

Resumen

La investigación presentada en este artículo pretende mostrar el grado en que los materiales didácticos diseñados específicamente para la docencia/enseñanza de la fraseología a alumnos de español como lengua extranjera recogen las unidades fraseológicas (UF) empleadas en el discurso oral coloquial. Con este fin se han analizado todas las UF documentadas en un corpus oral de 30.000 palabras creado a partir de las conversaciones mantenidas entre los concursantes del *reality show Gran Hermano*. Concretamente, se ha comprobado la presencia o ausencia de estas UF en cuatro tipos distintos de materiales ELE. Los datos aportados por este análisis sacan a la luz la conveniencia de basar en corpus orales las obras diseñadas específicamente para el aprendizaje de la fraseología española, así como la necesidad de replantearse la inclusión de refranes en este tipo de materiales didácticos.

Palabras clave

Fraseología española; ELE; materiales didácticos; discurso oral; español coloquial; *Gran Hermano*.

Abstract

The analysis presented in this paper aims to show the extent to which the learning materials for Spanish as a foreign language designed specifically for the purpose of learning Spanish phraseology collect the phraseological units (PhU) employed in informal discourse contexts. To achieve this, a 30,000-word corpus composed of conversations uttered among the contestants of the Spanish version of the reality show Big Brother was analysed. In particular, whether the PhU excerpted from the corpus were included in four types of learning materials was checked. A critical evaluation of the data obtained shows that it would be convenient to base phraseology-learning materials on oral corpora. Moreover, the inclusion of sayings in works of this sort should be reconsidered.

Key words

Spanish phraseology; Spanish as a foreign language; learning materials; oral language production; informal Spanish; *Big Brother*.

Fecha de publicación: 26/08/2020

DOI: <https://doi.org/10.37536/LINRED.2020.XVII.20>

1. Discurso espontáneo, corpus orales en español y *Gran Hermano*

El trabajo presentado en este artículo forma parte de un proyecto de carácter multidimensional¹ dirigido a sacar a la luz la configuración de la fraseología empleada en discursos orales y, muy especialmente, en contextos que cabría clasificar de informales, familiares o coloquiales, según las muy diversas definiciones y etiquetas empleadas por algunos de los mayores especialistas en el tema (*cf.* Casares 1950/1992; Seco 1973; Lázaro Carreter 1980; Beinhauer 1985). En este sentido, partimos de la siguiente hipótesis: no existe un registro netamente informal frente a otro formal, sino que en el marco de cualquier discurso oral espontáneo se pueden registrar rasgos más o menos formales en función de muy diversos factores. En consecuencia, somos partidarios de la visión dinámica defendida principalmente por Briz a este respecto, según la cual estamos ante una cuestión de carácter eminentemente contextual y gradual, donde no hay espacio para parámetros absolutos, de modo que “[...] cabe distinguir entre *situaciones de mayor o menor coloquialidad* y *situaciones de mayor o menor formalidad* (entre la *constelación comunicativa +/-coloquial* y la *+/-formal* [o, en otros términos, de] situaciones de más o menos inmediatez comunicativa” (Briz y Albelda 2013: 293).

En este sentido, consideramos un acierto la propuesta de Briz de una serie de características propias de la conversación coloquial prototípica, los llamados rasgos coloquializadores, en función de los cuales una conversación tendrá una naturaleza más o menos coloquial/formal. Los principales rasgos coloquializadores son los siguientes:

- + *relación social o funcional de igualdad entre los interlocutores*: acercamiento social o de los papeles comunicativos en un momento dado
- + *relación vivencial de proximidad entre estos*: saberes, experiencias y contextos compartidos
- + *marco interaccional familiar*: relación de cotidianidad de los participantes con el marco espacial en el que se sitúa la interacción
- + *cotidianidad temática de la interacción*: temas de la vida cotidiana, no especializados

(Briz 2010: 126)

A pesar de que en los últimos años se han elaborado varios corpus orales del español, en su mayoría las conversaciones o entrevistas que estos recogen no cumplen adecuadamente los cuatro rasgos coloquializadores arriba señalados. Sería, sin duda, el caso del proyecto más ambicioso de creación de un corpus oral del español, el *Proyecto para el estudio sociolingüístico del español de España y de América* (PRESEEA), basado en la elaboración de corpus orales extensos y variados, pero recogidos mediante entrevistas o diálogos semidirigidos. A causa de estos principios metodológicos, no se evita la conocida como “paradoja del observador” propuesta por Labov, de acuerdo con la cual, resulta muy complejo observar sistemáticamente el

¹ Quiero mostrar mi agradecimiento a la Facultad de Filosofía y Letras de la Universidad Palacký de Olomouc por el apoyo económico e institucional concedido al proyecto “Análisis multidimensional del uso actual de la fraseología española/Multidimenzionální analýza současného užívání španělské frazeologie” (FPVČ2017/04).

modo en que los hablantes se expresan de forma natural sin que su discurso se vea influido por el mismo hecho de saberse observados².

La gran excepción en este sentido entre los corpus orales del español (al menos peninsular) es el corpus creado por el grupo Val.Es.Co. (*Valencia, Español Coloquial*) y dirigido, en sus orígenes, por el propio Briz. En él, se trata de evitar esta “paradoja del observador” a través de la realización de grabaciones secretas en contextos con rasgos marcadamente coloquializadores. A pesar de los problemas técnicos que esto supuso, ante todo en relación con la calidad de los audios, el 90% de las 46 conversaciones y 120.246 palabras que conforman este corpus se realizaron de este modo; así, apenas en un 10% de las grabaciones los hablantes eran conscientes de que estaban siendo monitorizados³. Dadas sus características, este corpus parece ser hoy la fuente de material lingüístico oral espontáneo y de rasgos coloquializadores más fiable en español.

Sin embargo, nuestro análisis multidimensional de la fraseología española en contextos comunicativos espontáneos e informales no se puede basar en el material aportado por el grupo Val.Es.Co., principalmente porque, para los fines específicos de nuestro estudio, resulta necesario hacer un seguimiento detallado de los sujetos observados, lo que implica que el número de palabras por hablante debe ser elevado y que estos deben mantener conversaciones con distintas personas en distintos contextos comunicativos. Además, para nuestro análisis resulta insuficiente la información aportada respecto a los distintos hablantes por el grupo Val.Es.Co. en su corpus.

Por estas y otras razones, hemos decidido elaborar nuestro propio corpus oral de rasgos altamente coloquializadores, realizado a partir de las conversaciones espontáneas mantenidas entre los participantes del concurso *Gran Hermano*. Hemos escogido este *reality show*⁴, ante todo, porque consideramos que estas conversaciones cumplen a la perfección los cuatro rasgos coloquializadores de Briz anteriormente presentados, incluso mejor que muchas de las grabaciones secretas registradas por el grupo Val.Es.Co. Además, podemos monitorizar con una excelente calidad de audio (e incluso vídeo) diálogos realizados por un grupo cerrado de personas, mantenidos por prácticamente todas las combinaciones posibles de hablantes y en diversas situaciones comunicativas durante un periodo de aproximadamente tres meses.

Por otra parte, no nos pasa desapercibido que el uso de *Gran Hermano* como fuente de material lingüístico despierta desconfianza entre los especialistas, precisamente por tratarse de un programa de televisión en el que tan solo se *imita* la realidad. Así, la gran duda metodológica de nuestro proyecto es hasta qué punto podemos tomar por *naturales* las conversaciones mantenidas entre los concursantes. En este sentido, aceptamos que difícilmente quepa hablar de material lingüístico absolutamente natural, puesto que estamos ante un producto televisivo, en mayor o menor grado, dirigido. Efectivamente, los diálogos se desarrollan en un espacio artificial, en el que la realidad más bien se recrea o que, como indica Castañares (2003), se halla a medio camino entre el

² “[T]he aim of linguistic research in the community must be to find out how people talk when they are not being systematically observed; yet we can only obtain these data by systematic observation” (Labov 1972: 209).

