

Las pirámides etnolingüísticas en español: una nueva propuesta para la visualización de datos

FRANCESCO SPINOGLIO
ESADE / Universidad Ramon Llull
francesco.spinoglio@esade.edu

RESUMEN: Este artículo ofrece un estudio etnolingüístico que, a partir de una visualización de datos denominada *pirámides etnolingüísticas*, tiene como finalidad analizar la influencia de la cultura en la lengua española a partir de dos ámbitos culturales escogidos: el fútbol y los toros. Partiendo de la hipótesis de que es en las expresiones donde se puede apreciar la influencia de la cultura en la lengua, se ha creado un sistema de visualización de datos que consta de 4 pasos: el brainstorming (o tormenta de ideas), la obtención del listado de trabajo, la búsqueda en los corpus y la ordenación dentro de las pirámides.

Palabras clave: etnolingüística, pirámides, brainstorming, corpus, fútbol, toros.

ABSTRACT: This article offers an ethnolinguistic study based on a data visualization called *ethnolinguistic pyramids*. This study aims to analyze the influence of culture on the Spanish language from two cultural fields: soccer and bullfighting. Starting from the hypothesis that it is in expressions where the influence of culture on the language can be noticed, a data visualization system has been created consisting of 4 steps: brainstorming, work list elicitation, corpus search and pyramid ordination.

Keywords: ethnolinguistics, pyramids, brainstorming, corpus, football, bulls.

1. INTRODUCCIÓN

La idea de este artículo surgió a partir de una profunda inquietud por la etnolingüística, una rama de la lingüística a la que se le ha prestado poca atención en los últimos años y de la que podemos extraer valiosa información acerca de la influencia de la cultura en la lengua¹.

Según Germán Fernández Guizzetti (1957), la etnolingüística es como un puente que conecta dos mundos: el del idioma y el de la cultura. Hoy en día se suelen utilizar con relativa frecuencia los términos *etnolingüística*, *etnografía de la comunicación* y *antropología lingüística*. En consecuencia, es lícito plantearse cuál de los tres es el más adecuado para referirse a la influencia que ejerce la cultura en una determinada lengua. Según Giorgio Cardona (2006), los tres términos son sinónimos y se diferencian únicamente por el hecho de que *antropología lingüística* (*anthropological linguistics*) se empezó a utilizar en Estados Unidos a mediados de los años cincuenta del siglo XX, *etnografía de la comunicación* (*ethnography of speaking*) fue introducida por Dell Hymes (1964) también en EE.UU. y *etnolingüística* ganó terreno mayoritariamente en Europa. El propio Cardona aboga por utilizar este último, principalmente por su posible función de adjetivo y por ser una sola palabra. Por esa misma razón, creemos que *etnolingüística* es el término

¹ Este artículo es el resultado de mi tesis doctoral *Las pirámides etnolingüísticas, estudio contrastivo entre el español y el italiano basado en corpus*, defendida en la Universidad de Barcelona el 12 de diciembre de 2017. Disponible en: <https://www.tdx.cat/handle/10803/459071#page=1>.

menos ambiguo de los tres y el que mejor refleja, sin remitirnos a la antropología y a la comunicación, la idea de relación entre lengua y cultura que constituye el eje vertebrador de este artículo. Así pues, con el presente trabajo se aspira a elaborar un estudio sistemático en el campo de la etnolingüística, analizando cómo influye la cultura en la lengua española a partir de la representación visual de las pirámides etnolingüísticas, una nueva herramienta que facilita la visualización de datos. En este mundo dominado por la tecnología, la visualización es sin duda un instrumento muy efectivo, como expone Robert Spence (2007) en su libro *Information Visualization: Design for Interaction*, dedicado enteramente a cómo se puede optimizar el diseño de cara a una mejor visualización de los datos, y también como señala Colin Ware (2004), quien hace particular hincapié en la percepción que despierta el diseño. Por este motivo, hemos decidido apostar por un sistema de visualización de datos totalmente nuevo que hemos denominado *pirámides etnolingüísticas* a partir de un proceso de 4 fases que se detallará en los próximos apartados.

1. MARCO TEÓRICO Y PREGUNTAS DE INVESTIGACIÓN

El marco teórico de este artículo bebe de estudios, ensayos y tesis doctorales sobre la etnolingüística, la neurociencia y la semántica; pero, por encima de todo, aborda la relación lengua-cultura para tratar de entender hasta qué punto determinados factores pueden influir en la manera de hablar de un nativo (en particular, en el uso recurrente de determinadas expresiones), de ahí que se haya escogido la etnolingüística como marco teórico principal de esta investigación. Asimismo, desde el punto de vista metodológico, el estudio no hubiera sido posible sin la aportación de los corpus, que sirven para corroborar el uso real de una lengua, saber cuáles son las expresiones fraseológicas más frecuentes y conocer el verdadero alcance de una determinada palabra. Como apuntamos en la introducción la idea de este artículo surgió a partir de un interés por la etnolingüística, una rama de la lingüística que se centra en la influencia de la cultura en la lengua y que, hoy en día, todavía no ha recibido la suficiente atención por parte de la comunidad investigadora.

El prefijo *etno-* (del griego *ἔθνος* *ethnos*, ‘pueblo’ o ‘nación’) alude a una población humana en la cual los miembros se identifican entre ellos, normalmente con base en una real o presunta genealogía y ascendencia común, u otros lazos históricos. Las etnias están también normalmente unidas por unas prácticas culturales, de comportamiento, lingüísticas, o religiosas comunes, de ahí que se haya utilizado dicho prefijo para referirse a esta rama de la lingüística.

Hoy en día se suelen utilizar con relativa frecuencia los términos *etnolingüística*, *etnografía de la comunicación* y *antropología lingüística*. En consecuencia, antes de empezar este artículo nos hemos planteado cuál de los tres es el más adecuado para referirse a la influencia que ejerce la cultura en una determinada lengua, sobre todo porque representa el marco teórico de la investigación. Según Giorgio Cardona, los tres términos son sinónimos y se diferencian únicamente por el hecho de que *antropología lingüística* (*anthropological linguistics*) se empezó a utilizar en Estados Unidos a mediados de los años cincuenta del siglo XX, *etnografía de la comunicación* (*ethnography of speaking*) fue introducida por Dell Hymes (1964) también en EE.UU. y *etnolingüística* ganó terreno mayoritariamente en Europa (*Introduzione all’etnolingüística*, 9). El propio Cardona aboga por utilizar este último, principalmente por su posible función de adjetivo y por ser una sola palabra. Para esta tesis también creemos que *etnolingüística* es el término menos ambiguo de los tres y el que mejor refleja, sin remitirnos a la antropología y a la comunicación, la idea de relación entre lengua y cultura que constituye el eje vertebrador de este trabajo.

Finalmente, Eugenio Coseriu (1981) sostiene que hay dos grandes sentidos que proporcionan la justificación racional de la etnolingüística. En primer lugar, el lenguaje refleja la cultura no lingüística; en segundo lugar, no se habla solo con el lenguaje como tal (competencia lingüística), sino también con los saberes, las ideas y las creencias (competencia extralingüística) que influyen sobre la propia expresión y la determinan de alguna manera.

El ser humano siempre ha tenido la necesidad primaria de comunicarse con sus semejantes y ha encontrado en el lenguaje su mejor y más elaborado instrumento de comunicación, que destaca, según Casado Velarde (1988) por ser racional (es un producto de nuestro cerebro y de nuestra facultad de pensar) y universal (todos tenemos la capacidad de articular nuestro pensamiento elaborando signos). Además, el lenguaje es innato (todos nacemos con la capacidad de comunicarnos), aprendido (vamos conociendo los signos de la comunidad lingüística en la que nos criamos) y convencional (los signos adaptados por una determinada comunidad son aceptados por todos sus miembros). Casado Velarde habla de cinco rasgos relacionados con el carácter universal del lenguaje: *la semánticidad, la creatividad, la alteridad, la materialidad y la historicidad* (29-31). En primer lugar, el lenguaje siempre significa algo, por eso hablamos de semánticidad, y todo hecho material que no lleve asociado un contenido significativo no se considera lenguaje. Además, el lenguaje es una producción genuina e innata del ser humano, y por consiguiente está sujeto a la creatividad del individuo, quien lo va moldeando en función de sus necesidades y bajo la influencia de la cultura, la ideología y la cosmovisión. Pero no vivimos solos, y para transmitir nuestras ideas y nuestros pensamientos necesitamos al *otro*, por eso la alteridad es quizá uno de los rasgos más indispensables del lenguaje, y se ve plasmado en el diálogo. Con el concepto de materialidad se hace referencia a que el lenguaje se presenta en el mundo como un hecho que podemos percibir con algún sentido, por ejemplo el oído o la vista, y se realiza en formas históricamente determinadas que llamamos *lenguas*, de ahí el último rasgo: la historicidad. En las pirámides etnolingüísticas se prestará particular atención a los rasgos de creatividad e historicidad, indispensables para analizar la influencia de la cultura en la lengua.

Con respecto a la creatividad, es necesario subrayar la labor y las teorías del lingüista alemán Karl Vossler (1941), considerado como el padre del idealismo lingüístico. Fue influido en gran parte por el italiano Benedetto Croce, quien defendía los aspectos creativos del lenguaje², y también por Von Humboldt (1999), quien consideraba el lenguaje como una actividad espiritual del ser humano. A raíz de esto, Vossler (1941) sostenía que toda lengua cambia en función de las necesidades de una comunidad y le atribuyó al hablante el papel de artista creador. Para este autor, muchos fenómenos lingüísticos derivan de causas culturales, y pretendió demostrarlo en sus obras *Algunos caracteres de la cultura española* (1941) y *Cultura y lengua de Francia* (1955). Creía que el espíritu de un pueblo, que a menudo se ve plasmado en las obras literarias, es la causa de los cambios lingüísticos y refleja la mentalidad de un pueblo. En su libro *Cultura y lengua de Francia*, el romanista alemán va más allá y establece sorprendentes asociaciones entre rasgos lingüísticos y culturales, como la que relaciona el artículo partitivo en francés con el espíritu práctico y mercantil de los franceses, lo que le valió numerosas críticas. Estas ideas están vinculadas con el concepto de *cosmovisión*, ese conjunto de representaciones mentales compartidas por un grupo social que pretende explicar la totalidad del universo desde un punto de vista social y natural. La cosmovisión es más amplia que la ideología por lo que a contenidos se refiere, pues sus proposiciones no solo abarcan la realidad social del hombre, sino también el universo físico. A diferencia de la cultura, que engloba todo aquel contenido cognitivo y valorativo, la cosmovisión es compartida por grupos sociales

² Concebido como *enérgeia*, término griego utilizado por Aristóteles para referirse a algo que está actuando (Jaeger, 1946).

determinados en defensa de sus intereses y es mucho más arbitraria. Esto tiene un gran impacto en el lenguaje, y muchas veces una misma expresión esconde más información de la que expresa llana y superficialmente su significado, y aprender a leer entre líneas y a detectar la posible ideología o cosmovisión del *otro* nos puede ayudar a mejorar nuestra comunicación y a hacer que nos entendamos mejor en un contexto empático de respeto mutuo.

