

Fundamentación textual en el Corpus del español del siglo XXI (CORPES) del americanismo obsolecente. Variables de un estudio de caso

Textual foundation in the Corpus del español del siglo XXI (CORPES) of obsolescent Americanism. Variables of a case study

María Teresa Cáceres-Lorenzo

UNIVERSIDAD DE LAS PALMAS DE GRAN CANARIA

ESPAÑA

mteresa.caceres@ulpgc.es

Recibido: 10-XII-2017 / **Aceptado:** 04-III-2019

DOI: 10.4067/S0718-09342020000100144

Resumen

Los diccionarios utilizan la marca obsolescencia para indicar, además del potencial proceso de mortandad léxica, el uso restringido de ciertas palabras en determinados contextos discursivos. Todas estas posibilidades descubren un problema de investigación que se especifica en el desconocimiento de cómo se usan estos vocablos en los textos generados en distintas áreas lingüísticas de América. En este trabajo planteamos que es viable aportar datos sincrónicos que clarifiquen lo anterior a través de reconocer los parámetros con los que aparecen las voces obsoletas en los textos actuales. Con este fin se ha diseñado una investigación empírica de estudio de caso con un lexicon de vocablos obsoletos de América obtenidos del Diccionario de Americanismos. Posteriormente, se realizó una búsqueda textual de aquellas voces que fue posible documentar en el Corpus del español del siglo XXI. Se trata de noventa y ocho voces obsoletas que analizamos según las variables ficción/no ficción, geográfica, número de documentos y sus distintas frecuencias. Los resultados muestran que algunas unidades léxicas se utilizan exclusivamente en una determinada tipología textual, pero el grupo más numeroso aparece de manera general. También, se evidencia la similitud en los textos entre áreas lingüísticas vecinas, con la excepción de México, Centroamérica, Caribe Continental y Río de la Plata. Nuestra indagación es una contribución a la caracterización del americanismo léxico desde la fundamentación textual sincrónica.

Palabras Clave: Español de América, Obsolescencia léxica, Corpus del español del siglo XXI.

Abstract

Dictionaries use the obsolescence mark to indicate, in addition to the possible process of lexical mortality, the restricted use of certain words in certain discursive contexts. All these possibilities uncover a problem of investigation that is specified in the ignorance of how these words are used in the texts generated in different linguistic areas of America. In this work, we propose that it is possible to provide synchronic data that clarifies the above by recognizing the parameters with which the obsolescent voices appear in the current texts. To this end, an empirical case study research has been designed with a lexicon of obsolete American words obtained from the *Diccionario de Americanismos*. Subsequently, a textual search of those voices that were possible to document in the *Corpus del español del siglo XXI* was realized. These are ninety-eight obsolescent voices that we analyze according to the variables fiction / non-fiction, geographical, number of documents and their different frequencies. The results show that some lexical units are used exclusively in a textual typology, but the most numerous group appear in general way. In addition, the similarity in texts between neighboring linguistic areas is evident, with the exception of Mexico, Central America, Caribe and Río de la Plata. Our exploration is a contribution to the characterization of lexical Americanism from the synchronic textual foundation.

Key Words: Spanish of America, Lexical obsolescence, Corpus of the Spanish of the 21st century.

INTRODUCCIÓN

En una obra lexicográfica, la asignación de la acotación de obsolescencia a una unidad léxica informa al lector que los autores de dicho diccionario la consideran anticuada, inadecuada a las modas o a las necesidades comunicativas más frecuentes. Asimismo, puede indicar que una determinada palabra ha iniciado un proceso de mortandad léxica (Burrige & Bergs, 2016). La marca de obsolescencia es también una connotación estilística que hace referencia en el discurso a varias áreas léxicas: la terminología deportiva; la de distintos profesionales; grupos humanos marginales o con actividades específicas; las voces históricas que aún se utilizan, pero que se refieren a realidades de una época pasada; etc. (Lara, 1994; Alvar Ezquerro, 2003; San Martín Núñez, 2011; Rodríguez González, 2012; González, 2016). Además, dicha marca con valor cronológico que se adjudica a un determinado vocablo, raramente se hace explícita (Barrio Estévez & Torner Castells, 1994-1995).

Esta investigación se centra en un rasgo de la caracterización sincrónica del americanismo léxico y se fija en concreto en el léxico obsolescente, como unidades que precisan del apoyo documental. El objetivo de este artículo es comprobar si los vocablos considerados obsolescentes por el *Diccionario de Americanismos* (DA, 2010) se pueden encontrar en el *Corpus del español del siglo XXI* (CORPES, 2018) como indicador de las variables que acompañan su uso actual en textos escritos. La versión actual de CORPES presenta limitaciones con respecto a la inclusión de textos orales, por lo que nuestra propuesta de investigación se ciñe a lo escrito, aunque se registren

textos que han sido creado para la representación oral como las obras de teatro y los guiones.

Se busca aportar datos textuales que nos informen sobre la diversidad de factores subyacentes tras la marca de obsolescencia del DA. Dicha fundamentación proporciona a los estudiosos datos de las variables que acompañan a una pieza léxica obsoleta (preferencia por la ficción o no ficción, por ejemplo). Ahora bien, la marca de obsolescencia responde a una realidad pluridimensional que no siempre se pormenoriza, por lo que los ejemplos que encontramos en un diccionario pueden tener un valor relativo (Fajardo Aguirre, 1996-1997).

La definición del DA de la marca diacrónica de obsolescencia se ha identificado según lo siguiente: “las palabras de poca frecuencia de uso, pero correlacionada con la generación mayor del espectro etario son las obsoletas” (DA, 2010: xlv). En nuestra indagación nos ceñimos a esta definición, porque al utilizarla como fuente documental inicial de este trabajo, asumimos que la adjudicación de la marca de obsolescencia es correcta. El criterio seguido por DA para asignar dicha marca también ha sido utilizado en otros diccionarios como explica Garriga (1997), aunque Lara (2012) al referirse a este diccionario, cuestiona el proceso seguido para determinar que una voz sea obsoleta según su frecuencia de uso. Nuestro propósito es identificar determinadas variables del uso a través de datos estadísticos que promuevan otras investigaciones desde otras fuentes documentales, o concretar otras definiciones que complementen nuestros posibles resultados.

El mismo concepto pluridimensional de obsolescencia descubre un problema de investigación que esperamos resolver en parte con nuestro estudio de caso. Se trata de que no conocemos cómo se utilizan dichos vocablos en los textos de distintos puntos geográficos y a través de que tipología discursiva se registran. Nos referimos a la necesidad de reconocer mediante los parámetros que presenta el apoyo documental, si la marca de obsolescencia hace referencia a una voz que ya no se documenta en textos del siglo XXI o si es un vocablo utilizado únicamente en una tipología textual y temática de una de las áreas lingüísticas de América. El origen de esta laguna de información en los estudios dialectológicos, se encuentra en que los diccionarios académicos actuales no proporcionan fundamento documental de distintas sincronías (Álvarez de Miranda, 2000) y de que el léxico evoluciona con mucha celeridad por factores lingüísticos y extralingüísticos (Haensch, 1987; Gerding, Fuentes, Gómez & Kotz, 2012).

En la actualidad, la producción lexicográfica de las academias de la lengua española ofrece la posibilidad de completar los resultados alcanzados en los estudios dialectológicos del DA (2010) con los datos textuales reales del CORPES en su versión de 2018. La existencia de este corpus de referencia, aunque muestre limitaciones similares a otras bases de datos (Molina Salinas & Sierra Martínez, 2015)

hace viable el diseño de una investigación como estudio de caso. El CORPES tiene como objetivo el continuo incremento de su base textual (Rojo, 2017), por lo que, esta investigación pretende ofrecer una metodología para indagar sobre el problema de investigación que hemos planteado, ya que somos conscientes de que las posibles conclusiones resultantes no son definitivas y deberán completarse constantemente (Tognini-Bonelli, 2001).

Finalmente, este artículo pretende ser una contribución al estudio del americanismo léxico obsolecente, desde la perspectiva de las variables que se obtienen de la fundamentación documental.

