

Helene R. Gudmann

Lagunas de significado en los diccionarios españoles - Una investigación de los elementos de significado en diccionarios generales y monolingües de español para la recepción de textos

1. Introducción

Durante mis estudios de la lengua castellana experimenté a menudo que no podía encontrar las acepciones que buscaba y necesitaba para comprender debidamente una palabra textual en un texto español. Me encontraba con este problema tanto en situaciones en las que no tenía la menor idea del significado de la palabra textual como cuando tenía alguna sospecha o idea que quería confirmar, no consiguiendo ayuda ni los diccionarios bilingües de español-danés ni los diccionarios generales y monolingües de español destinados a la recepción de textos. El resultado de estas *lagunas de significado* fue una falta de comprensión, problema que puede tener consecuencias muy graves para el usuario de este tipo de diccionario, en especial si este es traductor. Y a nivel más teórico, se podría preguntarse que si el diccionario para la recepción de textos no satisface esta necesidad de recepción del usuario, ¿para qué sirve entonces?

Esta disyuntiva fue la motivación de la investigación de mi tesina de Máster tutorizada por Sven Tarp en la Universidad de Aarhus (véase Gudmann 2014). En aquel trabajo investigué mi hipótesis de que los diccionarios generales y monolingües de español para la recepción de textos tienen *lagunas de significado* en cierta medida, lo que significa que no incluyen la necesaria cantidad de ‘meaning items’ (en español: *elementos de significado*, tradicionalmente llamados *acepciones*) correspondiente al *objeto genuino* de dichos diccionarios. De este modo, se analiza si los diccionarios seleccionados cumplen con su objeto genuino, o sea de satisfacer las necesidades de recepción de textos de los usuarios previstos, postulando que este tipo de diccionarios generales no cumple con su objeto genuino si no incluye la mayor parte posible de los elementos de significado relevantes e identificados en textos de conocimientos generales.

En el presente artículo presentaré la parte de mi trabajo en la que se ha comparado la cantidad de elementos de significado que tienen 23 lemas incluidos en los diccionarios elegidos para el estudio, con la cantidad de elementos de significado que se han seleccionado mediante un análisis de corpus en el Internet. Los diccionarios estudiados son: *Clave. Diccionario de uso del español actual* (CLA), *Diccionario del español actual* (DEA), *Diccionario de la lengua española* (DRAE), *María Moliner: Diccionario de uso del español* (MAR) y *Diccionario de uso del español de América y España* (VOX). El punto de partida teórico del trabajo es la teoría funcional de la lexicografía desarrollada en el Centro de Lexicografía en Aarhus (véase p.ej. Tarp 2013). La comparación mencionada pretende verificar si los diccionarios tienen lagunas de significado o no.

2. Significado en la lexicografía

La lexicografía es una disciplina independiente que, por lo tanto, tiene su propia terminología que a veces se parece a la de la lingüística. En consecuencia, en esta contribución resulta necesario presentar aquí algunos términos esenciales en el contexto del estudio de las lagunas de significado.

2.1 La palabra

Según los teóricos funcionales Bergenholtz & Agerbo (2014c), hay que discriminar entre cinco conceptos que abarca el término *palabra*. Primero, el término puede referir a la *palabra ortográfica*, es decir, una serie de letras unidas entre dos espacios. Segundo, tenemos las *palabras textuales* que son unidades de letras concretas con ortografía, significado, gramática, etc. particulares del texto específico en cuestión. Estas son las palabras que ocasionan la consulta de un diccionario mediante la palabra ortográfica. Una *palabra gramatical* es una expresión con un morfema radical o gramatical que pertenece a un paradigma y que puede pertenecer a uno o más lexemas. En cuarto lugar, tenemos las *palabras lexicales* o *lexemas* que son abstracciones que abarcan una cantidad de palabras gramaticales con el mismo radical y paradigma. Si varía el radical, hay más lexemas ya que el radical se define como un signo lingüístico (término explicado más adelante). Además, se suele llamar *palabras* a los *lemas* que son abstracciones que incluyen una cantidad de palabras gramaticales y que representan la unidad que se busca en el diccionario (Bergenholtz & Agerbo 2014c:4).

2.2 La acepción y el significado

En la lexicografía tradicional, se hablan de *acepciones* y se dice, por ejemplo, que “este lema tiene polisemia ya que tiene tres acepciones y una subacepción”. Este discurso implica que las acepciones estén subordinadas al lema que se supone que representa el “significado nuclear”. Pero este punto de vista es semántico-lingüístico y no toma en consideración a los usuarios, factor imprescindible en la producción de herramientas lexicográficas. Por este motivo, es preciso reconsiderar el rol de la acepción y el lema en conexión con las necesidades lexicográficas del usuario previsto tal y como han hecho Bergenholtz & Agerbo (2014c). Según su argumentación, si uno estudia la teoría sobre el signo lingüístico de por ejemplo Hjelmslev (1943), se encuentra un hecho muy importante: Un signo lingüístico siempre está compuesto de dos partes, el *significado* y la *expresión*, que están integralmente unidas y que solamente juntas forman este determinado signo lingüístico. En consecuencia, si encontramos una expresión que tenga dos significados, estos son en realidad dos expresiones separadas que, por incidencia, tienen la misma forma ortográfica, acústica, etc. De este modo, no tiene sentido hablar de acepciones o subacepciones ya que estas unidades de ‘significado + expresión’ son signos en su propio derecho, cada uno con su propia relación con las *cosas* en el mundo (véase Bergenholtz & Agerbo 2014a:490). Esto también significa que, en el contexto lexicográfico, no se puede hacer ‘lumping’ y ‘splitting’ de las acepciones porque el punto de partida no es que todas las acepciones provienen del mismo lema (lo que realmente puede ser correcto, etimológicamente) sino que, en la lengua actual, representan su propio significado al que se relaciona una palabra textual que puede resultar incomprensible para un usuario. A este usuario no le importa el origen de la palabra, pues solo quiere comprender la palabra textual en cuestión (ibid.).

Según esta argumentación, las tradicionales explicaciones abstractas que pretenden abarcar varios elementos de significado al mismo tiempo tampoco asisten a este tipo de usuario ya que a este le interesa la relación de la palabra textual con las cosas en el mundo, relación que se esconde en el ‘lumping’ (en español: *agrupación*) de elementos de significado que en realidad se relacionan en uso, contexto, etc.

De este modo, la homonimia etimológica y semántica no tienen relevancia en este contexto lo que abre paso al uso de la homonimia gramatical para estructurar los datos de diccionarios de este tipo de una manera reconocible (véase Tarp 2009:301 y Bergenholtz & Agerbo 2014c:6 para una explicación más detallada). En resumen, el punto de partida de este trabajo ha sido que una acepción tiene el mismo valor que un lema en la valoración y compilación de un diccionario. Para poner hincapié en esta distinción teórica, las acepciones se llamarán *elementos de significado* en el presente trabajo.

2.3 Lagunas

En la compilación del diccionario es importante identificar y seleccionar el número adecuado de lemas y de elementos de significado para poder satisfacer las necesidades del usuario previsto. En el diccionario de anatomía para estudiantes de medicina que necesitan ayuda a la hora de comprender palabras textuales presentes en su libro de texto, hay que incluir todos los huesos del cuerpo y no solo los que supuestamente son difíciles para los estudiantes. En caso de que no se haga, se producirán las llamadas *lagunas léxicas* o, en el caso de elementos de significado, *lagunas de significado*. Esta situación es grave para el diccionario porque es muy probable que el usuario que repetidamente no encuentra ayuda en un diccionario particular no vuelva a utilizarlo y, definitivamente, no lo recomiende, problema que perjudica la venta del diccionario (Bergenholtz & Norddahl 2012). Las situaciones contrarias se llaman ‘lemma flooding’ y, consecuentemente, en el caso de elementos de significado, *abundancia de significado* (mi expresión), problemas que involucran la producción de lemas y elementos de significado no relevantes y, con esto, una pérdida de recursos temporales y económicos lo que hoy día supone grandes retos para el editorial.

2.4 El objeto genuino

Pero, ¿cómo se evitan las lagunas de significado? Pues, un análisis previo del objeto genuino del diccionario que se quiere compilar es esencial. En nuestro caso, se trata de un diccionario general y monolingüe del castellano destinado a la recepción de textos lo que significa que tiene el objeto genuino de ayudar al usuario hispanohablante a la hora de comprender el significado y las restricciones pragmáticas y culturales de palabras textuales específicas. Estas palabras textuales se encuentran en textos de gran variación que, generalmente, se caracterizan como textos de conocimientos generales, accesibles para el público general. De ahí, el diccionario debe presentar los datos correspondientes a este espectro tan amplio, o sea, idealmente debe incluir todos los elementos de significado presentes en textos de conocimientos generales, accesibles para el público general (véase Bergenholtz 1996:10). Además, los diccionarios para la recepción de textos deben incluir elementos de significado poco frecuentes porque justamente estos pueden conllevar problemas de comprensión (Bergenholtz 2012:27). Sin embargo, es oportuno minimizar la

selección de elementos de significado periféricos e irrelevantes y, así, la abundancia de significados por lo que Bergenholtz & Agerbo proponen observar un límite mínimo de 10 ‘tokens’ (en español: *presencias*) del elemento de significado en el corpus (Bergenholtz & Agerbo 2014b:135).

