la forma ‘lo que pasa es que’ siempre tiene relación estadística significativa de forma conjunta con la ciudad (Palma o Alcalá) y con alguno de los siguientes factores lingüísticos y extralingüísticos estudiados: función discursiva, edad, instrucción o grupo etnolingüístico. Todo lo anterior permite hablar de distintas velocidades o tendencias dentro del cambio lingüístico, pues, aunque en ambas ciudades está claramente establecido el uso de la forma gramaticalizada de ‘lo que pasa es que’ y sus variantes las formas más innovadoras solamente están presentes en la ciudad de Palma y son principalmente impulsadas por hablantes más jóvenes dominantes en catalán o bilingües con formación media/baja. En futuros trabajos sería interesante comprobar el funcionamiento de estas estructuras en otros grupos de hablantes de otras ciudades hispanohablantes (Camargo & Enrique-Arias, 2021). También sería fundamental tener en cuenta otras posibles variables independientes

como el conocimiento de lenguas extranjeras o la profesión y analizar estas formas gramaticalizadas en discursos orales más espontáneos, que vayan más allá de las entrevistas semidirigidas que componen los corpus PRESEEA.

REFERENCIAS BIBLIOGRÁFICAS

- Aaron, J. E. (2010). Pushing the Envelope: Looking beyond the Variable Context. *Language Variation and Change*, 22, 1-36.
- Arumí, M. & Esteve, O. (2006). Using Instruments Aimed at Self-Regulation in the Consecutive Interpreting Classroom: Two Case Studies. *Electronic Journal of Foreign Language Teaching*, 3, 158-189.
- Azurmendi, M. J., Bourhis, R. Y., Ros García, M. & García, I. (1998). Identidad etnolingüística y construcción de ciudadanía en las Comunidades Autónomas Bilingües (CAB) de España. *Revista de Psicología Social*, 13(3), 559-589.
- Blas-Arroyo, J. L. (2004). El español actual en las comunidades del ámbito lingüístico catalán. En R. Cano (Ed.), *Historia de la Lengua Española*. Vol. 2 (pp.1065-1086). Barcelona: Ariel.
- Blas-Arroyo, J. L. (2007). El contacto de lenguas como factor de retención en procesos de variación y cambio lingüístico. Datos sobre el español en una comunidad bilingüe peninsular. *Spanish in Context*, 4(2), 263-29.
- Blas-Arroyo, J. L. (2008). The Variable Expression of Future Tense in Peninsular Spanish: The Present (and Future) of Inflectional Forms in the Spanish Spoken in a Bilingual Region. *Language Variation and Change*, 20(1), 85-126.
- Briz, A., Pons, S. & Portolés, J. (2008). *Diccionario de partículas discursivas del español (DPDE)* [en línea]. Disponible en: www.dpde.es
- Camargo, L. (2021). *CORWAS. Corpus de Wasaps – Mallorca* (en proceso de digitalización).
- Camargo, L. & Enrique-Arias, A. (2021). La expresión de la adversatividad en el castellano hablado en Mallorca: Lo que adversativo y pero en posición final de cláusula. *Comunicación presentada en el XIX Congreso de la Asociación de Lingüística y Filología de la América Latina. Modalidad virtual*.
- Censabella, M. (2018). ¿Proceso de pragmaticalización o de gramaticalización? Contextos de uso del aplicativo debajo en qom/toba. *Cuadernos de la Alfal*, 10, 103-116.
- Collins, P. C. (1991). *Cleft and Pseudo-cleft Constructions in English*. Londres/Nueva York: Routledge.