1. Marco teórico y preguntas de investigación

Las voces obsoletas que analizamos remiten a la posibilidad de que estas piezas léxicas pertenezcan al vocabulario receptivo o pasivo de los individuos (es decir, son comprendidas, pero no se usan de forma cotidiana) y que, a pesar de su carácter obsoleto, en ocasiones han pervivido durante siglos (Buesa Oliver & Enguita Utrilla, 1992; Frago Gracia & Franco Figueroa, 2001) o son objeto de una ‘resurrección léxica’ en la que pueden adquirir nuevos significados (López Morales, 2006). Este repertorio léxico se complementa con el vocabulario productivo como parte de un ‘continuum de conocimiento’ que se amplía a lo largo del periplo vital de un individuo (Sala, 1970). Como consecuencia lógica, los adultos tendrán un vocabulario receptivo más rico que los jóvenes o adolescentes (Haensch, 1987). En este aspecto, no existe una dicotomía nítida entre los vocabularios activo y pasivo, porque los hablantes elaboran su propio lexicón mental, al mismo tiempo que es posible que ciertas voces mueran o no con la generación que las usa.

La obsolescencia puede coincidir con el concepto relativo de ‘arcaísmo’ en su sentido de voz considerada anticuada por una comunidad de hablantes, por lo que las conclusiones de Pascual (1997) pueden ser útiles para el estudioso. Su enumeración presenta siete tipos de arcaísmos recogidos en los diccionarios de la lengua española: voces inexistentes (variantes gráficas de otras palabras; derivados creados en distintos momentos de la historia de la lengua; voces de uso muy restringido desde su primera aparición); palabras en trance de desaparecer; tecnicismos históricos; arcaísmos que fueron neologismos; voces anticuadas en una determinada área lingüística y no en otra; léxico pasivo; y arcaísmos marcados como vulgarismos.

El argot o lenguaje especial entre individuos de un mismo oficio o actividad recoge muchas voces obsoletas porque estas permiten caracterizar a un determinado grupo. En este contexto, se detecta el registro de préstamos del francés o del inglés para dar entrada en los textos a léxico novedoso (Hediger, 1977; Giralt Latorre, 1991; Mateu Serra, 2012) o aquellas que se incorporaron en una determinada época por vacío denominativo, prestigio social, economía lingüística y preferencia de los

hablantes, pero que su uso ha decaído porque se han adoptado y no adaptado en los respectivos vocabularios (Moreno de Alba, 1992; García-Medall Villanueva & Aleza Izquierdo, 1992; Alba, 2007; Gerding, Fuentes & Kotz, 2012; Vásquez Amador, Lario & López, 2015).

Ante esta diversidad teórica, muchos estudiosos han mostrado la eficacia del apoyo textual, porque este refleja distintas situaciones de comunicación como un recurso del que obtener información ‘auténtica’ sobre la naturaleza poliédrica del cambio lingüístico (Frago Gracia & Franco Figueroa, 2001; Kabatek, 2005; Company, 2008; Rojo, 2017). En nuestro caso, el uso del CORPES hace factible datar o reconocer cómo se utiliza un determinado vocablo en los textos de ficción y no ficción. Nuestra hipótesis de trabajo asevera que dicho material textual puede ser una fuente de referencia para la dialectología sincrónica (Enguita Utrilla, 1998). De tal forma que los que describen sucesos y personajes imaginarios poseen un objetivo comunicativo diferente con respecto a los de no ficción, que a su vez se diversifican como consecuencia de la actuación lingüística de los hablantes.

El CORPES informa sobre el registro de un vocablo, ya sea en un medio escrito u oral, en un texto de ficción (‘textos de un guión, novela, relato y teatro’) y no ficción (‘documento de actualidad, ocio y vida cotidiana’; ‘artes, cultura y espectáculos’; ‘ciencias sociales, creencias y pensamiento’; ‘ciencias y tecnología’; ‘política, economía y justicia’; y, ‘salud’). Además de las frecuencias de aparición, absoluta y normalizada, su registro diatópico, tipos de textos, y soporte en el que se publica (Rojo, 2016). Dichos factores son muy útiles para caracterizar el americanismo léxico desde distintas perspectivas, pero en esta ocasión, limitaremos nuestra indagación a las que hemos considerado más útiles para conocer la naturaleza de los americanismos obsoletos: ficción/no ficción, área lingüística, frecuencias de uso y número de documentos. En toda esta información, se ha detectado como limitación de CORPES la imposibilidad de tener en cuenta el número de páginas o de palabras de los documentos.

Las indagaciones de la variable ficción/no ficción han concluido en que es factible que algunos vocablos del español americano se hayan convertido en un vocabulario específico de la ficción, como ya detalló Alvar (1960: 62) al concluir con respecto a los dialectalismos en la poesía española del siglo XX:

“Las voces que el diccionario académico da como anticuadas o regionales solo lo son con un estrecho criterio purista y, de nuevo, resurgen cargadas de un nuevo ímpetu en el uso dialectal de un poeta (Unamuno, Dámaso Alonso)”

En esta línea, Fajardo Aguirre (1991) concluye en su búsqueda del americanismo léxico de los textos narrativos argentinos (1940-1990) que el lenguaje coloquial puede ser un elemento fundamental en estas obras. A lo que Lèvêque (2011) en su estudio del léxico en la novela centroamericana (1940-1970) añade que el 75% del vocabulario

utilizado es propio de un nivel dialectal, el resto se corresponde al familiar, popular y vulgar.

Otro aspecto que aporta el género ficción es la menor utilización del indigenismo léxico o el anglicismo en correlación con la prensa escrita. Las conclusiones de Fajardo Aguirre (1991) revelan que en el análisis de cincuenta años de producción novelística en Argentina llega a computar 129 palabras amerindias. Algo parecido concluye Lèvêque (2011) que registra 65 indoamericanismos en el análisis de tres décadas de discurso novelístico. En contraposición, Prieto (2006) ha localizado 206 quechuismos en su análisis de treinta años de la prensa chilena. El contraste puede presentarse en otros contextos, como explica Ueda (1996) en el discurso de la transmisión televisiva en el que advierte diferencias en el uso del vocabulario entre los programas dedicados a las noticias con respecto a las telenovelas, programas de variedades y similares. En lo que se refiere a los anglicismos relacionados con el vocabulario político con un análisis diacrónico, García Godoy (2017) explica la importancia de la tipología textual. De hecho, el discurso periodístico documenta estos vocablos desde el siglo XIX y el texto literario lo usará una centuria más tarde.

Al mismo tiempo, los estudios sobre el caudal léxico de los individuos concluyen que un determinado número de términos se aprenden a través de la lectura, es lo que López Morales (2006) llama ‘palabras temáticas’ (asociadas a un asunto específico) del léxico disponible, por lo que algunas palabras obsoletas no están aún en proceso de desaparición, sino que pueden formar parte de un repertorio que pertenece al acervo léxico pasivo común de los hablantes cultos. Este vocabulario pasivo, en palabras de Reyes (1987: 1070):

“Los lenguajes sociales (dialectos históricos, sociales y geográficos, jergas profesionales, lenguas de generación o grupo, el lenguaje de la autoridad, el de la propaganda, etc.) se interesan en la novela, ‘dialogan’: se los comentan, se los evalúa, en cuanto lenguajes y en cuantos puntos de vistas sobre el mundo”.

La fundamentación textual en la no ficción en soporte libro o prensa arroja también otro tipo de información para el investigador, ya que plantea la posibilidad de que en estos textos siga utilizándose el vocabulario considerado obsoleto para referirse a cuestiones culturales. Aunque, también, es cierto que según Ávila (2003) y Company (2008) que, dado el carácter divulgativo de este tipo de documentos, los periodistas seleccionan vocablos comprensibles a un mayor número de destinatarios y se convierte en un género con una gran influencia en la obsolescencia o pervivencia de determinadas unidades lingüísticas.

La variable diatópica que acompaña a un vocablo obsoleto aparece en cualquier documento, independientemente del bloque temático en el que surge. Esto no ha sido indagado todavía de manera sistemática para todas las áreas lingüísticas,

aunque de manera indirecta, se han realizado comparaciones entre varios países de América (Fabre, 1976; Hediger, 1977; Fajardo Aguirre, 1991; Giralt Latorre, 1991; Navarro Carrasco, 2000; Prieto, 2006; Vásquez Amador et al., 2015).

