

Distribución y productividad de las construcciones resultativas con *de quicio*: estudio de corpus con macrodatos en español

Distribution and productivity of resultative constructions with “de quicio”: a corpus-based study with Spanish microdata

Gloria Corpas Pastor

IUITLM, Universidad de Málaga
España

ONOMÁZEIN | Número especial XIV

Variación(es), enseñanza y traducción: investigación(es) en fraseología: 07-29

DOI: 10.7764/onomazein.ne14.01

ISSN: 0718-5758



Gloria Corpas Pastor: Departamento de Traducción e Interpretación, Facultad de Filosofía y Letras, IUITLM, Universidad de Málaga, España. Orcid: 0000-0001-6688-1531. | E-mail: gcorpas@uma.es

Resumen

La fraseología está experimentando un cambio de paradigma provocado, en gran parte, por la aplicación de técnicas y conceptos procedentes de otras disciplinas cercanas. El presente trabajo combina el análisis de macrodatos (*big data*) con los principios de la gramática de construcciones, e incluye un protocolo de extracción automática para las construcciones resultativas fraseológicas con *de quicio*. Del análisis cuantitativo y cualitativo de los datos se extraen conclusiones sobre la adecuación de los corpus seleccionados para el análisis, las relaciones que se establecen entre estas construcciones y otras más esquemáticas, junto a diversos aspectos de distribución, frecuencia y productividad, con el foco puesto en el español “global” y sus distintas variedades.

Palabras clave: lingüística de corpus; gramática de construcciones; construcciones resultativas fraseológicas; español transnacional; variedades diatópicas; productividad.

Abstract

Phraseology is undergoing a paradigm shift largely due to the application of techniques and concepts from other related disciplines. This paper combines big data analysis with the principles of construction grammar. It includes an automatic extraction protocol for the resultative constructional idioms with *de quicio*. Quantitative and qualitative analysis of the data are performed in order to draw conclusions about the adequacy of the corpora selected for analysis, the relationships established between these constructions and more schematic ones, as well their distribution, frequency, and productivity, with a focus on “global” Spanish and its diatopic varieties.

Keywords: corpus linguistics; construction grammar; resultative constructional idioms; transnational Spanish; diatopic varieties; productivity.

1. Introducción

En los últimos años estamos asistiendo a un cambio de paradigma en la fraseología como disciplina. Por una parte, el análisis de corpus de grandes dimensiones y la aplicación de técnicas de procesamiento de lenguaje natural ha dotado de grandes volúmenes de datos a los investigadores, lo cual les ha permitido comprobar la naturaleza flexible y creativa de las unidades fraseológicas¹. Al mismo tiempo, el descubrimiento de regularidades ha llevado a profundizar en las relaciones que se establecen entre tales unidades y otras estructuras más abstractas y sistemáticas. El foco ha pasado de lo anecdótico, idiosincrásico e irregular a los principios cognitivos y constructivos que parecen sustentar la cuestión fraseológica. Central en este proceso, especialmente en los últimos años, han sido los desarrollos de la gramática de construcciones².

Por construcción se entiende la unidad simbólica biunívoca de forma y significado (semántico o funcional), a lo largo de una escala con diversos grados de abstracción y lexicalización. Serían construcciones desde morfemas a estructuras pasivas, pasando por sustantivos, compuestos, así como unidades fraseológicas de todo tipo. Especialmente relevantes son los esquemas fraseológicos y construcciones frasémicas (también denominadas fraseológicas o idiomáticas). Siguiendo la definición de Dobrovolskij (2016: 78):

Las *construcciones frasémicas* se pueden definir como estructuras que tienen un significado léxico en su conjunto, con ciertas posiciones lexicalizadas en su estructura sintáctica, mientras que otras corresponden a *slots* que deben ser rellenados con un componente léxico libre aunque con ciertas restricciones semánticas.

En este trabajo nos vamos a ocupar de construcciones fraseológicas con base resultativa y casillas (o *slots*) lexicalizadas (parcialmente), compatibles con la concepción de Dobrovolskij (2016). Concretamente nos ocuparemos de las construcciones resultativas con *de quicio*. Sin embargo, nuestra concepción va más allá de los esquemas fraseológicos o las unidades fraseológicas con casillas vacías y está más cercana a los postulados de Goldberg (1995) y, más concretamente, a la noción de *constructional idiom* de Taylor (2016: 464), que denominaremos construcción *fraseológica* en español:

[...] constructional idioms, that is, patterns (of varying degrees of productivity and schematicity) for the formation of expressions, but whose syntactic, semantic, pragmatic and even phonological properties cannot be derived from general principles, whether universal or language-specific.

-
- 1 A este respecto, véanse los trabajos recientes sobre fraseología computacional recogidos en Corpas Pastor y Colson (2020).
 - 2 Para una visión de conjunto sobre la familia de teorías constructivas y sus distintos enfoques, véase la selección de capítulos incluidos en Hoffmann y Trousdale (2013).

Partimos de la hipótesis de que solo el análisis de grandes volúmenes de datos (esto es, macrodatos o *big data*) permite a) estudiar estas construcciones teniendo en cuenta los anisomorfismos diatópicos y los rasgos transnacionales, y b) establecer las relaciones que mantienen este tipo de construcciones fraseológicas entre sí y con respecto a la construcción resultativa genérica, lo cual explica e incide en el funcionamiento de tales construcciones en el discurso. La primera parte de trabajo ofrecerá una breve caracterización teórica sobre las construcciones que nos ocupan. A continuación, se abordarán cuestiones de metodología, con especificación de los corpus utilizados, su tamaño y composición. Se presenta también un protocolo de extracción automática para el análisis de las construcciones resultativas fraseológicas con *de quicio*. Seguirá una tercera sección, donde se expondrán los resultados más relevantes del estudio a partir de los datos obtenidos de frecuencia, distribución y productividad, los cuales sustentan también nuestras reflexiones finales.

2. Construcciones resultativas fraseológicas

Las construcciones resultativas conforman patrones esquemáticos para la representación de un estado final causado por una determinada acción o evento. Desde esta perspectiva, el significado genérico de la construcción resultativa sería: A (agente) causa que B (paciente) pase a C (resultado/objetivo). Cuando la acción implica movimiento, las construcciones resultativas lexicalizan el modo, la trayectoria y el resultado mediante un adjetivo o frase preposicional (Levin y Rappaport Hovav, 2006), como, por ejemplo, *batir las claras a punto de nieve*, *pintar la puerta de verde*. Según Goldberg (2006), las construcciones resultativas se pueden considerar como una extensión metafórica de la construcción de movimiento causado (X causa que Y se desplace al lugar Z mediante la acción V). También se les denominan construcciones resultativas de propiedad (en inglés, *property resultative constructions*, en la terminología de Goldberg y Ray Jackendoff, 2004), dado que en ellas los estados mentales o propiedades se estructuran como ubicaciones o lugares físicos (espacio físico → espacio mental).

