



**Las nuevas tecnologías y la lexicografía
digital: Fundamentación metodológica para
la elaboración de un diccionario electrónico
de fraseologismos español-árabe**

By

Dr. Tarek Shaban Mohammad Salem

*Department of Spanish and Literature, Faculty of
Languages and Translation
Al-Azhar University, Cairo, Egypt*

New Technologies and Digital Lexicography: Methodological Foundation for The Elaboration of a Spanish-Arabic Electronic Dictionary of Phraseological Units

Tarek Shaban Mohammad Salem

Department of Spanish Language, Faculty of Languages and Translation, Al Azhar University, Cairo, Egypt.

E-mail: TarekSalem579.el@azhar.edu.eg

ABSTRACT: During the last few years, the Internet has become the key element of most of the personal and academic aspects of our lives. In the 21st century, the development of technology was enough to raise the construction of electronic dictionaries that can be used intelligently. For this reason, lexicographical creation has changed radically. So, we live in a time where digital techniques allow us to take important steps towards dictionaries that respond to the specific needs of users. This work aims to elaborate the design of what could become the first bilingual Spanish-Arabic electronic dictionary of idiomatic phraseological unit. The objective is, in addition to addressing a gap in the field of phraseology of the Spanish-Arabic pairing, to show some guidelines for the representation of phraseology in digital tools as well as to present a source of cultural data for learners of Spanish as a foreign language.

Keywords: Electronic Dictionary, Technology, Phraseology, Lexicography, Spanish-Arabic

التقنيات الجديدة والمعجمية الرقمية: تأسيس منهجي لإعداد قاموس إلكتروني إسباني-عربي للتعبيرات الاصطلاحية

طارق شعبان محمد سالم

قسم اللغة الإسبانية، كلية اللغات والترجمة، جامعة الأزهر، القاهرة، مصر.

البريد الإلكتروني: TarekSalem579.el@azhar.edu.eg

ملخص: على مدى السنوات القليلة الماضية، أصبح الإنترنت العنصر الأساسي في معظم الجوانب الشخصية والأكاديمية في حياتنا. وفي القرن الحادي والعشرين، كان التطور التكنولوجي كافياً لاقتراح صناعة قواميس إلكترونية يمكن استخدامها بذكاء. ولهذا، تغيرت صناعة المعاجم بشكل جذري. لذلك، نحن نعيش في عصر تتيح لنا فيه التقنيات الرقمية اتخاذ خطوات مهمة نحو القواميس التي تستجيب للاحتياجات المحددة للمستخدمين. ويهدف هذا العمل إلى تطوير تصميم ما يمكن أن يصبح أول قاموس إلكتروني ثنائي اللغة إسباني-عربي للتعبيرات الاصطلاحية. ويكمن الهدف من ذلك، بالإضافة إلى معالجة فجوة في مجال التعبيرات الإسبانية والعربية، إلى تأسيس بعض المبادئ المنهجية لتقديم التعبيرات الاصطلاحية من خلال الأدوات الرقمية، وكذلك تقديم مصدر للمعلومات الثقافية لمتعلمي اللغة الإسبانية كلغة أجنبية.

الكلمات المفتاحية: قاموس إلكتروني، تكنولوجيا، تعبيرات اصطلاحية، علم المعاجم، إسباني-عربي

1. Introducción

A comienzos del siglo XXI, la tecnología ya determina una parte esencial de nuestra existencia con un impacto y desarrollo nunca experimentado (Cañete González, 2006: 82). Este desarrollo lo podemos notar en el proceso de transformación de todo aquello que nos rodea. A modo ilustrativo, pasamos de la clase presencial a la de *online*, del libro en papel al electrónico, de la carta al correo electrónico, etc. Así que, Internet ya no se limita al cambio y desarrollo de determinados usos, sino que ha podido afectar a todos los aspectos de nuestra vida diaria, tal y como lo confirma Casanovas Catalá (2016: 78) en que “Internet se ha convertido en un elemento clave en muchos aspectos de nuestra vida, personal y académica”.

Al respecto, y en cuanto a que Internet ya ha podido tener relación estrecha con la investigación lexicográfica, existe un campo abierto de reflexión, donde el peso de una sólida tradición lexicográfica debe ser permeable a nuevas posibilidades de construcción, donde se dé más espacio a las necesidades de los usuarios. Por eso, este campo se abre ante el soporte electrónico para la creación de nuevos diccionarios y adaptación de los modelos existentes. Igualmente, el acceso generalizado al mundo digital y las expectativas de los usuarios de herramientas libres y abiertas han conducido al surgimiento de nuevos proyectos lexicográficos.

Por lo dicho, podemos decir que la lexicografía actual está atravesando un momento decisivo de cambio que afecta directamente a la elaboración de diccionarios por ayuda de herramientas de la tecnología que contienen un conjunto de características que permiten satisfacer específicamente las aspiraciones de los usuarios. Así que, la incorporación de la tecnología en la producción lexicográfica ha provocado la aparición de nuevos conceptos que se abren paso en los medios electrónicos (Huete García, 2019: 220).