En resumen, el examen del marco teórico con relación al objetivo principal de este trabajo de aportar datos textuales sincrónicos que informen sobre la diversidad de parámetros que subyacen en la marca de obsolescencia, nos indica que debemos plantear las siguientes preguntas de investigación: ¿qué tendencia muestran las áreas temáticas a la hora de incorporar léxico obsoleto en los textos americanos?; ¿en nuestro glosario existe una preferencia por las tipologías de ficción o no ficción?; y consecuentemente, ¿qué aporta el análisis de la correlación de las variables ‘cantidad de documentos’ y ‘frecuencias’, con respecto al registro diatópico?

2. Marco metodológico

Con la intención de dar respuestas a las preguntas propuestas, hemos diseñado una investigación que se ha desarrollado con una metodología cuantitativa de un estudio de caso en distintas fases y con los siguientes materiales.

El procedimiento comienza con la compilación en el DA de lemas con la marca de obsolescencia. Esta exploración nos proporciona el dato de 1.947 voces con esta acotación que constituyen nuestro material de trabajo preliminar. El DA (2010) recoge pesquisas dialectológicas realizadas en América entre 1975-2010 acorde a las distintas academias de la lengua española.

Posteriormente, se inicia la búsqueda de estas unidades léxicas en el CORPES en su versión 0.9 de 2018, en el que aparecen 277 millones de formas, 25 millones por año (2001-2012), con un 90% de textos en lengua escrita. Se ha comprobado que la acepción con la que se emplea la palabra en el vocablo documentado en el CORPES coincide con el DA. Este corpus de referencia se ha construido con la selección de 70% de documentos americanos que se presentan a través de los parámetros ‘lengua escrita/oral’; ‘soporte: libro o prensa’; ‘temático’; y ‘tipología textual’; ‘geográfico’. La posibilidad de recopilar datos de estos factores justifica su utilización como un material útil para examinar a través de diferentes enfoques el uso de una determinada palabra (Rojo, 2016). En cuanto a la representatividad de estos datos hay que tener en cuenta lo ya indicado por Molina Salinas y Sierra Martínez (2015) para otros corpus académicos, es decir, si bien estos nos ofrecen una cantidad importante de muestras reales sobre el uso de la lengua, la representatividad de la misma está en cuestión.

De hecho, nuestras pesquisas en el CORPES nos han demostrado que el diseño de la recopilación de materiales que se incluyen en dicho corpus nos aporta diferencias importantes en los periodos temporales. Es decir, no existe la proporcionalidad deseable desde la lingüística de corpus. Por ejemplo, entre 2001 y 2004 se incluyen 66.666 documentos, con 82.172.725 palabras registradas, de las que el 33% aparecen

en obras españolas, el resto en las americanas. En los cuatro años siguientes, del 2005 al 2008, el número de obras se eleva hasta los 100.064 documentos, con 95.106.413 palabras, y una proporción similar de términos recogidos en obras españolas, 30,71%. Esta disparidad temporal se repite en el resto de variables analizadas, por lo que las conclusiones que pueden extraerse del análisis de estos datos en el CORPES, deben tener en cuenta esta limitación en cuanto a la representatividad de los datos (Biber, 1993). Esto puede deberse a que estamos ante una versión provisional del CORPES.

Nuestro procedimiento ha tenido en cuenta exclusivamente, el subcorpus de textos de América de ficción y no ficción, además del seguimiento de los otros parámetros: registro diatópico, número de documentos y las distintas frecuencias de aparición sincrónica que proporciona el CORPES. Las áreas lingüísticas que estructuran dicho corpus son las que siguen: México y Centroamérica (Costa Rica, El Salvador, Guatemala, Honduras, México, Nicaragua y Panamá); Antillas (Cuba, Puerto Rico y R. Dominicana); Caribe (Colombia y Venezuela); zona Andina (Bolivia, Ecuador y Perú); Río Plata (Argentina, Paraguay y Uruguay); Chilena (Chile) y EEUU. La presencia o ausencia de un término en cada región se plasma con los valores 0 para la ausencia y 1 para la presencia.

Las frecuencias reconocidas por el CORPES son dos, una que nombra como frecuencia absoluta o conjunta (Freq.) y que refleja el número de veces que aparece un vocablo en un área lingüística; y la otra, denominada en este corpus como normalizada (Fnorm.) porque ofrece el número de ocurrencias por millón de palabras. Una vez obtenidos todos estos datos que aparecen en el Anexo, se ha procedido a un análisis cuantitativo que haga posible contestar las preguntas de investigación.

La búsqueda manual del léxico seleccionado ha explicitado otras limitaciones que, a su vez, han provocado que se hayan descartado aquellas voces que el DA muestra con distintos sentidos, uno obsoleto y otro no. Dicha exclusión viene motivada porque que el fragmento que nos presenta el CORPES no hace posible dilucidar su sentido. Por ejemplo, con respecto a ‘danzón’, el DA exhibe un primer significado de plena vigencia ya que no aparece con la susodicha marca: ‘música del danzón, en compás de dos por cuatro y ritmo lento’; al mismo tiempo que recoge el sentido obsoleto de ‘baile popular parecido a la habanera’; o lo que sucede con ‘abismar(se)’ ‘sorprender(se)’ con un uso actual en Chile, además del obsoleto ‘conmover(se)’. A esta dificultad, se añade el número de voces del DA no encontradas en el CORPES: ‘abarrotado’ ‘persona desbordada de trabajo’; ‘abrigadero’ ‘doblez de la ropa’, ‘acordiona’ ‘acordeón’, etc.

Finalmente, se volcaron todos los datos del CORPES de 2018 en una hoja de cálculo Excel, lo que permitió extraer diferentes tipos de información sobre los mismos: proporción del léxico obsoleto en cada área temática; examen del léxico de esta clase en cada tipología textual; el estudio de la dependencia entre variables, a

través del coeficiente de correlación de Pearson; y el análisis del grado de similitud entre los repertorios de este tipo de voces existentes en las distintas zonas consideradas en el CORPES. Este último estudio se lleva a cabo mediante la elaboración de una clasificación aglomerativa jerárquica (CAJ) que emplea el índice de similitud de Jaccard.

3. Resultados y discusión

La búsqueda en CORPES nos muestra 97 americanismos obsoletos de DA que pueden fundamentarse con textos del siglo XXI (5% de las 1.947 recogidos) y que se exponen en el Anexo 1. En el Gráfico 1 se exponen los porcentajes de las áreas temáticas de los textos de ficción (ítems del a. al d.) y no ficción (del e. al j.) en las que se han encontrado las piezas léxicas que estudiamos. Con estos guarismos se ratifica las preferencias por la novela, dentro de la ficción; y por las ‘ciencias sociales, creencias y pensamientos’ en la no ficción, mayoritariamente en los soportes de libro y prensa, en los que perviven un mayor número de las voces analizadas. El 58,76% de los 97 términos analizados, aparecen en textos novelísticos, y el 45,36% en obras dedicadas a las ciencias sociales, las creencias y el pensamiento.

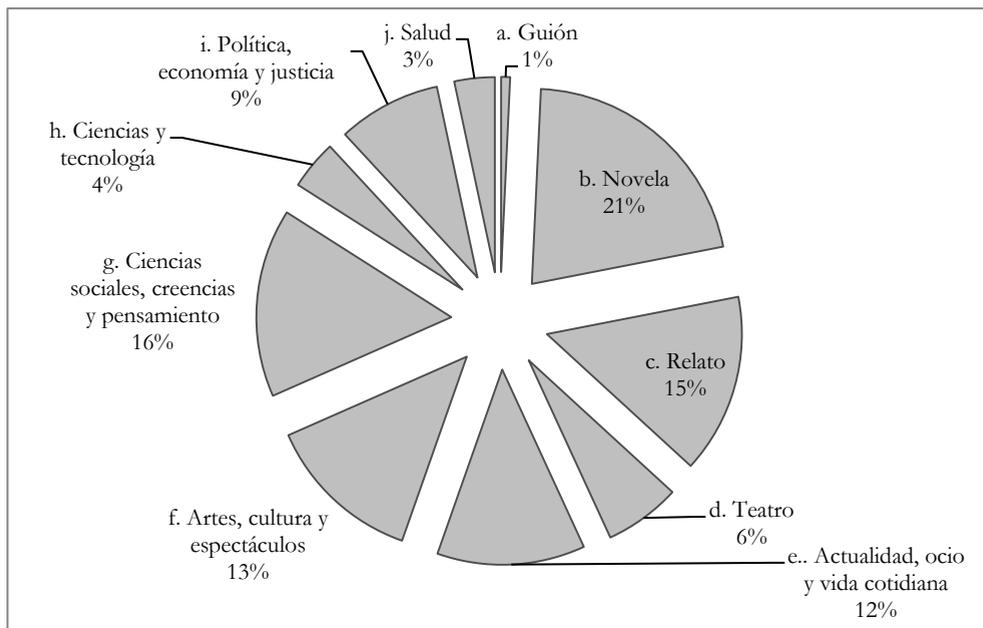


Gráfico 1. Porcentajes del léxico obsoleto en las áreas temáticas.