³ *cfr.* <http://valesco.es/?q=preguntas>.

⁴ No consideramos necesario detallar en estas páginas las características del concurso; en caso de querer profundizar en este aspecto, recomendamos consultar la obra de Castañares (2003), en español, o de Mathijs y Jones (2004), en inglés.



mundo de la vida real y el de la ficción. Sin embargo, en nuestra opinión, sí consideramos, por las razones que se expondrán a continuación, que las conversaciones que forman nuestro corpus deben considerarse una variante muy cercana –prácticamente idéntica, en realidad– a la lengua natural.

En cualquier caso, no somos los únicos especialistas que trabajan con el material lingüístico procedente de este popular concurso. De hecho, en noruego incluso existe un corpus desarrollado por el Departamento de Lingüística y Estudios Escandinavos de la Universidad de Oslo que recoge las conversaciones de los concursantes del *Gran Hermano* de este país.⁵ Sí es cierto que solo conocemos un trabajo que proponga un estudio lingüístico en español, concretamente el análisis del uso de acusaciones y conflictos interpersonales basado en las versiones española y argentina de *Gran Hermano* a cargo de Haugh y Sinkeviciute (2018). Aun así, debemos reconocer que su uso como fuente de estudios de lengua oral natural o espontánea es muy reducido, al menos si lo comparamos con el enorme interés que este concurso ha despertado desde la perspectiva sociológica y de los estudios de los medios de comunicación. Precisamente desde estas otras disciplinas no es infrecuente considerar que *Gran Hermano* tiene un indiscutible potencial como fuente de conocimiento *real*. Así, tal y como señalada Dovey:

“[Although we are watching] simulated situations that only exist because of the intervention of the TV production [...] the simulation, in my understanding, is not about unreal states [...] the simulation as a material practice produces real knowledge about real things in the real world and has real effects upon real lives” (Dovey 2004: 233).

De acuerdo con las palabras de Dovey, no parece del todo descabellado el uso de *Gran Hermano* con fines de análisis lingüístico. Además, consideramos que la hipotética tendencia de los concursantes a un discurso poco natural, causado por la convivencia con desconocidos y la presencia constante de micrófonos y cámaras, tiende a disminuir (e incluso desaparecer) según pasan los días de “confinamiento” en la casa. Por esa razón, las conversaciones que forman el corpus fueron grabadas a mediados de noviembre de 2015, es decir, tras más de dos meses de convivencia ininterrumpida en la casa y con miles de horas potenciales de diálogos a sus espaldas. En este sentido, parece que incluso en conversaciones semidirigidas recogidas mediante una grabadora a la vista de los hablantes de una duración de entre apenas 45 y 60 minutos, el grado de formalidad disminuía con el paso del tiempo, de modo que algunos participantes acababan usando palabras propias de un tono informal e incluso llegaban a comentar haber olvidado que la conversación estaba siendo grabada (Penadés Martínez 2004: 2227).

Queda abierta, por tanto, la cuestión de si es o no *Gran Hermano* una fuente adecuada para la investigación del lenguaje espontáneo e informal, si bien uno de los objetivos de esta investigación multidimensional (de la que este artículo forma parte) es precisamente comprobar hasta qué punto la configuración de la fraseología empleada en este concurso es o no comparable, en términos principalmente de presencia de rasgos coloquiales, a las conversaciones recogidas en el corpus Val.Es.Co.⁶

⁵ cfr. <http://www.tekstlab.uio.no/nota/bigbrother/english.html>.

⁶ Algunos resultados preliminares sobre esta investigación podrán pronto leerse en Gutiérrez Rubio (2020).

2. Fraseología en la enseñanza de español como lengua extranjera

En la tradición lingüística española la fraseología ha sido generalmente considerada, hasta hace muy poco tiempo, una subdisciplina secundaria o periférica. Y eso en el mejor de los casos, ya que la existencia de un término como *idiotismo* para denotar un tipo de UF parece mostrar más bien un marcado desprecio⁷. Afortunadamente los trabajos, tanto teóricos como prácticos (en forma de glosarios o diccionarios de fraseología), publicados en los últimos treinta años han supuesto un avance sin precedentes en este campo del saber en el contexto español y han colocado de forma definitiva a la fraseología en el espacio que personalmente creemos que se merece: como disciplina lingüística plena y autónoma⁸, algo de lo que hoy pocos especialistas dudan, tal y como podemos leer, a modo de ejemplo, en las siguientes palabras de Luque Nadal (2012: 9): “actualmente los fraseologismos no son vistos como algo inerte y periférico al lenguaje sino como algo central, vivo y creativo”. No es de extrañar que, si la propia disciplina vivía hundida hasta hace poco en un mar de dudas, sin contar con una mínima unidad conceptual, terminológica o tipológica, difícilmente se podía abordar su estudio e integración entre las ramas de la lingüística aplicada, incluida la enseñanza/aprendizaje de ELE. Sin embargo, conforme fueron consolidándose los estudios teóricos de fraseología, especialmente a lo largo de la última década del siglo XX, también fueron apareciendo trabajos cada vez más completos y sistemáticos sobre su aplicación a la enseñanza de aprendientes extranjeros de español. Buena muestra de ello es el breve libro publicado ya en 1999 por Penadés Martínez, *La enseñanza de las unidades fraseológicas*, en el que se llama la atención sobre las principales dificultades relacionadas con la enseñanza y el aprendizaje de la fraseología en el marco de las clases de ELE: por una parte, las causadas por dos de las principales características de las propias UF: su fijación formal e idiomatismo; por otra, las relacionadas con la escasez de materiales específicos adecuados.

No cabe duda de que, en los más de veinte años que han transcurrido desde la publicación de esta obra, han sido muchos los avances que se han logrado en este sentido, entre los que cabe destacar un diccionario de locuciones verbales para aprendientes de español publicado por la propia Penadés Martínez, donde esta pone en práctica muchas de las propuestas presentadas en *La enseñanza de las unidades fraseológicas*: en el diccionario se especifica el nivel que le corresponde a cada locución en la enseñanza/aprendizaje de ELE (intermedio, avanzado o superior), si se usa en situaciones informales o si tiene carácter vulgar, los elementos con los que se combina la locución (alguien, a alguien, para algo, etc.), un ejemplo de uso de la locución, otras locuciones sinónimas y antónimas, además de unos apéndices muy completos, donde, entre otras cuestiones, se presenta un listado de las UF de acuerdo a los tres niveles de enseñanza/aprendizaje de ELE arriba referidos (Penadés Martínez 2002).

⁷ La definición de la tercera entrada del *DRAE* aún lo muestra con meridiana claridad: “*Ling.* Giro o expresión propios de una lengua que no se ajustan a las reglas gramaticales; p. ej., *a ojos vistas*” (<<https://dle.rae.es>>, entrada “*idiotismo*”, última consulta: 18/2/2020). La UF *a ojos vistas* se documenta en el mismo diccionario con la etiqueta de locución adverbial.

⁸ Quizá no esté de más recordar que incluso Corpas Pastor (1996: 15) en el capítulo introductorio de su aclamado *Manual de fraseología española* trataba a la fraseología como una “subdisciplina de la lexicología”.