Además, hay que caracterizar el tipo de usuario del diccionario para asegurar que el contenido del diccionario realmente corresponda a sus necesidades. En cuanto al diccionario monolingüe para la recepción de textos, se pueden discriminar dos tipos de usuario: El tipo A que tiene una idea sobre qué podría ser el significado de la palabra textual en cuestión y que solo quiere confirmar la sospecha, y el tipo B que no tiene ni idea sobre el significado de la palabra (Bergenholtz & Agerbo 2014b:128).

Todos estos factores son importantes para el presente estudio y la evaluación de los cinco diccionarios elegidos. Pero antes de presentar los resultados del trabajo, se presentarán el método investigativo y los datos empíricos.

3. Método de investigación

El estudio abarca una comparación de la cantidad de elementos de significado de los cinco diccionarios elegidos con la cantidad resultante de mi propio análisis de cuántos elementos de significado sería posible seleccionar para este tipo de diccionario. No se puede caracterizar la investigación como representativa sino como exploratoria aunque se han utilizado métodos de selección bastante arbitrarias. No obstante, los resultados pueden indicar la existencia de fenómenos importantes y relevantes para la lexicografía española.

El método de mi propio análisis se basa en las ideas siguientes: Según Bergenholtz & Agerbo, la base de los datos de cualquier diccionario debe ser empírica, es decir, que los datos pueden provenir de corpus, blogs, páginas web, manuales e introspección, entra otras cosas (Bergenholtz & Agerbo 2014b). Para los diccionarios generales, los autores proponen la utilización de corpus principalmente, tipo de datos que durante los últimos años se ha vuelto la base generalmente preferida de los lexicógrafos. Rechazan, no obstante, la idea utópica de la *representatividad* del corpus (véase Kilgarriff & Grefenstette 2003) lo que abre paso al uso del Internet como corpus lexicográfico. Es más económico porque la redacción no tendrá que acumular los datos ni editarlos y, al mismo tiempo, el corpus queda actualizado y voluminoso. Sin embargo, su uso, por supuesto, pone mucha responsabilidad en el lexicógrafo que tendrá que averiguar la calidad y la utilidad de las fuentes. No obstante, por las razones ya mencionadas resulta un buen punto de partida para la identificación de elementos de significado para los diccionarios generales y monolingües destinados a la recepción de textos. Por lo tanto, Bergenholtz & Agerbo basan su método lexicográfico en el análisis de los datos procedentes de búsquedas en Google (Google Inc. 2014), el método que también se utiliza en este trabajo. En resumen, el método abarca cuatro o cinco pasos: El lexicógrafo:

1. busca el lema en Google, extrayendo colocaciones, ejemplos, sinónimos, etc. de los resultados de búsqueda lo que le ayuda a relacionar los elementos de significado con el mundo;

2. evalúa estos datos, identificando uno o más elementos de significado e intentando formular la primera versión de la *explicación lexicográfica* (tradicionalmente llamada la *definición lexicográfica*);
3. (pide ayuda a un experto para comprender correctamente el fenómeno, si el elemento de significado resulta ser de conocimiento más especializado);
4. encuentra y selecciona más ejemplos, etc., asignándolos al correspondiente elemento de significado identificado; y
5. formula la última versión de la explicación lexicográfica lo que puede llevar a la identificación de otros elementos de significado y el reinicio del proceso (Bergenholtz & Agerbo 2014b:136).

Este proceso es iterativo. Para asegurar la identificación y selección de elementos de significado relevantes en mi estudio, se fijaron que los *tokens* tenían que aparecer en textos de conocimientos generales y que diez *tokens* en el corpus eran el límite mínimo.

En el análisis comparativo de los resultados del ya mencionado análisis de corpus (AC) con los datos de los cinco diccionarios, se utiliza el *principio de sustituibilidad* (véase Hartmann & James 1998:133) para averiguar la necesidad y razón de ser de los elementos de significado seleccionados aunque sin apoyar la idea de que la explicación lexicográfica y la palabra de texto son equivalentes.

En el tratamiento de los elementos de significado de los diccionarios, se ha intentado categorizar estos elementos en grupos de elementos de significado que parecen hacer referencia al mismo significado. Esta técnica ha resultado en una comparación más sistematizada y menos redundante (véase la figura B más adelante).

4. Datos empíricos

Los cinco grandes diccionarios contemporáneos estudiados se pueden categorizar, según la teoría funcional, como diccionarios generales y monolingües para la recepción de textos en español, es decir, que el tema principal es de conocimiento no especializado sino general, el texto está escrito en castellano y los usuarios previstos son hispanohablantes nativos (más, en cuatro casos, hablantes no nativos) que necesiten ayuda para la comprensión de palabras textuales leídas en textos castellanos. En cuanto a CLA, DRAE y VOX, se han analizado los datos de la versión en Internet y CD-ROM cuyos datos corresponden a los de las versiones impresas. Además, se ha utilizado los datos de la versión de MAR en CD-ROM cuyas explicaciones lexicográficas deberán corresponder básicamente a la versión impresa y los datos de DEA que se hallan solo en papel.

La investigación se basa en el estudio de los elementos de significado de 23 lemas, esto es: 1 *aniversario* (m), 2 *aseo* (m), 3 *autocontrol* (m), 4 *biografía* (f), 5 *cuestionario* (m), 6 *dicotomía* (f), 7 *director* (m), 8 *explorador* (m), 9 *flecha* (f), 10 *frontera* (f), 11 *guerra* (f), 12 *intérprete* (m), 13 *limbo* (m), 14 *material* (m), 15 *móvil* (m), 16 *orgullo* (m), 17 *patrimonio* (m), 18 *puya* (f), 19 *riqueza* (f), 20 *ritmo* (m), 21 *serie* (f), 22 *tarjeta* (f) og 23 *técnica* (f). Para asegurar una selección bastante arbitraria que refleje problemas modernos, estos mismos elementos se han elegido porque se hallan en El País n° 1982 del 21 de septiembre 2014. Además, todos ellos están presentes en los

cinco diccionarios, todos son sustantivos y no se han discriminado a base de la homonimia del diccionario sino de si tenían la misma gramática. Para delimitar apropiadamente el análisis a mis habilidades, se han rechazado acepciones con marcos de una región lingüística no ibérica, como por ejemplo “(Bol.)”. También se han descartado acepciones marcadas con “anticuado” o etiqueta parecida porque no son relevantes para una investigación de diccionarios contemporáneos.

Como ya se mencionó, los datos del análisis de corpus son los resultados de búsqueda mediante Google desde octubre hasta noviembre de 2014.

5. Resultados

La comparación de los elementos de significado de los diccionarios con los del AC tiene dos partes: Primero, se comparan las cantidades totales de elementos de significado y, luego, se comparan los elementos específicos que se han seleccionado en cada diccionario con los del AC.