En la preferencia por estas temáticas encontramos algún ejemplo de préstamos léxicos del francés¹ y del inglés, lo que corrobora en parte, lo planteado por Giralt

Latorre (1991) para las voces galas que aparecen en la prensa panameña, y por Alba (2007) y Gerding et al. (2012), quienes consideran que en la prensa los préstamos pueden tener un carácter testimonial, porque certifican los vocablos que se usan por moda, impacto periodístico, precisión, etc. En los ejemplos recabados, estos préstamos con distinto grado de consolidación se han adaptado a la fonética de la lengua española y parecen estar vinculados a determinadas temáticas. Si consideramos las áreas lingüísticas en las que fundamentamos documentalmente los anglicismos ('cachascón', 'cotelé', 'motormán' y 'shotear') y que se muestran en el Anexo 1, nuestros resultados no aparecen en textos antillanos, pero sí en los de México, Centroamérica y Sudamérica. Esto corrobora las conclusiones de Hediger (1977) en su estudio de la novela hispanoamericana.

El glosario proveniente de nuestras pesquisas confirma lo propuesto por Haensch (1987), quien afirma que gran parte del glosario que examinamos designa objetos del pasado, tal como ejemplifica la aparición en nuestra investigación de pasacintas 'magnetófono', radiocasetera 'aparato de radio y casetes de audio', radiola 'mueble con radio y tocadiscos', crinolina o fustán 'enagua', etc. Además, se evidencia algún ejemplo en trance de desaparecer del discurso, porque los hablantes ya no reconocen el significado primigenio (López Morales, 2006) y que únicamente registramos en el CORPES como topónimo. Es el caso del indigenismo colca 'depósito' en río Colca, cañón del Colca, valle del Colca, etc. Asimismo encontramos el valor relativo de la marca de obsolescencia (Fajardo Aguirre, 1996-1997) como sucede en un artículo de prensa que utiliza fachentear 'actuar con vanidad o jactancia' para traducir un vocablo en inglés², lo que nos hace suponer que se trata del término más comprensible para los lectores (Ávila, 2003).

El análisis de los 97 vocablos revela creaciones americanas, tanto para aquellas que aparecen solo en la ficción como para las que se muestran únicamente en la no ficción, además de las que registramos en las dos tipologías³. Estas voces peculiares de las áreas americanas son más abundantes que las voces amerindias⁴, y en nuestro estudio de caso presentan un índice muy bajo si lo comparamos con los trabajos de Fajardo Aguirre (1991) y Prieto (2006).

Entre nuestros vocablos obsoletos, aparecen voces del lunfardo que hemos inventariado tanto en la ficción como en la no ficción. Dado el carácter argótico de este lenguaje especializado, nuestra datación en distintos contextos corrobora lo expuesto por San Martín Núñez (2011) y González (2016) sobre la popularización de este vocabulario en la literatura. De igual forma, el lenguaje especializado tiene un gran número de ejemplos a la hora de describir la delincuencia, la prostitución o el deporte⁵. Estos datos confirman lo expuesto por Enguita Utrilla (1998) en su análisis léxico-semántico de los americanismos en *Cien años de soledad*, con los resultados de Alba (2007); Gerding et al. (2012); Vásquez Amador et al. (2015) con respecto al

lenguaje deportivo, y los de Rodríguez González (2012) para el vocabulario de la droga.

En la Tabla 1 se muestra el cómputo de los términos que aparecen en cada tipología. La diferencia en la cantidad de palabras en las dos clases no es muy significativa, pero sí el porcentaje de documentos. En este punto se debe tener en cuenta que la media del número de textos que registran voces obsoletas en la no ficción (10,19%) es más del doble en relación con la otra tipología (4,851%).

Tabla 1. Datos estadísticos.

Textos	n° términos	n° documentos medio (%)	frecuencia absoluta media (%)	Frecuencia normalizada media (%)
No ficción	80	10,19	15,5	0,123
Ficción	74	4,851	7,57	0,131
Textos	n° términos	n° documentos desviación (%)	frecuencia absoluta desviación (%)	frecuencia normalizada desviación (%)
No ficción	80	23,89	36,9	0,3
Ficción	74	7,383	15,8	0,315

Ahora bien, el examen de la desviación, nos indica que en ambos casos es superior a la media, 23,89% y 7,383%, respectivamente. Es decir, los porcentajes iniciales no son explicativos, ya que la variabilidad dentro de cada grupo es mayor que la propia media, es decir, no son grupos homogéneos. Esta constatación nos llevó a continuar nuestra indagación en el análisis de las variantes que acompañan a un vocablo en su fundamentación textual, añadiendo en el análisis, la variable geográfica.

Tabla 2. Número total en cada zona.

Regiones lingüísticas	A: México y Centroamérica							B: Antillas			C: Caribe		D: Andina			E: Río Plata			F: Chilena	G: EEUU
N° de términos obsoletos	40							21			42		30			47			23	5
Países	CR	ES	Gu	Ho	Mx	Ni	Pa	Cu	PR	RD	Co	Ve	Bo	Ec	Pe	Ar	Py	Ur	Ch	EU
N° de términos obsoletos	7	9	11	10	25	10	3	9	6	14	30	26	17	11	14	36	15	10	23	5

Esta variable se estudia tanto desde el punto de vista de las regiones hispanoamericanas reseñadas en el CORPE, como en los distintos países donde se emplea cada término. Con respecto a esta variable, conocemos el número total de términos de este tipo que aparecen en cada región y en cada país analizado (Tabla 2), y la presencia o ausencia de cada vocablo en cada zona (Anexo 1).

Es importante resaltar que un buen número de términos aparecen en un único país (46), por lo que el grado de similitud entre los repertorios de cada nación es en general muy bajo.

Con esta información realizamos un estudio sobre la correlación existente entre las variables anteriores (tipología textual, frecuencia absoluta, etc.) con la diatópica, manteniendo esta última al nivel de región (Tabla 3) y además, obtenemos el grado de semejanza existente entre los territorios estudiados, tanto regiones como países (Gráfico 2).

Tabla 3. Correlación de Pearson de las variables seleccionadas del CORPES.

Variables	n° documentos	frecuencia absoluta	Frecuencia normalizada
n° documentos	1	0,965	0,769
frecuencia absoluta	0,965	1	0,858
Frecuencia normalizada	0,769	0,858	1
A: México y Centroamérica	0,342	0,345	0,311
B: Antillas	0,234	0,214	0,137
C: Caribe	0,325	0,313	0,318
D: Andina	0,254	0,298	0,312
E: Andina	0,103	0,108	0,091
F: Chilena	0,239	0,263	0,220
G: EEUU	0,464	0,461	0,282

Los valores en negrita son diferentes de 0 con un nivel de significación $\alpha=0,05$

Esta correlación entre las variables, no identifica ninguna preferencia por el vocabulario que buscamos en los textos de ficción o no ficción, a pesar de los porcentajes anteriores. En cuanto al resto de variables examinadas, es la zona caribeña la que posee un mayor número de documentos en los que aparecen este tipo de términos (correlación entre zona geográfica y n° de documentos = 0,342), mientras que es la zona andina la que parece tener una mayor frecuencia de aparición de términos, tanto de manera absoluta como normalizada (0,108 y 0,091 respectivamente).

El estudio del factor diatópico y su correlación con los otros parámetros tal como aparecen en la Tabla 2, muestran la similitud entre distintas áreas. El índice de similitud existente entre las distintas regiones consideradas nos proporciona una clasificación de estas zonas, tal como se observa en el Gráfico 2. Este resultado permite relacionar las áreas lingüísticas con respecto a la tendencia seguida en los textos a la hora de utilizar las unidades léxicas consideradas obsoletas por DA.