Por último, cierra el marco teórico un aspecto de neurociencia relacionado con la empatía y la imitación, lo que nos sirve para explicar por qué el ser humano reproduce determinadas expresiones y por qué algunas de ellas triunfan en una comunidad lingüística y en otras no. Todo proceso imitativo parte primero de un proceso empático, y eso es posible gracias a la función de las neuronas espejo, descubiertas en 1991, en Parma, por el equipo de Giacomo Rizzolatti. Estos científicos, tras numerosas pruebas, decidieron colocar electrodos en la corteza frontal inferior de un macaco con el fin de estudiar las neuronas especializadas en el control de los movimientos de la mano al realizar una determinada acción. Durante cada experimento, registraban la actividad en el cerebro del simio mientras le facilitaban tomar trozos de alimento, de manera que los investigadores pudieran medir la respuesta de la neurona a tales movimientos. Lo curioso fue que algunas de las neuronas del mono reaccionaron incluso cuando el animal no se había movido, y eso demostró la existencia de unas neuronas que reconocían una acción y la reproducían en la mente. Estas células están ubicadas en una zona del cerebro conocida como F5, situada en la corteza premotora del cerebro, en la denominada Área de Broca, responsable de la producción motora del lenguaje y también de la comprensión sintáctica de la oración. El hecho de que las neuronas espejo se encuentren justamente en el Área de Broca explica, por ejemplo, la tendencia de los seres humanos a repetir estructuras gramaticales fijas o a reproducir muletillas que escuchamos en nuestro entorno. En palabras del científico Vilayanur Ramachandran, “el descubrimiento de las neuronas espejo ha sido de tal envergadura que podría hacer por la neurociencia lo mismo que el descubrimiento del ADN hizo por la biología”³ (Ramachandran, 2000: 29).

Algunos estudios realizados por Giacomo Rizzolatti (2006) y Marco Iacoboni (2009), quienes estuvieron tomando registros en neuronas de pacientes humanos, apuntan a que las neuronas espejo constituyen el sustrato cerebral de la tendencia automática a imitar, permitiendo relacionar tanto las acciones propias como las ajenas y dotarlas de un significado en nuestra mente. Avanzando con las investigaciones, incluso se ha llegado a sugerir que las neuronas espejo serían el mecanismo neuronal básico para el desarrollo del lenguaje; en concreto, el neurocientífico computacional Michael Arbib (2005) demostró que la pantomima desempeña un papel fundamental en el avance evolutivo que va desde el sistema de neuronas espejo simple del mono hasta el sistema neuronal mucho más sofisticado del lenguaje humano. Somos conscientes porque tenemos lenguaje y esto nos permite pensar, relacionarnos con los otros, originar cultura y entender toda la abstracción relacionada con el propio lenguaje, como podría ser, por ejemplo, el significado de una metáfora. Imitamos lo que vemos y escuchamos, y este aprendizaje por imitación constituye una de las bases de la cultura. Pero resulta que, pese a reproducirlo en nuestra mente, a veces nos negamos a pronunciarlo. ¿Qué es lo que actúa como inhibidor lingüístico? Según Marco Iacoboni (2009), existen dos grandes tipos de neuronas espejo: las normales y las superneuronas espejo. Parece ser, en efecto, que en los sistemas cerebrales más complejos, como el de los seres humanos, existe una capa de neuronas en el lóbulo frontal que se encarga de modular y controlar la actividad de las neuronas espejo “simples”. Lo interesante es que muchas veces ese control y esa modulación están fuertemente influenciados por la cosmovisión del hablante.

³ La traducción es propia.

Gracias a la visualización de las pirámides etnolingüísticas, en el presente trabajo se tratará de demostrar que muchas expresiones fraseológicas son aceptadas por una comunidad lingüística de acuerdo con la cosmovisión y el sistema de creencias de sus hablantes, y las neuronas espejo desempeñan un papel fundamental en su aceptación porque el uso de una determinada expresión y el hecho de que se asiente en la lengua depende mucho de la imitación. De alguna manera, primero las aceptamos y luego las añadimos a nuestro vocabulario activo. Pensamos que todo el proceso de imitación no se debe solo a una cuestión social, como se ha defendido en muchas ocasiones, sino que influye de manera determinante el aspecto biológico del cerebro. Dicho de otra manera, estamos programados para imitar y reproducir todo lo que dictaminen nuestras neuronas espejo, y su actividad cerebral es directamente proporcional al tiempo que invertimos para relacionarnos con los demás.

A partir de estos conceptos, se han formulado las dos siguientes preguntas de investigación:

1. ¿Cómo se podría utilizar la etnolingüística para analizar la influencia de la cultura en la lengua española?
2. ¿Cómo se manejaría la información cuantitativa y cualitativa para visualizar los resultados obtenidos, consiguiendo una serie de expresiones fraseológicas a partir de un listado de palabras?

Para dar respuesta a la primera pregunta, se ha diseñado un proceso de creación que se ajusta a cuatro fases bien delimitadas: *brainstorming*, obtención del listado de trabajo, búsqueda en los corpus y ordenación dentro de la pirámide.

Para contextualizar la segunda pregunta, en cambio, se optó por utilizar una metodología mixta basada en la triangulación de datos y siguiendo el modelo de Zoltan Dörnyei (2007), quien propone combinar métodos cuantitativos y cualitativos para luego aplicarlos con un plan de acción. Para que este tipo de metodología resulte efectiva, es fundamental atenerse al concepto de *investigación-acción*, sobre todo para tener acceso a la información cuantitativa (González-Lloret, 2014). En este caso concreto, el primer paso de la triangulación de datos empieza con la realización de un *brainstorming* para obtener el mayor número posible de palabras relacionadas con dos ámbitos culturales que se han escogido por su marcada influencia en la cultura española: el fútbol y los toros. Cabe señalar que se ha decidido utilizar el *brainstorming* porque partimos de la hipótesis de que necesitamos primero un listado de palabras que nos permita realizar una búsqueda posterior en los corpus y llegar así a las expresiones pluriverbales.

Tras el *brainstorming*, se suele disponer de información cuantitativa para empezar a trabajar, pero normalmente muchas palabras apenas tienen relación con el ámbito cultural en cuestión y su aparición se debe a factores relacionados con la cosmovisión de los participantes y su capacidad de asociar ideas afines. Aquí es donde cobra importancia la metodología mixta y la combinación de los enfoques cuantitativos y cualitativos. Así pues, lo primero que se llevará a cabo es una criba de los listados de palabras (fase 2 del proceso de creación de las pirámides) para descartar, en base a unos criterios fijados previamente, todas aquellas palabras que no guarden una relación directa con el ámbito cultural seleccionado. El segundo paso de la triangulación consiste en apoyarse en los corpus lingüísticos para poder buscar todas las palabras (fase 3 del proceso de creación) y pasar de las ocurrencias a las expresiones pluriverbales, lo que representaría un puente entre la información cuantitativa y la cualitativa. La triangulación de datos se cierra con la visualización de todas las expresiones pluriverbales que aparecen tras la búsqueda en los corpus y su clasificación dentro de las pirámides ateniéndose a dos criterios principales: el de frecuencia y el de rentabilidad lingüística. Estos criterios son fundamentales para ordenar tanto la información cuantitativa como cualitativa y ofrecer una clara visualización de datos.

Estas dos preguntas de partida se han materializado en dos grandes objetivos que servirán de guía para el presente artículo. El primer gran objetivo (A) consistirá en encontrar el mayor número posible de expresiones fraseológicas relacionadas con el ámbito cultural en cuestión, ya que se parte de la hipótesis de que es en las expresiones donde más se puede apreciar la influencia de la cultura en la lengua. Para ello, se necesitará primero tener un listado de palabras con el que trabajar en los corpus, y por eso se realizará un *brainstorming* con hablantes nativos. El proceso, que representará la fase 1 de creación de las pirámides etnolingüísticas, consistirá en reunir a cinco participantes y entregarles tres hojas en blanco. Cada hoja será un ámbito cultural y se dispondrá de veinte minutos para hacer una tormenta de ideas y escribir todas las palabras (evitando expresiones pluriverbales) que vengan a la mente y que estén relacionadas, de forma directa o indirecta, con el fútbol y los toros. Una vez hecho esto, se someterá cada listado a unos criterios de criba (fase 2 de creación de las pirámides etnolingüísticas) para quedarnos solo con aquellas palabras que guarden un vínculo directo con el ámbito cultural seleccionado y que formarán parte de los listados definitivos. El tercer paso consistirá en introducir todas las palabras de cada listado en los corpus (fase 3 del proceso de creación de las pirámides etnolingüísticas) con el fin de encontrar el mayor número posible de expresiones pluriverbales.

Para poder presentar mejor las expresiones fraseológicas encontradas en los corpus (fase 3), hemos fijado un segundo objetivo (B) que consistirá en ordenar y visualizar toda esta información cuantitativa y cualitativa dentro de unas pirámides para que luego sea posible realizar un primer análisis etnolingüístico. Así pues, se establecerán dos criterios fundamentales para ordenar las expresiones dentro de las pirámides (fase 4 del proceso): el de frecuencia y el de rentabilidad comunicativa. El primero tiene que ver con el número de ocurrencias encontradas tras la búsqueda en el corpus (fase 3), mientras que el segundo se aplicará en base a las marcas que aparecen en los diccionarios para determinar si una expresión se considera coloquial o poco apropiada en según qué contexto. Así pues, todo el proceso de creación quedaría reflejado de la siguiente manera:

1. *Brainstorming* (fase 1) – Resultado: listado
2. Criba de los resultados (fase 2) – Resultado: listado definitivo
3. Búsqueda en el corpus de unidades pluriverbales (fase 3) – Resultado: expresiones fraseológicas y frecuencia.
4. Selección de las expresiones y visualización de datos (fase 4) – Resultado: pirámide etnolingüística.

2. METODOLOGÍA

Como el interés de esta investigación se centra en las expresiones fraseológicas, ya que partimos de la hipótesis de que son vehículos de la cultura (véase apartado 2), al principio se barajó la posibilidad de apoyarse en los diccionarios fraseológicos; pero luego dicha idea se descartó y se prefirió utilizar el Data-driven, un método que ofrece un enfoque mucho más inductivo con el que se pretende formular y corroborar una teoría a partir de la recopilación de datos. Según Hilary Mason y Dhanurjay “DJ” Patil (2015), el análisis de los datos es fundamental en todo proceso investigador, y fomentar una cultura basada en los datos va a ser una de las claves del éxito en el siglo XXI, de ahí que el Data-driven constituya uno de los puntos de partida de este estudio.

El manejo de la información cualitativa y cuantitativa a menudo puede resultar complicado; y es por ello que se ha utilizado una metodología mixta basada en la triangulación de datos y siguiendo el modelo de Zoltan Dörnyei (2007), quien propone combinar métodos cuantitativos y cualitativos para luego aplicarlos con un plan de acción. Evidentemente, los métodos mixtos

implican una serie de suposiciones que guían hacia la recolección de datos y su posterior análisis con enfoques cualitativos y cuantitativos en combinación para dar respuesta a los objetivos propuestos. Para que este tipo de metodología resulte efectiva, es fundamental atenerse al concepto de *investigación-acción*, sobre todo para tener acceso a la información cuantitativa (González-Lloret, 2014). En nuestro caso, el primer paso de la triangulación de datos empieza con la realización de un *brainstorming* para obtener el mayor número posible de palabras relacionadas con los dos ámbitos culturales seleccionados: el fútbol y los toros. Para ello, es preciso encontrar unos informantes nativos españoles para someterlos a la prueba y conseguir una gran cantidad de palabras. Tras el *brainstorming*, se suele disponer de información cuantitativa para empezar a trabajar, pero normalmente muchas palabras apenas tienen relación con el ámbito cultural en cuestión y su aparición se debe a factores relacionados con la cosmovisión de los participantes y su capacidad de asociar ideas afines. Aquí es donde cobra importancia la metodología mixta y la combinación del enfoque cuantitativo y cualitativo. Así pues, lo primero que se llevará a cabo es una criba de los listados de palabras (fase 2 del proceso de creación de las pirámides) para descartar, en base a unos criterios fijados previamente, todas aquellas palabras que no guarden una relación directa con el ámbito cultural seleccionado. El segundo paso de la triangulación consiste en apoyarse en los corpus lingüísticos para poder buscar todas las palabras (fase 3 del proceso de creación), y pasar de las ocurrencias a las expresiones pluriverbales, lo que representaría un puente entre la información cuantitativa y la cualitativa.