5.1 La cantidad de elementos de significado

La cuenta de los elementos de significado para los 23 lemas del estudio revela los siguientes números: CLA tiene 58 elementos de significado en total, VOX tiene 75, DRAE tiene 76, DEA está más arriba con sus 91 mientras MAR es el “ganador” con sus 97 elementos de significado. En comparación, el AC llevado a cabo con los métodos y criterios ya mencionados resultó en nada menos que un total de 191 elementos de significado, es decir, que se han seleccionado entre 94 y 133 elementos de significado más que los diccionarios. Los números se distribuyen así:

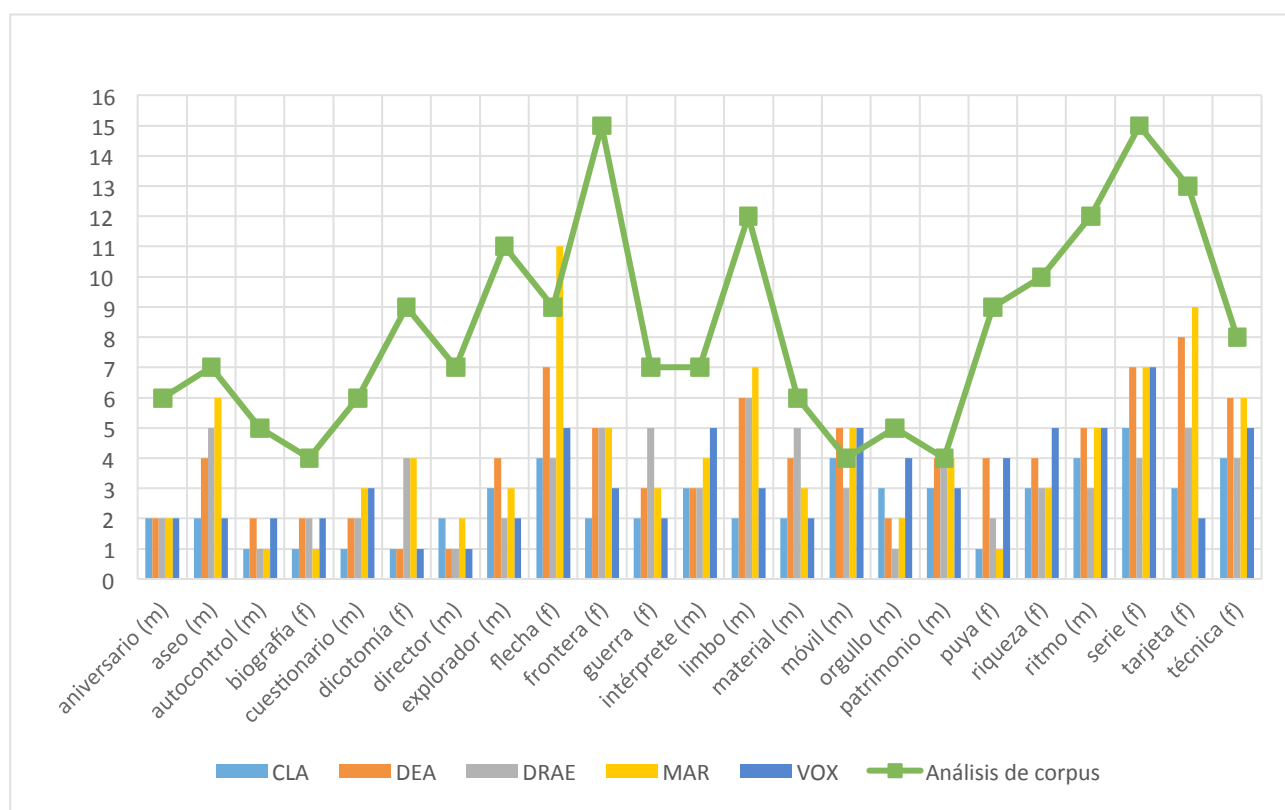


Figura A – Distribución del número total de elementos de significado

Como se ve, el AC ha resultado en más elementos de significado que todos los diccionarios para 20 de los 23 lemas. En un caso, o sea el de *patrimonio*, DEA, DRAE, MAR y yo hemos seleccionado el mismo número de elementos, DEA MAR y VOX tienen uno más que yo en cuanto a *móvil* y MAR ha seleccionado dos más que yo para *flecha*. Sin embargo, la diferencia es muy notable.

5.2 La necesidad de los elementos de significado

Pero una cantidad grande de elementos de significado no es una meta en sí si estos no son relevantes y útiles para el usuario previsto del diccionario. Aunque los métodos utilizados en el AC deberían haber asegurado esto, me gustaría destacar el razonamiento bajo la selección de elementos de significado que no están en los diccionarios, por lo que presentaré aquí la necesidad y la razón de ser de estos elementos. El punto de partida de este estudio es los dos tipos de usuarios (A y B) que, en correspondencia con las descripciones de los cinco diccionarios, son hispanohablantes nativos. Se han elegido los lemas *aseo*, *dicotomía*, *director* y *orgullo* para esta presentación porque ninguno de los diccionarios contiene elementos de significado para estos que no son seleccionados por el AC.

La razón de ser de algunos elementos de significado parece muy lógica. Por ejemplo, resulta difícil disputar la selección del elemento de significado de *aseo* con la explicación

“dispositivo que, como mínimo, comprende una taza con asiento, una cisterna, una salida para las aguas residuales y un sifón; normalmente, está hecho de cerámica o acero y se utiliza para poder llevar a cabo las necesidades corporales, como orinar y evacuar”,

en particular al leer los siguientes colocaciones y ejemplos reales del corpus:

- *Un autobús con aseo y ducha*
- *Tiene un aseo, lavabo y baño con ducha.*
- *Conecte el enchufe y ya tiene un aseo funcionando.*
- *La taza del aseo estaba suelta, había perdido el sellado al suelo.*
- *Sabiendo que no tiene nada que temer porque ya está sentado en el aseo, le pedimos al paciente que intente retrasar todo lo que pueda el inicio de la micción.*

Ninguno de los diccionarios tiene este elemento de significado aunque todos tienen el de “Cuarto de aseo” (MAR). El elemento de significado aparece en textos de conocimientos generales y tiene más de diez presencias en el corpus, así que no queda duda sobre la necesidad de incluir este elemento de significado en el diccionario general. Lo mismo ocurre con *orgullo* con el siguiente significado:

“suceso que consiste en que un grupo de personas se reúne para manifestar su apoyo a los derechos de los homosexuales, bisexuales, transexuales y personas con otras sexualidades alternativas a la hetera; normalmente, incluye discursos políticos, desfiles por las calles y fiestas con música y baile”.

Esta explicación lexicográfica no se parece a ninguna de las de los diccionarios aunque tiene muchas presencias en Google.

Pero no todos los elementos de significado necesarios y adecuados se revelan tan fácilmente por lo que haremos uso del principio de sustituibilidad. Tomemos el ejemplo del corpus de *dicotomía*:

Marte muestra una dicotomía entre el norte y el sur, así que mientras el sur muestra una superficie antigua, llena de cráteres, el norte muestra una superficie joven con llanuras bajas.

Como ejemplo ponemos las explicaciones de este lema en DRAE ya que tiene 4 elementos de significado a diferencia de CLA, DEA y VOX que solo tienen uno:

- “División en dos partes.”
- “*Fil.* Método de clasificación en que las divisiones y subdivisiones solo tienen dos partes.”
- “*Bot.* Bifurcación de un tallo o de una rama.”
- “Práctica condenada por la recta deontología, que consiste en el pago de una comisión por el médico consultante, operador o especialista, al médico de cabecera que le ha recomendado un cliente.”

Se ve que el contexto indicado en las explicaciones solo nos deja la primera explicación para la prueba de sustitución. Al sustituir *dicotomía* en el ejemplo con “División en dos partes”, tendremos esta oración:

Marte muestra una [división en dos partes] entre el norte y el sur, así que mientras el sur muestra una superficie antigua, llena de cráteres, el norte muestra una superficie joven con llanuras bajas.

Esta sustitución nos indica que la explicación en cuestión simplemente está dedicada a otro elemento de significado ya que el ejemplo no quiere decir que Marte haya sido dividido en dos partes sino que está mostrando una diferencia entre el sur y el norte del planeta. Por lo tanto, el elemento de significado para *dicotomía* del AC, “Cualidad de ser distintas dos cosas por ejemplo con respecto a su aspecto o sus propiedades”, es necesario en este tipo de diccionario para la recepción de textos.

5.3 La posibilidad de ‘lumping’

Los críticos quizás dirían que muchos de los elementos de significado seleccionados mediante el corpus están relacionados entre sí de modo que podrían haber sido ‘lumped’, o sea *agrupados*, en pocos elementos de significado. Seguramente, sería posible hacerlo pero miremos otros ejemplos. En el AC, he seleccionado estos tres elementos de significado “relacionados” para *director*:

- 1) “varón que se dedica a gestionar una organización, grupo o cosa parecida, tomando decisiones en lo que respecta a los objetivos y problemas de este y, normalmente, administrando a la plantilla y las tareas diarias”

2) “varón que se dedica a dirigir obras dramáticas como películas, obras teatrales o cosa parecida, fijando los aspectos principales de la obra, dictando cómo deben actuar los actores, tomando decisiones decisivas sobre iluminación, efectos, decorados, etc.”