Las construcciones resultativas, en general, suelen ser bastante esquemáticas. Sin embargo, algunas construcciones de verbo específico pueden ser parcialmente sustantivas respecto a la acción verbal y, hasta cierto punto, al cambio de espacio (físico/mental) que expresan. De acuerdo con Hoffmann (2017), las construcciones resultativas son unidades simbólicas que llevan aparejadas una determinada forma y función (significado incluido), las cuales se pueden describir de la siguiente manera:

- FORMA [X V Y Z] ↔ SIGNIFICADO: 'X causa que Y pase a Z mediante V', donde Z es generalmente un adjetivo o una construcción prepositiva.

En un trabajo anterior (Corpas Pastor, 2021) nos hemos ocupado, entre otras, de construcciones resultativas en español (y en inglés) que expresan perturbación, falta de juicio o

irracionalidad (estado mental o inducción a dicho estado), como, por ejemplo, *volver a alguien loco* o *poner a alguien tarumba*. Otras instancias fraseológicas de construcciones resultativas similares son *sacar de sus casillas*, *sacar de quicio*, *poner fuera de sí*, *poner contra las cuerdas*, *poner contra la pared*, etc. Por limitaciones de espacio, en este trabajo nos vamos a centrar en un subtipo de construcciones resultativas ligadas a estados mentales o emocionales cuya casilla principal está ocupada por la construcción o frase preposicional *de quicio*.

En las construcciones fraseológicas, la forma/función no es enteramente compositiva. Por una parte, estamos ante construcciones de verbo específico (Croft, 2012: 392), en las que solo pueden darse algunos de los verbos pertenecientes a una determinada clase (y no todos), como, por ejemplo, *volver/poner/traer a alguien loco* (**llevar/desplazar/conducir a alguien loco*). Ello quiere decir que este tipo de construcciones presenta lexicalización de la casilla verbal, instanciado mediante una serie limitada (y arbitraria) de verbos concretos (*fillers*), al estilo de las colocaciones (principalmente, *sacar + de quicio*, en el caso que nos ocupa).

De acuerdo con Goldberg (1995), las construcciones se integran en una red compleja dentro de la cual interactúan a través de cuatro relaciones de herencia: enlace por polisemia (*polisemy links*), enlace de subpartes (*subpart links*), extensión metafórica (*metaphorical extensión links*) y vínculo de instancia (*instance links*). En este sentido, las construcciones idiomáticas de verbo específico heredan las relaciones sintáctico-semánticas de la construcción resultativa más genérica y desarrollan un significado específico similar en el cual X mediante V hace que Y, en su desplazamiento, pase por un estado emocional intenso Z que afecta a su propio control mental, causando irritación, entre otras sensaciones y emociones negativas. En dicho proceso, los *fillers* de tales construcciones se interpretan en este marco general mediante adaptaciones producidas, en parte, por un proceso cognitivo de coerción y, en parte, por las relaciones conceptuales que modulan el semantismo básico de la construcción resultativa. En este caso, se trata de una relación metafórica que estructura los estados mentales o emocionales en términos de localizaciones físicas, o, lo que es lo mismo, el cambio resultante (mental/emocional) como movimiento físico causado, y, en segundo lugar, de una relación de instancia, donde el componente léxico verbal (*sacar*) en esta construcción resultativa restringe la interpretación del argumento a 'irritado', 'alterado', cuya aparición, a su vez, viene condicionada por la construcción prepositiva (*de quicio*) que forma parte como *filler* de la construcción idiomática resultativa.

Siguiendo la notación de Hoffmann (2017), estos casos especiales de construcciones resultativas se podrían describir mediante su correspondiente esquema, distinguiendo entre paciente (Y) animado y no animado, que se ilustran con ejemplos extraídos del Corpus del español del siglo XXI (CORPES):

- FORMA: [X SACAR Y_{anim} DE QUICIO] ↔ SIGNIFICADO: 'X causa que Y se sienta, sea percibido o actúe de forma Z (alterada, irritada, exasperada) {intensificación, implicaciones metafóricas, prosodia negativa}':

Ej. 1: *Oye, no me hagas caso. Soy un viejo gruñón, un cascarrabias, pero es que, ya lo sabes, ciertos ambientes me sacan de quicio.*

- FORMA: [X SACAR Y DE QUICIO] ↔ SIGNIFICADO: 'X causa que Y sea percibido de forma Z (alterada, forzada, exagerada), {intensificación, implicaciones metafóricas, prosodia negativa}':

Ej. 2: *Las cosas se han sacado de quicio. Yo sólo quiero entrar en el Libro Guinness de Récorde como la navegante más joven.*

3. Metodología

Por lo general, las unidades fraseológicas (y, por ende, cualquier tipo de construcción fraseológica) son fácilmente reconocibles por los hablantes gracias a la saliencia cognitiva que les confiere su ubicación en el lexicón mental. Sin embargo, este rasgo no siempre trae aparejado una alta frecuencia de uso (frecuencia de aparición). De ahí que el estudio de dichas unidades precise de corpus de enormes dimensiones que permitan extraer datos suficientes para el análisis. En ese apartado presentamos los corpus seleccionados y las secuencias de búsqueda protocolizadas que vamos a emplear en nuestro estudio.

3.1. Protocolo de extracción automática

Con objeto de analizar en detalle las construcciones que nos ocupan, hemos diseñado un protocolo de extracción automática que emplea seis secuencias de búsqueda, las cuales combinan patrones formados por lemas, categorías gramaticales y huecos posibles para casillas (*slots*) no especificadas:

- 1) [SACAR + {entre 0 y 3 palabras} + prep + QUICIO]
- 2) [V (filtro: - SACAR) + {entre 0 y 3 palabras} + prep + QUICIO]
- 3) [SACAR + {entre 0 y 3 palabras, (filtro N= + cosas)} + prep + QUICIO]
- 4) [V (no SACAR) + {entre 0 y 3 palabras, (filtro N= + cosas)} + prep + QUICIO]
- 5) [SACAR + {entre 0 y 3 palabras, (filtro N= - cosas)} + prep + QUICIO]
- 6) [V (no SACAR) + {entre 0 y 3 palabras, (filtro N= - cosas)} + prep + QUICIO].

Los dos primeros patrones permiten extraer de forma automática concordancias, datos de frecuencia y distribución relativas a construcciones resultativas con *de quicio*, tanto con el *filler* central de la casilla verbal (*sacar*) como con otros verbos que pudieran entrar a formar parte de dicha construcción. Se ha establecido una ventana de hasta un máximo de

3 posiciones para determinar *fillers* posibles (Y) y se ha empleado la etiqueta POS V (con filtrado negativo de *sacar*) con el objetivo de poder extraer de forma automática datos y otras instancias (ejemplares) de esta construcción.

A fin de refinar las búsquedas, en los cuatro patrones siguientes hemos empleado la etiqueta POS N (con filtrado positivo y negativo de cosas). Así, los patrones tercero y cuarto extraen de forma automática instancias de la construcción resultativa fraseológica que nos ocupa, donde una de las tres palabras es necesariamente un sustantivo, de manera que la posición o casilla Y está ocupado por el *filler* central cosas, tanto si el verbo específico de la construcción es *sacar* como si ocurre con otros verbos. Finalmente, los patrones quinto y sexto han servido para extraer otras instancias de esa misma construcción con N distinto de cosas.