- Cuadros, R. (2015). Lo único (que): Un incipiente marcador discursivo. *LEA*, 37(1), 149-171.
- Declerck, R. (1984). The Pragmatics of It-Clefts and Wh-Clefts. *Lingua*, 64, 251-289.
- Declerck, R. (1994). The Taxonomy and Interpretation of Clefts and Pseudo-clefts. *Lingua*, 93, 183-220.
- Dufter, A. (2008). Evolución pragmática de las oraciones hendidas en español: El papel de los usos no focalizadores. En C. Company Company & J. Moreno de Alba (Eds.), *Actas del VII Congreso Internacional de la Asociación de Historia de la Lengua Española* (pp. 1763-1779). Madrid: Arco Libros.
- Eckert, P. & McConnell-Ginet, S. (2007). Putting Communities of Practice in their Place. *Gender & Language*, 1(1), 27-37.
- Enrique-Arias, A. (2008). Materiales para la historia del castellano hablado en Mallorca: El archivo epistolar de Pedro de Santacilia. En C. Sinner & A. Wesch (Eds.), *El castellano en las tierras de habla catalana* (pp. 219-232). Frankfurt/Madrid: Vervuert/Iberoamericana.
- Enrique-Arias, A. (2014). Efectos del contacto de lenguas en el castellano de Mallorca: Una perspectiva histórica. En A. Enrique-Arias, M. J. Gutiérrez, A. Landa & F. Ocampo (Eds.), *Perspectives in the study of Spanish language variation* (pp. 271-297). Santiago de Compostela: Universidad.
- Enrique-Arias, A. & Méndez Guerrero, B. (2020). On the Effects of Catalan Contact in the Variable Expression of Spanish Future Tense. A Contrastive Study of Alcalá de Henares, Madrid and Palma, Majorca. En L. A. Ortiz López, R. E. Guzzardo Tamargo & M. González-Rivera (Eds.), *Hispanic Contact Linguistics. Theoretical, methodological and empirical perspectives* (pp. 316-334). Amsterdam: John Benjamins.
- Feldhausen, I. & Vanrell, M. M. (2015). Oraciones hendidas y marcación del foco estrecho en español: Una aproximación desde la teoría de la optimidad estocástica. *Revista Internacional de Lingüística Iberoamericana*, 26, 39-59.
- Gallego, A. J. (2006). A Note on lo que [en línea]. Disponible en: <http://seneca.uab.es/ggt/membres/professors/gallego.html>.
- García Mouton, P. (2003). *Así hablan las mujeres curiosidades y tópicos del uso femenino del lenguaje*. Madrid: La Esfera de los Libros.
- Garrido, M. C. (2006). Gramaticalización y marcadores discursivos: Los contraargumentativos. *Estudios Humanísticos. Filología*, 28, 9-26.

- Guerrero, S., Gajardo, C., Ibarra, D., González, J. & Reyes, A. (2020a). Lo que pasa es que la política se ha puesto farandulera: Justificaciones atenuantes de aserciones de opinión en el corpus Preseca de Santiago de Chile. *Literatura y lingüística*, 42, 247-273.
- Guerrero, S., Gajardo, C., Ibarra, D., González, J. & Reyes, A. (2020b). Variación genérica de las justificaciones atenuantes en español chileno. *Nueva Revista del Pacífico*, 73, 88-113.
- Gutiérrez, M. (1995). On the Future of the Future Tense in the Spanish of the Southwest. En C. Silva-Corvalán (Ed.), *Spanish in Four Continents: Studies in Language Contact and Bilingualism* (pp. 214-226). Washington, DC: Georgetown University Press.
- Gutiérrez, M. (2002). Innovación en el español de los Estados Unidos: ¿Progreso o decadencia? *Revista de Lingüística Teórica y Aplicada*, 40, 113-128.
- Gutiérrez Bravo, R. (2021). Las oraciones pseudohendidas en español: Sintaxis y propiedades informativas. En M. Leonetti & M. V. Escandell Vidal (Eds.), *La estructura informativa* (pp. 359-392). Madrid: Visor.
- Heine, B. & Narrog, H. (2010). Grammaticalization and Linguistic Analysis. En H. Narrog & B. Heine (Eds.), *The Oxford Handbook of Linguistics Analysis* (pp. 401-424). Oxford: Oxford University Press.
- Hopper, P. (1991). On Some Principles of Grammaticalization. En E. Traugott & B. Heine (Eds.), *Approaches to Grammaticalization* (pp. 17-35). Ámsterdam: John Benjamins.
- Hopper, P. & Traugott, E. C. (1993). *Grammaticalization*. Cambridge: Cambridge University Press.
- Johanson, M. (2001). Clefts in Contrast: A Contrastive Study of It Clefts and Wh Clefts in English and Swedish Texts and Translations. *Linguistics*, 39, 547-582.
- Krassenberg, L. & Lahousse, K. (2018). The Information Structure of French il y a Clefts and c'est Clefts: A Corpus-Based Analysis, *Linguistics*, 56, 513-548.
- Lambrecht, K. (2001). A Framework for the Analysis of Cleft Constructions. *Linguistics*, 39, 463-516.
- Méndez Guerrero, B. (2015). Corpus oral juvenil del español de Mallorca (COJEM). *Lingüística en Red*, 13.
- Molina, I. (2006). La moda del ¿sabes? en Madrid: Un análisis sociolingüístico. *Filología y lingüística: Estudios ofrecidos a Antonio Quilis*, 1, 1045-1056.