3) “varón que se dedica a dirigir grupos musicales, es decir, que por ejemplo dicta el ritmo y el tempo y que les ayuda a jugar a sus instrumentos o cantar de manera sincrónica y armonizada; normalmente, en conciertos, indica estas cosas utilizando las manos”

Seguramente sería posible agrupar estos tres elementos, lo que por ejemplo ha hecho VOX con esta explicación: “Persona que dirige algo **ej** director de orquesta; director de cine; directora general de una empresa”. Pero esta explicación no le presenta al usuario casi ningún dato nuevo que le pueda ayudar a relacionar el elemento de significado con el mundo. Si el usuario del tipo B, por ejemplo, lee *director* en este texto: “Ha sido director de los Cursos Internacionales de esa Universidad, así como fundador de las Academias literarias renacentistas y de los Encuentros de escritores de las lenguas de España en Verines”, ya le queda claro que se trata de una persona que hace algo. Pero si este usuario no comprende *dirigir* – que solo es una derivación de *director* – no recibirá más ayuda consultando *dirigir* en VOX porque tiene no menos de ocho acepciones y, ¿cómo saber cuál es la más adecuada? Además, como indican las explicaciones de arriba, las tareas de los tres tipos de directores son muy distintas y este tipo de datos explicativos son imprescindibles para el usuario en su búsqueda de respuestas. Además, otros ejemplos del uso de *director* revelan que los autores de estos ejemplos suponen que el lector, mediante el contexto, sabrá de qué tipo de director se está hablando ya que no especifican, por ejemplo, que es un director “de orquesta” o “de cine” etc.:

La productora Fourgrounds ha realizado un vídeo divertidísimo donde imagina cómo son las navidades en la mente de directores como Steven Spielberg, Wes Anderson, Martin Scorsese, Lars Von Trier, Stanley Kubrick, Woody Allen, Sergei Eisenstein, Werner Herzog, Michael Moore o Baz Luhrmann.

Si a la mayoría de músicos se le preguntase cuál es el director más importante de la actualidad, seguramente muchos de ellos, si no la mayoría, responderían que Claudio Abbado.

De esta manera, los autores suponen que el contexto le ayudará al lector a ligar la palabra textual a un elemento de significado ya establecido y corriente. En consecuencia, los elementos de significado que cumplen con los criterios establecidos son necesarios en el diccionario general y monolingüe para la recepción de textos, cada uno debiendo aparecer separadamente para ofrecer la mejor ayuda posible al usuario.

5.4 Los elementos de significado de los diccionarios y del corpus

Como ya he mencionado, evidentemente, ha sido posible seleccionar más elementos de significado mediante el AC que los que están incluidos en los diccionarios. Pero la diferencia numérica no significa que se hayan seleccionado los mismos elementos de significado ni que se hayan

seleccionado los más relevantes por lo que estas diferencias merecen un análisis más profundo para obtener una imagen más global.

La figura B ejemplifica el método utilizado para las siguientes comparaciones:

Lema	N°	CLA	DEA	DRAE	MAR	VOX	Análisis de corpus
orgullo	1	Exceso de estimación propia o sentimiento que hace que una persona se considere superior a los demás: <i>Su orgullo le impide pedir perdón.</i>	Alta estima de sí mismo o de las cosas propias, frec. con sentimiento de superioridad sobre los demás. \\ Laforet <i>Mujer</i> 288: Miguel era más guapo y más alto que el chico de Joaquín, que tenía su edad. Eso colmaba de orgullo a Mariana. Asensi <i>Cotón</i> 25: Uno de mis grandes pecados es el orgullo, lo reconozco.	Arrogancia, vanidad, exceso de estimación propia, que a veces es disimulable por nacer de causas nobles y virtuosas.	<i>reprobatorio (Tener; de)</i> Sentimiento y actitud del que se considera superior a los otros y les muestra desprecio o se mantiene alejado de su trato: 'Tiene orgullo de clase'. ~Altanería, altivez, arrogancia, ~soberbia.	Exceso de estimación hacia uno mismo y hacia los propios méritos por los cuales la persona se cree superior a los demás ej los hijos, a los que entregó su vida, ahora la pagan con ingratitude, desatención y orgullo.	sentimiento de una persona de que tiene gran valor lo que causa que no quiere hacer algo específico porque cree que perdería algo de este valor haciéndolo; suele combinarse con el sentimiento de valer más que los demás y estarles superior
	2	Satisfacción grande que siente una persona por algo propio que considera muy bueno o digno de mérito: <i>Enseñaba con orgullo la foto de sus nietos.</i>			<i>no reprobatorio (Sentir, Tener; por)</i> Sentimiento de *satisfacción de alguien por cosas propias a las que atribuye mérito o por cualidades propias que considera superiores a las de otros: 'Siente un legítimo orgullo por su obra. Sus hijos le llenan de orgullo'.	Sentimiento de satisfacción hacia algo propio o cercano a uno que se considera meritorio ej el hasta ayer obispo se despidió en un sencillo acto en el que expresó su orgullo por haber presidido la diócesis.	sentimiento que tiene una persona o grupo de personas de que algo, con que tiene que ver, vale mucho, que siente alegría de tener algo que ver con ello y que quiere que otros lo sepan; por ejemplo puede ser animado porque la persona ha producido la cosa en cuestión, porque ha ayudado hasta llegar a su éxito o porque es parte de las propias cualidades de la persona
	3	Amor propio, valoración y estima que se tiene uno mismo, esp. como merecedora de un mínimo respeto o consideración: <i>No le ofrecí dinero para no herir su orgullo.</i>				Amor propio o pundonor ej tiene todavía un resto de orgullo y de sentido de lo que se debe hacer; ha peleado hasta el final y ha perdido, pero con la cabeza alta, con orgullo y con una clase de la que muchos dudaban hace muy pocos meses.	sentimiento de tener uno mismo mucho valor, de amarse a sí mismo, de que otros debían respetarle a uno y de que uno no puede equivocarse sin perder algo de este valor
	4		Pers. o cosa que es causa de orgullo [1] [para alguien] (<i>compl de posesión</i>). \\ <i>Salom Vuelo</i> 149: Eres el orgullo de la familia. <i>PReverte Trofalgar</i> 36: El majestuoso Santísima Trinidad ... orgullo de la Marina española.			Persona o cosa que produce este sentimiento ej el orgullo de la familia; los templos representaban para los griegos la culminación de su mundo espiritual y el orgullo mayor de su creación artística.	cosa, circunstancia o persona de la que alguien se siente orgulloso, es decir, que siente que esa persona, circunstancia o cosa tiene valor y por ello se siente contento y lo manifiesta para que otros lo sepan
							suceso que consiste en que un grupo de personas se reúne para manifestar su apoyo a los derechos de los homosexuales, bisexuales, transexuales y personas con otras sexualidades alternativas a la hetero; normalmente, incluye discursos políticos, desfiles por las calles y fiestas con música y baile
Total	4		3	2	1	2	4

Figura B – Ejemplo del sistema metodológico de los grupos de elementos de significado

En la columna derecha, se hallan los resultados del AC ordenados según los grupos de elementos de significado que ya se habían establecido en el análisis preliminar del contenido de los diccionarios. Como se ve, cada diccionario tiene su propia columna con todos los datos de cada elemento de significado. En este caso, se observa que DEA y el AC tienen los elementos de significado 1 y 4 en común pero que el AC además tiene los elementos 2, 3 y 5, no presentes en DEA.

El esquema de la figura B demuestra también otro asunto interesante identificado durante otra parte de mi investigación no presentada aquí, a saber el hecho de que existen grandes diferencias en cuanto a los elementos de significado que se han estimado relevantes en el mismo tipo de diccionario con la misma función. Aunque el tamaño de este artículo no nos deje profundizar este tema más, la observación en sí nos indica que hay lagunas de significado en los diccionarios y acentúa las siguientes conclusiones.

5.4.1 CLA versus el análisis de corpus

Para los 23 lemas relevantes, CLA ha seleccionado 58 elementos de significado frente a los 191 del AC. En la figura C, se ve la distribución de los elementos de significado únicamente seleccionados en el AC (parte azul de las columnas), los únicamente seleccionados por CLA (parte amarilla) y los que coinciden en el diccionario y los resultados del AC (parte rayada):