Finalmente, las concordancias obtenidas a partir del protocolo de búsqueda se han exportado a un archivo Excel que se ha ordenado y filtrado manualmente para localizar instancias relevantes que no se han extraído de forma automática. Si bien el porcentaje de instancias localizadas de esta manera ha sido bajo (<5 %), se han extraído Y complejos no detectados automáticamente. También se han eliminado concordancias repetidas (por fallos de deduplicación o por solapamiento de los resultados de algunos patrones de búsqueda), así como algunos ejemplos que no constituyen instancias de este tipo de construcciones fraseológicas (ej. 3):

Ej. 3: *Otros quedaron atrapados al encajarse las puertas y ventanas, y algunos se salvaron en los huecos y umbrales, mientras intentaban arrancarlas de quicio.*

Otras limitaciones de la extracción automática que no se han podido solventar con la extracción manual son la presencia de faltas de ortografía, problemas de lematización y fallos de anotación y etiquetado morfosintáctico, especialmente en los megacorpus esTenTen18 y Web Dialectos (cfr. Corpas Pastor, 2018). A continuación, se describen los corpus analizados en este estudio.

3.2. Corpus analizados

Para el presente trabajo hemos seleccionado tres corpus de español “global” o transnacional: un corpus de referencia, creado por la Real Academia de la Lengua Española (RAE), y dos corpus de macrodatos (*big data*) compilados de forma semiautomática mediante motores de búsqueda a partir de Internet y tratados mediante técnicas de procesamiento de lenguaje natural (PLN).

El primero de ellos, el Corpus del español del siglo XXI (CORPES)³, alberga documentos desde 2001 (textos orales y escritos) a 2020, que se van actualizando e incrementando

3 <https://apps2.rae.es/CORPES2/org/publico/pages/consulta/entradaCompleja.view>.

periódicamente. Actualmente, contiene unos 395 millones de palabras (versión 1.0, junio 2023) y 23 variedades, incluidas las variedades de español habladas en Guinea Ecuatorial y en Filipinas. Aproximadamente el 35.90 % de los documentos corresponden a la variedad europea o peninsular, el 63.50 % restante se dedica a las variedades americanas y un 0.59 %, a las otras dos variedades.

La compilación del CORPES obedece a criterios de diseño equilibrados y representativos, en línea con los principios de la lingüística de corpus tradicional. Sin embargo, presenta ciertas desventajas por su tamaño, que es relativamente pequeño, la inestabilidad del propio sistema de gestión a través del cual se accede a los datos y las limitaciones del lenguaje de interrogación del corpus (cfr. Corpas Pastor, 2015).

En cambio, los otros dos megacorpuses empleados en este estudio contienen documentos (textos escritos) rastreados en la red de forma automática, los cuales han sido localizados y descargados de sus respectivos sitios web automáticamente. En la construcción de dichos corpus se aplican técnicas especializadas de compresión de datos para eliminar copias duplicadas de datos repetidos (*deduplication*), se limpian automáticamente códigos y etiquetas (*boilerplate removal*), y se asigna a los documentos la variedad diatópica correspondiente a partir del dominio de nivel superior o TLD (del inglés *top-level domain*): .ar, .cu, .es, .mx, etc. (cfr. Suchomel y Pomikálek, 2012).

El corpus EsTenTen18 comprende 17 500 millones de palabras de español transnacional. Fue compilado de forma automática en 2018. Contiene un componente de español europeo o peninsular (1 992 millones de palabras; 11,38 %) y un componente de español americano (6 375 millones de palabras; 36,42 %), además de otros documentos con TLD .com, .org, etc., que no están asignados a ninguna variedad en particular (9 141 millones de palabras). Cada variedad diatópica ha sido rastreada y compilada de forma independiente como subcomponente del corpus, al tiempo que se ha computado la similitud y proximidad de cada una de ellas. Por su propia configuración, este megacorpuse se puede dividir en 19 subcorpuses más pequeños de variedades diatópicas. El 75 % del corpus americano procede, en orden descendente, de Argentina (2 447 millones), México (1 470 millones) y Chile (859 millones), según la descripción de Kilgarriff y Renau (2013). El EsTenTen18 está disponible a través de SketchEngine⁴.

Le sigue en tamaño el corpus Web Dialectos, que contiene algo más de 2 000 millones de palabras. Se nutre de documentos rastreados y compilados de forma automática a partir de sitios web de 21 países hispanohablantes: desde España (21,91 % del corpus) a otras variedades americanas (78,08 %), que incluyen Estados Unidos, México o Argentina, entre otros.

4 <https://www.sketchengine.eu/>.

Se puede acceder a este megacorporus a través del portal y sistema de gestión de corpus desarrollados por Mark Davies⁵.

En la tabla 1 se ofrece una comparativa del tamaño (aproximado) de los tres corpus utilizados en este estudio, expresado en millones de palabras. Para la estimación del tamaño de los megacorporus nos hemos basado en la descripción general que se encuentra en SketchEngine respecto al EsTenTen18⁶, y a las cifras por variedades que se ofrecen en el propio sitio de Mark Davies para el Web Dialectos, pero indicando solo el número de millones, sin incluir centenas ni decenas. En el caso del CORPES, el tamaño de los distintos subcorpus se ha calculado a partir de los datos de frecuencia normalizada y absoluta por países que muestra la opción 'estadística'. Para ello se ha empleado una fórmula de cálculo inversa: frecuencia absoluta multiplicada por 1 000 000 y dividida por la frecuencia normalizada. El resultado se expresa de forma reducida en número de millones también. Por ejemplo, según nuestros cálculos, el número de palabras que contiene el subcorpus mexicano del CORPES es 37 344 398, lo cual se consigna como 37 millones en la citada tabla.

TABLA 1

Tamaño de los corpus: datos globales y por variedades

SUBCORPUS	CORPES	WEB DIALECTOS	ESTENTEN
ESPAÑOL EUROPEO	121	459	1 992
ESPAÑOL AMERICANO	214	1 635	6 375
Argentina	30	183	2 447
Bolivia	6	43	47
Chile	20	70	859
Colombia	25	180	371
Costa Rica	4	31	47
Cuba	11	67	211

- 5 <https://www.corpusdelespanol.org/> (interfaz en inglés) y <https://www.corpusdelespanol.org/xs.asp> (interfaz en español). Hemos descartado el otro megacorporus sincrónico del español accesible desde este mismo portal, el Corpus del Español NOW, de unos 7 200 millones de palabras actualmente, dado que es un corpus dinámico que se actualiza y amplía mensualmente. Incluye documentos procedentes de prensa escrita (periódicos y revistas) desde 2012 hasta la fecha, subdivididos por variedades diatópicas.
- 6 Sketch Engine, "esTenTen – Spanish corpus from the web", <https://www.sketchengine.eu/estenten-spanish-corpus/>.