Existen muchos tipos de corpus disponibles en internet que nos pueden ser de gran ayuda para obtener información lingüística relevante, pero no se puede perder de vista el concepto de *representatividad*, fundamental para poder extraer conclusiones fiables a partir de los datos estadísticos (Data-driven). Por eso, aunque el *CREA* (*Corpus de Referencia del Español Actual*), un banco de datos que actualmente recoge textos en lengua española desde 1975 hasta 2004, parecía *a priori* el motor de búsqueda ideal, precisamente por aspirar a ser “de referencia”, al final se optó por el *CORPES XXI*, disponible actualmente en la red en versión beta y donde el reparto de pesos entre España y América pasa a ser del 30/70 (en lugar del 50/50 del *CREA*), con lo que resulta más representativo de la distribución de los hablantes. Según la teoría de Joan Torruella y Joaquim Llisterra (1999), no siempre un corpus será representativo solo por ser mayor y contener muchas palabras.

Para la elección del corpus como motor de búsqueda se tuvo también en consideración el problema de la lematización; es decir, el hecho de que en el corpus se relacionen con el *lema* (por ejemplo, el verbo en infinitivo) todas sus *formas* (cada una de las posibilidades flexionadas de ese verbo). La falta de lematización dificulta bastante tanto la búsqueda en sí como las listas de frecuencia, ya que si por ejemplo buscamos el verbo *tener*, solo nos aparecerá esa forma de infinitivo y no todas sus formas flexionadas (*tengo, tenía, tuve*, etc.).

La triangulación de datos se cierra con la visualización de todas las expresiones pluriverbales que aparecen tras la búsqueda en el *CORPES* y su clasificación dentro de las pirámides ateniéndose a dos criterios principales: el de frecuencia y el de rentabilidad lingüística. Estos criterios son fundamentales para ordenar tanto la información cuantitativa como cualitativa y ofrecer una visualización de datos clara (objetivo B) que facilitará el análisis etnolingüístico de las pirámides en español.

3. CREACIÓN DE LAS PIRÁMIDES ETNOLINGÜÍSTICAS

El concepto de *pirámide* surgió a partir de una doble necesidad: por un lado, jerarquizar toda la información cuantitativa y cualitativa; por otro, ofrecer una imagen visual para que dicha

información esté más ordenada y resulte más clara. Además, pese a la jerarquización de las expresiones, las *pirámides etnolingüísticas* son una propuesta muy flexible, ya que pueden variar constantemente en función de los parámetros cronológicos, diatópicos e incluso diastráticos que se les apliquen. Siguiendo las directrices que se expondrán a continuación, veremos que es posible realizar cualquier tipo de pirámide etnolingüística en la lengua seleccionada, y luego hacer lo mismo con otra lengua en caso de buscar una comparación. La única condición fundamental es que se pueda acceder al menos a un corpus de referencia de la lengua en cuestión para obtener un resultado mucho más preciso y riguroso.

Dependiendo del tipo de pirámide que se quiera crear, el proceso puede ser más o menos laborioso, aunque se ajusta a cuatro fases bien delimitadas que se detallarán a continuación: *brainstorming*, obtención del listado de trabajo, búsqueda en los corpus y ordenación dentro de la pirámide. Todo el proceso de creación quedaría reflejado tal como se indica en los siguientes subapartados.

3.1. Brainstorming. La técnica del *brainstorming* fue propuesta por Alex Faickney Osborn en su libro *Applied imagination*, publicado por primera vez en 1953 y traducido al castellano en 1957. Se trata de romper las limitaciones habituales del pensamiento y producir un conjunto de ideas entre las que poder escoger. El *brainstorming* es útil para generar una gran cantidad de ideas y el mismo Osborn propuso cuatro reglas básicas para ser lo más efectivo posible: en primer lugar, hay que suspender el juicio para eliminar toda crítica consciente, ya que según el autor es muy complicado crear y juzgar al mismo tiempo y es mejor dejar que nuestro pensamiento fluya sin interferencias. Esta primera regla, en realidad, es quizá la más complicada de las cuatro, pues estamos tan acostumbrados a analizar todo lo que hacemos que no resulta fácil dejarse llevar. Esto nos conduce a la segunda regla, que consiste en pensar libremente y en no descartar ninguna idea *a priori*, por muy inimaginable, confusa e ilógica que parezca. La tercera regla tiene que ver con la cantidad, puesto que uno de los objetivos principales es concentrarse para sacar el mayor número posible de ideas, aunque luego se descarten. Osborn (1957) sitúa el objetivo ideal entre 50 y 100 ideas, y el objetivo mínimo por encima de las 25 ideas. Por último, hay que reunir todas las ideas buscando el denominado efecto multiplicador; es decir, buscar la combinación de todas las ideas y tratar de mejorarlas. Por eso la cantidad es importante y es directamente proporcional al efecto multiplicador: cuantas más ideas tengamos, más posibilidades tenemos de pulirlas y de dar con la solución más idónea para nuestro propósito. En cuanto al grupo, se estableció que no puede superar los doce miembros, situándose entre cuatro y siete el número ideal, aunque lógicamente el *brainstorming* también se puede realizar de manera individual. Como herramientas solo necesitamos un bolígrafo y una hoja en blanco, idealmente de tamaño DIN A3 que DIN A4, ya que cuanto más espacio en blanco tengamos, mayor será el alcance de nuestra creatividad. Además de todo esto, es fundamental que el ambiente del grupo sea ameno y que los participantes estén relajados y sin presión. Todo tiene que fluir de manera natural y despreocupada, y es importante además que haya un moderador para controlar el proceso, un papel que en el aula puede desempeñar el profesor. Respecto a la duración de la sesión, Osborn (1957) recomendó que no superara los treinta minutos para no correr el riesgo de que los participantes perdieran el interés, razón por la que nos inclinamos a realizar sesiones incluso más breves, de veinte minutos máximo, dedicando luego entre diez y quince minutos para evaluar todas las ideas, esta vez de manera lógica y analítica.

Así pues, en 2016 realizamos el primer experimento en español con cuatro profesores del departamento de Filología Hispánica de la Universitat de Barcelona: el Dr. Antonio Torres Torres, la Dra. Paloma Arroyo Vega, el Dr. Juan Pablo García-Borrón y la Dra. Mar Cruz Piñol, a quienes

agradecemos su paciencia y disponibilidad⁴. Se optó por escoger a cinco participantes españoles al decidirse limitar la investigación a la variedad diatópica de España, y al ser recomendable que los participantes fuesen personas nativas del país objeto de análisis. El *brainstorming* duró alrededor de veinte minutos y cada profesor lo realizó de manera individual y aislada en su despacho. Se ha de subrayar este detalle porque esta técnica normalmente se implementa en grupo; sin embargo, esta vez pensamos que sería mejor que cada profesor no se sintiese de alguna manera influenciado por la presencia de un colega de departamento. Se les entregaron a los participantes tres hojas en formato DIN-A3, una correspondiente al ámbito del fútbol y otra al de los toros. Como se trataba de dejar fluir las ideas, los participantes podían escribir libremente saltando de una hoja a otra o cubriendo tema por tema; la idea, básicamente, era que cada sujeto se sintiera lo más a gusto posible, libre de tensiones, esquemas o directrices, y se entregase por completo al proceso creativo. Respecto a la categoría gramatical, podían poner tanto sustantivos como adjetivos, verbos y adverbios (intentando evitar las unidades pluriverbales), ya que el objetivo principal del *brainstorming* es buscar el efecto multiplicador (sin limitaciones) para luego realizar la criba correspondiente.

Evidentemente, se nos abrieron muchas opciones para realizar el *brainstorming*, pudiendo elegir el número y perfil de los participantes, el lugar y la duración de la prueba. Por ello, esa misma tarde se llevó a cabo el mismo ejercicio también con otro informante, quien no forma parte del mundo de la filología, para ver si podía obtener información adicional y complementaria a los otros cuatro *brainstorming*. El resultado final fue satisfactorio y con cinco personas se consiguieron un gran número de palabras que se enumerarán a continuación por ámbito cultural y por orden alfabético.

3.1.1. Listado obtenido tras la Fase 1 para el ámbito del fútbol. *Apuestas, árbitro, Argentina, arquero, Atlético de Madrid, balompié, balón, banderas, banquillo, Barça, Bayern, Benítez, Boca Junior, bola, Borussia Dortmund, bota, botas, Brasil, Buffon, cabezazo, camiseta, campo, cantar, cánticos, capitán, celebrar, Celta, césped, champions, chut, chutar, club, colores, compañerismo, competencia, competición, copa, Copa del Rey, córner, correr, crack, Cruyff, culé, deporte, Deportivo, despejar, dinero, diversión, ejercicio, El Chiringuito, el Molinón, entradas, entrenador, entrenamiento, entrevistas, equipo, escuadra, estadio, estrellas mediáticas, Eurocopa, éxito, falta, fichajes, Fiorentina, fisio, fondo, futbolista, gol, goleada, grada, himno, hincha, histeria colectiva, Iker Casillas, Inter, internet, Italia, juego, jugador, Juventus, larguero, Lazio, lesión, lesiones, liga, liguilla, linier, madridista, mediático, Messi, Mestalla, mi padre, Milán, Mourinho, multitud, mundial, niño, norte, países, palco, parar, partido, pase, pasión, patadas, patrocinador, Pelé, pelota, penalti, Piqué, portería, portero, prensa, presidente, prima, publicidad, público, puerta, quiniela, radio, Rajoy, Real Madrid, red, reglas, representantes, reventa, Riazor, rival, rivales, rivalidad, River Plate, Roma, Ronaldo, rotación, seguidor, Sergio Ramos, Sevilla, silbato, streaming, Suárez, sur, tarjeta amarilla, tarjeta roja, televisión, vestuario, violencia.*

3.1.2. Listado obtenido tras la Fase 1 para el ámbito de los toros. *Abanico, agua, albero, alternativa, animal, arena, arrastrar, arrastre, banda de música, banderilla, banderillas, banderillero, barrera, burladero, caballo, callejón, camisa, campo, capote, color, corbata, cornada, cortar, coso, crítica, crueldad, cuadrilla, cuerno, dedicatoria, dehesa, El Cordobés, El Guerra, entrada, espectáculo, estampitas, faena, faenar, Feria de Sevilla, fiesta, finca, fuerza, furgoneta, Goya, herida, indulto, La Monumental, Las Ventas, libertad, lidia, Lorca, luces, lucha, Maestranza, manolelina, mantilla, mi abuelo Pepe, miedo, montera, morlaco, muerte, mujeres,*

⁴ Se indican sus nombres porque contamos con su consentimiento para ello.

mulillas, orejas, pasado, paseillo, pasión, pasodoble, picador, Picasso, plaza, presidente, público, puerta grande, puntilla, pureza, rabo, rejoneador, reventa, rueda, sangre, sangría, Sevilla, sol, sombra, taleguilla, tendido, torear, torero, toro, tradición, traje de luces, valentía, vino, violencia.

3.2. Obtención del listado de trabajo. El proceso creativo del *brainstorming* refleja, en buena medida, la información guardada en el subconsciente que está sujeta a nuestro personal sistema de creencias. Observando las palabras obtenidas, se puede notar el enorme caudal de información que contiene; por ejemplo, resulta curioso que en los dos campos aparezca la palabra *violencia*, que está relacionada con los temas de forma indirecta y refleja más bien una visión subjetiva. Otro aspecto interesante del *brainstorming* es la concatenación de ideas, que puede darse en forma de nombres propios o de objetos afines entre sí; es decir que, por ejemplo, cuando a un participante se le ocurre la palabra *Mourinho*, es muy probable que en su cerebro aparezcan nombres de otros entrenadores o de equipos de fútbol concretos y que se sucedan durante unos segundos.