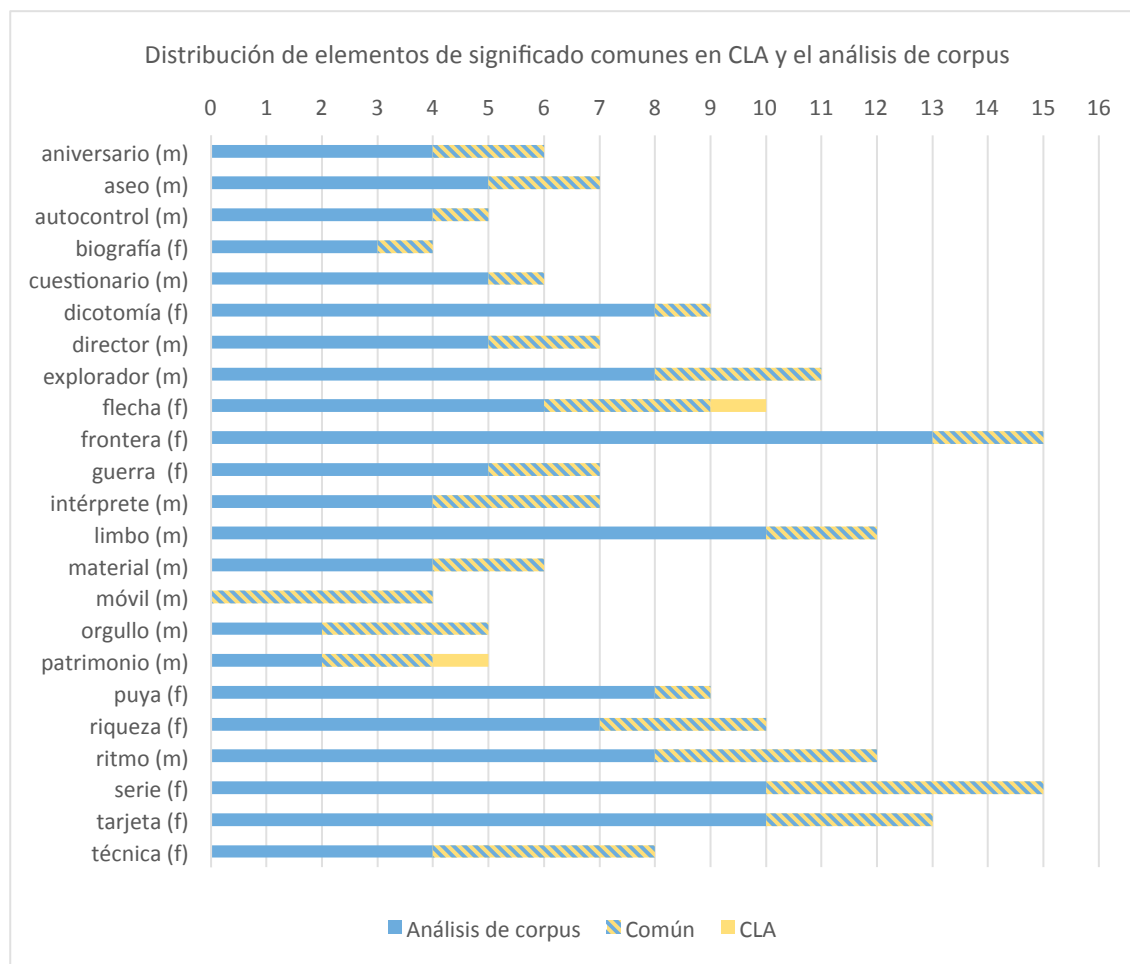


Figura C – Distribución de elementos de significado comunes en CLA y el análisis de corpus

Se han seleccionado entre 2 y 13 elementos de significado más que CLA para 20 de los 23 lemas. En cuanto a *móvil*, se han seleccionado los mismos elementos, siendo solo dos las veces que CLA ha seleccionado un elemento que no se seleccionó en el AC, precisamente para *flecha* y *patrimonio*. El primero, de *flecha*: “En zonas del español meridional, intermitente de un vehículo: *Siempre que se vaya a adelantar, hay que utilizar la flecha*“, ni siquiera ha sido identificado en el corpus y, por lo tanto, queda difícil determinar si esta falta del AC se debe a un corpus deficiente o un lexicógrafo no calificado o si este elemento de significado ya forma parte de otro elemento de significado explicado así en el AC:

“signo que indica algo específico con la dirección de su punta; típicamente, tiene forma de un cuadrado o línea en cuyo extremo delgado está un triángulo cuya punta está en el extremo opuesto a la otra figura; sirve, por ejemplo, para marcar la existencia y posición de algo, para indicar la dirección en que va algo o para hacer visible alguna cosa especial entre muchas otras cosas”

El segundo para *patrimonio* (“En economía, diferencia entre los valores económicos que pertenecen a una persona o a una entidad y las deudas u obligaciones de que responde: *Las últimas deudas contraídas por la empresa hicieron disminuir notablemente su patrimonio*”), tampoco ha sido identificado en el AC y, aunque se encuentra también en VOX, según fuentes económicas, simplemente parece explicar un elemento de significado para el lema *patrimonio neto* (véanse p.ej. Martín Mateos 2014 o Blanco Richart 2008). Por tal motivo, se estima que estos elementos no son necesarios en nuestro tipo de diccionario.

En resumen, las lagunas de significado de CLA son muy notables e indisputables.

5.4.2 DEA versus el análisis de corpus

En total, DEA ha seleccionado 91 elementos de significado. En la figura D, se ve la distribución de los tipos de elementos en comparación con los del AC:

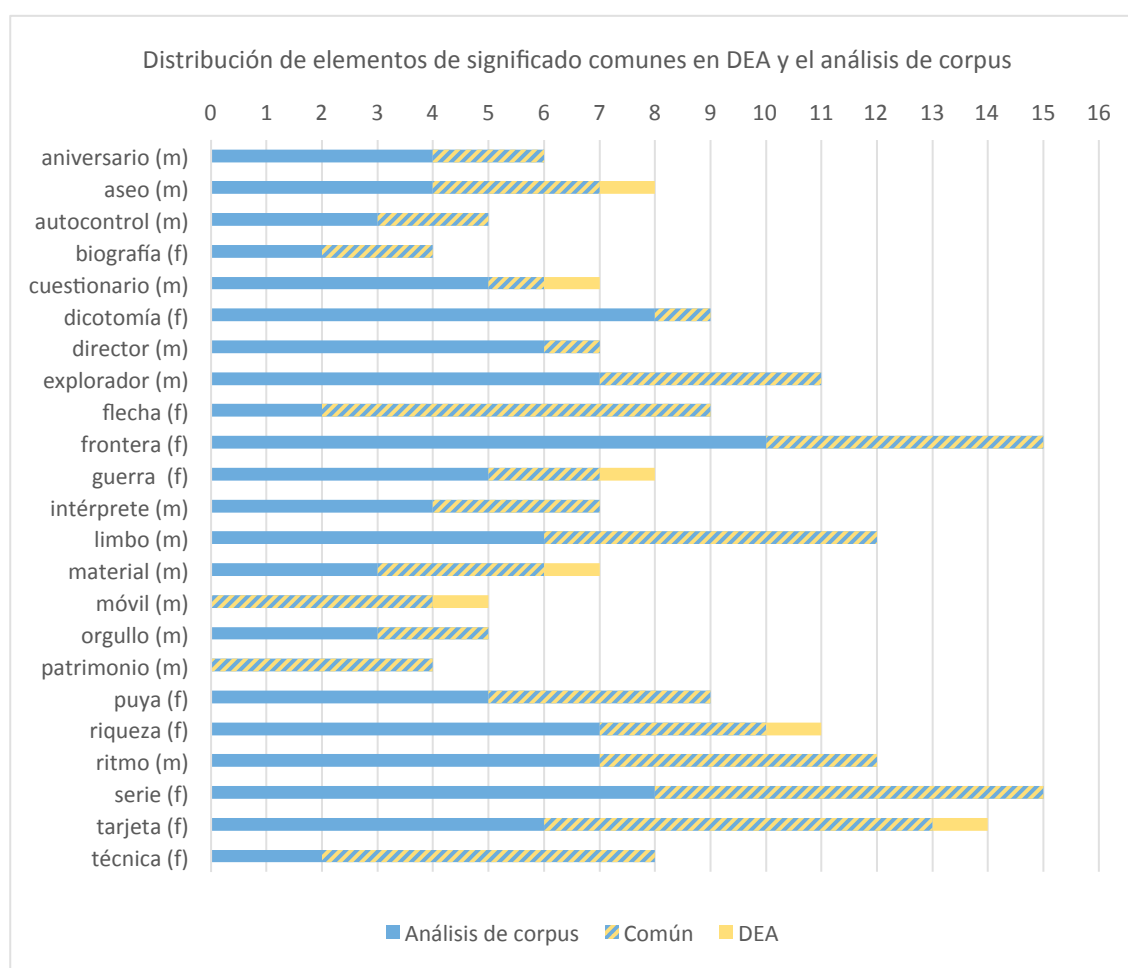


Figura D – Distribución de elementos de significado comunes en DEA y el análisis de corpus

De ahí se desprende que para 15 de los lemas en el AC se han seleccionado los mismos elementos de significado que los contenidos en DEA además de entre 2 y 10 no incluidos en DEA. En cuanto a

patrimonio, coinciden DEA y el AC y, en siete casos, ha seleccionado DEA un elemento de significado no seleccionado en el AC. Cuatro de ellos ni siquiera se identificaban en el AC:

- *guerra*: “(col) Alboroto grande. *Frec en constr* como **amarse la ~** o **ser la ~**. [...]”
- *material*: “Cuero curtido. *Nar* 6.77, 11: Los instrumentos auxiliares que utiliza [el alfarero] son: "tiradera" .. "Alpañata". Trozo de material para suavizar las paredes” (también en DRAE y MAR)
- *móvil*: “(hoy raro) [Sello o timbre] que se pega en el papel. *Tb n m.* *Zvicente Traque* 130: Luego, la factura. Eso sí, con detalle y timbre móvil” (también en MAR y VOX)
- *riqueza*: “Cualidad o condición de rico (que tiene mucho dinero o bienes)” (también en MAR)

Otra vez, esta falta puede deberse al corpus, al lexicógrafo o un ejemplo de la selección de elementos de significado anticuados, periféricos o irrelevantes por parte de DEA. En cuanto a *material*, también puede ser que el significado se haya cambiado y ampliado hasta comprender textil o cosa parecida empleada en la fabricación de ropa, elemento de significado que sí se incluye en el AC.