República Dominicana	6	37	43
Ecuador	8	57	64
El Salvador	4	39	27
Estados Unidos	4	180	-
Filipinas	0.14	-	-
Guatemala	4	59	27
Guinea Ecuatorial	0.90	-	-
Honduras	4	39	8
México	37	260	1 470
Nicaragua	4	34	53
Panamá	2	24	15
Paraguay	7	33	51
Perú	12	115	253
Puerto Rico	4	35	-
Uruguay	8	41	156
Venezuela	15	108	218
TOTAL	337	2 094	17 500

El EsTenTen18 incluye 19 variedades por TLD: Argentina (.ar), Bolivia (.bo), Chile (.cl), Colombia (.co), Costa Rica (.cr), Cuba (.cu), República Dominicana (.do), Ecuador (.ec), España (.es), Guatemala (.gt), Honduras (.hn), México (.mx), Nicaragua (.ni), Panamá (.pa), Paraguay (.py), Perú (.pe), El Salvador (.sv), Uruguay (.uy) y Venezuela (.ve). Este macrocorpus incluye también textos rastreados de dominios .com, .org, .net, etc., que no van asignados a ninguna variedad diatópica en concreto. El corpus Web Dialects incluye, además de las anteriores, subcorpus de español de Puerto Rico (.pr) y EE.UU. (.us), además de textos en español con dominios como los anteriores (.com, .net, etc.). Por su parte, el CORPES incluye esas mismas 21 variedades, así como un pequeño porcentaje de español hablado en Guinea Ecuatorial y Filipinas.

En este trabajo tendremos en cuenta la variedad de español general o global (que integra todas las variedades incluidas en cada corpus), la variedad europea o peninsular frente a la variedad americana (que comprende todos los subcorpus de variedades de español americano incluidas en cada corpus) y las 18 variedades de español americano coincidentes en el CREA, EsTenTen18 y Web Dialectos.

4. Resultados y discusión

En esta sección se analizan y presentan los datos obtenidos para las construcciones resultativas con *de quicio* en cada uno de los corpus analizados. En primer lugar, se realiza un análisis cuantitativo respecto a los datos de frecuencia absoluta y normalizada por variedades diatópicas y en español general. Se contrastan los resultados obtenidos, a fin de poder determinar la cobertura y representatividad de cada variedad de español transnacional. En la segunda parte se realiza un estudio cualitativo de las concordancias analizadas a partir de los datos obtenidos mediante el protocolo de extracción automática y filtrado manual posterior. Esta parte aspira a caracterizar la variabilidad de los *fillers* específicos que ocupan las casillas centrales de estas construcciones, y cómo ello explica su productividad e interacción con otras construcciones.

4.1. Datos de frecuencia y distribución

Los datos que arrojan los corpus indican que la construcción con *de quicio* es relativamente más frecuente en la variedad peninsular o europea que en la variedad de español americano, e, incluso, que en español general o global. La tabla 2 ofrece un resumen cuantitativo con datos de frecuencia absoluta (FA) y normalizada (FN), es decir, frecuencia por millón de palabras, en cada corpus y subcorpus analizado.

La diferencia en frecuencia normalizada entre la variedad peninsular y la americana es de 1,96 puntos en el CORPES, 1,04 en el Web Dialectos y de 0,28 en el EsTenTen18, cifras que resultan inversamente proporcionales al tamaño de los tres corpus analizados. La diferencia se acorta si se comparan frecuencias normalizadas de estas dos variedades respecto al español global o general: español global *versus* variedad peninsular (CORPES: -1,27; Web Dialectos: -0,77; EsTenTen: -0,03) y español global *versus* variedad americana (CORPES: 0,69; Web Dialectos: 0,27; EsTenTen: 0,31). Estas diferencias podrían obedecer, por un lado, a una mejor representación de las variedades diatópicas del español en el CORPES, o bien al mayor tamaño del componente peninsular (35,90 %). También podría deberse al porcentaje de documentos que no se encuadran en ninguna de estas dos grandes variedades (52,2 %) en el EsTenTen18, pero que, teniendo en cuenta sus URL de procedencia, muchos de ellos bien podrían ser sitios web de España (mallorcadiario.com, elpais.com, fcbarcelonanoticias.com, elmundodecordoba.com, cervantesvirtual.com, etc.).

En cuanto a la distribución intervarena del español americano, el EsTenTen18 no permite extraer conclusiones relevantes, dado que las frecuencias normalizadas por variedad se mantienen por debajo del 0,19, con valores medios entorno a 0,09 o incluso 0 (Guatemala, El Salvador), salvo en el caso de las variedades argentina (0,22) y mexicana (0,29), los dos subcorpus de mayor tamaño (2 447 y 1 470 millones de palabras, respectivamente). Por su parte, el CORPES ofrece una mejor representación de variedades americanas, aunque los

TABLA 2Distribución de construcciones resultativas con *de quicio*

(Sub)corpus	CORPES		WEB DIALECTOS		ESTENTEN18	
	FA	FN	FA	FN	FA	FN
GLOBAL	543	1,76	1 583	0,82	10 129	0,50
ESPAÑOL EUROPEO	314	3,03	676	1,59	1 667	0,47
ESPAÑOL AMERICANO	230	1,07	907	0,55	1 087	0,19
Argentina	42	1,53	102	0,60	392	0,22
Bolivia	8	1,42	15	0,39	10	0,15
Chile	19	1,06	34	0,52	108	0,17
Colombia	19	0,81	94	0,56	41	0,06
Costa Rica	0	0	29	0,97	4	0,04
Cuba	15	1,38	19	0,30	29	0,12
Rep. Dominicana	10	1,55	17	0,50	3	0,08
Ecuador	2	0,27	47	0,90	11	0,13
Guatemala	9	1,99	17	0,32	0	0
Honduras	2	0,48	8	0,23	1	0,07
México	47	1,37	154	0,62	424	0,29
Nicaragua	0	0	9	0,28	3	0,07
Paraguay	6	0,89	10	0,34	6	0,1
Perú	20	1,88	82	0,77	36	0,16
El Salvador	4	0,72	15	0,40	0	0
Uruguay	2	0,25	16	0,43	10	0,07
Venezuela	15	1,07	54	0,55	7	0,07

valores medios se sitúan por debajo del 1,07, siendo de 0 en el caso de Costa Rica y Nicaragua. Si bien la frecuencia normalizada de la construcción resultativa con *de quicio* se sitúa por encima de la media americana para las variedades argentina y mexicana (1,53 y 1,37, respectivamente), las cifras más altas se dan en las variedades guatemalteca (1,99) y peruana (1,88), mientras que otras variedades también se presentan muy igualadas entre ambas franjas: dominicana (1,55), boliviana (1,42), cubana (1,37) y nicaragüense (1,37).

La mejor representación en cuanto a variedades diatópicas americanas se encuentra en Web Dialects. Todas presentan frecuencias normalizadas superiores a cero. Por encima de la media, en este corpus (0,55) se sitúan, en orden decreciente, las variedades costarricense (0,97), ecuatoriana (0,90), peruana (0,77), mexicana (0,62), argentina (0,60) y colombiana (0,56).