Otra aproximación teórica, también con vertiente práctica asociada, es la obra *La enseñanza-aprendizaje de fraseología en ELE* a cargo de Velázquez Puerto (2018). Se trata de un trabajo que refleja los importantes avances logrados en las dos décadas transcurridas entre la publicación de esta obra y la de Penadés Martínez arriba mencionada (1999). Su intención ya no es tanto poner el foco sobre la falta de materiales adecuados (algo hasta cierto punto ya superado), como sobre otras cuestiones más centradas en el propio proceso de enseñanza/aprendizaje: los mecanismos de comprensión y producción de las UF, los criterios de su selección e incluso las unidades didácticas⁹.

Junto a estas dos obras monográficas sobre el tema que hemos querido destacar aquí, existe un importante número de artículos centrados en cuestiones concretas del proceso de enseñanza/aprendizaje de la fraseología en la clase de ELE, ya sea desde una perspectiva más teórica, como desde una más práctica. Sin embargo, no es nuestra intención hacer aquí un repaso pormenorizado de los trabajos dedicados a este tema, sino apenas llamar la atención sobre la necesidad de seguir avanzando en este aspecto aún insuficientemente tratado del campo del español como lengua extranjera. Esperamos que la investigación presentada en este trabajo sirva de ayuda en este sentido.

3. Objetivos y principios metodológicos

El objetivo de la investigación presentada en este artículo no es otro que mostrar el grado en que cuatro materiales diseñados específicamente para la enseñanza/aprendizaje de la fraseología a alumnos de español como lengua extranjera recogen las UF empleadas en el discurso oral espontáneo. Concretamente, las UF documentadas en un corpus de 30.041 palabras elaborado a partir de varias horas de conversaciones mantenidas entre los concursantes del *reality show Gran Hermano 16*, emitido entre septiembre y diciembre de 2015 por la cadena privada de televisión Telecinco.

Los fragmentos que contienen las conversaciones que conforman el corpus no se obtuvieron de los programas de televisión, sino de la web que emitía el concurso de forma ininterrumpida. Concretamente, se trata de distintos fragmentos grabados de forma aleatoria en los días 65^º, 67^º y 76^º de la 16^ª edición del *reality show*. Resulta importante recalcar que se descartaron las interacciones mantenidas con miembros del equipo técnico o con los presentadores del programa de televisión, es decir, que se consideraron tan solo las conversaciones llevadas a cabo entre los propios concursantes, que en los días del programa arriba especificados eran diez (de los dieciocho originales), concretamente seis hombres y cuatro mujeres. Aunque el número de palabras recogidas en el corpus para cada uno de los hablantes es desigual, todos ellos pronuncian al menos 1.500 palabras (ver Tabla 1). La edad media de los concursantes era de 22,9 años y, en general, cabe hablar de

⁹ Suya también es la reciente publicación de uno de los pocos manuales diseñados explícitamente para la enseñanza-aprendizaje de fraseología dirigidos a los profesores (especialmente a los alemanes) de español como lengua extranjera: *¡Dale a la lengua! Integrando el aprendizaje de fraseología en ELE* (Velázquez Puerto 2017). Otros manuales enfocados a ELE son *Hablar por los codos* (Vranic 2004) y *Locuciones y refranes para dar y tomar* (Gómez González y Ureña Tormo 2014).

que pertenecen a la clase social media o media-alta, como demostraría que tres de ellos cursaban estudios universitarios al comenzar el concurso (Han, Sofía y Marta).

Nombre	Sexo	Profesión	Edad	Número de palabras
Carlos Rengel	Hombre	Hostelero	25	5.617
Niedziela Raluy	Mujer	Artista de circo	22	4.826
Aritz Castro	Hombre	Artesano y guía turístico	24	3.505
Marta Peñate	Mujer	Estudiante de Periodismo	24	2.769
Han Wang	Hombre	Estudiante de Ingeniería Eléctrica	21	2.677
Marina Landaluce	Mujer	Auxiliar administrativa	28	2.556
Sofía Suescun	Mujer	Estudiante de Psicología	19	2.552
Suso Álvarez	Hombre	Encargado de tienda	22	2.278
Ricky Natalicchio	Hombre	Estudiante de Nutrición	26	1.739
Daniel Vera	Hombre	Empleado de restaurante de comida rápida	23	1.522

Tabla 1. Características de los hablantes del corpus de *Gran Hermano*¹⁰

Por otra parte, las conversaciones analizadas no fueron transcritas. Una vez grabados los fragmentos objeto de análisis de la web que emitía el programa de forma ininterrumpida, los visionamos personalmente en, al menos, tres ocasiones y extrajimos de ellos todas las expresiones que podrían, por sus características, ser calificadas como UF. Después, y tras la consulta de varias fuentes (*DiLEA*, *DFDEA*, *DFEM*, *DELE* y *DRAE*), se decidió si se trataba o no de UF¹¹.

En total, se documentaron 1.501 ocurrencias y 550 UF distintas¹² en el corpus de *Gran Hermano*¹³. Así, la media de ocurrencias por UF es de 2,7, una cifra que, en realidad, resulta engañosa, dado que hay algunas UF que aparecen con grandísima frecuencia en el corpus, como sería el caso de *es que*, documentada en hasta 113

¹⁰ La mayoría de los datos los hemos obtenido de Wikipedia, entrada: "Gran Hermano (España)" (última consulta: 12/2/2020).

¹¹ Aunque no es un tema fundamental de este análisis en concreto, las UF se dividen en tres grupos: a) las UF recogidas en los diccionarios de referencia y que pueden ser clasificadas como UF sin lugar a dudas, como en el caso de *a la defensiva* o *manos a la obra*; b) aquellas que, a juzgar por su relativamente elevada presencia en Internet y mi propia experiencia como hablante nativo, han sido aceptadas sin dudas como UF, a pesar de no documentarse en nuestras fuentes de referencia, como las UF *a machete* o *en cero coma*; c) aquellas expresiones que, aunque al final han sido aceptadas como UF válidas, nos generan algún tipo de duda. Es el caso principalmente de un buen número de fórmulas rutinarias no documentadas en ninguno de los diccionarios de referencia, como *¡qué fuerte!* o *¡es broma!*.

¹² Todas las UF documentadas en el corpus se recogen, junto con su número de ocurrencias, en los tres anexos que cierran este artículo.

¹³ Dadas las características del artículo, no queremos entrar a discutir aquí en detalle la problemática cuestión de la taxonomía fraseológica empleada en el análisis. Creemos suficiente aclarar que se han dividido las UF en tres tipos principales: locuciones, fórmulas rutinarias y refranes.

ocasiones, u o sea con 69 ocurrencias. En el polo contrario se hallan las 350 UF recogidas en una única ocasión en el corpus (ver anexos).