Dos elementos de significado solo aparecían en textos de los años 1960 y 1970, hecho que los hacen innecesarios en nuestra selección contemporánea:

- *cuestionario*: “Programa o conjunto de temas que se establece obligatoriamente para un curso de una asignatura o para un examen. *Bachillerato* 1967 24: El Ministerio de Educación y Ciencia publicará los cuestionarios de cada asignatura” (también en MAR)
- *tarjeta* som “(hoy raro) Pieza gralm. Pequeña y rectangular de cartulina o plástico, en que se almacena información mediante perforaciones o codificación magnética.[...]”

El último elemento de significado único de DEA es para *aseo*: “*En un lugar público*: Servicio o retrete. *Umbral Mortal* 65: Algunas veces me quedo dentro de unos servicios públicos, aseos los llaman en algunos sitios”. Como era el caso para CLA, parece que este elemento ya está abarcado por la explicación de DEA de “Cuarto de aseo [4] *Frec* contrapuesto a BAÑO o CUARTO DE BAÑO [...]”, pero si de verdad queda una diferencia lo tomaré por mi cuenta.

Sin embargo, la diferencia queda bastante grande, indicando así lagunas de significado en DEA y, tal vez, en el AC.

5.4.3 DRAE versus el análisis de corpus

La distribución de los 76 elementos de significado y los 191 del AC se ve en la figura E:

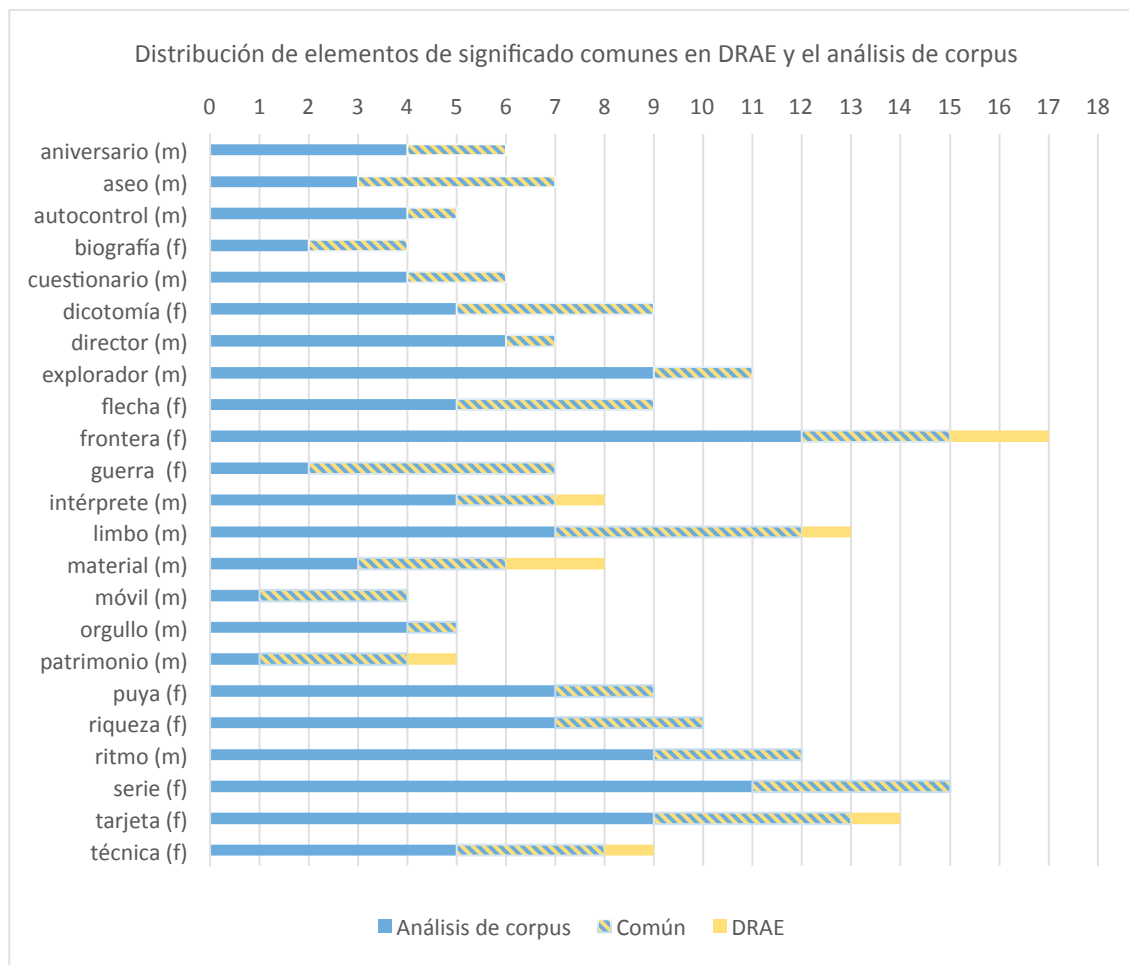


Figura E – Distribución de elementos de significado comunes en DRAE y el análisis de corpus

Como se puede observar, se han seleccionado más elementos de significado en el AC que los comunes para todos los lemas, la cantidad extra alcanzando entre 1 y 12 elementos. Para 16 lemas, el AC no tiene competencia pero en los siete casos restantes, DRAE sí ha seleccionado nueve elementos de significado que no se incluyen en el AC. Otra vez, seis de ellos sencillamente no se identificaban en el AC, probablemente, por las razones ya mencionadas:

- *intérprete*: “Cosa que sirve para dar a conocer los afectos y movimientos del alma”
- *limbo*: “Borde de una cosa, y especialmente orla o extremidad de la vestidura” (también en MAR)
- *material*: “Cuero curtido” (también en DEA y MAR)
- *material*: “Elemento que entra como ingrediente en algunos compuestos”
- *tarjeta*: “Membrete de los mapas y cartas” (también en MAR)
- *técnica*: “Habilidad para ejecutar cualquier cosa, o para conseguir algo” (también en MAR)

Otros tres de ellos se pueden considerar anticuados o periféricos porque sí aparecían en el AC pero con pocos *tokens* y en textos del siglo XVIII:

- *frontera*: “Cada una de las fajas o fuerzas que se ponen en el serón por la parte de abajo para su mayor firmeza” (también en MAR)

- *frontera*: “frontis (|| fachada)” (también en MAR)
- *patrimonio*: “Conjunto de los bienes y derechos propios adquiridos por cualquier título” (también en MAR)

Tal vez, existan estas lagunas de significado en el AC pero también se podría preguntar si los elementos de significado extra de DRAE son útiles y necesarios en un diccionario contemporáneo.