Los datos anteriores de frecuencia se refieren a todas las instancias de construcciones resultativas con *de quicio* encontradas en los corpus. Si bien la mayor parte de tales instancias corresponden al primer tipo (X SACAR Y_{anim} DE QUICIO), se observa un porcentaje importante del segundo (X SACAR Y DE QUICIO), donde Y suele ir ocupado por *cosas*, especialmente la variedad peninsular. Apenas se observan casos en las distintas variedades diatópicas americanas, por lo que el número de instancias encontradas en el español global posiblemente se deba al peso de dicha construcción en la variedad hablada en España. Por ejemplo, el EsTenTen18 contiene 651 instancias en total, de las cuales 317 corresponden al subcorpus del español hablado en España, 5 al subcorpus mexicano, 4 al subcorpus argentino (una de las cuales está utilizada metalingüísticamente para incidir en su “sabor” eminentemente peninsular, ej. 4), una a la variedad panameña y otra a la variedad colombiana:

Ej. 4: ... *de hacer polemica aqui ni nada de eso y tampoco quiero crearme enemigos ya que soy nuevo jeje, pero creo que se esta sacando las cosas de quicio (por si no lo entienden la expresion española significa exagerando mucho)* (EsTenTen18, TLD: .ar).

Posiblemente, el resto de instancias de esta construcción se hallen en el subcorpus formado por documentos con TLD no asignados a ninguna variedad en concreto, el cual queda fuera de este estudio (pero que posiblemente incluya un porcentaje importante de documentos producidos en España).

Los otros dos corpus parecen confirmar también esta tendencia. Así, la segunda construcción (X SACAR Y DE QUICIO), donde el *filler* principal de Y es *cosas*, se encuentra mayoritariamente en el subcorpus de español peninsular: 37 instancias en el CORPES y 86 instancias en el Web Dialectos. La presencia de dicha construcción es mucho menor en el español americano en ambos corpus: en el CORPES (1 instancia en los subcorpus uruguayo y peruano, respectivamente) y en el WebDialects (3 instancias en el subcorpus argentino, 2 en el subcorpus colombiano, 1 en el chileno, 2 en el mexicano, 3 en el peruano y 2 en el salvadoreño). La cifra total de instancias de la segunda construcción en la variedad americana aumenta ligeramente en el Web Dialectos por el número de instancias presentes en el subcorpus de español estadounidense (19), una variedad no contemplada en este estudio.

Dado que los subcorpus de español americano apenas contienen instancias del segundo tipo de construcción con *de quicio*, en la tabla 3 solo se recogen datos de frecuencia (absoluta) para las variedades peninsular o europea, americana y en total, así como el porcentaje de cada una ellas en cada variedad.

TABLA 3Construcciones resultativas con *de quicio* 1 y 2 (español global, europeo y americano)

	X SACAR Y _{-anim} DE QUICIO					
	CORPES		Web Dialectos		EsTenTen18	
GLOBAL	504	92,81 %	1 465	92,54 %	9 478	93,57 %
ESPAÑOL EUROPEO	277	88,21 %	590	87,27 %	1 350	80,98 %
ESPAÑOL AMERICANO	228	99,13 %	875	96,47 %	1 076	98,98 %
	X SACAR Y DE QUICIO					
	CORPES		Web Dialectos		EsTenTen18	
GLOBAL	39	7,18 %	118	7,45 %	651	6,42 %
ESPAÑOL EUROPEO	37	11,78 %	86	12,72 %	317	19,01 %
ESPAÑOL AMERICANO	2	0,86 %	32	3,52 %	11	1,01 %

Tomando como referencia el corpus más extenso (EsTenTen18), de las 10 129 instancias de construcciones resultativas con *de quicio*, el 93,57 % corresponde a la construcción resultativa 1 (X SACAR Y_{-anim} DE QUICIO X) y el 6,42 % a la construcción resultativa 2 (X SACAR Y DE QUICIO X). Por variedades, la segunda construcción con Y no animado es más propia de la variedad peninsular o europea (19,01 %) que de la americana (1,01 %).

Datos similares arrojan los otros dos corpus. Respecto al español global, ambos varían tan solo en 0,27 puntos (tanto para la construcción 1 como para la construcción 2). En cambio, el EsTenTen18 varía en + 1,03 respecto al Web Dialectos y en + 0,76 respecto al CORPES para la construcción 1, y, para la construcción 2, arroja la misma diferencia, pero en negativo (-0,76 respecto al CORPES y -1,03 respecto al Web Dialectos).

La distancia es mayor si se consideran las dos variedades principales. En español peninsular o europeo, CORPES y Web Dialectos presentan una diferencia de 0,94 (construcción 1 y 2), mientras que el EsTenTen18 varía -6,29 (construcción 1) y +6,29 (construcción 2) respecto a Web Dialectos, y -7,23 (construcción 1) y +7,23 (construcción 2) respecto al CORPES. En cuanto al español americano, CORPES y Web Dialectos de nuevo muestran una distancia similar (2,66 puntos para ambas construcciones), a diferencia del EsTenTen18, que se distancia del Web Dialectos en -2,51 (construcción 1 y 2), pero mucho menos del CORPES: -0,95 (construcción 1) y + 0,15 (construcción 2). Ello podría deberse a una menor representación de ambas construcciones (especialmente la segunda) en CORPES y EsTenTen18, con un perfil más plano de variedades diatópicas.

En líneas generales, se puede afirmar que la construcción resultativa con *de quicio* es más frecuente en español peninsular que en español americano, con una diferencia que varía en cada uno de los tres corpus analizados, debido a su diferente configuración (variedades incluidas, componentes y tamaño de los distintos subcorpus). En cuanto a su distribución en las distintas variedades americanas, se observan coincidencias respecto a los datos de frecuencia, que apuntan a una mayor presencia en las variedades argentina y mexicana, así como la peruana (salvo en EsTenTen18).

De los tres corpus analizados, Web Dialectos y CORPES presentan una mejor cobertura de variedades diatópicas que EsTenTen18. De los tres, Web Dialectos parece ofrecer la mejor cobertura para el estudio de las construcciones que nos ocupan por variedades. Sin embargo, y aunque el número de ejemplos que ofrece es mayor, sería preciso contar con corpus de variedades diatópicas de tamaño aún mucho mayor. Por otro lado, los datos de frecuencia y distribución de las construcciones analizadas en los diversos subcorpus muestran que la variedad americana, como tal, constituye, más bien, un conjunto no homogéneo de variedades nacionales que apuntan a conglomerados que comparten cierta uniformidad, algunas de las cuales se presentan más cercanas a la variedad europea o peninsular.

4.2. Productividad y variación fraseológica

Si bien Web Dialectos y CORPES parecen reflejar mejor aspectos diferenciales entre variedades diatópicas, el tamaño de los distintos subcorpus resulta insuficiente para poder realizar estudios fiables sobre la construcción que nos ocupa en las distintas variedades del español por separado, más allá de datos de distribución. Con todo, EsTenTen18 y Web Dialectos ofrecen una mayor riqueza en general, al tratarse de megacorpora de miles de millones de palabras que contienen infinidad de instancias de estas dos construcciones. Por ejemplo, el corpus Web Dialectos triplica el número de ejemplares de las dos construcciones resultativas con *de quicio* en el CORPES, mientras que el EsTenTen18 multiplica los resultados casi por 19.