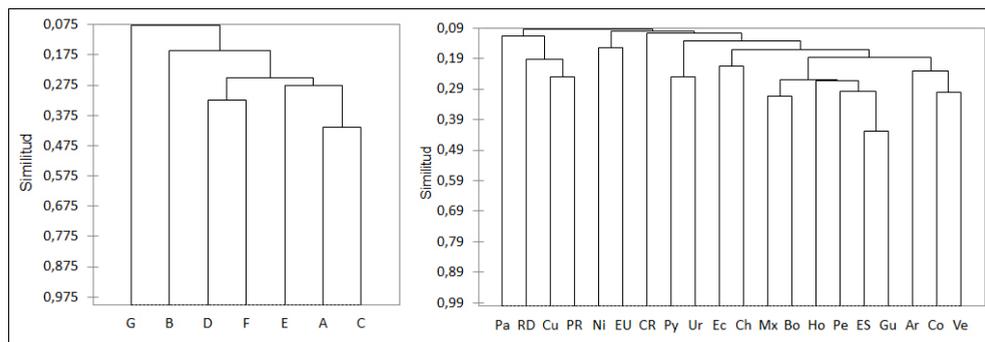


Gráfico 2. Asociaciones de las áreas lingüísticas y países en la fundamentación documental, según la CAJ realizada. A: México y Centroamérica (Costa Rica, El Salvador, Guatemala, Honduras, México, Nicaragua y Panamá); B: Antillas (Cuba, Puerto Rico y R. Dominicana); C: Caribe (Colombia y Venezuela); D: Andina (Bolivia, Ecuador y Perú); E: Río Plata (Argentina, Paraguay y Uruguay); F: Chilena (Chile) y G: EEUU.

Los grupos creados comparten el uso de los términos obsoletos recogidos en el Anexo. El índice de similitud más alto lo presentan las áreas mexicana y centroamericana con el Caribe Continental. Un valor considerable de similitud se presenta entre las áreas Andina y Chilena. En cuanto a la similitud por países, es interesante el grupo formado por las Antillas (Cuba, República Dominicana y Puerto Rico) y Panamá. También llama la atención la similitud existente entre Argentina, Colombia y Venezuela.

El alto índice de similitud existente entre las zonas ribereñas del Caribe se debe fundamentalmente a su situación de vecindad, de la misma manera que ocurre con las zonas andinas y chilenas. De ahí que el agrupamiento entre las regiones mesoamericanas y las del Río de la Plata cobren aún mayor importancia por su aparente lejanía.

A pesar de que no existen muchos estudios que analicen la variable diatópica desde esta perspectiva, hemos encontrado que parte de nuestros resultados confirman la indagación de Fajardo Aguirre (1991) con novelas argentinas en la que se muestra la similitud entre México y Río de la Plata en el uso del americanismo léxico en general. En esa ocasión, se justificó este dato porque los textos de México han sido tradicionalmente, más estudiados. Por su parte, Navarro Carrasco (2000) publica un lexicón recopilado en el texto peruano *La tía Julia y el escribidor* que coincide parcialmente con Chile y Río de la Plata.

CONCLUSIONES

La indagación empírica de un estudio de caso ha mostrado con datos cuantitativos la tendencia seguida por los textos escritos, a la hora de utilizar voces con la marca de obsolescencia en el DA. El lexicón obtenido de la búsqueda en el CORPES constituye el 5% de lo que se ha recopilado inicialmente en el DA. Esto no nos debe hacer

conjeturar de manera tajante que el resto de las voces ya se han perdido, porque el CORPES aún sigue incrementando sus fuentes.

Las preguntas de investigación que han enfocado nuestro estudio se responden a la luz de los resultados:

- i. ¿Qué tendencia muestran las áreas temáticas a la hora de incorporar léxico obsolecente en los textos americanos?

Los datos que proporcionan el CORPES 0.9, en los que se tienen en cuenta los números de documentos y las frecuencias, indican que en la ficción la novela y el relato son los materiales discursivos preferidos en el uso de las voces obsolescentes, en detrimento del teatro y los guiones. En la no ficción, la `ciencia y tecnología´ seguida por la salud parecen que son las áreas temáticas que menos han utilizado estas unidades léxicas frente a las otras que contienen información de tipo cultural (`ciencias sociales, creencias y pensamiento´; `actualidad, ocio y vida cotidiana´; `artes, cultura y espectáculos´; y, `política, economía y justicia´). Es posible concluir que los textos escritos (novela, relato) emplean más las voces obsolescentes que los documentos redactados para su uso oral: obras de teatro y guiones. En las áreas temáticas, los lenguajes especiales y las creaciones americanas caracterizan el vocabulario obsolecente como léxico que se comprende, pero no se usa porque se refieren a cuestiones muy específicas.

- i.i. ¿En nuestro glosario existe una preferencia por las tipologías de ficción o no ficción?

Desde el punto de vista estadístico, no hay ninguna tendencia frecuente. Lo que sí se evidencia, como se observa en el Anexo 1, es que algunas voces parecen haberse especializado en la ficción (alicroje, batola, cácaro, chimbiroca, etc.), otras en la no ficción (agrafe, babeador, bambinazo, etc.) y el número mayor de ejemplos (adoratorio, amansador, anoticiarse, burén, etc.) aparecen indistintamente, en las dos tipologías. Igualmente, se constata que algunas lexías tienen presencia en muchas áreas temáticas de las dos tipologías: discar, discado, perimido, yagueté, reprisar, lo que las convierte en voces con mucho uso.

- i.i.i. ¿Qué aporta el análisis de la correlación de las variables `cantidad de documentos´ y `frecuencias´, con respecto al registro diatópico?

El resultado desvela que los textos de distintas áreas lingüísticas no se comportan de igual forma a la hora de emplear el vocabulario que analizamos. México, Centroamérica y los países del Río de la Plata se revelan como las regiones en las que aparecen un mayor número de voces obsolescentes distintas. El resto de regiones se comportan de maneras muy diferentes: la zona chilena y andina poseen un léxico obsolecente muy parecido, como fruto de su fuerte vínculo cultural y geográfico; los

escasos términos utilizados en la zona antillana (22) se utilizan también en la caribeña, que posee una mayor riqueza en este tipo de términos (44). La semejanza más significativa es la de México y el Caribe continental, que, además de ser las zonas con mayor riqueza léxica de voces obsoletas, asimismo guardan una gran similitud entre sus repertorios (índice de similitud de Jaccard = 0,4 aproximadamente), a pesar de su lejanía geográfica.

Este estudio de caso se ha realizado con las voces obsoletas, por lo que nuestros resultados son una contribución a un problema de investigación más amplio: la necesidad de fundamentar textualmente los americanismos léxicos sincrónicos. Nuestros resultados dialectológicos con apoyo documental, complementan a otros obtenidos a través de distintos procedimientos (análisis de textos orales; resultados de encuestas, etc.). El grupo de voces obsoletas que hemos encontrado, muestra un gradiente que va desde los ejemplos que solo aparecen en una determinada tipología textual hasta aquellas que se utilizan en todas. En futuros trabajos se podrá confirmar esta tendencia con otro tipo de vocablos y su relación con diferentes tipologías textuales.

REFERENCIAS BIBLIOGRÁFICAS

- Alba, O. (2007). Integración fonética y morfológica de los préstamos: Datos del léxico dominicano del béisbol. *Revista de lingüística teórica y aplicada*, 45(2), 89-109.
- Alvar, M. (1960). Los dialectalismos en la poesía española del siglo XX. *Revista de Filología Española*, XLIII, 57-59.
- Alvar Ezquerro, M. (2003). *Nuevo diccionario de voces de uso actual*. Madrid: Arco-Libro.
- Álvarez de Miranda, P. (2000). Palabras y acepciones fantasmas en los diccionarios de la Academia. En J. C. Chevalier (Ed.), *La fabrique des mots. La neologie ibérique* (pp. 87-106). París: Presses Universitaires.
- Ávila, R. (2003). La lengua española y sus variantes en los medios de comunicación masiva. En G. Wotjak (Ed.), *Pautas y pistas en el estudio del léxico hispano (americano)* (pp. 11-26). Vervuert: Iberoamericana.
- Barrio Estévez, S. & Torner Castells, L. (1994-1995). La información diacrónica en el Diccionario de la Lengua Española de la Real Academia (vigésima primera edición). *Revista de lexicografía*, 1, 29-54.
- Biber, D. (1993). Representativeness in corpus design. *Literary and Linguistic Computing*, 8(4), 243-257.
- Buesa Oliver, T. & Enguita Utrilla, J. M. (1992). *Léxico del español de América: Su elemento patrimonial e indígena*. Madrid: MAPFRE.