5.4.4 MAR versus el análisis de corpus

MAR toma el primer puesto con sus 97 elementos de significado y la distribución comparativa con el AC se ve en la figura F:

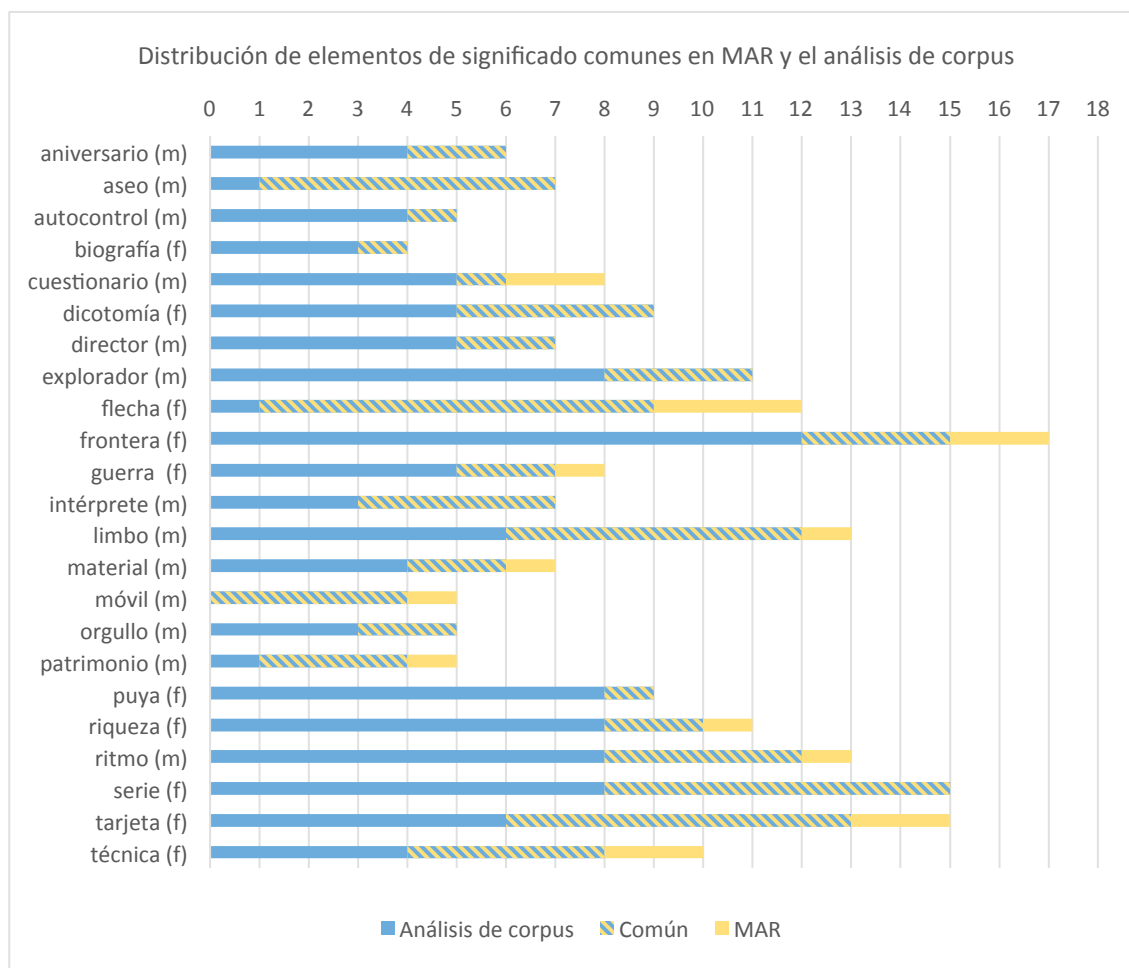


Figura F – Distribución de elementos de significado comunes en MAR y el análisis de corpus

Para 11 de los 23 lemas, no se han seleccionado elementos de significado en MAR que no sean *comunes* y, al mismo tiempo, se han seleccionado entre 1 y 8 elementos extra en el AC. Solo para *móvil*, el AC no ha llevado a más elementos que los comunes, caso en que MAR ha seleccionado el elemento de significado ya tratado arriba de “Se aplica a los *sellos o timbres que no se estampan en el papel, sino que están impresos y se pegan en él”. En el resto de los casos, MAR incluye entre

1 y 3 elementos de significado no seleccionados en el AC en tanto que, contrariamente, el AC contiene entre 1 y 12 elementos no presentes en MAR.

Ya se han tratado los nueve de los 17 elementos de significado no seleccionados en el AC:

- *cuestionario*: “Relación de las cuestiones que comprende un curso de cierta materia, de temas para unas oposiciones, etc. ~ Programa”
- *frontera*: “Arq. Frontispicio” y como ” Se aplica a las fajas con que se refuerzan las *seras y *espuertas por debajo”
- *limbo*: “*Borde de una cosa; particularmente, de una vestidura talar”
- *material*: “Hablando de *calzado, *cuero”
- *patrimonio*: “Bienes adscritos a un ordenando como título para su ordenación”
- *riqueza*: “Cualidad de *rico”
- *tarjeta*: ”Membrete de mapas, cartas, etc.”
- *técnica*: “inf. A veces, se emplea con el significado de «habilidad» o «táctica»: ‘No tienes técnica para convencer al jefe’”

Seis de los demás elementos ni siquiera se han identificado en el AC:

- *cuestionario*: “Lista de temas o cuestiones que hay que tratar, por ejemplo en una reunión” (también en VOX)
- *flecha*: “Fort. Obra compuesta de dos caras y dos lados que se construye en tiempo de guerra, para dificultar los aproches, en los ángulos entrantes y salientes del glacis”, “Bordado que se hace al final de una costura o de una abertura de una prenda, para que no se descosa o abra más” y “El mismo bordado hecho aparte en trabajo de *pasamanería, y aplicado al sitio en donde es necesario. > *Bordar”
- *guerra*: “Cierta juego de *billar”
- *técnica*: “Persona dedicada a la técnica: ‘Hay que avisar al técnico porque la caldera no funciona bien. El mundo está en manos de los técnicos’”

Otra vez, esto puede deberse a un corpus deficiente, al lexicógrafo o a la selección de elementos periféricos y anticuados por parte del diccionario, razón que parece bastante plausible considerando los primeros cuatro elementos de significado.

Uno de los dos elementos de significado restantes de MAR es *ritmo* como “Adaptación de las divisiones de que es susceptible un movimiento, una acción, una sucesión de sonidos, etc., a intervalos *regulares de tiempo”, explicación tan abstracta que este lexicógrafo no podía determinar su significado y, en consecuencia, si vendría en el grupo de otros elementos de significado sí seleccionados en el AC. El último elemento de significado “único” de MAR es *tarjeta* como “Cartulina en que va *impreso el nombre, señas, actividad, etc., de una entidad o un establecimiento o empresa, que se da a las personas a quienes puede interesar” que parece redundante ya que MAR también tiene “Tarjeta de visita: ‘Me dio su tarjeta’” para el mismo lema.

De todos modos, estos datos indican que incluso MAR tiene lagunas de significado de cierta medida que excede la extensión de las lagunas del AC. Los datos también revelan que quizá se pudiera eliminar algunos elementos de significado periféricos de MAR.

5.4.5 VOX versus el análisis de corpus

VOX ha seleccionado 75 elementos de significado para los lemas en cuestión. La distribución de su coincidencia con los elementos del AC está ilustrada en la figura G:

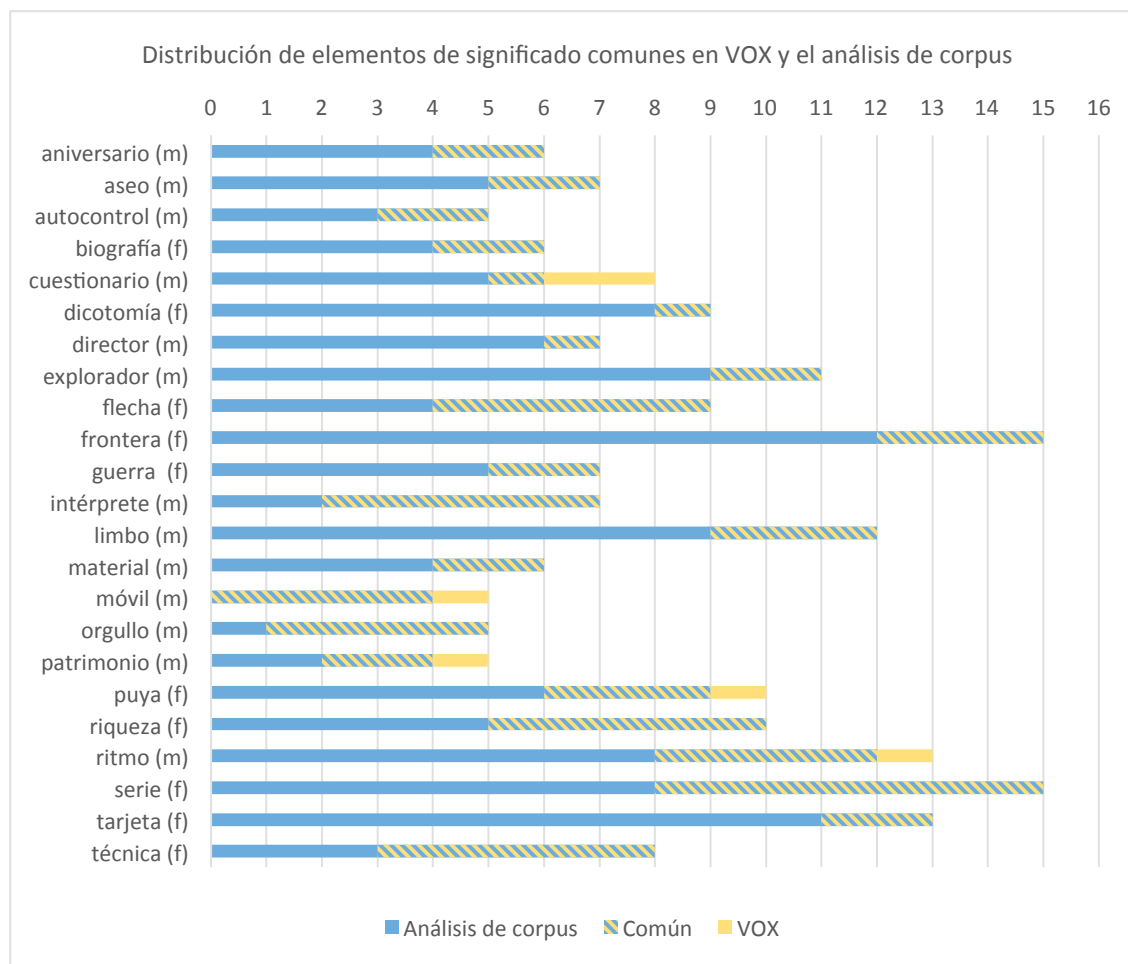


Figura G – Distribución de elementos de significado comunes en VOX y el análisis de corpus