Así pues, los dos megacorpora seleccionados para este estudio, por su enorme tamaño, son especialmente útiles para analizar construcciones fraseológicas de menor frecuencia de aparición en el español global (e, incluso, en el peninsular). En lo relativo al español americano (y sus variedades), se precisa de corpus todavía más amplios, puesto que los subcorpus disponibles de variedades habladas en los distintos países del continente americano aún no cuentan con el tamaño necesario para poder estudiar este tipo de construcciones en profundidad. Además, como se ha visto en la sección anterior, las variedades americanas distan de conformar un todo monolítico u homogéneo. El análisis de las construcciones fraseológicas de diverso tipo apunta a la existencia de conglomerados, o a variedades más próximas a la variedad europea (cfr. Corpas Pastor, 2015, 2018). Por los motivos anteriores, en este apartado nos vamos a centrar en el español global o transnacional.

Ambos tipos de construcciones resultativas fraseológicas comparten la estructura semántica y sintáctica de la construcción resultativa más abstracta, con la cual interactúan mediante varias relaciones de herencia, a saber, extensión metafórica, enlace de subpartes y vínculo de instancia. Así, ambas construcciones se conjugan libremente en todos los tiempos, personas y formas verbales (“sacaba/sacaban/sacaron/sacando/sacado”, etc.), con cierta preferencia por la tercera persona (singular y plural), lo cual guarda relación con el tipo de prosodia de ambas construcciones, como veremos más adelante. Al igual que la construcción genérica, admite cuantificadores y adverbios (sobre todo de intensidad, pero también de tiempo u otros), como *tanto, tan, muy, mucho, completamente, ligeramente, bastante, algo, pelín, pelo, a veces*, etc. (ej. 5), e, incluso, toma la forma no personal como participio pasado con función adjetival (ej. 6):

Ej. 5: *La mayoría de las veces, el elefante es una bestia tranquila y sagaz, pero este fenómeno sin precedentes de un animal ruidoso y desconocido en el cielo había sacado completamente de quicio a la manada entera* (Web Dialectos).

Ej. 6: *Este estado de preguerra civil parece lo bastante sacado de quicio como para hacer popular a Trump, cuyas medidas cuentan ya todas con más partidarios que detractores* (EsTenTen18).

Por otra parte, el contenido semántico-funcional de ambas presenta semejanzas por la construcción resultativa abstracta de la cual son instancias. Al mismo tiempo, se dan diferencias e interactúan entre sí mediante el vínculo de instancia, el enlace de subpartes y el enlace de polisemia.

La construcción con Y animado implica personas, circunstancias o acontecimientos (X) que fuerzan a Y (normalmente una persona y, en menor medida, ente vivo u objeto animado) a desplazarse metafóricamente de su lugar (físico) a otro lugar (mental o emocional), lo cual tiene un efecto negativo en Y, que se altera y reacciona de manera desproporcionada e irracional. La construcción con Y no animado pone el foco específicamente en la manera en la que se percibe o presenta Y por desplazamiento inducido de su ubicación normal. Véase el ejemplo 7, en el que esta circunstancia se aprovecha para crear un efecto por yuxtaposición y contraste semántico entre ambas construcciones, acentuando la responsabilidad del actor (X) mediante la construcción ergativa con *salirse*:

Ej. 7: *Y quien saca las cosas de quicio, señor Procurador, creo que es usted... no que las saca, que se sale usted de quicio cuando es incapaz -tal vez por cobardía política- de no defender Aranda de Duero y sus intereses* (EsTenTen18).

El significado de las dos construcciones no es enteramente composicional, por la relación metafórica que estructura el espacio mental o emocional en términos de espacio físico (extensión metafórica). A este respecto, nótese el uso del adverbio *literalmente* en los ejemplos que siguen a continuación (ejs. 8 y 9):

Ej. 8: *¡Cómo se le echa de menos en el Parlamento catalán! Cuando ocupaba su escaño, no había nadie como él para sacar literalmente de quicio a Jordi Pujol. Daba gusto oír hablar a Vidal-Quadras y ver cómo Papá Pitufo se iba poniendo de los nervios* (EsTenTen18).

Ej. 9: *Hasta el momento en que la crisis sacó literalmente de quicio las cosas [...], la articulación de la idea de la España democrática fue en efecto inseparable de los conceptos de ‘normalización’ y ‘normalidad’* (EsTenTen18).

El uso del adverbio *literalmente* refuerza la intensidad implícita en ambas construcciones precisamente porque contribuye a acentuar la extensión metafórica de base⁷. El carácter intensificador de ambas construcciones se pone de manifiesto en el análisis de concordancias, a través de las relaciones colocacionales y los contextos en los que ocurren (*ponerse de mala leche, tener más paciencia que el santo Job, desquiciado, loco, fuera de sí, enloquecer, perder la paciencia, perturbar, enfurecer, irritar, poner a prueba, pretender, crisis nerviosas, acorralar, acosar, etc.*). Véanse los ejemplos 10 y 11:

Ej. 10: *Que no es la primera cerveza que toman, Raquel, eso quiero decir. No te pongas histérica y no lo saques todo de quicio* (CORPES).

Ej. 11: *Rubalcaba, quizás fascinado por el agresivo estilo de la periodista Ana Pastor, de Televisión Española, que logró acosar y sacó de quicio a algunos dirigentes de la derecha con sus entrevistas, asumió el mismo papel y quiso desestabilizar a Rajoy* (EsTenTen18).

Por otra parte, y en relación a lo anterior, ambas construcciones suelen presentar prosodia negativa⁸. La construcción con Y no animado tiende a aparecer en oraciones negativas, imperativas negativas, con adverbios de negación, adverbios de cantidad que indican exceso, y similares: por ejemplo, *no conviene, no saquemos, no quieras sacar, sacar demasiado, inconveniencia de sacar las cosas de quicio* (cfr. ej. 12). También suele colocarse con palabras que implican una interpretación incorrecta o exageración de la percepción/presentación de Y (*malinterpretar, exagerar, tergiversar, etc.*). En el caso de la construcción con Y animado la prosodia afecta principalmente a X, pues se suele asociar al hecho de manipular a alguien (Y) de forma intencionada hasta alterarlo, o bien se trata de una persona o situación insufrible o cansina que enerva a Y (ej. 13). Son colocaciones frecuentes *propósito, idea, intención de + sacar a alguien de quicio; tener talento, maña, habilidad para + sacar de*

7 Cfr. construcciones resultativas literales, como *sacar la puerta del quicio* o *sacar una puerta de su quicio* frente a una secuencia como **sacar una puerta de quicio*.