Las obras de ELE diseñadas específicamente para la enseñanza/aprendizaje de la fraseología que han sido analizadas para este estudio fueron elegidas, ante todo, por representar distintos tipos de materiales didácticos: un diccionario fraseológico monolingüe, otro bilingüe, un glosario electrónico trilingüe con ejercicios y, por último, un manual:

- a) El *Diccionario de locuciones verbales para la enseñanza de español (DLVEE)* compilado por Penadés Martínez (2002), al que ya nos hemos referido en el apartado 2, es una obra muy completa y extensa que aporta gran cantidad de información sobre 1.942 locuciones verbales (LV) del español peninsular actual. Aunque no se señale explícitamente que recoge LV comúnmente empleadas en conversaciones naturales o coloquiales, el hecho de que no haya una marca de registro *literario* o *arcaico*, pero sí *informal* y *vulgar*, nos lleva a pensar que, efectivamente, la intención última de la autora no es recoger UF empleadas en registros formales, sino centrarse en aquellas que se usan en el día a día.
- b) El breve diccionario fraseológico bilingüe español-checo compilado por Bajger (2003) que lleva por título *Locuciones idiomáticas español-checo, Španělsko-české idiomatické vazby (ŠČIV)* en el original checo. Sus 78 páginas recogen 539 UF en español, su equivalente (fraseológico o no) en checo y un ejemplo en las dos lenguas. Al diccionario le falta rigor lexicográfico (o fraseográfico, para ser más exactos) y presenta una importante cantidad de erratas. De acuerdo con la escasa información que se aporta en su prólogo, las UF que recoge este diccionario habrían de ser las más comúnmente empleadas en las conversaciones cotidianas en español y servir de complemento para el estudio de esta lengua (Bajger 2003: 3).
- c) Un glosario trilingüe compilado por Raymonde Lamothe y René Gingras accesible de forma gratuita a través de la web *100 imágenes en la punta de la lengua* (http://expressions.ccdmd.qc.ca/repertoire_es.php). El glosario cuenta con 683 UF, incluidos refranes. Sin embargo, dado que algunas UF aparecen repetidas y otras, por el contrario, vienen acompañadas de una o dos variantes (como *no hay ni un gato / no hay ni un alma*), el número definitivo con el que vamos a trabajar es 739 UF. El repertorio bilingüe presenta numerosos errores y muy escasa información sobre las UF. Además, esta obra no es una fuente útil para cualquier aprendiente de ELE, sino únicamente para aquellos que tengan el francés o el inglés como lengua materna, ya que en la web no se especifica el significado o uso de las UF recogidas en él, sino tan solo un ejemplo en español, un archivo de audio y la categoría semántica a la que se adscribe la UF, como *anatomía*, *animal*, *objeto*, etc. Eso sí, al tratarse de un repertorio trilingüe, cada UF se relaciona con la ficha de su equivalente en francés y en inglés (si bien algunas UF incluyen el equivalente en solo una de estas dos lenguas). Además, la web incluye ejercicios, juegos y animaciones encaminados a facilitar el proceso de aprendizaje.

- d) El manual ELE *Hablar por los codos* (Vranic 2004), diseñado específicamente para la enseñanza/aprendizaje de fraseología. En su introducción no se hace referencia al nivel de ELE de las UF recogidas (175 locuciones y 45 refranes), ni si pertenecen al registro coloquial o formal. Esta información sí se desprende del subtítulo del libro, *Frases para un español cotidiano*, así como del texto que aparece en la contraportada, donde se especifica que se trata de expresiones “de uso frecuente en el español cotidiano actual” (Vranic 2004: contraportada). Aunque no se especifique el tipo de locuciones que incluye el manual, se trata ante todo de locuciones verbales (150 de las 175 locuciones que registra).

Por último, resulta imprescindible especificar que nuestro análisis se centrará en dos parámetros que habrán de orientarnos sobre el grado de eficacia de las obras estudiadas. Por una parte, un número absoluto que refleje el porcentaje de UF usadas por los concursantes de *Gran Hermano* y recogidas en los cuatro materiales didácticos analizados. Sin embargo, este dato dependerá en gran medida del tamaño de los diccionarios y del manual y, por tanto, no nos permite realizar un estudio comparativo eficaz de las distintas fuentes. Así, para conocer la medida en que sus autores han acertado al escoger las UF propias de las conversaciones coloquiales (el objetivo explícito o implícito de las cuatro obras estudiadas), resultará necesario conocer un segundo número, relativo en este caso: el porcentaje de UF presentes en el corpus en relación con el número total de entradas de las cuatro obras analizadas, que van desde las 220 UF de *Hablar por los codos* hasta las 1.942 del *DLVEE*.

4. Resultados del análisis

4.1 Diccionario de locuciones verbales para la enseñanza de español (DLVEE)

Dado que el *DLVEE* se centra exclusivamente en el estudio de las locuciones verbales (LV), a la hora de interpretar los datos obtenidos del análisis solo se tendrá en cuenta este subtipo de UF. Así, hay que primero precisar que entre las 550 UF documentadas en el corpus de *Gran Hermano* se contabilizan 96 LV distintas y 148 ocurrencias (1,5 ocurrencias de media). De estas 96 LV, el *DLVEE* recoge 39, que suman un total de 74 ocurrencias (1,9 ocurrencias de media). Estos datos señalan que el diccionario no solo recoge el 41% de todas las LV documentadas en el corpus, sino que, además, estas son algunas de las más frecuentemente empleadas por los concursantes de *Gran Hermano* (exactamente el 50% de las ocurrencias). Un dato que puede ayudarnos a entender hasta qué punto se trata de un diccionario completo (y útil) se refiere a que de estas 96 LV del corpus, el diccionario más completo publicado hasta el momento (la segunda edición del *DFDEA*), recoge 80, es decir, poco más del doble que el *DLVEE*, con 128 ocurrencias (1,6 de media). El hecho de que la media de ocurrencias por UF documentada sea 0,3 veces superior en el *DLVEE* respecto al *DFDEA* habla en favor de que la elección de LV que su compiladora realizó es adecuada en vista de los objetivos del diccionario (enseñanza de ELE), pues las descartadas son menos frecuente en el español, al menos en el empleado en conversaciones espontáneas enmarcadas en situaciones altamente coloquializadoras. Por otra parte, todas las UF sin excepción documentadas en el corpus y recogidas en el *DLVEE* se incluyen también en el *DFDEA*. Otro dato interesante se

refiere al nivel de ELE que Penadés Martínez relaciona con cada una de las UF documentadas. De entre las 39 UF documentadas 18 LV (y 48 ocurrencias) pertenecen al nivel intermedio (2,7 ocurrencias por UF), 20 LV (y 25 ocurrencias) al avanzado (1,25 ocurrencias por UF) y tan solo una UF con una única ocurrencia al nivel superior. En este sentido, hay que tener en cuenta que en el diccionario la distribución de UF pertenecientes a los tres niveles es muy desigual: mientras que las UF marcadas con la etiqueta de nivel intermedio no llegan al 10% del total, al nivel avanzado pertenece aproximadamente el 53% de las LV del diccionario y, por último, aproximadamente el 37% lo hace al nivel superior. Así, si el propósito del diccionario es situar en el nivel intermedio (el más bajo que recoge) las LV más frecuentes en el español empleado en conversaciones “naturales”, no puede haber duda de que el *DLVEE* cumple sobradamente con su objetivo, ya que a este nivel pertenece el 37,5% de todas las LV documentadas en el corpus (48 ocurrencias de 128) con menos del 10% del total de LV, es decir, una media de aproximadamente 0,3 ocurrencias por cada LV de nivel intermedio. En el extremo opuesto, la única ocurrencia correspondiente al nivel superior supone aproximadamente 0,002 ocurrencias por cada LV de este nivel de ELE incluida en el diccionario. Además, hay que señalar que ninguna de las LV marcadas como del nivel avanzado o superior cuenta con más de dos ocurrencias en el corpus. Por el contrario, entre las del nivel más bajo (intermedio) destacan tres LV por su elevada frecuencia: *hacer falta* (5 ocurrencias), *dar la razón* (7) y *ser/dar igual* (10).

Por último, de cara a poder establecer una comparación con los otros tres materiales didácticos analizados en este trabajo, hay que estudiar cuántas de las 1.942 LV que recoge el diccionario se usan en las conversaciones llevadas a cabo entre los concursantes del *reality show*. Así, las 39 LV significan exactamente el 2% de las 1.942 UF incluidas en el diccionario.