- Burridge, K. & Bergs, A. (2016). *Understanding language change*. Londres: Routledge.
- [CORPES] Real Academia Española (2018). *Corpus del Español del Siglo XXI* [en línea]. Disponible en: <http://www.rae.es>
- Company, C. (2008). Gramaticalización, género discursivo y otras variables en la difusión del cambio sintáctico. En J. Kabatek (Ed.), *Sintaxis histórica del español y cambio lingüístico. Nuevas perspectivas desde las tradiciones discursivas* (pp. 17-52). Frankfurt-Madrid: Iberoamericana-Vervuert.
- Diccionario de Americanismos (DA) (2010). *Asociación de Academias de la Lengua Española*. Madrid: Santillana.
- Enguita Utrilla, J. M. (1998). Americanismos léxicos y textos literarios. *La Torre: Revista de la Universidad de Puerto Rico*, 3(7-8), 381-397.
- Fabre, F. (1976). *Glosario de la novela hispanoamericana actual*. Puerto Rico: Editorial Cordillera.
- Fajardo Aguirre (1991). *Americanismos léxicos en la narrativa argentina contemporánea*. Madrid: Universidad Complutense.
- Fajardo Aguirre, A. (1996-1997). Las marcas lexicográficas: Concepto y aplicación práctica en la lexicografía española. *Revista de lexicografía*, 3, 31-58.
- Frago Gracia, J. A. & Franco Figueroa, M. (2001). *El español de América*. Cádiz: Universidad.
- García Godoy, M. T. (2017). La diferenciación léxica del español de América: Anglicismos jurídicos e institucionales en la Colonia tardía. *Hispania*, 100, 65-78.
- García-Medall Villanueva, J. A. & Aleza Izquierdo, M. (1992). Observaciones diacrónicas sobre el léxico utilizado por Gabriel García Márquez. En M. Ariza Viguera (Ed.), *Actas del II Congreso Internacional de Historia de la Lengua española* (pp. 303-312). Sevilla: Pabellón de España.
- Garriga, C. (1997). Las marcas de uso. En los diccionarios del español. *Revista de Investigación Lingüística*, 1, 75-110.
- Gerding, C., Fuentes, M. & Kotz, G. (2012). Anglicismos y aculturación en la sociedad chilena. *Onomázein: Revista de lingüística, filología y traducción de la Pontificia Universidad Católica de Chile*, 25, 139-162.
- Gerding, C., Fuentes, M., Gómez, L. & Kotz, G. (2012). El préstamo en seis variedades geolectales del español: Un estudio en prensa escrita. *Revista Signos. Estudios de Lingüística*, 45(80), 280-299.

- Giralt Latorre, J. (1991). Algunos préstamos en el español de Panamá. *Estudios de lingüística*, 7, 137-158
- González, D. S. (2016). El lunfardo: ¿Un habla de delincuentes que quedó en el pasado? *Filología y Lingüística de la Universidad de Costa Rica*, 41(2), 107-118.
- Haensch, G. (1987). La lexicografía hispanoamericana entre la teoría y la práctica. En M.T. Vaquero (Ed.), *Actas del I Congreso Internacional sobre el español de América* (pp. 555-578). Madrid: Academia Puertorriqueña de la Lengua Española.
- Hediger, H. (1977). *Particularidades léxicas en la novela hispanoamericana contemporánea*. Frankfurt: Peter Lang.
- Kabatek, J. (2005). Tradiciones discursivas y cambio lingüístico. *Lexis*, XXIX(2), 151-177.
- Lara, L. F. (1994). Tipos de definición lexicográfica en el *Diccionario del español de México*. En B. Garza Cuarón (Ed.), *II Encuentro de lingüistas y filólogos de España y México* (pp. 153-164). Salamanca: Universidad de Salamanca.
- Lara, L. F. (2012). Diccionario de americanismos. *Panacea*, XIII, 352-355.
- Lèvéque, D. (2011). *Hablantes y hablas en la novelística social centroamericana (1940-1970). Para una tipificación dialectológica del discurso literario regional*. Alemania: Edición Lingüística.
- López Morales, H. (2006). La obsolescencia léxica. A propósito del Diccionario usual del español de Puerto Rico. En R. C. Gonzalo (Ed.), *Corcillvm: Estudios de traducción, lingüística y filología dedicados a Valentín García Yebra* (pp. 907-916). Salamanca: Universidad de Salamanca.
- Mateu Serra, R. (2012). Pragmática y argot: Una visión de los diccionarios. En N. Vila Rubio (Ed.), *El argot, entre España y Colombia: Estudios léxicos y pragmáticos* (pp. 99-116). Lleida: Universidad de Lleida.
- Molina Salinas, C. & Sierra Martínez, G. (2015). Hacia una normalización de la frecuencia de los corpus CREA y CORDE. *Revista Signos. Estudios de Lingüística*, 48(89), 307-331.
- Moreno de Alba, J. G. (1992). *Diferencias léxicas entre España y América*. Madrid: MAPFRE.
- Navarro Carrasco, A. I. (2000). *La Academia y los americanismos de La tía Julia y el escribidor*. Alicante: Servicio de Publicaciones de la Universidad de Alicante.

- Pascual, J. A. (1997). La caracterización de los arcaísmos en un diccionario de uso. En VVAA (Eds.), *Cicle de Conferències 95-96: Lèxic, corpus i diccionaris* (pp. 9-30). IULA, Barcelona.
- Prieto, L. (2006). Quechuisms en el léxico de la prensa de Santiago de Chile. *Boletín de Filología*, XLI, 97-196.
- Reyes, G. (1987). Polifonía textual. Variedades lingüísticas en la narrativa hispanoamericana reciente. En M. T. Vaquero de Ramírez & H. López Morales (Eds), *Actas del I Congreso Internacional sobre el español de América* (pp. 1067-1076). San Juan: Academia Puertorriqueña de la Lengua Española.
- Rodríguez González, F. (2012). Anglicismos en el mundo del deporte: Variación lingüística y sociolingüística. *Boletín de la Real Academia Española*, 92(306), 317-341.
- Rojo, G. (2016). *Citius, maius, melius*: Del CREA al CORPES XXI. En J. Kabatek (Ed.), *Lingüística de corpus y lingüística histórica iberorrománica* (pp. 197-212). Berlín: de Gruyter.
- Rojo, G. (2017). Sobre la configuración estadística de los corpus textuales. *Lingüística*, 33, 121-138.
- Sala, M. (1970). Arcaísmos e innovaciones en el léxico del español americano. En C. H. Magis (Ed.), *Actas del Tercer Congreso Internacional de Hispanistas* (pp. 779-785). México: Asociación de Hispanistas.
- San Martín Núñez, A. (2011). Voces de origen lunfardo en el registro festivo del diario chileno La Cuarta. *Onomázein: Revista de lingüística, filología y traducción de la Pontificia Universidad Católica de Chile*, 23, 105-147.
- Tognini-Bonelli, E. (2001). *Corpus linguistics at work*. Amsterdam: John Benjamins.
- Ueda, H. (1996). Análisis demolingüístico del léxico variable español. *Lingüística Hispánica*, 19, 63-98.
- Vásquez Amador, M., Lario, C. & López, P. (2015). Los anglicismos en la prensa deportiva de los 50. *Estudios filológicos*, 55, 157-176.

NOTAS

¹ Estos préstamos son en el género ficción, rastacuero; en la no ficción, agrafe, canevá, caramayola y coqueluche; en las dos tipologías, crinolina (también puede proceder del italiano) ‘enagua’ paletó y reprisar.