8 La construcción resultativa abstracta no conlleva prosodia negativa ni intensificación. No se trata de relaciones heredadas.

quicio; *lograr*, *querer*, *conseguir*, *disfrutar*, *terminar (de/por)*, *acabar (por)*, *ingeniárselas (para)*, *amenazar (con)* + *sacar a alguien de quicio*, entre otras:

Ej. 12: *Está bien que sea llorón pues eso es más realista, todo tiene un límite, y Yukiteru era un puto calzonazos y al final te acaba sacando de quicio, aunque no quieras y aunque no lo odies (EsTenTen18).*

Ej. 13: *Es INSOPORTABLE. Estoy hasta la polla de tus insultitos, Meta... si lo que pretendes es sacar de quicio a la gente, conmigo lo estás consiguiendo. Espero que seas muy feliz con ello (esTenTen18).*

La prosodia negativa también puede expresarse mediante imperativos que instan a resistirse a la manipulación o presión (ej. 14). Cuando la prosodia es positiva se tiende a explicitarlo (ej. 15):

Ej. 14: *No dejes que tus enemigos laborales te saquen de quicio, se educado siempre y mantente cerca de quienes te ayudan (Web Dialectos).*

Ej. 15: *Me enamoré de uno de sus personaje, del personaje de la chica, Lisbeth. En algunos aspectos no la entiendo y me saca de quicio, pero es un sacar de quicio que te engancha y te atrae (EsTenTen 18).*

Además de idiomaticidad semántica, en la segunda construcción se observa un mayor grado de restricción formal (idiomaticidad sintáctica), que se manifiesta mediante la marca del rasgo no animado, y la limitación del inventario de posibles *fillers* léxicos (generalmente, cosas) que aluden, de manera global y casi anafórica, a algún asunto o tema tratado.

No obstante, esta construcción no está completamente lexicalizada, dado que presenta cierta productividad, la cual está en relación al tamaño de los datos analizados. El *filler* central o principal en los tres corpus es *cosas*, generalmente precedido del determinante *las* (CORPES: 38; Web Dialectos: 99; EsTenTen18: 735). Sin embargo, se detectan otros con frecuencia de aparición >1, que presentamos por orden decreciente: *todo* (CORPES); *asunto*, *tema*, *todo*, *algo* (Web Dialectos); *cosa*, *todo*, *eso*, *tema*, *asunto*, *cuestión*, *cuestiones*, *debate*, *libertad*, *situación*, *relaciones* (EsTenTen18). Junto a estos, ocupan esa casilla otros sustantivos que solo ocurren una vez en los corpus (*happax legomena*), como referencia genérica (del tipo *debate*, *comentario*), o bien relacionado con cualquier otro tema. La productividad de esta construcción se manifiesta de manera más palpable cuanto más extenso sea el corpus analizado: *cuento*, *poema* (CORPES); *esto*, *debate*, *palabras*, *tema*, *conceptos* (Web Dialectos); *anécdota*, *argumentaciones*, *chorradas*, *ejemplo*, *situación*, *fuzzy logic*, *prerrogativas*, *objeto*, *problema*, *universo*, *especulación*, *subrayado incisivo*, *sucesos*, *riesgo*, *economía*, *vicios*, *pasiones*, *broma*, *conceptos*, *lo que voy a decir*, *todo lo que hace y lo que dice*, *el peso de Irak*, *la Semana Santa*, etc. (EsTenTen18).

La construcción con Y animado es mucho más productiva en cuanto a los *fillers* posibles para ocupar las casillas X e Y, para las cuales no parece haber ningún tipo de restricciones, más allá de cierta preferencia por la tercera persona, debido a la prosodia semántica negativa de esta construcción. Además, se encuentran instancias que evidencian cierta productividad verbal: *traer, quitar, salir + de quicio; sacarse + de quicio* (ejs. 16, 17, 18):

Ej. 16: *Yo esto de las varas de medir me traen de quicio, y los jetas más* (EsTenTen18)⁹.

Ej. 17: *Si hay algo que me quita de quicio es la defensa que se hace del salario mínimo, cuando es una medida que ataca a las personas más desfavorecidas* (EsTenTen18).

Ej. 18: *Hizo una breve pausa, volviendo a pensar las palabras para responder a lo demás que le había dicho sin volver a sacarse de quicio* (EsTenTen18).

Las construcciones con verbos ergativos, como en el ejemplo anterior, parecen interactuar con otras construcciones intransitivas o ergativas a partir del enlace de subparte (ejs. 19 y 20), de manera que *de quicio* hereda el contenido semántico de la construcción más amplia de la cual forma parte:

Ej. 19: *Después en persona me vino a repetir lo mismo, quizá con más contundencia que nos habíamos salido de quicio* (Web Dialectos).

Ej. 20: *A raíz del accionar del adversario, uno de los puntos de mayor polémica durante el Concilio Vaticano II, fue el tema de la Colegialidad Episcopal. A punto tal que se hizo necesaria una intervención directa del Papa para que las cosas no salieran de quicio*¹⁰ (Web Dialectos).

Por otra parte, *de quicio* también presenta cierta alternancia con otra construcción con la que interactúa a través del enlace de subparte: *fuera de quicio*. Se atestiguan instancias con *sacar* (ejs. 21 y 22), pero también con otros verbos, como *poner, tener* y *dejar* (ej. 23):

Ej. 21: *Tenemos tendencias a exagerar o sacar fuera de quicio los problemas que nos afectan* (EsTenTen18).

Ej. 22: *De ahí nació la polémica sobre el pacifismo democrático, polémica desfigurada, sacada fuera de quicio, como todas las desarrolladas en el curso de los últimos años, y que por eso introdujeron tanta perturbación en la conciencia de los partidos de la Internacional Comunista* (EsTenTen18).

9 Cfr. *traer de cabeza*.

10 Obsérvese el uso de *las cosas* (Y) en la posición X. Ello podría dar lugar a una nueva construcción intransitiva o ergativa.

Ej. 23: *La turba de jovencitos que lo recepcionó en el aeropuerto de Corpac fue demasiado entusiasta con Bill Haley a quien despojó de sus anteojos, su pañuelo, su sombrero, su cartera con 300 dólares y documentos personales. Suceso que le hizo pasar un buen susto y puso fuera de quicio a su representante* (Web Dialectos).

La interacción entre construcciones relacionadas permite explicar también el cambio semántico y los procesos de gramaticalización asociados. Por ejemplo, la construcción *fuera de quicio* es más compositiva y más antigua en el tiempo. Los primeros ejemplos en el Corpus Diacrónico del Español (CORDE)¹¹ la sitúan entre 1445-1480, mientras que sacar *Y_{anim} de quicio* data de mediados del siglo XVI (1535). Por su parte, *sacar Y de quicio* aparece por primera vez en el CORDE a finales del XIX (1847)¹² y, algo más tarde, en 1869, se atestigua el primer ejemplo con *sacar las cosas de quicio*.

5. Conclusiones

El análisis de macrodatos indica que las construcciones resultativas con *de quicio* presentan mayor frecuencia normalizada en la variedad de español peninsular o europeo, seguido del español global y la variedad americana en general, en orden decreciente. No obstante, los corpus analizados no permiten extraer datos concluyentes sobre la distribución de estas construcciones en las distintas variedades de español americano. De los dos tipos de construcciones principales, [X SACAR *Y_{anim} DE QUICIO*] es la más frecuente y la más esquemática. Por su parte, [X SACAR *Y DE QUICIO*] muestra, además, mayor restricción diatópica y más especificidad léxica, por lo que resulta menos productiva, aunque no está completamente lexicalizada. Cuando el componente léxico *cosas* ocupa la casilla Y, la construcción se da casi exclusivamente en la variedad hablada en España.