4.2 Locuciones idiomáticas español-checo, *Španělsko-české idiomatické vazby (ŠČIV)*

Este breve diccionario fraseológico bilingüe español-checo recoge quince UF de entre las documentadas en el corpus de *Gran Hermano*: seis locuciones verbales, cinco adverbiales, una nominal, una adjetiva/comparativa, una conjuntiva y una prepositiva. Además, tres de ellas presentan dos ocurrencias: *como una cabra*, *dar pie* y *de golpe*. En total, por tanto, se registran 18 ocurrencias, lo que supone una media de 1,2 ocurrencias por UF.

Si tomamos ahora en consideración que en el corpus se documentan 345 locuciones –es decir, sin tener en cuenta las fórmulas oracionales y los refranes, tipos estos de UF que no se incluyen en el *ŠČIV*–, las quince UF recogidas en el diccionario suponen que un usuario encontrará en él un 4,3% de las locuciones empleadas en las conversaciones que conforman el corpus, así como el 1,9% de sus ocurrencias. Aunque hay que considerar algunos factores relativos a la metodología de la compilación de este diccionario –como la ausencia de UF vulgares en el *ŠČIV* (que suponen más del 6% de las locuciones registradas en el corpus), así como de muchas de las UF menos idiomáticas, especialmente de las locuciones marcadoras y conjuntivas– consideramos que las cifras relativas a las UF documentadas en el corpus y recogidas por el *ŠČIV* son excesivamente bajas para un diccionario que, aunque breve, asegure en su prólogo recoger aquellas UF que forman la base de cualquier conversación cotidiana.

Por otra parte, de cara al análisis comparativo hay que señalar que las quince UF del *ŠČIV* que se documentan en el corpus de *Gran Hermano* suponen el 2,8% de las 539 UF compiladas en el diccionario. Esta cifra relativa es

superior a la obtenida para el *DLVEE*, lo que significa que, aunque se trata de un diccionario excesivamente breve y con grandes carencias metodológicas, al menos la elección de las UF en él incluidas es, a la vista de los resultados, adecuada.

4.3 Repertorio trilingüe 1000 imágenes en la punta de la lengua

En el glosario fraseológico de esta web hemos documentado diez UF: cinco locuciones verbales, dos adverbiales, una nominal, una pronominal y una adjetiva/comparativa. Esta última, la UF *como una cabra*, presenta dos ocurrencias, mientras que el apelativo *mi vida* se documenta en tres ocasiones entre las conversaciones de *Gran Hermano*. En total, lógicamente, se registran trece ocurrencias, lo que supone una media de 1,3 ocurrencias por UF. El corpus recoge 550 UF y 1.501 ocurrencias (incluidos refranes), de modo que las diez UF recogidas por el repertorio trilingüe implican que sus usuarios solo hallarían información sobre el 1,8% de las UF empleadas en las conversaciones (y apenas un 0,9% de las ocurrencias). Nuevamente, como ya ocurriera con el diccionario español-checo, juzgamos este un porcentaje extremadamente bajo.

Por último, hay que añadir que, dado que el glosario recoge 739 UF, en el corpus documentamos el 1,3% de las UF propuestas en el repertorio trilingüe, un porcentaje menor que el de las otras dos obras analizadas. Esta divergencia posiblemente se deba a que, a diferencia de los otros dos diccionarios, el repertorio trilingüe recoge un importante número de refranes, un tipo de UF sin apenas presencia en el corpus de *Gran Hermano*.

4.4 Hablar por los codos

Este manual recoge cuatro UF de entre las documentadas en el corpus de *Gran Hermano*, todas ellas con una única ocurrencia: *andarse por las ramas*, *irse por los cerros de Úbeda*, *Írsele la olla* y *un cero a la izquierda* (la única de las UF que no es una LV). Por otra parte, los 45 refranes registrados en el manual no coinciden en ningún caso con los cinco refranes documentados en las conversaciones de *Gran Hermano*.

Si tenemos en cuenta tan solo las locuciones y los refranes —es decir, excluyendo de los datos del corpus las 200 fórmulas oracionales y sus 544 ocurrencias, un tipo de UF que no recoge el manual— las cuatro UF de *Hablar por los codos* suponen el 1,1% de las UF recogidas en el corpus y apenas el 0,4% de sus ocurrencias. Esto refleja a las claras la enorme dificultad que supone la enseñanza/aprendizaje de la fraseología: incluso si un estudiante extranjero lograra aprender de memoria las 175 locuciones y 45 refranes del libro (algo poco probable), no entendería el 99% de las UF usadas en las conversaciones naturales que conforman el corpus (y eso sin tener en cuenta las 200 fórmulas oracionales). Cabría pensar que esta cifra sería mucho más baja si la autora del manual hubiera optado por UF más comúnmente empleadas. Sin embargo, los primeros puestos respecto a la frecuencia de uso en el corpus lo ocupan UF con muy poca carga figurativa, como *es que*, *o sea*, *a lo mejor*, *un poco*, *un montón* o *por ejemplo*. Dado que en *Hablar por los codos* a cada UF le acompaña un dibujo que ilustra el uso o significado de la UF, así como, en muchos casos, la explicación de su origen, no hay duda de que este tipo de UF —en su mayoría locuciones marcadoras, aunque también hay un buen número de locuciones nominales— no tiene cabida en un manual de estas características. En cualquier caso, consideramos que el uso de corpus orales debería ser un punto de partida imprescindible previo a la elaboración de este tipo de materiales didácticos, dado que varias de las UF

incluidas en este manual –aun presentando un origen que cabría calificar de interesante y que, por tanto, podría despertar el interés del aprendiente extranjero– no se usan con frecuencia en el español actual, como, a modo de ejemplo, *cargarle el mochuelo* o *atar perros con longanizas*.

Acabamos esta sección señalando que las 4 UF que se recogen en el corpus suponen el 1,8% de todas las UF recogidas en *Hablar por los codos* (si excluyéramos los 45 refranes, el porcentaje subiría hasta el 2,3%).

4.5 Recapitulación de los datos obtenidos para los cuatro materiales didácticos

En la Tabla 2 se recogen los principales datos que nuestro análisis ha sacado a la luz:

Título del material ELE	Número de UF compiladas	Tipo de UF compiladas	UF recogidas (% de su tipo)	Ocurrencias recogidas (% de su tipo)	% respecto al total de UF compiladas
<i>DLVEE</i>	1.942	solo locuciones verbales	39 (41%)	74 (50%)	2%
<i>ŠČIV</i>	539	todo tipo de locuciones	15 (4,3%)	18 (1,9%)	2,8%
<i>Repertorio trilingüe</i>	739	locuciones, fórmulas y refranes	10 (1,8%)	13 (0,9%)	1,3%
<i>Hablar por los codos</i>	220	locuciones y refranes	4 (1,1%)	4 (0,4%)	1,8%

Tabla 2. Principales datos obtenidos del análisis de los cuatro materiales didácticos