² En un texto de la prensa de Nicaragua en 2009 registrado en el CORPES como de ‘ciencias sociales, creencias y pensamiento’ (ítem g. del Gráfico 1) se lee: “El hecho de tener este referente -que se centra en recursos monetarios- es reconocido por el vestuario -que resulta llamativo- y por los furniture (muebles y equipos electrodomésticos provenientes de Estados Unidos) en las casas, Son más pobres los que no tienen una familia en Estados Unidos. Es pobre el que no tiene ese referente que le mande mensualmente para vivir mejor... to show off (para fachentear), afirma Cunningham”.

³ En la ficción registramos alicrejo (animal + cangrejo); cácaro (creación local), conversona (conversar), hembraerío (hembra), jesusear (de Jesús), radiocasetera (radio + casete), y vandalaje (vandalismo + bandidaje). En la no ficción, los ejemplos son los siguientes (entre paréntesis la palabra o frase de la que proviene): babeador (babero), cabinero (cabina de avión), centenarista (de centenario), cordiona (acordeón), esclavatura (esclavitud), escofieta (escofía), fachentear (fachenda), sucedencia (suceso), tercena ‘carnicería’ (con cambio de significado con respecto al español europeo en el que tiene el sentido de ‘estanco’). En ambas tipologías, tal como se presentan en el Anexo 1 encontramos adoratorio, amansador, anoticiarse, apeligrar, arriostre, batola, bividí, cachascán, candombe, cantonal, carrilano, chacabana, correntoso, cotelé, crinolina, dentistería, discado, discar, ecolecuá, entecado, fachento, fajero, faldellín, fustán, gachupín, hatero, leprosario, marconigrama, mertiolate, pampino, pasacintas, perimido, pisacorбата, radiola, rocola y toddy.

⁴ Los indoamericanismos son calchaquí, colca, sanguaraña, tianguéz y zonzoneco que muestran un uso más restringido que burén, tocuyo y yagareté.

⁵ Las unidades léxicas del lunfardo en textos de no ficción, esputsa y esquenún; aunque también hay ejemplos para las dos tipologías, estrilo, fyingo, pastenaca, pútero y rantifuso. Con respecto a lenguajes específicos: en la ficción hemos inventariado: bambinazo (beisbol), krumiro, con la grafía con <k> de crumiro (voz del léxico del sindicato) y el vocablo sanguaraña. Las formas léxicas encontradas exclusivamente en la no ficción de distintos lenguajes especializados son chimbiroca, macuteno, mabil, manfloro, manflorón, menega, miñanga, rastacuerismo, rífí y reperpero. Otros ejemplos que aparecen indistintamente en cualquier tipo de textos son también del argot de la delincuencia, bartolina, cafiche (cafisho), chacabana, chinchel, cocó, crispeta, marramuncia, y del deporte: umpire (beisbol).

ANEXO

Inventario de voces. N° (número de documentos); Freq (frecuencia absoluta media); Fnorm (Frecuencia normalizada media); Áreas (temáticas): a. guion; b. novela; c. relato; d. teatro; e. actualidad, ocio y vida cotidiana; f. artes, cultura y espectáculos; g. ciencias sociales, creencias y pensamiento; h. ciencias y tecnología; i. política, economía y justicia; j. salud. A: México y Centroamérica (Costa Rica, El Salvador, Guatemala, Honduras, México, Nicaragua y Panamá); B: Antillas (Cuba, Puerto Rico y R. Dominicana); C: Caribe (Colombia y Venezuela); D: Andina (Bolivia, Ecuador y Perú); E: Río Plata (Argentina, Paraguay y Uruguay); F: Chilena (Chile) y G: EEUU.

Lexías obsoletas	N°	Freq	Fnorm	No ficc	Ficc	Áreas
adoratorio `altarcito`	11	36	0,14	1	0	b, c, d
adoratorio `altarcito`	5	8	0,03	0	1	e, f, g
agrafe `pieza de metal`	3	8	0,03	1	0	e
alcrejo `feo y débil`	1	1	0	0	1	c
amansador `domador`	2	3	0,01	1	0	i
amansador `domador`	3	3	0,01	0	1	b
anoticiarse `enterarse`	13	17	0,06	1	0	e, f, g, h, i, j
anoticiarse `enterarse`	5	7	0,02	0	1	b
apeligrar `arriesgar`	9	9	0,03	1	0	e, g, h, i
apeligrar `arriesgar`	1	1	0	0	1	d
arriostre `pieza`	1	1	0	1	0	h
arriostre `pieza`	1	1	0	0	1	b
babeador `babero`	1	1	0	1	0	e
bambinazo `batazo`	113	133	0,53	1	0	e
bartolina `celda`	30	42	0,16	1	0	e, f, g, i
bartolina `celda`	15	24	0,09	0	1	b, c,
batola `prenda`	10	13	0,05	0	1	a, b, c, d
bividí `camiseta`	1	1	0	1	0	e
bividí `camiseta`	4	6	0,02	0	1	b, c
burén `plancha`	1	1	0	1	0	f, h
burén `plancha`	1	1	0	0	1	b
cabínero `azafata`	2	3	0,01	1	0	e, i
cácaro `proyeccionista`	6	5	0,02	0	1	b
cachascán `lucha libre`	1	1	0	1	0	f
cachascán `lucha libre`	2	2	0	0	1	b, c
cafique `proxeneta`	1	1	0	1	0	f
cafique `proxeneta`	11	22	0,08	0	1	b, c, d
calchaquí `indio`	18	56	0,22	1	0	g
candombe `música`	78	119	0,48	1	0	e, f, g, i
candombe `música`	9	15	0,06	0	1	b, d
canevá `tela`	3	3	0,01	1	0	e
cantonal `del cantón`	156	244	0,98	1	0	e, f, g, h, i, j

cantonal 'del cantón'	5	8	0,03	0	1	b, c
caramayola 'recipiente'	1	1	0	1	0	e
carrilano 'ferroviario'	1	1	0	1	0	i
carrilano 'ferroviario'	3	7	0,02	0	1	b, c
centenarista 'ideario'	4	7	0,02	1	0	f, g, i
chacabana 'camisa'	6	6	0,02	1	0	e, f, g
chacabana 'camisa'	3	5	0,02	0	1	b
chimbiroca 'prostituta'	1	3	0,01	0	1	b
chinchel 'cantina'	1	1	0	1	0	f
chinchel 'cantina'	1	1	0	0	1	b
cocó 'cocaína'	2	2	0	1	0	g
cocó 'cocaína'	2	4	0,01	0	1	b
colca 'depósito'	18	59	0,23	1	0	g, i
conversona 'diálogo'	1	1	0	0	1	b
coqueluche 'dolencia'	8	16	0,06	1	0	e, j
cordiona 'acordeón'	1	1	0	1	0	f
correntoso 'torrentoso'	11	11	0,04	1	0	e, f, g, h
correntoso 'torrentoso'	10	14	0,05	0	1	b, c
cotelé 'fiesta'	1	1	0	1	0	e
cotelé 'fiesta'	6	8	0,03	0	1	b, c
crinolina 'enagua'	3	4	0,01	1	0	e, f, g
crinolina 'enagua'	19	26	0,1	0	1	b, c
crispeta 'palomitas'	6	7	0,02	1	0	e, f, j
crispeta 'palomitas'	8	8	0,03	0	1	b, c, d
dentistería 'odontología'	3	7	0,02	1	0	f, i, j
dentistería 'odontología'	4	4	0,01	0	1	b, c
discado 'marcado'	13	18	0,07	1	0	e, f, g, h, i
discado 'marcado'	5	5	0,02	0	1	b
discar 'marcar'	26	31	0,12	1	0	e, f, g, h, i, j
discar 'marcar'	50	74	0,29	0	1	a, b. c. d
ecolecuá 'aprobación'	3	2	0,01	0	1	b
entecado 'delgado'	2	2	0	1	0	j
esclavatura 'esclavitud'	2	2	0	1	0	g
escofieta 'gorrito'	1	1	0	1	0	f
esputsa 'mal olor'	1	1	0	1	0	g
esquenún 'holgazán'	1	1	0	1	0	g
estrilo 'mal humor'	1	1	0	1	0	f
estrilo 'mal humor'	2	2	0	0	1	b
fachentear 'pompear'	1	1	0	1	0	g
fachento 'engreído'	1	1	0	1	0	g
fachento 'engreído'	2	2	0	0	1	c, d
fajero 'venda'	1	3	0,01	1	0	f
fajero 'venda'	1	1	0	0	1	b
faldellín 'faldón'	3	5	0,02	1	0	i
faldellín 'faldón'	1	1	0	0	1	b
fiyingo 'arma'	1	1	0	1	0	g
fiyingo 'arma'	1	1	0	0	1	d
fustán 'prenda'	1	2	0	1	0	i
fustán 'prenda'	7	7	0,02	0	1	b, c
gachupín 'español'	6	9	0,03	1	0	e, f, g, h
gachupín 'español'	13	19	0,07	0	1	b, c, d
hatero 'hacendado'	1	1	0	1	0	i
hatero 'hacendado'	1	1	0	0	1	b