El siguiente paso, antes de proceder a buscar las expresiones pluriverbales en los corpus, será realizar una criba para obtener un listado más reducido con el que se trabajará en la fase 3. Dicho de otra manera, se trata de analizar consciente y racionalmente toda esa información cuantitativa fruto de la parte más irracional.

En primer lugar, salta a la vista la gran cantidad de nombres propios y topónimos que aparecen en los dos temas. Desde un punto de vista semántico, presentan mucha intensión y poca extensión, por lo que se han establecido los siguientes criterios para todos los nombres propios:

- 1) Se excluirán aquellos nombres propios de personas que no estén directamente relacionados con la temática, como Rajoy (fútbol); Goya, Lorca, mi abuelo Pepe y Picasso (toros).
- 2) Se excluirán los topónimos por su falta de extensión, como *Brasil, Italia, Roma y Sevilla* (fútbol y toros).
- 3) Se excluirán los nombres de los estadios y de las arenas (siempre por su falta de extensión), como *el Molinón, Mestalla, Riazor* (fútbol); *Feria de Sevilla, La Monumental, Las Ventas y Maestranza* (toros).
- 4) En el caso concreto del fútbol, se prescindirá de cualquier nombre de programa televisivo y de los nombres de clubes (falta de extensión), como *Atlético de Madrid, Barça, Bayern, Boca Junior, Borussia Dortmund, Celta, Deportivo, El Chiringuito, Fiorentina, Inter, Juventus, Lazio, Milán, Real Madrid, River Plate y Roma*.

Por lo tanto, nos decantaremos principalmente por los nombres propios que en algún momento hayan podido tener una ampliación de significado y se hayan asentado en la lengua.

Respecto a la criba de los sustantivos, verbos y adjetivos se ha decidido excluir todas aquellas palabras que sean demasiado genéricas y no presenten una relación directa con el ámbito cultural en cuestión. Por ello, se han establecido los siguientes criterios:

- 1) Con los sustantivos o nombre comunes, se excluirán aquellos que no tengan una relación directa con el ámbito cultural y cuya aparición en el *brainstorming* se deba a factores emotivos, personales, etc. Para el fútbol, se prescindirá de *colores, compañerismo, competencia, competición, dinero, diversión, ejercicio, entradas, entrevistas, éxito, fisio, fondo, himno, internet, multitud, niño, norte, países, pasión, patrocinador, prensa, presidente, prima, publicidad, público, puerta, radio, reglas, representantes, rotación, streaming, sur, televisión y violencia*. Por último, con los toros se eliminará del listado *agua, animal, banda de música, camisa, campo, color, corbata,*

crítica, crueldad, dedicatoria, entrada, fiesta, finca, furgoneta, libertad, miedo, muerte, mujeres, pasado, pasión, presidente, tradición, valentía, vino y violencia.

- 2) Respecto a los verbos, se aplicará un criterio muy similar al anterior, aunque vemos que su presencia en el *brainstorming* es mucho más limitada. Se ha decidido prescindir, por su falta de relación directa, de *cantar, celebrar y correr*, todos pertenecientes al ámbito del fútbol; en cambio, se optó por incluir el verbo *despejar* porque está muy vinculado al ámbito futbolístico.
- 3) En cuanto a los adjetivos, pasa algo similar a los verbos, ya que su presencia no es muy elevada. Una vez más, se seguirá el mismo criterio aplicado anteriormente y en este caso solo se prescindirá de *mediático* (fútbol).
- 4) Por último, vamos a ver cómo se tratarán los compuestos pluriverbales formados por un sustantivo y un adjetivo, como por ejemplo *tarjeta amarilla, tarjeta roja* (fútbol) y *puerta grande* (toros). Se ha decidido incluirlos por su relación directa con el ámbito cultural, ya que a menudo el adjetivo suele aportar una mayor intensidad al nombre común que acompaña y, en los casos mencionados, acota su significado, proporcionándole una mayor intensidad directamente relacionada con el tema cultural. De los compuestos pluriverbales aparecidos en el *brainstorming*, solo se excluirán *estrellas mediáticas e histeria colectiva* (fútbol) por no ajustarse a los criterios propuestos.

Dicho esto, y antes de procederse a analizar cuál sería el resultado de los listados tras aplicar la criba (fase 2), habría que algunos incisos respecto a los sustantivos o nombres comunes, puesto que representan el grueso del *brainstorming*. Por lo que se refiere al fútbol, la mayoría de los sustantivos se caracterizan por una gran extensión y podrían pertenecer a otros deportes, como es el caso de *apuestas, banderas, camiseta, campo, capitán* y un largo etcétera. La idea de incluirlas se debe al hecho de que constituyen gran parte del *brainstorming* del fútbol, y quitarlas sería muy limitante de cara a las búsquedas en los corpus. Está claro, por ejemplo, que un *seguidor* puede ser también de un equipo de baloncesto o de cualquier otro deporte, pero se quiere ver si en el corpus aparecen expresiones que contengan *seguidor* y que se refieran solamente al ámbito futbolístico, de ahí la decisión de incluirla. Por último, la criba de los toros ha sido un poco más sencilla y nos encontramos con muchas palabras técnicas y específicas del mundo taurino que no dejaban lugar a dudas. Solo cabría mencionar la inclusión de las palabras *herida y lucha* que, pese a su gran extensión, se consideró que podrían ser de ayuda en la búsqueda en los corpus (fase 3).

Tras aplicar estos criterios, se indica a continuación en los dos siguientes subapartados cómo quedarían las dos listas.

3.2.1. Listado definitivo tras la Fase 2 para el ámbito del fútbol. *Apuestas, árbitro, arquero, balompié, balón, banderas, banquillo, Benítez, bola, bota, botas, Buffon, cabezazo, camiseta, campo, cánticos, capitán, césped, champions, chut, chutar, club, copa, Copa del Rey, córner, crack, Cruyff, culé, deporte, despejar, entrenador, entrenamiento, equipo, escuadra, estadio, eurocopa, falta, fichajes, futbolista, gol, goleada, grada, hincha, Iker Casillas, juego, jugador, larguero, lesión, lesiones, liga, liguilla, linier, madridista, Messi, Mourinho, mundial, palco, parar, partido, pase, patadas, Pelé, pelota, penalti, Piqué, portería, portero, quiniela, red, reventa, rival, rivales, rivalidad, Ronaldo, seguidor, Sergio Ramos, silbato, Suárez, tarjeta amarilla, tarjeta roja, vestuario.*

3.2.2. Listado definitivo tras la Fase 2 para el ámbito de los toros. *Abanico, albero, arena, arrastrar, arrastre, banderilla, banderillas, banderillero, barrera, burladero, caballo, callejón,*

capote, cornada, cortar, coso, cuadrilla, cuerno, dehesa, El Cordobés, El Guerra, estampitas, faena, faenar, herida, indulto, lidia, luces, lucha, manoletina, mantilla, montera, morlaco, mulillas, orejas, paseillo, pasodoble, picador, plaza, puerta grande, puntilla, rabo, rejoneador, ruedo, sangre, sangría, sol, sombra, taleguilla, tendido, torear, torero, toro, traje de luces.

3.3. Búsqueda en el CORPES: de las ocurrencias a las expresiones. Para pasar a la fase 3, es necesario elegir un corpus lingüístico donde realizar la búsqueda de las palabras obtenidas tras la criba. En este sentido, y como ya comentamos en el apartado 3, al final se optó por el *CORPES XXI*, ya que el reparto de pesos entre España y América pasa a ser del 30/70 (en lugar del 50/50 del *CREA*), con lo que resulta más representativo de la distribución de los hablantes.

Con respecto a la elección del corpus como motor de búsqueda se tuvo también en consideración el problema de la lematización; es decir, el hecho de que en el corpus se relacionen con el *lema* (por ejemplo, el verbo en infinitivo) todas sus *formas* (cada una de las posibilidades flexionadas de ese verbo). Además, el *CORPES* presenta una interfaz relativamente sencilla e intuitiva, y lo único que hay que hacer para completar la fase 3 es acotar la búsqueda al español de España desplegando la pestaña de *subcorpus* y luego introducir una por una las palabras de los listados en la casilla del *lema*. Respecto al tiempo cronológico, no es necesario aplicar ninguna marca, pues la casi totalidad de los textos va desde 2001 a 2015, lo que agiliza todavía más la búsqueda.

Esta fase 3 es crucial en la elaboración de las pirámides etnolingüísticas y nos permite pasar de las ocurrencias a las expresiones. A continuación, veremos los resultados tras la búsqueda en el *CORPES* de todas las palabras de los dos listados.

3.3.1. Resultados en el ámbito del fútbol.

A

Apuestas: 2.661 casos en 1.914 documentos. En el corpus, aparecen ocurrencias tanto con el sustantivo como con el verbo *apostar*. Se han analizado una por una, pero no se ha encontrado ninguna expresión significativa relacionada con el fútbol.

Árbitro: 1.880 casos en 1.107 documentos. Ninguna expresión en el corpus.

Arquero: 101 casos en 74 documentos. Ninguna expresión relacionada con el fútbol en el corpus.

B

Balompie: 99 casos en 59 documentos. Ninguna expresión en el corpus.

Balón: 4.174 casos en 2.148 documentos. Destaca en el corpus la expresión *echar/lanzar/despejar*⁵ *balones fuera* (54 ocurrencias), utilizada cuando alguien trata de evitar una situación comprometida.

Bandera: 9.955 casos en 1.625 documentos⁶. Ninguna expresión en el corpus relacionada directamente con el fútbol.

Banquillo: 1.316 casos en 1.032 documentos. Solo encontramos la expresión *calentar banquillo* (3 ocurrencias), utilizada cuando alguien se queda mucho tiempo de suplente⁷. Cabe mencionar también la variante más coloquial *chupar banquillo* (RAE, 2014), de la que no se encontró ninguna ocurrencia en el *CORPES*.

⁵ Véase más adelante también la entrada *despejar*.

⁶ La mayoría de las expresiones que se encontraron están relacionadas con el ámbito marítimo y militar.

⁷ Se trata de una expresión muy acotada al ámbito futbolístico, pero se puede utilizar también con significado metafórico para indicar que alguien se queda en segundo plano, sin tener la posibilidad de actuar y demostrar su valor.

Benítez: 261 casos en 141 documentos buscando por *forma* al tratarse de un nombre propio. Aparecen en el corpus diferentes personas con este apellido, no solo el conocido entrenador de fútbol. De todas formas, no hay ninguna expresión ni casos de ampliación de significado.

Bola: 2.519 casos en 1.031 documentos. Se trata de una palabra polisémica con muchas expresiones idiomáticas, pero la mayoría de ellas pertenecen al español de América (RAE, 2014) y en España son menos frecuentes, como indica el resultado de la búsqueda en el *CORPES*, donde aparecen únicamente las expresiones *no dar pie con bola* (24 ocurrencias, con significado de ‘no acertar/no tener un buen día’) y *dar bola* [a alguien] (8 ocurrencias, con significado de ‘prestar atención’).

Bota: 1.477 casos en 729 documentos. Destaca en el corpus la expresión *colgar las botas* (19 ocurrencias) con significado principal de ‘dejar de jugar al fútbol’, pero utilizada metafóricamente con el significado de ‘dejar una actividad’. También se encontraron muchas ocurrencias de *ponerse alguien las botas* con significado de ‘comer mucho’ o ‘aprovecharse de algo’, que se descarta por no estar vinculada con el fútbol.

Buffon: 71 casos en 41 documentos buscando por *forma* al tratarse de un nombre propio. Ninguna expresión en el corpus.