- Morales, A. (2005). Recursos expresivos, condicionamiento pragmático y variación dialectal: Las oraciones hendidas. *Spanish in context*, 2(1), 77-92.
- Moreno Cabrera, J. C. (1983). Las perífrasis de relativo. En F. Lázaro (Ed.), *Serta philologica* (pp. 455-467). Madrid: Cátedra.
- Moreno Cabrera, J. C. (1999). Las funciones informativas: Las perífrasis de relativo y otras construcciones perifrásticas. En I. Bosque & V. Demonte (Eds.), *Gramática descriptiva de la lengua española* (pp. 4245-4304). Madrid: Espasa Calpe.
- Moreno, F. (2021). Metodología del Proyecto para el estudio sociolingüístico del español de España y de América (PRESEEA), Documentos PRESEEA de investigación, Documentos de trabajo 1. DOI: 10.37536/PRESEEA.2021.doc1
- Narbona, A. (1986). Problemas de sintaxis coloquial andaluza. *Revista de la Sociedad Española de Lingüística*, 16, 229-275.
- Orozco, R. (2015). Castilian in New York City: What Can We Learn from the Future? En S. Sessarego & M. González-Rivera (Eds.), *New Perspectives on Hispanic Contact Linguistics in the Americas* (pp. 347-372). Frankfurt: Iberoamericana / Vervuert.
- Osborne, S. D. (2008). *Variable Future Tense Expression in Andalusian Spanish*. Master's Thesis. Athens, GA: University of Georgia.
- Portolés, J. (2004). *Pragmática para hispanistas*. Madrid: Síntesis.
- Prince, E. (1978). A Comparison of Wh-Clefts and It-Clefts in Discourse. *Language*, 54, 883-906.
- Reig, A. (2011). The Pragmatic Meaning of the Spanish Construction lo que pasa es que. *Journal of Pragmatics*, 43, 1435-1450.
- Romera, M. (2008). La formación de una unidad funcional discursiva: Lo que en el castellano de Mallorca. En C. Sinner & A. Wesch (Eds.), *El castellano en las tierras de habla catalana* (pp. 199-218). Frankfurt/Madrid: Vervuert Verlagsgesellschaft/ Iberoamericana.
- Sedano, M. (1990). *Hendidas y otras construcciones con SER en el habla de Caracas*. Caracas: Universidad Central de Venezuela.
- Sedano, M. (2007). Future Tense Expressions in Several Spanish Corpora. En G. Parodi (Ed.), *Working with Spanish Corpora. Research in Corpus and Discourse* (pp. 132-144). Londres: Continuum.

- Wesch, A. (1997). El castellano hablado de Barcelona y el influjo del catalán. Esbozo de un programa de investigación. *Verba*, 24, 287-312.
- Wesch, A. (2002). La investigación sobre variedades del español hablado en contacto con el catalán (particularmente en Cataluña y Baleares): Estado de la cuestión y perspectivas para el futuro. En M. T. Echenique & J. P. Sánchez (Eds.), *Actas del V Congreso Internacional de Historia de la Lengua Española* (pp. 1857-1874). Madrid: Gredos.
- Zimmerman, K. (2002). La variedad juvenil y la interacción verbal entre jóvenes. En F. Rodríguez (Ed.), *El lenguaje de los jóvenes* (pp. 37-164). Barcelona: Ariel.

NOTAS

¹ Según explica Morales (2005), las construcciones pseudo-hendidadas que comienzan con la fórmula de relativo (lo que) pueden representar hasta el 85% de las construcciones de este tipo en algunas variedades del español.

² La gramaticalización explica la interacción existente entre el uso y la estructura de la lengua (Hopper & Traugott, 1993). Desde el punto de vista de la teoría de la gramaticalización, se considera que los mecanismos o parámetros del fenómeno van desde la extensión o surgimiento de nuevos significados gramaticales en contextos nuevos; la dessemanticización o pérdida de contenido semántico; la decategorialización o pérdida de propiedades morfosintácticas y la erosión o pérdida de sustancia fonética (Hopper, 1991; Heine & Narrog, 2010; Censabella, 2018).

³ Como indica Romera (2008), es frecuente encontrar una alternancia entre las formas gramaticalizadas plenas 'lo que pasa es que' y 'lo que ocurre es que' en los mismos contextos, así como una serie de variantes sin cópula ('lo que pasa/ocurre que') y de expresiones que han perdido además la perífrasis de relativo ('lo único que', 'lo que') con los mismos usos.

⁴ Gallego (2006) considera que no puede hablarse de verdadera elipsis en el caso de la reducción de 'lo que pasa [es que]' a 'lo que', puesto que no existe una estructura antecedente en la que se sustente el proceso de elisión. Por tanto, parece más apropiado hablar de ella como reducción, tal y como propone Romera (2008).