Para 18 de los 23 lemas, VOX solo incorpora los elementos de significado que también se incluyen en el AC, casos en que se han seleccionado entre 2 y 12 elementos extra en el AC.

Como antes, en el AC no se identificaban cinco de estos seis elementos extra seleccionados por VOX, esto es:

- *cuestionario*: "Lista o relación de las cuestiones que se han de tratar **ej** el cuestionario de una reunión" (también en MAR) y "Documento impreso o mecanografiado donde se recogen cuestiones **ej** las contestaciones a los 1 200 cuestionarios enviados hace más de un mes para la selección previa aún se están clasificando"
- *móvil*: "[sello, timbre] Que está impreso y se pega en el papel **ej** estaba estampando un móvil en el primero de los papeles; gastaba todo el día rodando de ventanilla en ventanilla,

de mesa en mesa, pidiendo firmas y avales, rellenando impresos, pegando sellos móviles y dando las gracias” (también en MAR y DEA)

- *puya*: “Herida o daño producidos por este golpe. **sin** puyazo”
- *ritmo*: “Alegría, entusiasmo o vigor con que se realiza determinada cosa **ej** a este programa le falta un poco de ritmo; la lambada es un baile con mucho ritmo”

Y como ya se explicó para CLA, el último, es decir, *patrimonio* como “*economía* Diferencia entre los valores económicos pertenecientes a una persona o a una empresa y sus deudas u obligaciones” parece incorrecto.

Por falta de más datos empíricos sobre estos elementos de significado resulta difícil determinar si de verdad representan lagunas de significado del análisis de corpus o si representan elementos anticuados o periféricos. De todos modos, como se ve en la figura G, VOX tiene lagunas de significado.

6. Conclusiones

Desde la punta de vista de la teoría funcional de la lexicografía que mantiene que un diccionario debe incluir todos los datos que pueda necesitar el tipo de usuario previsto, lo que en este caso, idealmente, son todos los elementos de significado presentes en textos de conocimientos generales, la indicación de este estudio de cinco diccionarios generales y monolingües de español destinados a la recepción de textos es que todos ellos sí tienen lagunas de significado.

La falta de datos necesarios y relevantes para el usuario previsto comprobada en esta investigación, se debe, probablemente, al hecho de que los diccionarios de papel siempre se han sentido restringidos por su formato que no debe ser muy voluminoso por razones comerciales y prácticas. Sin embargo, esto no explica su inclusión de elementos de significado anticuados o periféricos en vez de elementos perfectamente actuales y necesarios, hecho que hemos observado en esta investigación. De este modo, los resultados de la investigación motivan la discusión sobre qué elementos de significado son realmente relevantes en este tipo de diccionario con este objeto genuino y sobre la necesidad de los elementos de significado ya incluidos en los diccionarios que, a veces, parecen ser muy periféricos. Otra cuestión es si ¿todavía se puede defender estas restricciones con las tecnologías actuales, teniendo en cuenta que cuatro de los diccionarios ya utilizan los medios digitales que, en realidad, no tienen las limitaciones del diccionario de papel? Con la competencia creciente de Wikipedia y los diccionarios gratuitos de Internet, ¿no es tiempo de repensar el contenido y el formato del diccionario moderno para asegurar su supervivencia? El formato electrónico de estos cuatro diccionarios sí que abre la puerta a una cantidad de elementos de significado más adecuada y adaptada a las necesidades de los usuarios previstos, solo hay que renovarse y explotar estas grandes posibilidades que ofrezca el formato electrónico. De esta manera, los diccionarios podrían seguir siendo relevantes y útiles para sus usuarios, a beneficio de su propia economía pero a cierto también de los usuarios y profesiones que dependen tanto de ellos.

7. Bibliografía

- BERGENHOLTZ, Henning (1996): «Korpusbaseret leksikografi», *LexicoNordica* 3, pp. 5-18.
- BERGENHOLTZ, Henning (2012): «What is a Dictionary?», *Lexikos* 22, pp. 20-30.
- BERGENHOLTZ, Henning y Heidi AGERBO (2014a): «Extraction, selection and distribution of meaning elements for monolingual information tools», *Lexicographica* 30, pp. 488-512.
- BERGENHOLTZ, Henning y Heidi AGERBO (2014b): «Meaning identification and meaning selection for general language monolingual dictionaries», *Hermes. Journal of Language and Communication in Business* 52, pp. 125-139.
- BERGENHOLTZ, Henning y Heidi AGERBO (2014c): «There is no need for the terms polysemy and homonymy in lexicography», *Lexikos* 24, pp. 1-9.
- BERGENHOLTZ, Henning y Bjarni NORDDAHL (2012): «Ordbogsartikler, som ingen læser», *LexicoNordica* 19, pp. 207-223.
- BLANCO RICHART, D. E. R. (2008): *Contabilidad y fiscalidad: Capítulo 2: El patrimonio empresarial*. <http://www.eumed.net/libros-gratis/2008b/396/El%20Patrimonio%20Empresarial.htm> (27-11-2014).
- DACOSTA ESTEBAN, Joaquín, ed., (2008): *MARÍA MOLINER Diccionario de uso del español – Edición electrónica*. Madrid: Editorial Gredos, S.A.U.
- EDICIONES SM (2012): *Diccionario Clave: Diccionario de uso del español actual*. <http://clave.smdiccionarios.com/app.php> (16-09-2014).
- GOOGLE INC. (2014): *Búsqueda de Google*. <https://www.google.es/> (09-12-2014).
- GUDMANN, Helene R. (2014): *Betydningshuller i spanske ordbøger. En undersøgelse af betydningsenheder i spanske monolingvale almene receptionsordbøger*. (Tesina de Máster). Aarhus: Aarhus Universitet, BSS, Institut for Erhvervskommunikation. <http://pure.au.dk/portal-asb-student/da/studentprojects/betydningshuller-i-spanske-ordboeger%28a5ac59ac-5924-4c7a-b9d3-bf6b7da19f05%29.html> (16-02-2015).
- HARTMANN, R. R. K. y Gregory JAMES (1998): *Dictionary of lexicography*. London y Nueva York, Routledge.
- HJELMSLEV, Louis (1943): *Omkring sprogteoriens grundlæggelse*. Copenhagen, Bianco Lunos Bogtrykkeri.
- KILGARRIFF, Adam & Gregory GREFENSTETTE (2003): «Introduction to the Special Issue on the Web as Corpus», *Computational Linguistics* 29, pp. 333-347.
- MARTÍN MATEOS, Raquel (2014): *Patrimonio neto*. <http://www.expansion.com/diccionario-economico/patrimonioneto.html> (27-11-2014).
- REAL ACADEMIA ESPAÑOLA (2014): *Diccionario de la lengua española*. <http://www.rae.es/recursos/diccionarios/drae> (17-09-2014).
- SECO, Manuel, Olimpia ANDRÉS y Gabino RAMOS (2011): *Diccionario del español actual*, 2nd edition. Madrid, Aguilar, Grupo Santillana de Ediciones.
- TARP, Sven (2009): «Homonymy and polysemy in a lexicographic perspective», *Zeitschrift für Anglistik und Amerikanistik* 57 (3), pp. 289-306.
- TARP, Sven (2013): «Necesidad de una teoría independiente de la lexicografía: El complejo camino de la lingüística teórica a la lexicografía práctica», *Círculo de Lingüística Aplicada a la Comunicación* 56, pp. 110-154.
- VOX (2003): *Diccionario de uso del español de América y España*. Barcelona, SPES.