5. Conclusiones

Los datos presentados en la sección anterior dejan pocas dudas respecto a la eficacia de algunos de los materiales ELE diseñados específicamente para el proceso de enseñanza/aprendizaje de la fraseología. Así, la única fuente entre las cuatro analizadas que sería útil para entender una parte sustancial de las UF empleadas en las conversaciones realizadas en contextos altamente coloquializadores es el *Diccionario de locuciones verbales para la enseñanza de español (DLVEE)*, que recoge el 41% de las LV empleadas en las conversaciones de nuestro corpus. Además, en él se registran muchas de las más frecuentes (1,9 ocurrencias de media frente al 1,5 del total de LV documentadas en el corpus) y acierta, a nuestro entender, a la hora de clasificar las LV recogidas en los tres niveles de ELE. Sin embargo, no debemos olvidar que se trata de un material de consulta, un diccionario al uso, no de una obra diseñada para fomentar activamente el aprendizaje de UF. Además, centra su atención en una parte muy pequeña del total de UF documentadas en el corpus, ya que las LV suponen tan solo el 17,4% de las UF del corpus (y el 9,9% del total de ocurrencias). Por su parte, las otras tres obras analizadas (de mucho menor extensión) muestran cifras extremadamente bajas en este sentido: el 4,3 % de las UF empleadas (y el 1,9% de las ocurrencias) en el *ŠČIV*, el 1,8% (y el 0,9%) en el repertorio trilingüe y apenas de 1,1% (y el 0,4%) en *Hablar por los codos*. Ninguna

de estas tres fuentes supone, por tanto, una ayuda significativa para los aprendientes extranjeros que pretendan entender conversaciones similares a las registradas en *Gran Hermano*. Otro factor para tener en cuenta es que el objetivo primario de estas tres obras, especialmente en el caso del glosario de Internet y del manual *Hablar por los codos*, no es tanto recoger el mayor número posible de UF como ser una herramienta útil en el proceso de enseñanza/aprendizaje de la fraseología en ELE. Aun así, creemos que el glosario de Internet debería presentar cifras más similares, e incluso superiores, a las del *ŠČIV*, dado que recoge un número de UF 1,4 veces mayor (739 UF frente a 539 UF).

El último dato que queremos interpretar aquí es el referido al porcentaje de UF presentes en el corpus en función del número total de entradas de los materiales didácticos, lo que habría de darnos una idea algo más objetiva de hasta qué punto los autores de las cuatro fuentes consultadas han realizado una labor adecuada a la hora de seleccionar las UF. En este sentido, la obra más *eficaz* sería el breve diccionario español-checo *ŠČIV*, dado que el 2,8% de sus UF están documentadas en las conversaciones que conforman el corpus. A continuación se situarían el *DLVEE* con un 2% de sus 1,942 LV y *Hablar por los codos* con un 1,8% de sus 220 locuciones y refranes. A la cola se hallaría el repertorio bilingüe con apenas el 1,3% de sus UF. Consideramos que los datos relativos al manual y al glosario de Internet se han visto especialmente perjudicados por la inclusión de refranes, ya que estos son extremadamente infrecuentes en el corpus con solo cinco entradas, lo que supone apenas el 0,9% de las UF documentadas y el 0,3% de las ocurrencias totales.

Una vez presentados e interpretados los datos obtenidos del análisis de las cuatro obras objeto de estudio, consideramos que los futuros materiales didácticos elaborados con el fin de mejorar el proceso de enseñanza/aprendizaje de la fraseología en ELE deberían partir de corpus orales. Además, el hecho de que la frecuencia de uso de los refranes en las conversaciones reales –al menos en las llevadas a cabo en contextos coloquializadores entre los miembros de una generación joven, de entre veinte y treinta años– se haya demostrado tan extremadamente baja, nos lleva a pensar que los autores de dichos materiales especializados deberían replantearse la inclusión de este tipo de expresiones en sus obras. Además, a la luz de los datos obtenidos, queda claro que la elaboración de diccionarios y glosarios con apenas unos cientos de UF no parece ser una herramienta eficaz para sus usuarios y su publicación solo se justificaría si vienen acompañados de materiales complementarios, como ejercicios o ilustraciones que ayuden a su comprensión y memorización.

Por último, subrayar que el análisis ha demostrado la enorme dificultad que supone el dominio (aunque sea pasivo) de las UF por parte de aprendientes de ELE, como mostraría el hecho de que, en el improbable supuesto de que un estudiante aprendiera de memoria el uso de las 175 locuciones y 45 refranes incluidos en el manual *Hablar por los codos*, apenas entendería cuatro de las 1.501 ocurrencias registradas en el corpus de *Gran Hermano*, es decir, el 0,3% del total de UF empleadas en varias horas de conversaciones.

Enrique Gutiérrez Rubio

ORCID: 0000-0001-8877-4446

enrique.gutierrez@upol.cz

University Palacký Olomouc - Czech Republic



Referencias bibliográficas

- Bajger, Kryštof (2003): *Španělsko-české idiomatické vazby (SČIV)*, Ostrava: Impex.
- Beinhauer, Werner (1985, 3.ª ed. española): *El español coloquial*, Madrid: Gredos.
- Briz, Antonio y Albelda, Marta (2013): “Una propuesta teórica y metodológica para el análisis de la atenuación lingüística en español y portugués. La base de un proyecto en común (ES.POR.ATENUACIÓN)”, *Onomázein. Revista semestral de lingüística, filología y traducción*, 28, pp. 288-319.
- Briz, Antonio (2010): “Lo coloquial y lo formal, el eje de la variedad lingüística”, Rosa María Castañer Martín y Vicente Lagüéns Gracia (coord.), *De moneda nunca usada: Estudios dedicados a José Mª Enguita Utrilla*, Zaragoza: Instituto Fernando El Católico, CSIC, pp. 125-133.
- Casares, Julio (1950/1992): *Introducción a la lexicografía moderna*, Madrid: CSIC.
- Castañares, Wenceslao (2003): “La vida en directo, el último espectáculo televisivo”, Gabriela Gómez Rodríguez y María Martha Renero Quintanar, *Tv Global y espectáculos locales*, Guadalajara (Méx.): Universidad de Guadalajara, pp. 127-145.
- Corpas Pastor, Gloria (1996): *Manual de fraseología española*, Madrid: Gredos.
- Dovey, Jon (2004): “It’s only a game show: Big Brother and the theatre of spontaneity”, Ernest Mathijs y Janet Jones (eds.), *Big Brother International. Formats, Critics and Publics*, London & New York: Wallflower Press, pp. 232-248.
- Gómez González, Alba y Ureña Tormo, Clara (2014): *Locuciones y refranes para dar y tomar. El libro para aprender más de 120 locuciones y refranes del español*, Alcalá de Henares: Servicio de Publicaciones de la Universidad de Alcalá.
- Gutiérrez Rubio, Enrique (2020): Realidad, *reality* y ficción. Análisis de la fraseología española en tres contextos comunicativos informales, *Lingüística y Literatura*, 79 (en prensa).
- Haugh, Michael y Sinkeviciute, Valeria (2018): “Accusations and interpersonal conflict in televised multi-party interactions amongst speakers of (Argentinian and Peninsular) Spanish”, *Journal of Language Aggression and Conflict*, 6 (2), pp. 248-270.
- Labov, William (1972): *Sociolinguistic patterns*, Philadelphia: University of Pennsylvania Press.
- Lamothe, Raymonde y Gingras, René (2009): Repertorio trilingüe. 100 imágenes en la punta de la lengua (http://expressions.ccdmd.qc.ca/repertoire_es.php).
- Lázaro Carreter, Fernando (1980): *Estudios de lingüística*, Barcelona: Crítica.
- López Simó, Mireia (2016): *Fórmulas de la conversación. Propuesta de definición y clasificación con vistas a su traducción español-francés, francés-español*. Tesis doctoral inédita. Universidad de Alicante.
- Luque Nadal, Lucía (2012): *Principios de culturología y fraseología españolas. Creatividad y variación en las unidades fraseológicas*, Frankfurt Am Main: Peter Lang.
- Martínez López, Juan Antonio y Jørgensen, Annette Myre (2009): *Diccionario de expresiones y locuciones del español (DELE)*, Madrid: Ediciones de la Torre.
- Mathijs, Ernest y Jones, Janet (2004): “Introduction: Big Brother International”, Ernest Mathijs y Janet Jones (eds.), *Big Brother International. Formats, Critics and Publics*, London & New York: Wallflower Press, pp. 1-8.
- Penadés Martínez, Inmaculada (1999): *La enseñanza de las unidades fraseológicas*, Madrid: Arco/Libros.