C

Cabezazo: 297 casos en 248 documentos. Ninguna expresión en el corpus.

Camiseta: 2.207 casos en 1.177 documentos. Destaca en el corpus la expresión *sudar la camiseta* (9 ocurrencias) con significado de ‘esforzarse y poner mucho empeño en algo’.

Campo: 17.692 casos en 7.597 documentos. Se trata de una palabra muy común, lo que justifica el número de casos. Nos apoyamos también en el *Diccionario de la Real Academia Española* y se encontraron muchas expresiones (RAE, 2014), pero ninguna relacionada con el ámbito futbolístico.

Cántico: 383 casos en 213 documentos. Ninguna expresión relacionada directamente con el fútbol.

Capitán: 4.017 casos en 1.361 documentos. Ninguna expresión relacionada directamente con el fútbol.

Césped: 1.045 casos en 661 documentos. Ninguna expresión en el corpus.

Champions: 862 casos en 642 documentos buscando por *forma* al tratarse del nombre de una competición deportiva. Ninguna expresión en el corpus.

Chutar: 105 casos en 86 documentos. Se encontró en el corpus la expresión *ir [algo] que chuta* (5 ocurrencias), usada para indicar que algo va muy bien⁸.

Club: 9.015 casos en 3.522 documentos. Ninguna expresión en el corpus.

Copa: 8.577 casos en 3.202 documentos. Ninguna expresión en el corpus.

Copa del Rey: 772 casos en 531 documentos buscando por *forma* al tratarse del nombre de una competición deportiva. Ninguna expresión en el corpus.

Córner: 378 casos en 290 documentos. Ninguna expresión en el corpus.

Crack: 401 casos en 256 documentos. Con más de 200 ocurrencias, destaca en el corpus la expresión *ser un crack* para indicar que alguien es tan bueno en algo que literalmente ‘rompe moldes’, de ahí la onomatopeya con la palabra inglesa *crack*⁹.

⁸ El verbo proviene de la palabra inglesa *shoot*, que significa ‘tiro o disparo fuerte de pelota’.

⁹ Esta expresión se empezó a utilizar en el ámbito deportivo, especialmente en el fútbol, para designar jugadores que sobresalían sobre los demás, pero actualmente ha ampliado su significado y puede representar a cualquier persona que destaque por sus capacidades. Parece un claro ejemplo de la trascendencia que tiene el fútbol en la sociedad moderna, especialmente en España.

Cruyff: 93 casos en 62 documentos buscando por *forma* al tratarse de un nombre propio. Ninguna expresión en el corpus.

Culé: 221 casos en 171 documentos. Ninguna expresión en el corpus.

D

Deporte: 5.969 casos en 2.746 documentos. Se trata de un hiperónimo de *fútbol*, por lo que su extensión es muy grande. No se encontró ninguna expresión destacable en el corpus.

Despejar: 1.079 casos en 821 documentos. Aparece en el corpus la expresión *despejar balones fuera* (1 ocurrencia), como ya se ha visto en la entrada balón¹⁰.

E

Entrenador: 4.419 casos en 2.260 ocurrencias. Ninguna expresión relevante en el corpus.

Entrenamiento: 3.183 casos en 1.669 documentos. Ninguna expresión en el corpus.

Equipo: 33.180 casos en 13.120 documentos. Destaca en el corpus la expresión *ser un equipo* (65 ocurrencias), para indicar que hay que colaborar y trabajar juntos por el bien del grupo¹¹.

Escuadra: 440 casos en 345 documentos. Ninguna expresión en el corpus.

Estadio: 2.883 casos en 1.754 documentos. Se trata de una palabra polisémica con diferentes acepciones; sin embargo, la búsqueda en el corpus no desveló ninguna expresión relacionada con el fútbol.

Eurocopa: 640 casos en 345 documentos. Ninguna expresión en el corpus.

F

Falta: 14.571 casos en 7.571 documentos. Se trata de una palabra muy frecuente, pero no se encontró en el corpus ninguna expresión relacionada con el fútbol.

Fichaje: 1.169 casos en 750 documentos. Se encontró en el corpus la expresión *ser un buen fichaje* (6 ocurrencias) para indicar que una persona ha sido una incorporación positiva en un determinado ámbito. También aparece la interjección *¡menudo fichaje!* (3 ocurrencias), usada para delatar cierta decepción frente a la incorporación de alguien nuevo.

Futbolista: 2.771 casos en 1.726 documentos. Ninguna expresión en el corpus.

G

Gol: 7.923 casos en 3.193 documentos. Destaca en el corpus la expresión *meter/colar un gol [a alguien]* (6 ocurrencias) con el significado metafórico de que alguna persona es engañada por otra a causa de un descuido.

Goleada: 354 casos en 319 documentos. Destaca en el corpus la expresión *ganar por goleada* (60 ocurrencias), utilizada a menudo con significado metafórico cuando la victoria o la diferencia entre dos conceptos que se comparan es muy abultada.

Grada: 1.019 casos en 726 documentos. Ninguna expresión en el corpus.

H

Hincha: 197 casos en 134 documentos. Ninguna expresión en el corpus.

I

Iker Casillas: 160 casos en 134 documentos buscando directamente por *forma* al tratarse de un nombre propio. Ninguna expresión en el corpus.

¹⁰ Según el resultado de la búsqueda, son más frecuentes las versiones con *echar* y *lanzar*.

¹¹ La mayoría de las veces en el corpus aparece en primera persona plural del presente de indicativo (56 ocurrencias, somos un equipo).

J

Juego: 20.063 casos en 7.889 documentos. Es una palabra muy frecuente en la lengua, pero solo se encontró una expresión en el corpus relacionada con el fútbol y es *pillar* [a alguien] *en fuera de juego* (13 ocurrencias), para indicar que una persona no se encuentra en una posición correcta para contestar una determinada pregunta o resolver un problema¹².

Jugador: 14.323 casos en 5.378 documentos. Ninguna expresión en el corpus relacionada directamente con el fútbol.

L

Larguero: 276 casos en 239 documentos. Ninguna expresión en el corpus.

Lesión: 4.411 casos en 2.278 documentos. Ninguna expresión en el corpus.

Liga: 6.796 casos en 3.413 documentos. Se trata de otra palabra polisémica, y no se encontró en el corpus ninguna expresión relacionada con el fútbol a excepción de *haber liga* [en algún lugar] (1 ocurrencia), con el significado de que ‘en ese sitio también hacen lo mismo’.

Liguilla: 129 casos en 96 documentos. Ninguna expresión en el corpus.

Linier: 36 casos en 16 documentos. Ninguna expresión en el corpus.

M

Madridista: 886 casos en 586 documentos. Ninguna expresión en el corpus.

Messi: 738 casos en 347 documentos buscando por *forma* al tratarse de un nombre propio. No se encontró ningún caso de ampliación de significado en el corpus.

Mourinho: 654 casos en 269 documentos buscando directamente por *forma* al tratarse de un nombre propio. Ningún caso de ampliación de significado en el corpus.

Mundial: 12.161 casos en 6.275 documentos. Ninguna expresión relevante en el corpus.

P

Palco: 509 casos en 334 documentos. Ninguna expresión en el corpus.

Parar: el problema con esta palabra es que al buscar por *lema*, el *CORPES* en principio debería buscar todas las formas lematizadas del verbo, pero en este caso aparece también la preposición *para*, con lo que se obtendrían 501.629 casos en 65.734 documentos, algo que entorpecería demasiado nuestra búsqueda¹³.

Partido: 32.066 casos en 11.631 documentos. Relacionada con el fútbol, solo se encontró en el corpus la expresión *hacerse dueño del partido* (9 ocurrencias), para indicar que alguien se ha hecho con el control de una determinada situación.

Pase: 2.281 casos en 1.606 documentos¹⁴. Relacionada con el fútbol, solo se encontró en el corpus la expresión *hacer un pase a un periodista* (1 ocurrencia), con el significado de ‘pasar una noticia a la prensa, normalmente de forma exclusiva’.

Patadas: 865 casos en 510 documentos. Ninguna de las expresiones que aparecen en el corpus (a patadas/dar la patada a alguien) está relacionada directamente con el fútbol y se consideran demasiado generales como para incluirlas en la pirámide.

¹² También puede significar que alguien ha sido pillado con las manos en la masa.

¹³ Se trata de un error llamativo a la vez que inesperado, aunque desde el principio se ha tenido en cuenta el hecho de que fuera una versión beta y es posible que en la actualidad el problema se haya resuelto. Pese a estos pequeños imprevistos, seguimos confiando en nuestra decisión de utilizar la versión beta del *CORPES* en lugar del *Corpus del español* por tener una interfaz más sencilla y sobre todo por no necesitar registro. Para agilizar el proceso, se ha decidido buscar en el *CORPES* el verbo *parar* como *forma* (introduciendo las flexiones) y no se encontró ninguna expresión relacionada directamente con el fútbol.

¹⁴ Además de ser una palabra polisémica, aquí hay casos de homógrafos (*pase* sustantivo/*pase* forma flexionada del verbo *pasar*), por lo que la búsqueda ha resultado más complicada de lo previsto.

Pelé: 115 casos en 69 documentos buscando por *forma* al tratarse de un nombre propio. Curiosamente, aparecen 5 ocurrencias en el corpus con la expresión *ser el Pelé* [de algo], un claro caso de ampliación de significado para indicar, una vez más, que alguien es muy bueno en lo que hace. Sin embargo, hay que subrayar que todas aparecen en el mismo documento, un detalle que se tendrá en cuenta para ordenar esta palabra en la pirámide, ya que en este caso la frecuencia está influenciada por el idiolecto del autor¹⁵.

Pelota: 2.348 casos en 1.145 documentos¹⁶. Se encontraron en el corpus algunas expresiones idiomáticas como *devolver la pelota* [a alguien] (15 ocurrencias), con significado de ‘rechazar algo y devolvérselo al remitente’ y *dejar la pelota en el tejado* [de alguien/algo], (6 ocurrencias), con significado de que ‘el éxito de un negocio es todavía incierto’. No tienen una relación directa con el fútbol, sino con el juego de la pelota en general, por lo que no se van a tener en cuenta.

Penalti: 1.454 casos en 830 documentos. Se encontró en el corpus la expresión *casarse de penalti* (3 ocurrencias), con significado de ‘casarse como obligación por haberse quedado embarazada la mujer’¹⁷.

Piqué: 574 casos en 284 documentos buscando por *forma* al tratarse de un nombre propio¹⁸. No se encontró en el corpus ninguna ampliación de significado con el nombre del futbolista.

Portería: 1.085 casos en 815 documentos. Ninguna expresión en el corpus.

Portero: 2.217 casos en 1.223 documentos. Ninguna expresión en el corpus relacionada con el fútbol.

Q

Quiniela: 214 casos en 156 documentos. Ninguna expresión en el corpus.

R

Red: 17.389 casos en 6.536¹⁹. No se encontró ninguna expresión en el corpus relacionada con el fútbol.

Reventa: 59 casos en 42 documentos. Ninguna expresión en el corpus.

Rival: 5.455 casos en 3.657 documentos. No se encontró ninguna expresión en el corpus relacionada directamente con el fútbol.

Rivalidad: 509 casos en 326 documentos. Ninguna expresión en el corpus.

Ronaldo: 1.237 casos en 557 documentos buscando por *forma* al tratarse de un nombre propio. La mayoría de los casos se refieren al famoso jugador brasileño y los otros al más joven jugador portugués. Se encontró un solo caso de ampliación de significado en la expresión *ser el Ronaldo* [de algo] para indicar que alguien es muy bueno en lo que hace, como ya se vio con la entrada Pelé²⁰.