Parte del significado y función de las construcciones resultativas con *de quicio* está motivada por la construcción genérica que les sirve de base, con la cual interactúan a través de diversas relaciones de herencia (extensión metafórica, enlace de subpartes y vínculo de instancia). Ambas están vinculadas, además, por una relación conceptual de polisemia, que modula el foco de la alteración heredada de la construcción genérica. Al mismo tiempo, estas construcciones fraseológicas presentan propiedades semánticas, léxicas, sintácticas y pragmáticas propias. Además, pueden entrar a formar parte de otras construcciones, en un proceso recursivo. Estas redes complejas afectan a la productividad y a los usos creativos de las construcciones, lo cual tiene importantes consecuencias

11 <https://www.rae.es/banco-de-datos/corde>.

12 Nótese que el verbo *desquiciar* se atestigua con posterioridad a *sacar Y de quicio* en el CORDE: 1724 (objeto animado) y 1874 (objeto no animado).

para los procesos de gramaticalización y cambio semántico. Así, la productividad verbal de ambas construcciones (especialmente en el caso de la segunda) y la interacción con otras construcciones, como *fuera de quicio*, *poner fuera de quicio*, *salirse de quicio*, etc., hacen pensar en *de quicio* como construcción en sí misma, en relación recursiva e interactuando con el resto.

Los resultados obtenidos confirman la necesidad de contar con corpus extensos de miles de millones de palabras para estudiar la variabilidad y productividad de las construcciones fraseológicas. En este trabajo hemos utilizado el corpus de referencia CREA y los megacorpora Web Dialectos y EsTenTen18. Ninguno de ellos cubre suficientemente cada una de las variedades transnacionales del español. Sin embargo, la configuración del CREA y, sobre todo, del Web Dialectos les permite ofrecer una cobertura diatópica más equilibrada. Por el contrario, el EsTenTen18 ofrece más posibilidades para el análisis de construcciones fraseológicas que se centran en el español general o global, y no tanto en las distintas variedades de español habladas en el mundo. En cualquier caso, los resultados obtenidos del análisis de la construcción fraseológica con *de quicio* a partir de los macrodatos disponibles ponen de relieve que la denominada “variedad americana”, lejos de ser un conjunto homogéneo, se vertebra mediante conglomerados de variedades que presentan distintos grados de proximidad respecto a la variedad europea o peninsular.

Finalmente, si bien se podría considerar que la calidad de los dos megacorpora seleccionados se ve comprometida por el evidente desequilibrio de su diseño, la automatización indiscriminada de su compilación, y los documentos que los integran, la principal ventaja que supone su uso es, precisamente, su enorme tamaño, que compensa las limitaciones mencionadas y, en el caso de la fraseología, ofrece mejor cobertura para su análisis (cfr. Corpas Pastor, 2021), como se ha puesto de relieve en este estudio. Incluso el más pequeño de ellos sextuplica el tamaño del CORPES (Web Dialectos), mientras que el EsTenTen18 es casi 52 veces mayor que el CORPES. Otras ventajas de usar megacorpora como los anteriores es la robustez y versatilidad de sus respectivos sistemas de gestión de corpus, frente a las limitaciones del sistema de análisis del CORPES. La otra cara de la moneda, en cambio, son los errores de procesamiento (deduplicación, lematización, anotación morfosintáctica, etc.), frecuentes en este tipo de megacorpora, así como la presencia de errores gramaticales y ortotipográficos en sus documentos. Por tanto, el empleo de megacorpora para este tipo de análisis requiere también el refinamiento de los sistemas de extracción automática de construcciones (genéricas y fraseológicas).

La lingüística de corpus, el procesamiento de lenguaje natural y la gramática de construcciones tienen en la fraseología un magnífico campo de pruebas. Pero el beneficio es también mutuo: el avance de los estudios fraseológicos se nutre de las nuevas técnicas, conceptos y procedimientos provenientes de dichas disciplinas.

6. Bibliografía citada

CORPAS PASTOR, Gloria, 2015: "Translating English Verbal Collocations into Spanish: on Distribution and other Relevant Differences related to Diatopic Variation", *Linguisticæ Investigationes* 38 (2), 229-262.

CORPAS PASTOR, Gloria, 2018: "Laughing one's head off in Spanish subtitles: a corpus-based study on diatopic variation and its consequences for translation" en Pedro MOGORRÓN y Vicent MARTÍNES (eds.): *Fraseología, Diatopía y Traducción*, Ámsterdam: John Benjamins, 54-106.

CORPAS PASTOR, Gloria, 2021: "Constructional idioms of 'insanity' in English and Spanish: a corpus-based study", *Lingua* 254, 103013, doi: 10.1016/j.lingua.2020.103013.

CORPAS PASTOR, Gloria, y Jean-Pierre COLSON (eds.), 2020: *Computational Phraseology*, Ámsterdam: John Benjamins.

CROFT, William, 2012: *Verbs: Aspect and Causal Structure*, Oxford: Oxford University Press.

DOBROVOLS'KIJ, Dimitrij, 2016: "Fraseología y Gramática de Construcciones", *Language Design* (18), 71-106.

GOLDBERG, Adele E., 1995: *Constructions: A Construction Grammar Approach to Argument Structure*, Chicago: Chicago University Press.

GOLDBERG, Adele E., 2006: *Constructions at Work: The Nature of Generalization in Language*, Nueva York: Oxford University Press.

GOLDBERG, Adele E., y Ray JACKENDOFF, 2004: "The English resultative as a family of constructions", *Language* (80), 532-568.

HOFFMANN, Thomas, 2017: "From Constructions to Construction Grammars" en Barbara DANCYGIER (ed.): *The Cambridge Handbook of Cognitive Linguistics*, Cambridge: Cambridge University Press, 284-309.

HOFFMANN, Thomas, y Graeme TROUSDALE (eds.), 2013: *The Oxford Handbook of Construction Grammar*, Oxford: Oxford University Press.

KILGARRIFF, Adam, e Irene RENAU, 2013: "esTenTen, a Vast Web Corpus of Peninsular and American Spanish", *Procedia Social and Behavioral Sciences* 95, 12-19.

LEVIN, Beth, y Malka RAPPAPORT HOVAV, 2006: "Argument Realization", *Computational Linguistics* 32 (3), 447-450.

SUCHOMEL, Vít, y Jan POMIKÁLEK, 2012: "Efficient web crawling for large text corpora" en Adam KILGARRIFF y Serge SHAROFF (eds.): *Proceedings of the seventh web as corpus workshop (WAC7)*, Lyon, 39-43.

TAYLOR, John R., 2016: "Cognitive Linguistics" en Keith ALLAN (ed.): *Routledge Handbook of Linguistics*, Londres y Nueva York: Routledge, 455-469.