- Penadés Martínez, Inmaculada (2002): *Diccionario de locuciones verbales para la enseñanza de español (DLVEE)*, Madrid: Arco/Libros.
- Penadés Martínez, Inmaculada (2004): “¿Caracterizan las locuciones el registro coloquial?”, Milka Villayandre Llamazares (ed.), *Actas del V Congreso de Lingüística General, León, 5-8 de marzo de 2002*, III, Madrid, pp. 2225-2235.
- Penadés Martínez, Inmaculada (2019): *Diccionario de locuciones idiomáticas del español actual (DiLEA)* (<http://www.diccionariodilea.es/diccionario>).
- Real Academia Española (2014, 23ª ed.): *Diccionario de la lengua española (DRAE)* [versión 23.3 en línea] (<https://dle.rae.es>).
- Seco, Manuel; Andrés, Olimpia; Ramos, Gabino y Domínguez, Carlos (2017, 2ª ed.): *Diccionario fraseológico documentado del español actual: locuciones y modismos españoles (DFDEA)*, Madrid: Aguilar.
- Seco, Manuel (1973): “La lengua coloquial: ‘Entre visillos’, de Carmen Martín Gaité”, Emilio Alarcos et al., *El comentario de textos*, Madrid: Castalia, pp. 361-379.
- Varela, Fernando y Kubarth, Hugo (1994): *Diccionario fraseológico del español moderno (DFEM)*, Madrid: Gredos.
- Velázquez Puerto, Karen (2017): *¡Dale a la lengua! Integrando el aprendizaje de fraseología en ELE*, CreateSpace Independent Publishing Platform.
- Velázquez Puerto, Karen (2018): *La enseñanza-aprendizaje de fraseología en ELE*, Madrid: Arco Libros - La Muralla.
- Vranic, Gordana (2004): *Hablar por los codos: Frases para un español cotidiano*, Madrid: Edelsa.

Anexo I. Locuciones documentadas en el corpus

a base de, a bocajarro, a gusto (2), a la altura de, a la defensiva, a la vez, a lo bestia, a lo mejor (17), a lo/la que, a machete, a mano, a medida que, a muerte (3), a pulso, a punto de, a que [+ pregunta] (2), a saco, a solas, a su bola, a su* manera (3), a su* rollo, a todas horas, a todo esto, a tope (4), a través de, a última hora, a unas malas, a veces (16), a ver [más proposición interrogativa/L. conj.] (12), a ver quién (es el guapo/valiente que), a ver si [deseo/L. conj.] (7), a ver [L. marc.] (18), a/de malas, abrir los ojos, acabarse la fiesta, ahora mismo/mismito (2), al / a un lado (3), al cabo de, al cien por cien, al contrario, al final (14), al menos, al oído, al principio (de) (5), al revés (6), alguno que otro, amor odio, andarse/irse por las ramas, aparte de, así que, aun (y) así / incluso/siquiera así, aunque sea (2), bordarlo, cada dos por tres, cada uno (4), cada vez que, cagarla, coger coraje, coger el truco, cogerla/pillarla, comerle la polla, comerse la cabeza, como (la) mierda / como si fuera una mierda (2), como (muy) tarde, como (un) loco (2), como el perro y el gato, como en casa [sentirse], como mucho (2), como que, como quien no quiera la cosa, como si (3), como un cerdo, como una casa, con el corazón, con todo mi amor, correr riesgo, dar (o meter) caña (2), dar cosa, dar fe, dar guerra, dar gusto, dar la razón (7), dar las gracias, dar pie (2), dar por hecho, dar yuyu (2), dar/pegar para el pelo (2), dar/pegar un corte, darle [a algo], darle algo [a alguien], darlo todo, darse cuenta [de algo], darse de cabezazos (contra la pared) (2), dársele [algo + valoración], de acuerdo, de antemano, de bote, de cachondeo, de cojones, de compras [ir/salir/estar], de frente, de golpe (2), de lado (2), de lo lindo, de lujo, de marca, de mierda (3), de momento (5), de muerte, de nuevo, de parte (de), de puta madre (3), de rebote, de repente, de risa, de siempre, de sobra, de todo (3), de todo (2), de verdad [L. adv.] (4), de verdad [L. adj.], de viaje, de/con mil amores, de/en broma (3), dejar tirado (2), dejarlo para última hora, delante de, dentro de, día a día, echar de menos, echar la papa/pota, el tema es que, en (repetidas) ocasiones, en abundancia, en cabeza, en cero coma, en contra, en directo, en el culo, en el fondo, en el último momento, en frente, en las mismas, en medio (4), en mi vida (2), en opinión, en plan (13), en realidad, en serio (8), en su salsa, en teoría, en todo momento, en verdad (4), en vez de (3), entre horas (2), es decir, es que (113), es verdad que (4), está claro que, estar ahí (ahí) (2), estar como una (puta) chota, estar como una cabra (2), estar con / tener las hormonas revolucionadas, fundírsele los plomos, haber que ver [algo], hacer caso, hacer falta (5), hacer gracia, hacer ilusión, hacer un feo, hacerse el loco, hagas lo que hagas, hasta el gorro, hasta la polla (2), hasta la saciedad, hecho (una) mierda/ como una mierda (3), hoy día / en día, ir/salir de marcha, irse de las manos, irse/salir por los cerros de Úbeda, írsele la olla, jugársela (2), la verdad (2), liarla, llevar la contraria, llevar(lo) [+ valoración], llevarse [+ valoración] (3), lo mismo [=a lo mejor] (2), lo mismo [=igual] (6), lo peor, lo que pasa/ocurre es que (2), lo que sea (3), lo único es que (2), mala leche, mala vida, más bien, más claro que el agua, más o menos (5), mentiría si (te) dijera que, meter la pata, meter la tijera (2), meter mano, mi amor (3), mi vida (3), mirada perdida, mitad y mitad, morirse/partirse de risa, nada de nada, ni de coña, ni mucho menos, ni nada (del/por el estilo) (4), ni por el forro, ni que, ni siquiera (2), (ni) un pelo, ni... ni hostias, no poder más, no por nada, no sé qué [L. adj.], no sé qué [L. marc.] (16), no sé qué [L. pron.], no sé (11), no sea que /no vaya a ser que (2), no tener (ni) (puta, pajolera) idea (6), o algo (5), o así, o sea / o séase (3), o sea [muletilla] (69), o tal, oler que alimenta, otra vez (3), para nada (2), parecerle [+ valoración] (2), pasarlo [+ valoración], pasarse por la cabeza, pedazo de (2), perder el tiempo (2), perico el de los palotes, pero bueno (7), pero claro (2), perro/ito faldero (2), poca vergüenza, poco a poco (2), poner/hacer ojitos, ponerse burro, por A por B, por ahí [localización] (2), por ahí [aproximadamente], por ciento (8), por debajo de, por dentro, por ejemplo (14), por encima (4), por*