¹⁵ Según el índice en el *CORPES*, los ejemplos proceden del libro de Pradera, Máximo: *¿De qué me suena eso? Paseo informal por la música clásica*. Madrid: Santillana, 2005.

¹⁶ Palabra polisémica con diferentes acepciones, como indica el *DRAE* <http://dle.rae.es/?id=SPdVcPX|SPeziQe>.

¹⁷ Evidentemente, hoy en día esta expresión es mucho menos frecuente, como indica la búsqueda en el *CORPES*, por razones socio-culturales evidentes. Aventuramos la teoría de que vaya desapareciendo en los próximos años y deje de utilizarse.

¹⁸ En la búsqueda aparece también el sustantivo llano *pique*, ya que el *CORPES* no distingue las palabras con acento.

¹⁹ Se trata de otra palabra polisémica muy frecuente en la lengua, lo que justifica el gran número de casos.

²⁰ Según el *CORPES*, aparece en un artículo de prensa de José Comas titulado “Escándalo en Alemania por el paso de Schröder al sector privado”. *El País.com*. Madrid: elpais.com, 14/12/2005.

S

Seguidor: 2.190 casos en 1.398 documentos. Ninguna expresión en el corpus.

Sergio Ramos: 210 casos en 163 documentos buscando por *forma* al tratarse de un nombre propio²¹. Ningún caso de ampliación de significado en el corpus.

Silbato: 109 casos en 83 documentos. Ninguna expresión en el corpus.

Suárez: 443 casos en 197 documentos buscando por *forma* al tratarse de un nombre propio. No se encontró ningún caso de ampliación de significado relacionado con el futbolista.

T

Tarjeta amarilla: 129 casos en 123 documentos buscando por *forma* al tratarse de un compuesto pluriverbal. Se encontró en el corpus la expresión idiomática *sacar tarjeta amarilla* (2 ocurrencias), con el significado de ‘amonestar a alguien por algo considerado poco correcto’.

Tarjeta roja: 74 casos en 68 documentos buscando por *forma* al tratarse de un compuesto pluriverbal. Se encontró en el corpus la expresión *sacar/mostrar tarjeta roja* (8 ocurrencias), con significado de ‘aplicar un castigo severo por una falta grave’.

V

Vestuario: 1.987 casos en 1.299 documentos. Ninguna expresión en el corpus.

Tras la búsqueda en el corpus, se puede colegir que el fútbol tiene una discreta influencia en la lengua española, y eso se ve plasmado en un buen número de expresiones pluriverbales de uso común. Además, y pese a que se sigue la metodología Data-driven, se han consultado, *a posteriori*, el *Diccionario del fútbol*, de Antoni Nomdedeu Rull (2009), y la tesis de licenciatura *Terminología futbolística*, de Joaquim M.^a Puyal Ortiga (1972), actual comentarista de fútbol en la radio catalana; y se ha podido corroborar que la mayoría de las expresiones citadas coinciden con las que aparecen en la búsqueda realizada en el *CORPES*, lo que sugeriría que tanto el *brainstorming* (fase 1 del proceso de creación) como la búsqueda en el corpus (fase 3 del proceso de creación), han dado resultados hasta el momento satisfactorios para pasar de las ocurrencias a las expresiones fraseológicas (objetivo A).

3.3.2. Resultados en el ámbito de los toros

A

Abanico: 871 casos en 659 documentos. Ninguna expresión en el corpus.

Albero: 69 casos en 56 documentos. Ninguna expresión en el corpus.

Alternativa: 4.144 casos en 2.558 documentos. Se encontró en el corpus la expresión *tomar la alternativa* (2 ocurrencias), con significado metafórico de ‘titularse o graduarse en algo’). Es una expresión que procede del mundo taurino, ya que se refiere a la ceremonia en la que el novillero pasa a ser matador de toros, y para ello es autorizado a alternar en un festejo con el torero que lo apadrina.

Arena: 3.401 casos en 1.548 documentos. No se encontró en el corpus ninguna expresión relacionada con el mundo taurino.

Arrastrar: 4.164 casos en 1.888 documentos. Ninguna expresión en el corpus.

Arrastre: 259 casos en 188 documentos. Destaca en el corpus la expresión *estar para el arrastre* (21 ocurrencias), para referirse a alguien que se encuentra al límite de sus fuerzas y no puede más. La expresión viene de la imagen del toro que, tras ser estoqueado, es arrastrado por el ruedo hacia el desolladero.

²¹ Se trata de un apellido muy común y aparecen varios personajes públicos en el resultado de la búsqueda.

B

Banderilla: 134 casos en 87 documentos. Ninguna expresión en el corpus.

Banderillero: 63 casos en 36 documentos. Ninguna expresión en el corpus.

Barrera: 2.408 casos en 1.508 documentos. Ninguna expresión en el corpus relacionada con el mundo taurino.

Burladero: 45 casos en 37 documentos. Ninguna expresión en el corpus.

C

Caballo: 4.720 casos en 1.632 documentos. Ninguna expresión en el corpus relacionada con los toros²².

Callejón: 660 casos en 392 documentos. Ninguna expresión en el corpus.

Capote: 318 casos en 179 documentos. Se encontró en el corpus la expresión *echarle/ lanzarle un capote* [a alguien] (28 ocurrencias), con el significado de ‘ayudar a alguien’. **Cornada:** 94 casos en 63 documentos. Ninguna expresión en el corpus.

Cornada: 94 casos en 63 documentos. Ninguna expresión en el corpus.

Cortar: 7.583 casos en 2.820 documentos. Se encontró en el corpus la expresión *cortarse la coleta* (12 ocurrencias), expresión metafórica utilizada cuando alguien abandona una profesión, ya que era costumbre entre los toreros lucir una coletilla de pelo natural que se cortaban cuando se retiraban.

Coso: 168 casos en 107 documentos. Ninguna expresión en el corpus.

Cuadrilla: 371 casos en 223 documentos. Ninguna expresión en el corpus.

Cuerno: 726 casos en 403 documentos. Relacionada con los toros, se encontró en el corpus la expresión *coger/tomar/agarrar el toro por los cuernos* (51 ocurrencias), con significado de ‘enfrentarse resueltamente a una dificultad’²³.

D

Dehesa: 333 casos en 165 documentos. Ninguna expresión en el corpus.

E

El Cordobés: 137 casos en 57 documentos buscando por *forma* al tratarse de un nombre propio. Ningún caso de ampliación de significado en el corpus.

El Guerra: 11 casos en 4 documentos buscando por *forma* al tratarse de un nombre propio. Ningún caso de ampliación de significado en el corpus.

Estampitas: 34 casos en 29 documentos. Ninguna expresión en el corpus.

F

Faena: 862 casos en 492 documentos. Destaca la expresión en el corpus *rematar/terminar/acabar la faena* (73 ocurrencias, la mayoría con *rematar*), con significado de ‘llevar a cabo un trabajo de forma brillante’. Procede del momento final en el que el torero tiene que matar al toro para coronar su actuación.

²² Según el Diccionario de la Real Academia Española, existe una expresión procedente del mundo de la tauromaquia y es *sacar bien/limpio el caballo* para indicar que sale bien de una disputa o un peligro (<http://dle.rae.es/?id=6OP-HODx>). Sin embargo, en la búsqueda realizada en el CORPES no aparece ninguna ocurrencia de esta expresión, por lo que se ha decidido no tenerla en cuenta para la pirámide de los toros (Data-driven).

²³ Se trata de una expresión curiosa, ya que nadie durante una corrida de toros suele coger al animal por los cuernos, y es posible que su origen tenga que ver con la tradición de los *forcados* en Portugal, un grupo de ocho personas que al final de la corrida intentan inmovilizar el toro agarrándolo por los cuernos y resistiendo la embestida.

Faenar: 261 casos en 172 documentos. Se encontró en el corpus la expresión *salir/volver a faenar* (51 ocurrencias), pero no procede directamente del mundo taurino y su significado es más general²⁴, por lo que no se tendrá en cuenta de cara a la pirámide de los toros.

H

Herida: 3.909 casos en 2.117 documentos. Ninguna expresión en el corpus relacionada con los toros.

I

Indulto: 259 casos en 126 documentos. Ninguna expresión en el corpus.

L

Lidia: 270 casos en 56 documentos²⁵. Con referencia al mundo taurino, solo se encontró en el corpus la expresión *estar cabreado como un toro de lidia* (1 ocurrencia), para indicar un gran enfado.

Luces: 2.802 casos en 1.332 documentos buscando por forma para evitar el singular luz. Ninguna expresión en el corpus.

Lucha: 7.043 casos en 4.094 documentos. Ninguna expresión en el corpus relacionada con los toros.

M

Manoletina: 43 casos en 27 documentos. Ninguna expresión en el corpus.

Mantilla: 180 casos en 89 documentos. Ninguna expresión en el corpus.

Montera: 129 casos en 88 documentos. Destaca en el corpus la expresión *ponerse el mundo por montera* (28 ocurrencias), con significado de ‘llevar a cabo algo sin reparar en las consecuencias’.

Morlaco: 31 casos en 23 documentos. Ninguna expresión en el corpus.

Mulilla: 9 casos en 5 documentos. Ninguna expresión en el corpus.

O

Oreja: 3.028 casos en 1.174 documentos. Ninguna expresión en el corpus relacionada con los toros.

P

Paseíllo: 101 casos en 68 documentos. Ninguna expresión en el corpus.

Pasodoble: 170 casos en 102 documentos. Ninguna expresión en el corpus.

Picador: 198 casos en 74 documentos. Ninguna expresión en el corpus.

Plaza: 14.043 casos en 5.676 documentos. Ninguna expresión en el corpus relacionada con el ámbito de los toros.

Puerta Grande: 118 casos en 108 documentos buscando directamente por *forma* al tratarse de un nombre propio. Destaca en el corpus la expresión *salir por la puerta grande* (25 ocurrencias), utilizada para designar a quien obtiene un gran triunfo.

Puntilla: 539 casos en 394 documentos. Se encontraron en el corpus las expresiones de *puntillas*²⁶ (317 ocurrencias, con significado de ‘sigilosamente, pisando con la punta de los pies’) y *dar/poner la puntilla* (128 ocurrencias, con significado de ‘rematar/acabar algo’).

²⁴ Como se puede ver en el *DRAE* <http://dle.rae.es/?id=HUBDq0R>.

²⁵ La mayoría de las veces aparece el nombre propio femenino.

²⁶ Esta expresión no se tendrá en cuenta para la pirámide porque no procede directamente del mundo taurino.

R

Rabo: 532 casos en 259 documentos. Se encontró en el corpus la expresión *hasta el rabo todo es toro* (4 ocurrencias), para indicar que no se puede bajar la guardia hasta que haya finalizado la tarea.

Rejoneador: 69 casos en 37 documentos. Ninguna expresión en el corpus.

Ruedo: 432 casos en 248 documentos. Destaca en el corpus la expresión *salir al ruedo* (9 ocurrencias)²⁷, con significado metafórico de ‘entrar en escena’.

S

Sangre: 8.649 casos en 2.871 documentos. Ninguna expresión en el corpus relacionada directamente con el mundo taurino.

Sangría: 251 casos en 177 documentos. Ninguna expresión en el corpus.

Sol: 9.757 casos en 2.956 documentos. Destaca en el corpus la expresión *hacer un brindis al sol* (23 ocurrencias), utilizada a menudo para indicar que alguien hace falsas promesas para mantener contento al público.

Sombra: 4.348 casos en 1.757 documentos. Ninguna expresión en el corpus relacionada con los toros.

T

Taleguilla: 13 casos en 12 documentos. Ninguna expresión en el corpus.

Tendido: 285 casos en 197 documentos. Ninguna expresión en el corpus.

Torear²⁸: 471 casos en 221 documentos. No se encontró ninguna expresión pluriverbal en el corpus.