⁵ Otro tipo de reducciones fonéticas que suelen producirse son 'lo que pasa/ocurre es' y 'lo que pasa/ocurre que'. Romera (2008) incluye también las fórmulas 'lo único que' y 'lo único' como reducciones de este tipo de construcciones procedimentales gramaticalizadas.

⁶ Actualmente, no está claro que se considere un marcador del discurso. De hecho, todavía no aparece mencionada como tal en la bibliografía sobre las partículas del discurso en español (Portolés, 2004, 2008; Briz et al., 2008; Reig, 2011). Sin embargo, como se ha visto, la forma de la construcción gramaticalizada es invariable en la mayoría de los casos y además su significado es procedimental, es decir, guía la interpretación del segmento del discurso que introduce.

Estas propiedades llevan a pensar que estos elementos podrían estar muy cerca, desde un punto de vista funcional, de algunos de los marcadores discursivos que se utilizan con frecuencia en español (Cuadros, 2015). Romera (2008: 4) va un paso más allá y la considera una unidad funcional, que define como “aquellas partes del discurso cuya misión es realizar léxicamente las relaciones de coherencia que existen entre los segmentos discursivos que transmiten información proposicional”.

⁷ Romera (2008) obtuvo frecuencias de un 91,3% en las formas gramaticalizadas y de tan solo un 8,7% en las formas pseudo-hendidadas en su estudio de las ciudades de Palma y Soria (España).

⁸ No se han encontrado en los estudios realizados hasta la fecha casos de formas reducidas (‘lo que pasa’, ‘lo que’, etc.) con valor léxico pleno. De este hecho se deduce que la expansión funcional, la desemantización y la reducción fonética son procesos posteriores vinculados a la gramaticalización y a la pérdida de autonomía sintagmática y paradigmática de estas expresiones (Romera, 2008).

⁹ Las variantes contempladas son las siguientes: lo que, lo (único) que pasa (es que) y lo que ocurre es que (vid. Tabla 2). Dichas formas se corresponden con las consideradas en estudios previos sobre esta construcción (Romera, 2008; Reig, 2011; Cuadros, 2015).

¹⁰ Otros trabajos, como el presentado por Enrique-Arias y Méndez Guerrero (2020), ya han analizado los dos corpus en contraste y han determinado que el grupo etnolingüístico (castellanohablantes, catalanohablantes y bilingües equilibrados) tiene incidencia en algunos fenómenos lingüísticos en los grupos de hablantes manejados como, por ejemplo, la expresión de futuro

¹¹ Si se atienden los datos del estudio de Romera (2008) por ciudades, la frecuencia de la forma con mayor reducción fonética ‘lo que’ aumenta notablemente (26,6%). Aunque el corpus de Romera no resulta representativo desde un punto de vista sociolingüístico por la cantidad de muestra analizada, sí permite predecir la tendencia del cambio lingüístico en la parte centro-norte peninsular frente a lo que sucede en Mallorca.

¹² Para realizar su estudio en la ciudad de Palma, Romera (2008) recurrió a dos corpus orales: el Corpus Palma con 16 entrevistas sociolingüísticas de entre 40 y 60 minutos y el Corpus Palma Adicional con 39 conversaciones espontáneas de la variedad del castellano balear. En total, registró 82 casos de la forma gramaticalizada de ‘lo que pasa es que’ en alguna de sus variantes. La limitada cantidad de muestra localizada en los corpus, como explica la autora, hace necesario nuevos estudios más amplios como el que se presenta en esta investigación.

¹³ Resulta llamativo que algunas de estas formas tengan una mayor incidencia en el grupo etnolingüístico ‘dominantes en catalán’, ya que la forma ‘lo’ no existe en catalán normativo. No obstante, es frecuente en el catalán oral hablado en Mallorca el uso de ‘lo’ y por ello no es de extrañar que se atestigüe también en el COSER y en el CORWAS en hablantes dominantes en catalán, especialmente en mujeres jóvenes (Enrique-Arias & Camargo, 2021).

AGRADECIMIENTOS

Este trabajo se enmarca en los Proyectos FFI2017-83899-P financiado por MCIN/AEI /10.13039/501100011033/ FEDER “Una manera de hacer Europa” y PDR2020/51 - ITS2017-006, financiado por la CAIB a través de la Direcció General de Política Universitària i Recerca. Quiero agradecer a Laura Camargo, Andrés Enrique-Arias y a los dos evaluadores anónimos las observaciones realizadas sobre la versión previa de este trabajo, pues sus aportaciones han sido de gran valor para mejorar el resultado final del artículo. Los errores que existan son responsabilidad de la autora.