la cara (3), por lo bajo, por lo menos, por los pelos/un pelo, por mucho/más que, por poner un ejemplo, por si (acaso), por voluntad propia, porque sí, pues eso (6), pues mira (7), pues sí, (pues) sí que (3), (pues) ya está, qué es de, que es un gusto/da gusto (2), que pela, ¡qué sé yo! / ¡yo qué sé!, que sepas que, que si [+ repetición del verbo anterior], que si tal/que tal, [SN] que te cagas (3), que todas las cosas, quedarse frito, quieras que no (2), romper/partir/dar pana, saber mal (4), salir del coño/moño, salirle de los huevos / del nabo, seguir el hilo (de la conversación), seguir el rollo, según qué cosas, ser el/lo/la que es, ser la polla, ser una máquina, ser/dar igual (10), ser/dar lo mismo, siempre que, sin embargo, sin más (2), sin querer, sin tanta hostia, sobre todo, soltarse la melena (2), su madre, sudársela/sudar la polla (2), tal cual (5), tal vez, tal y cual (3), tanto... como..., te guste o no, tener (un par de) cojones, tener huevos, tener la gracia en el culo, tener que ver (4), tener/llevar razón (3), tener/ser *tu día, tirar caña (2), tirar del carro, tocar/chupar la polla (3), todas las cosas, todo cristo, todo el día (9), todo el mundo (6), todo el rato (2), tú a lo tuyo, un cero a la izquierda, un euro, un huevo (4), un montón [L. adv.] (9), un montón [L. nom.] (10), un poco [L. adv.] (18), un poco [L. nom.] (10), un tanto, una especie (de), una gota, una mierda (3), una pasta, una vez que, una/otra cosa (distinta) es que (2), valer/merecer la pena, vamos a ver, vender la moto, venir de lejos/de atrás (2), venirse arriba, y eso / todo eso (2), y mira que, y nada (2), y tal (11), y todo (11), (y) ya está (12), ya (te) digo (3).*

Anexo II. Fórmulas rutinarias documentadas en el corpus¹⁴

a cascarla, a la mierda, a la una... a las dos... (y) a las tres / una, dos y tres (4), a lo mejor, ¿a que sí? (7), a ver (qué remedio) (3), a ver [fórmula de relación interpersonal] (8), a ver [fórmula personal] (12), (además) de verdad, adivina adivinanza, ¿ah sí? (5), ah, bueno (4), ah, no (2), ahí está, ahí lo dejo, ay señor, ay sí (2), buenas noches, buenos días (3), como lo oyes / lo que oyes, ¿cómo lo ves?, como quien dice, da/es igual (12), de nada, ¿de qué vas?, de verdad (10), déjame en paz (2), déjate de historias, desde luego, despacito y buena letra, ¿en qué mundo vives?, ¿en serio? (18), ¡es acojonante!, ¡es broma! (3), ¡es increíble!, es lo que hay, es verdad [fórmula impersonal] (10), es verdad [fórmula personal] (2), escucha/mira una cosa (2), eso es (2), eso espero, eso no te lo crees ni tú*, está claro / claro está, ¡está guapo!, gracias por la parte que me toca / por lo que me toca, hasta mañana, hay gente para todo, hijo mío, hostia puta, la has cagado, la madre que le* parió/trajo (2), ¿lo entiendes?, lo más seguro, lo que estábamos diciendo, lo que te digo, lo que yo te diga, lo siento (mucho) (3), ¿lo ves?, madre mía / de dios (4), manos a la obra, más y mejor, me cago en la virgen / la leche / la puta (7), me importa un bledo, me la pela, me la suda (2), menos mal (2), menuda (puta) mierda, mira a ver (2), mira qué bien, mira que me extraña, muchas gracias (2), ¡muy bien! (2), ¡muy guapo!, ¡muy jeví!, no (lo) sé [fórmula personal cognitiva] (5), no (lo) sé [fórmula personal sintomática] (24), ¡no (me) jodas! (2), ¿no crees? (2), no es justo, no importa, no me digas, no me hagas esto (2), ¡no me lo puedo creer!, no mientas (2), no pasa nada (6), no sé yo, no seas putas, no te preocupes (4), (no tengo) ni (puta, pajolera) idea (3), no veas (2), no volveré a hacerlo, ¿o qué? (15), otra vez, pero bueno, pero esto qué es, pero hombre, pero qué va a ser [alguien algo], por eso (te digo) (4), por favor [fórmula personal] (7), por favor [fórmula de relación interpersonal] (15), puede ser (5), pues eso (2), pues nada, pues no, pues sí, (pues) ya está [fórmula personal] (8), (pues) ya está [fórmula de*

¹⁴ Dentro de las fórmulas rutinarias, seguimos la clasificación propuesta por López Simó (2016).

relación interpersonal] (4), *¡qué (puto) asco!* (5), *¿qué (te) pasa?* (8), *¡qué alegría!*, *¡qué arte!*, *¡qué bien!*, *¡qué bueno!* (3), *¡qué cabrón!* (3), *¡qué chulo!*, *¡qué coñazo!*, *¡qué cosa(s)!* (2), *¡qué divertido!*, *¡qué envidia!*, *¡qué estrés!*, *¡qué fuerte!* (6), *¡qué gracia / vaya (una) gracia!* (2), *¡qué guapo!* (9), *¡qué guay!* (12), *¡qué horror!*, *¡que le* den por el culo!*, *¡qué mal (me parece)!*, *¡qué maravilla! ¿qué me cuentas / estás contando?*, *¡qué miedo!*, *¡qué mono!*, *¡qué movida!*, *¡que no!* [fórmula personal] (2), *¡que no!* [fórmula de relación interpersonal] (3), *¡qué pena!* (2), *¡qué pesadilla!*, *¡qué pesado!*, *¡qué peste!* (3), *¿qué quieres que (le) haga?*, *¿qué quieres que te diga?*, *¡qué rabia!*, *¡qué raro!*, *que se prepare*, *¡qué sé yo! / ¡yo qué sé!* (8), *que sí* (3), *¡qué sueño!*, *¡qué suerte!*, *¿qué tal?* (4), *¡que te den por el culo!*, *¡qué triste!*, *¡qué va!* (11), *(qué) más quisieras*, *¿sabes lo que pasa?* (2), *¿sabes lo que te digo / quiero decir?* (8), *¡será cabrón!*, *¿sí o no?* (2), *si tú supieras*, *sí hombre*, *sí pero no*, *sí ya*, *¿sí, no?*, *tal cual te lo digo*, *tal para cual*, *te digo una cosa*, *te lo digo (yo)* (3), *¿te lo puedes creer?*, *te pido perdón*, *¡te voy a matar!*, *te* lo juro/prometo* [fórmula personal] (10), *te* lo juro/prometo* [fórmula de relación interpersonal] (1), *¡te* vas a cagar!* (2), *ten cuidado*, *(todo) bajo control*, *tres... dos... uno...* (5), *¿tú crees?*, *¿tú te crees (que es normal/posible)?*, *tú tranquilo*, *(vamos) digo yo* (2), *vas a flipar*, *¡vaya asco!* (2), *¡vaya mierda!*, *¡vaya putada!*, *¡vaya tela!*, *venga a [+ infinitivo]*, *venga va/vamos* (19), *¡venga, hombre!*, *¡vete a la mierda!* (2), *(vete) a saber* (2), *vete/anda por ahí* (2), *y lo sabes*, *¿y eso?*, *y fuera*, *¿y qué le voy a hacer?* (2), *¿y qué?*, *ya (lo) veremos* (2), *ya está* (9), *ya lo sé* (5), *ya te digo* (4), *ya verás / verás tú / y verás* (2), *ya ves* (3).

Anexo III. Refranes documentados en el corpus (en la variante recogida en las conversaciones)

Cada ladrón se piensa que son de su condición, mejor una verdad que duela que una mentira duradera, no cantes victoria antes de tiempo, una cosa es una cosa y otra cosa es otra cosa, una vez metido nada de lo prometido.