Torero: 1.535 casos en 580 documentos. Se encontró en el corpus la expresión *saltarse* [algo] *a la torera* (31 ocurrencias), con significado de ‘soslayar una obligación’.

Toro: 4.216 casos en 1.239 documentos²⁹. Destacan en el corpus las expresiones *coger/tomar/agarrar* el toro por los cuernos (51 ocurrencias), como ya se vio en la entrada *cuerno*, *a toro pasado* (36 ocurrencias), con significado de ‘tras dejar pasar la oportunidad’, la interjección *¡al toro!* (9 ocurrencias), utilizada como incitación para enfrentarse a un asunto, *pillar el toro* [a alguien] (5 ocurrencias), con significado de ‘no conseguir terminar algo en el plazo fijado’ o ‘llegar tarde’, *ver los toros desde la barrera* (3 ocurrencias), con significado de ‘presenciar algo sin correr peligro’ y *estar cabreado como un toro de lidia* (1 ocurrencia), como ya se vio en *lidia*³⁰.

Traje de luces: 26 casos en 21 documentos buscando directamente por *forma* al tratarse de un compuesto pluriverbal. Ninguna expresión en el corpus.

A primera vista, parece que este segundo listado ha proporcionado un buen número de expresiones pluriverbales, incluso más que con el fútbol, y eso tal vez se debe a que el mundo taurino es un ámbito cultural que se remonta al siglo XII, como señala Bartolomé Benassar en su obra *Historia de la tauromaquia: una sociedad del espectáculo* y que desde entonces ha sido muy

²⁷ En español también existen las variantes *tirarse al ruedo* y *lanzarse al ruedo*, pero no han aparecido en el corpus, por lo que no se tendrán en cuenta (Data-driven).

²⁸ El mismo verbo procede del ámbito taurino y puede ser utilizado metafóricamente con el significado de lidiar con alguien o con alguna dificultad. No se tendrá en cuenta para la pirámide de los toros al no tratarse de una expresión pluriverbal.

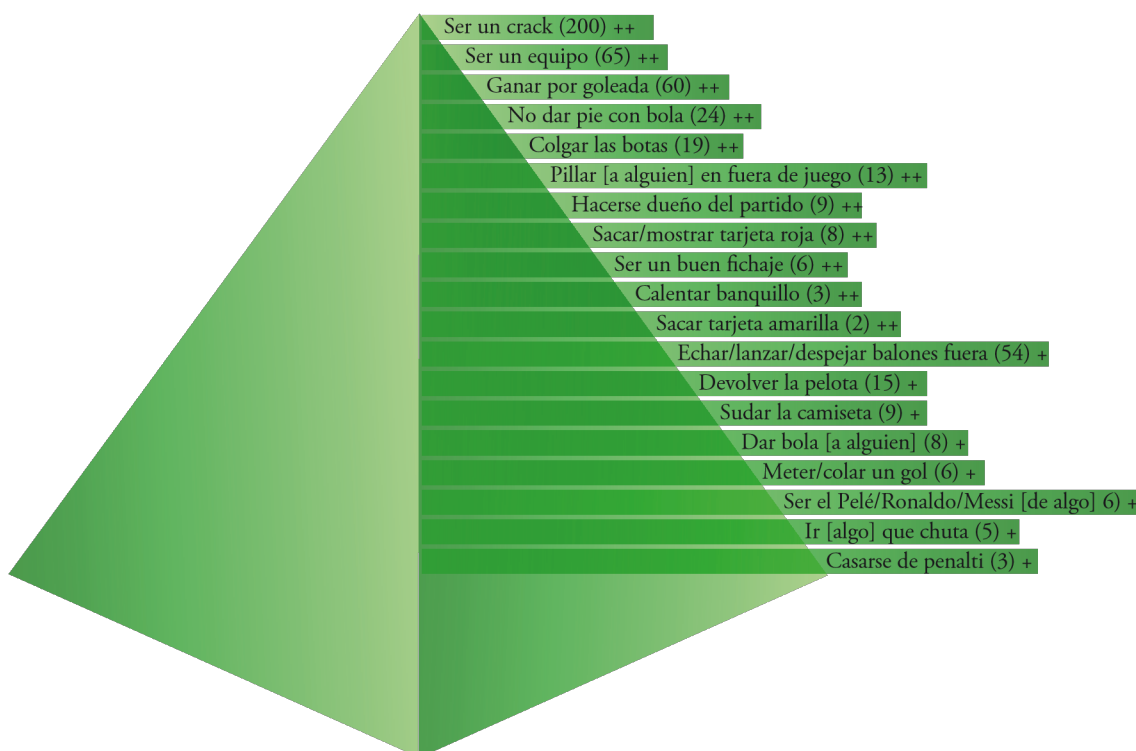
²⁹ Como cabía suponer desde un principio, se trata de la palabra más importante del listado y es la que nos brinda más expresiones.

³⁰ El Diccionario de la Real Academia Española muestra varias expresiones más (<http://dle.rae.es/?id=a6ZyZYZ|a6aLYnG>), pero estas son las únicas que se han encontrado en el CORPES y que por tanto se tendrán en cuenta en la pirámide española de los toros (Data-driven).

popular en España, con lo que su impacto en la lengua ha sido mayor. Al final de la búsqueda, se han cotejado estos resultados con los que aparecen en algunos estudios sobre el lenguaje taurino, como el ensayo *¡Derecho al toro! El lenguaje taurino y su influencia en lo cotidiano* de Carlos Abella (1996) y el interesante artículo de Francisco Reus Boyd-Swan (2004) titulado *El léxico taurino en la vida cotidiana*. Asimismo, la gran mayoría de las expresiones coinciden con las que se encontraron en esta fase 3, por lo que se considera que tanto el *brainstorming* como la búsqueda en el corpus, primeros dos pasos de la triangulación mixta, han servido para alcanzar el primer objetivo (de las ocurrencias a las expresiones). A continuación, se verá cómo manejar y ordenar todas estas expresiones en las pirámides (fase 4 del proceso de creación) para ofrecer una visualización de toda la información cuantitativa (tercer paso de la triangulación metodológica mixta) en base a los criterios de frecuencia y rentabilidad comunicativa.

3.4. Ordenación y visualización de los datos. En primer lugar, las expresiones se ordenarán según el número de ocurrencias; es decir, ateniéndose únicamente a su frecuencia. En segundo lugar, se aplicará el criterio de la rentabilidad comunicativa para destacar aquellas expresiones que se pueden utilizar en la mayoría de los contextos comunicativos sin que resulten vulgares o demasiado coloquiales. Para ello se asignará un valor de ++ a aquellas expresiones que no estén marcadas como coloquiales por el DRAE, un valor + a las expresiones coloquiales y un valor – a las expresiones marcadas como vulgares. Además, tanto para la pirámide del fútbol como para la de los toros, se decidió prescindir de las interjecciones por su escasa carga etnolingüística y también de aquellas expresiones que solo presenten una ocurrencia. Teniendo en cuenta estos dos criterios y las expresiones encontradas en el *CORPES*, así es como se visualizarían las pirámides:

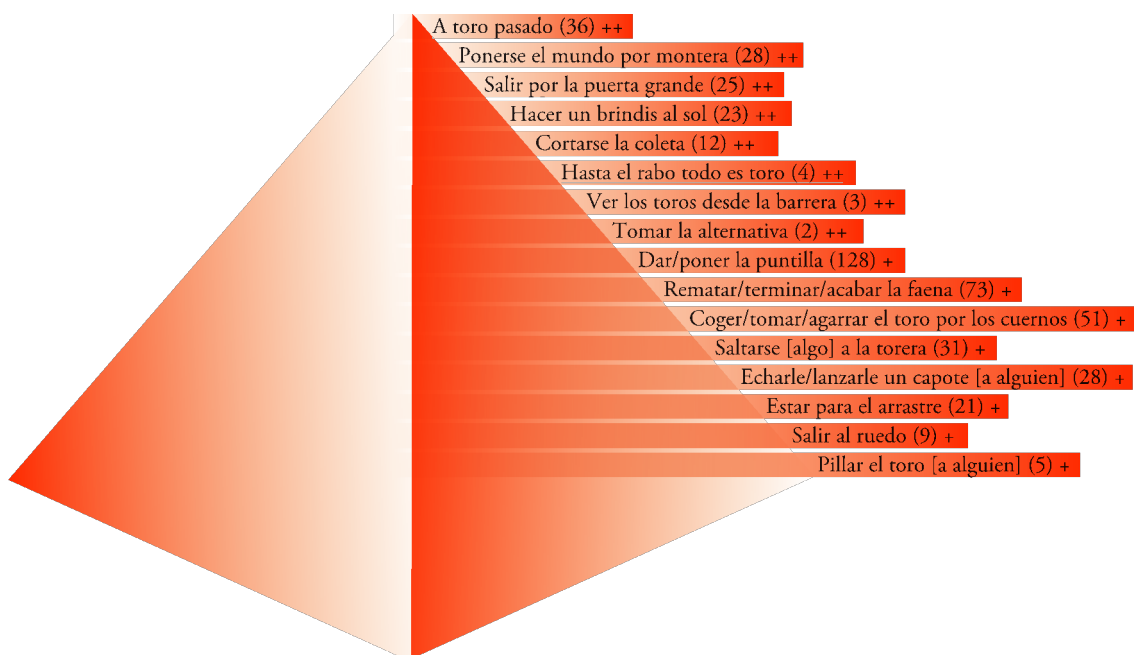
3.4.1. Pirámide etnolingüística del fútbol.



El resultado final es una pirámide con 19 expresiones, las cuales constituyen un número suficientemente reducido que ofrece una visualización clara y concisa de cómo influye el fútbol en la lengua real, mostrando qué expresiones se utilizan en España en el período seleccionado (2001-2015). Finalmente, es posible que esta pirámide sufra cambios de cara a los próximos años,

puesto que el fútbol tiene un impacto social cada vez mayor y no es de extrañar que se incorporen a la lengua nuevas expresiones que se vuelvan virales en un breve espacio de tiempo y que incluso reemplacen algunas de las que tenemos ahora. Un claro ejemplo es la expresión *ser un crack*, reina indiscutible de la pirámide. También cabe destacar otras expresiones que han ampliado su significado más allá del ámbito futbolístico, como *ganar por goleada* (60, ++), *colgar las botas* (19, ++) y *despejar balones fuera* (54, ++). En cambio, la expresión *casarse de penalti*, antaño bastante frecuente, parece que ha perdido fuelle (solo 3 ocurrencias) en los últimos quince años, y eso se debe sin duda a un cambio sociocultural que afecta directamente a la lengua, ya que cada vez hay menos personas que se casen a raíz de un embarazo.

3.4.2. Pirámide etnolingüística de los toros.



El resultado es una pirámide con 16 expresiones, un número muy parejo a la del fútbol.

Curiosamente, vemos que aquí las expresiones más frecuentes son las más coloquiales; sin embargo, ocupan la parte inferior en la pirámide por ser consideradas menos rentables. Como podemos apreciar, hay varias expresiones que reflejan claramente las huellas de la cultura en la lengua, como por ejemplo *ponerse el mundo por montera* (28, ++), en la que se asocia el gesto del torero que se pone la montera antes del combate con la idea de tomar una decisión sin importar lo que digan los demás. Se podría considerar como un caso de metonimia por contigüidad, donde el gesto del torero representaría la valentía en la vida real. Otros casos destacables son *echarle/lanzarle un capote [a alguien]* (28, +), que ha ampliado su significado pasando a designar cualquier tipo de ayuda, incluso fuera del ruedo, y *estar para el arrastre* (21, +), una hipérbole muy frecuente en la lengua que se utiliza para expresar un gran estado de cansancio, y en la que el hablante se compara con el toro muerto al final de la lidia, estando listos los dos para ser arrastrados fuera de la arena. Por último, vale la pena mencionar *salir al ruedo* (9, +), utilizada cuando alguien se dispone a pasar a la acción. Es difícil prever lo que pueda pasar dentro de unos años y queda por ver si los futuros usuarios de la lengua seguirán utilizando estas expresiones pese a la constante decadencia de la tauromaquia y al escaso interés que despierta en la sociedad actual. Algunas de ellas están muy fosilizadas en la lengua y es posible que sigan existiendo, aunque su permanencia dependerá en muchas ocasiones de factores impredecibles como los memes, las modas y las tendencias de los hablantes.