hembreiro 'mujeres'	1	1	0	0	1	b
incásico 'de los incas'	6	9	0,03	1	0	e, f, g
jesusear 'invocar'	1	1	0	0	1	c
krumiro 'esquirol'	1	1	0	1	0	g
leprosario 'hospital'	2	2	0	1	0	g
leprosario 'hospital'	11	11	0,04	0	1	b, c, d
mabil 'prostíbulo'	3	3	0,01	0	1	c
macuteno 'ladrón'	1	1	0	0	1	b
manfloro 'homosexual'	2	3	0,01	0	1	b, c
manflorón 'homosexual'	1	1	0	0	1	b
marconigrama 'telegrama'	2	2	0	1	0	f, g
marconigrama 'telegrama'	2	3	0,01	0	1	c
marramuncia 'treta'	3	3	0,01	1	0	g, i
marramuncia 'treta'	3	3	0,01	0	1	c, d
mayorengo 'policía'	2	4	0,01	1	0	g
menega 'dinero'	1	1	0	0	1	c
mertiolate 'antiséptico'	1	1	0	1	0	e
mertiolate 'antiséptico'	1	1	0	0	1	b
miñanga 'de poco valor'	1	1	0	0	1	c
mótorman 'conductor'	6	6	0,02	1	0	e, f, g, i
mótorman 'conductor'	1	1	0	0	1	b
paletó 'chaqueta'	1	2	0	1	0	g
paletó 'chaqueta'	2	4	0,01	0	1	b
pampino 'de la pampa'	14	20	0,08	1	0	e, f, g, i
pampino 'de la pampa'	4	19	0,07	0	1	b, c
pasacintas 'magnetófono'	1	1	0	1	0	e
pasacintas 'magnetófono'	6	8	0,03	0	1	b, c
pastenaca 'tonto'	2	2	0	1	0	f, g
pastenaca 'tonto'	1	1	0	0	1	c
perimido 'anticuado'	30	30	0,12	1	0	e, f, g, h, i, j
perimido 'anticuado'	4	4	0,01	0	1	b, c
petitero 'con gustos refinados'	4	6	0,02	1	0	e, f
petitero 'con gustos refinados'	4	6	0,02	0	1	b, c, d
pisacorbata 'pasador'	1	1	0	1	0	e
pisacorbata 'pasador'	1	1	0	0	1	b
radiocasetera 'aparato de radio y casetes de audio'	3	3	0,01	0	1	b, c
radiola 'radio y tocadiscos'	4	7	0,02	1	0	f, g
radiola 'radio y tocadiscos'	12	16	0,06	0	1	b, c, d
rantifuso 'sucio'	1	1	0	1	0	f
rantifuso 'sucio'	1	1	0	0	1	b
rastacuerismo 'fatuo'	2	2	0	0	1	b
reperpero 'desorden'	3	3	0,01	0	1	b, c
reprisar 'reestrenar'	13	15	0,06	1	0	e, f, g, i
reprisar 'reestrenar'	1	1	0	0	1	c
reprise 'reposición'	6	7	0,02	1	0	e, f, g, i
rífifi 'petulante'	2	2	0	0	1	b
rocola 'gramófono'	18	36	0,14	1	0	f, g
rocola 'gramófono'	48	101	0,4	0	1	b, c, d
sanguaraña 'danza'	1	1	0	1	0	g

shotear `golpear`	1	1	0	1	0	e
shotear `golpear`	1	1	0	0	1	c
soberado `altillo`	3	3	0,01	1	0	f, g, i
soberado `altillo`	1	1	0	0	1	c
sucedencia `suceso`	1	1	0	1	0	g
tercena `carnicería`	1	1	0	1	0	g
tianguéz `mercado`	3	8	0,03	1	0	g
tocuyo `tela`	53	98	0,39	1	0	f
tocuyo `tela`	10	29	0,11	0	1	b, c
toddy `cacao`	5	5	0,02	1	0	e, i
toddy `cacao`	2	2	0	0	1	b, c
umpire `arbitro`	53	31	0,21	1	0	e, g
umpire `arbitro`	2	2	0	0	1	d
vandalaje `pillaje`	1	1	0	0	1	b
yaguareté `jaguar`	6	9	0,03	1	0	e, f, g, h, i
yaguareté `jaguar`	2	2	0	0	1	c, d
zonzoneco `tonto`	1	1	0	1	0	g

Lexías obsoletas	A							B			C		D			E			F	G	
	CR	ES	Gu	Ho	Mx	Ni	Pa	Cu	PR	RD	Co	Ve	Bo	Ec	Pe	Ar	Py	Ur	Ch	EU	
adoratorio `altarcito`			X		X			X		X	X		X		X					X	
agrafe `pieza de metal`											X						X				
alcrejo `feo y débil`			X																		
amansador `domador`										X	X	X				X					
anoticiarse `enterarse`												X	X			X					
apeligrar `arriesgar`							X														
apeligrar `arriesgar dinero`							X														
arriestre `pieza`								X													
babeador `babero`											X										
bambinazo `batazo`		X	X		X		X	X	X	X	X	X									X
bartolina `celda`		X	X	X	X	X		X			X		X								
batola `prenda`										X	X	X									
bividi `camiseta`															X						
burén `plancha`										X											
cabinero `azafata`						X					X										
cácaro `proyeccionista`					X																
cachascán `lucha libre`															X						
cafique `proxeneta`														X	X	X				X	
calchaquí `relativo a los indio`					X						X					X					X
candombe `música`		X		X	X	X			X		X		X		X	X	X	X	X	X	X
canevá `tela`											X					X					
cantonal `del cantón`	X	X	X		X						X	X	X	X	X					X	

incásico 'de los incas'					X							X	X						X
jesusear 'invocar'						X													
krumiro 'esquirol'														X					X
leprosario 'hospital'				X	X				X	X				X					X
mabil 'prostíbulo'										X									
macuteno 'ladrón'					X														
manfloro 'homosexual'														X					
manflorón 'homosexual'														X					
marconigrama 'telegrama'									X					X					
marramuncia 'treta'										X									
mayorengo 'policía'														X					
menega 'dinero'																		X	
mertiolate 'antiséptico'										X									
miñanga 'de poco valor'																		X	
mótorman 'conductor'														X	X				
paletó 'chaqueta'					X								X						X
pampino 'de la pampa'														X	X		X	X	
pasacintas 'magnetófono'										X				X					
pastenaca 'tonto'														X					
perimido 'anticuado'					X					X				X	X	X	X		
petitero 'con gustos refinados'											X			X	X				
pisacorbata 'pasador'					X	X													
radiocasetera 'aparato de radio y casetes de audio'					X														X
radiola 'radio y tocadiscos'		X	X		X		X	X	X					X					
rantifuso 'sucio'														X	X				
rastacuerismo 'fatuó'					X													X	
reperpero 'desorden'								X	X										
reprisar 'reestrenar'											X	X					X		
reprise 'reposición'					X	X				X				X	X				
rififi 'petulante'								X		X								X	
rocola 'gramófono'	X	X	X	X	X		X		X	X	X	X	X	X	X				
sanguaraña 'danza'																			X
shotear 'golpear'																			X
soberado 'altillo'									X				X	X					
sucedencia 'suceso'										X									
tercena 'carnicería'	X									X									

tianguéz 'mercado'				X										X					
tocuyo `tela'			X		X					X	X	X	X		X				X
toddy `cacao'							X		X	X	X	X			X	X			
umpire `arbitro'	X				X	X		X	X	X	X	X			X				
vandalaje `pillaje'												X							
yaguareté `jaguar'																X	X	X	
zonzoneco `tonto'																			X