4. CONCLUSIÓN

En este artículo se ha querido proponer un nuevo sistema de visualización de datos para presentar expresiones culturalmente marcadas a partir de unos conceptos teóricos sobre etnolingüística y comunicación. Las pirámides etnolingüísticas se pueden crear para cualquier ámbito cultural, y representan una nueva herramienta de visualización de datos que permite ahondar más en la relación lengua-cultura, detectando patrones que pueden ser de ayuda tanto en el ámbito contrastivo como en la enseñanza de una L2. Para llevar a cabo esta propuesta, ha desempeñado un papel fundamental la metodología escogida; por un lado, se decidió utilizar una metodología mixta basada en la triangulación de datos (*brainstorming*, búsqueda en los corpus y visualización en las pirámides); por otro, se optó por seguir el método inductivo Data-driven, con el que se pretendía demostrar la hipótesis inicial a partir de la recopilación y análisis de datos lingüísticos, y sin recurrir a diccionarios fraseológicos en busca de las expresiones. Así pues, las cuatro fases de creación de las pirámides han dado resultados satisfactorios y han permitido alcanzar los objetivos propuestos: por un lado, encontrar el mayor número posible de expresiones fraseológicas relacionadas con el mundo del fútbol y de los toros (A); por otro, ordenar y visualizar esas expresiones dentro de unas pirámides (B). Es evidente que hoy en día vivimos un auge tecnológico sin precedentes, y el aspecto visual ha ido ganando mucho terreno en todos los ámbitos, por eso se decidió enmarcar las expresiones seleccionadas dentro de unas pirámides y ordenarlas de acuerdo con dos criterios principales: el de frecuencia y el de rentabilidad comunicativa. Además, se pudo comprobar que, para realizar las búsquedas, se necesita al menos un corpus general que sea lo más representativo posible de la lengua en cuestión y que, a ser posible, esté lematizado, por eso se escogió el *CORPES*, el corpus de la Real Academia Española que, pese a encontrarse todavía en versión beta en el período en que se realizaron las búsquedas (2016), está lematizado y presenta una interfaz muy cómoda y sencilla.

En conclusión, pensamos que la etnolingüística presenta nuevos recursos con los que se podría ahondar mucho más en la relación lengua-cultura, ofreciendo puntos de vista novedosos y esclarecedores sobre el impacto de la cultura en una determinada comunidad lingüística. Por consiguiente, el presente artículo constituye un primer paso en la creación de nuevas herramientas etnolingüísticas para el análisis de la influencia de la cultura en la lengua y viceversa. De momento, los objetivos propuestos han compartido la voluntad de demostrar que es posible servirse de un *brainstorming* y de los corpus con fines etnolingüísticos para detectar la influencia real de un ámbito cultural concreto en la lengua, y que esa influencia se puede apreciar gracias a la visualización de las pirámides.

A modo de cierre, cabe subrayar que la cultura y la lengua no solamente ejercen entre ellas una influencia bidireccional que se ve plasmada en el uso real de los hablantes, sino que además existe un factor impredecible relacionado con el papel del artista creador y con la aceptación de una determinada expresión que depende muchas veces de factores sociales, y por supuesto de la empatía, dirigida en todo ser humano por las neuronas espejos. Sin duda, la mejor manera de aprender y, sobre todo, de comprender una lengua es con una gran inmersión lingüística y cultural que permita entender más a fondo la cosmovisión de sus hablantes. La etnolingüística presenta nuevos recursos con los que se podría ahondar mucho más en la relación lengua-cultura, ofreciendo puntos de vista novedosos y esclarecedores sobre el impacto de la cultura en una determinada comunidad lingüística. Asimismo, se debe recalcar, una vez más, el papel fundamental de la neurociencia en este trabajo de investigación. El cerebro humano sigue siendo, en ciertos aspectos, un gran desconocido para el ser humano, y los avances de las últimas décadas subrayan la importancia de algunos mecanismos cognitivos como la imitación o la actividad del sistema

límbico, verdaderos motores del aprendizaje y de la integración sociocultural. Como comentó el Dr. Francisco Mora en la conferencia sobre Neuroeducación que impartió en la Universitat de Barcelona el 10 de marzo de 2017³¹, “no hay pensamiento sin el fuego emocional que lo alimente, ni creatividad sin equivocación, y la verdadera curiosidad estará siempre unos pasos más allá de las limitaciones impuestas en los mapas”.

BIBLIOGRAFÍA

- ABELLA, C. (1996), *¡Derecho al toro! El lenguaje taurino y su influencia en lo cotidiano*, Madrid, Anaya y Mario Muchnik.
- ARBIB, M. (2005), “From monkeylike action recognition to human language: An evolutionary framework for neurolinguistics”, en *Behavioural and Brain Science*, 28, 105-124.
- BENASSAR, B. (2000), *Historia de la tauromaquia: una sociedad del espectáculo*, Valencia, Editorial Pre-Textos.
- BOAS, F. (ed.) (2004), *Anthropology and Modern Life*, Abingdon, Taylor Francis.
- CARDONA, G. (2006), *Introduzione all’etnolinguistica*, Novara, De Agostini Scuola SpA.
- (1985), *I sei lati del mondo. Linguaggio ed esperienza*, Roma-Bari, Laterza.
- (1973), *La linguistica antropologica*, Roma, Parole e método.
- CASADO VELARDE, M. (1978), *Lengua e ideología*, Pamplona, Eunsa.
- (1988), *Lengua y cultura. La etnolingüística*, Madrid, Síntesis.
- CHEN, C. (2006), *Information Visualization: Beyond the Horizon*, 2nd ed. London, Springer.
- CHOMSKY, N. (1986), *Knowledge of Language*, New York, Praeger.
- COSERIU, E. (1981), “La socio- y la etnolingüística: sus fundamentos y sus tareas”, en *Anuario de Letras*, XIX, 5-29.
- COSERIU, E. (1982), *Teoría del lenguaje y lingüística general*, Madrid, Gredos.
- CRUZ PIÑOL, M. (2012), *Lingüística de corpus y enseñanza de español como L2*, Madrid, Arco Libros.
- DE SAUSSURE, F. (1983), *Curso de lingüística general*, Madrid, Alianza.
- DORNYEI, Z. (2007), *Research Methods in Applied Linguistic: Quantitative, Qualitative, and Mixed Methodologies*, Oxford, Oxford University Press.
- FERNÁNDEZ GUIZZETTI, G. (1957), “La etnolingüística: del mundo del idioma al mundo de la cultura”, en *Revista de antropología*, Sao Paolo, 5, 1, 75-93.
- (1960), *Guillermo de Humboldt, padre de la etnolingüística. Esquemas para una historia de la etnolingüística*, Buenos Aires, “Cuadernos del Instituto Nacional de Investigaciones Folklóricas”, 1, 229-245.
- GARCÍA GARCÍA, E. (2008), “Neuropsicología y educación. De las neuronas espejo a la teoría de la mente”, en *Revista de Psicología y Educación*, 1, 3, 69-90. Disponible en: <<http://eprints.ucm.es/9972/>>.
- GOLEMAN, D. (2012), *El cerebro y la inteligencia emocional: nuevos descubrimientos*, Barcelona, Ediciones B.
- GONZÁLEZ-LLORET, M. (2014), “Investigación-Acción (II): La Investigación Cuantitativa”, en *DidactiRed*. Centro Virtual Cervantes. Disponible en: <https://cvc.cervantes.es/aula/didactired/anteriores/diciembre_12/17122012.htm>.

³¹ Seminario Educación, comunicación y gramática II, en ocasión del II Foro Universidad de Barcelona-SGEL.

- GOODELL, J. (2007), *Approaching 100% by 2014: Using Data-driven Technology, scientifically proven Practices, and cultural Change to meet the no Child left behind Challenge*, Estados Unidos, Sound Image Inc.
- HUMBOLDT, W. (ed.) (1999), *On Language. The Diversity of human Language-Structure and its Influence on the mental Development of Making*, Cambridge, Cambridge University Press.
- HYMES, D. (1972), *Models of the Interaction of Language and social Life*, J. Gumperz y D. Hymes (dirs.): *Directions in Sociolinguistics: the Ethnography of Communication*, New York, Holt, Rinehart and Winston, Inc.
- IACOBONI, M. (2009), *Las neuronas espejo*, Buenos Aires, Katz Editores.
- ISIDORO, Santo, Arzobispo de Sevilla (ed.) (2004), *Etimologías*, Madrid, Biblioteca de Autores Cristianos.
- JAEGER, W. (1946), *Aristóteles. Bases para la historia de su desarrollo intelectual*, versión española de José Gaos, Fondo de Cultura Económica, México
- MASON, H. y PATIL, D. (2015), *Data Driven. Creating a Data Culture*, O'Reilly Media.
- MORA, F. (2007), *Neurocultura*, Madrid, Alianza Editorial.
- MORA, F. (2013), *Neuroeducación*, Madrid, Alianza Editorial.
- NOMDEDEU RULL, A. (2009), “Diccionario de fútbol”, en *Anexos de Revista de Lexicografía*, A Coruña, Universidade da Coruña.
- OSBORN, A. (1957), *Imaginación aplicada*, Madrid, Ediciones Velflex.
- PUYAL ORTIGA, J. (1972), *Terminología futbolística*, tesis de licenciatura, Barcelona, Universitat de Barcelona.
- RAMACHANDRAN, V. (2000), “Mirror neurons and imitation learning as the driving force behind the Great Leap Forward in human evolution”, *Edge*, 69, 29-36.
- REAL ACADEMIA ESPAÑOLA (2014), *Diccionario de la lengua española. Edición del tricentenario*, Madrid, Espasa Libros.
- REUS BOYD-SWAN, F. (2004), “El léxico taurino en la vida cotidiana”, *Revista de la Sociedad Española de Estudios Literarios de Cultura Popular*, 4, 239-263.
- RIZZOLATTI, G y SINIGAGLIA, C. (2006), *Las neuronas espejo: los mecanismos de la empatía emocional*, Barcelona, Paidós.
- SAPIR, E. (1949), *The selected writings of Edward Sapir*, Berkely, University of California.
- SPENCE, R. (2007), *Information Visualization: Design for Interaction*, Harlow, Addison Wesley.
- TERUEL SAEZ, A. (2007), *Vocabulario de fútbol*, Gijón, Trea.
- TORRUELLA, J. y LLISTERRI, J. (1999), “Diseño de corpus textuales y orales”, J. Blecua, G. Clavería, C. Sánchez y J. Torruella (eds.): *Filología e informática. Nuevas tecnologías en los estudios lingüísticos*, Barcelona, Seminari de Filologia i Informàtica, Departament de Filologia Espanyola, Universitat Autònoma de Barcelona – Editorial Milenio, 45-77.
- ULLMANN, S. (1976), *Semántica. Introducción a la ciencia del significado*, Madrid, Aguilar.
- VIVAS HOLGADO, J. (1999), *El fútbol: léxico, deporte y periodismo*, Cáceres, Universidad de Extremadura.
- WARE, C. (2004), *Information visualization: perception for design*, San Francisco, Morgan Kaufman.