Por otro lado, la habilidad lingüística en general se mide por la capacidad y fluidez del hablante en el uso de las expresiones conocidas como unidades fraseológicas (*UFs*). Por consiguiente, el uso apropiado de estas expresiones se puede considerar una señal del buen conocimiento de la lengua. Así que, la fraseología representa un elemento fundamental en el patrimonio cultural de cualquier sociedad y forma parte del que llaman “ámbito de la tradición oral y de las particularidades lingüísticas” (Aznárez Mauleón y Santazilia Salvador, 2016: 14).

Esto se debe a que la fraseología consiste en expresiones fijas transmitidas como tales de generación en generación y, en total, pertenece a lo que Coseriu (1981:

297) denomina *discurso repetido*. Es de aludir a que los estudios en este campo; desde perspectiva sincrónica y contrastiva, son cada vez más frecuentes.

Hemos de comprender que un diccionario no se trata de producto meramente lingüístico, sino también es una obra ideológica, social, cultural e histórica. Al respecto, coincidimos con Tarp (2019: 8) en que la aplicación de las nuevas tecnologías tiene grandes consecuencias para la lexicografía como práctica cultural. De ahí que, un diccionario para estudiantes de español como lengua extranjera/segunda lengua (L2) debe ser una herramienta que contenga información sobre los componentes culturales de la lengua, ya que el aprendiz del español le importa aprender cultura junto con la lengua.

En este sentido, y para los usuarios no nativos, el empleo de estas expresiones supone un reto, ya que contiene una enorme complejidad y conserva un saber sociocultural y lingüístico que suele ser diferente a la lengua materna. A este respecto, Corpas Pastor (2003: 337) apunta acertadamente que la competencia del elemento fraseológico se concibe como uno de los indicadores más claros del uso nativo de un idioma. Por eso, y en comparación con los usuarios nativos, estas expresiones introducen conocimientos lingüísticos, sociales, culturales, comunes y conservados en su memoria a través de la experiencia; por lo cual, representan “parte integrante del lexicón mental de cualquier hablante nativo” (Zuccalà, 2015: 271).

En la actualidad, merecen especial indicación que la fraseología se construye como una rama multidisciplinar con peculiaridades propias. De cara a la comprensión de la temática que vamos a tratar en este trabajo, importa aún señalar un aspecto de esta disciplina: se ocupa de las *UFs* en sentido estricto: básicamente locuciones, como en sentido amplio: colocaciones, aforismos, refranes, fórmulas, etc., (Ruiz Gurillo, 1997: 56-63). Para delimitar el proceso de selección de las *UFs*, hemos expuesto el primer criterio para ello, basándonos en el punto de vista funcional. Coincidimos con los planteamientos de Manuel Seco *et al.* (2004: XIII), a pesar de que existen algunas divergencias en casos concretos. De ahí, en este trabajo, el concepto de fraseología se entiende como no podía esperarse de otro modo; es decir, desde una concepción estrecha que trata de las locuciones.

Históricamente hablando, pese a que la informática surgió como resultado del trabajo conjunto de matemáticos, estadísticos, ingenieros y sociólogos (Ortego Antón y Fernández Nistal, 2015: 44), pero poco a poco afectó a todas las especialidades, sobre todo las ciencias de humanidades. Por consiguiente, la revolución tecnológica afectó directamente a la filología en general y a la lexicografía en particular que no pudo permanecer estática en relación con los desarrollos que se están produciendo,

por lo que tuvo que adaptarse a las nuevas realidades tecnológicas. De hecho, “la llegada de Internet ha tenido varias repercusiones importantes en el campo de la lexicografía” (Fuertes-Olivera y Bergenholtz, 2018: 6).

Por esa razón, el lenguaje humano está en constante actualización; dado que el léxico es el que representa el reflejo como la parte más mutable de la lengua. Por tanto, los repertorios léxicos exponen una realidad que no puede ser limitada por la rigidez inmutable y fija de la letra impresa. De ahí, y con vista de la revolución de la informática, creemos que un diccionario en línea es el recurso más adecuado para gestionar el conocimiento que está en constante cambio. En tal sentido, coincidimos en que “las obras lexicográficas deben caminar a la par de los enormes progresos realizados por la ciencia y la tecnología” (Llanillo Gutiérrez, 2008: 715).

Al otro lado, es importante describir una faceta sustancial del estado actual del panorama lexicográfico español-árabe para ilustrar los objetivos globales de nuestra propuesta. La ausencia de serios materiales lexicográficos, hasta el momento, relacionados con la fraseología del binomio español-árabe nos pone ante una posición difícil a la hora de ofrecer estrategias de composición de diccionarios de fraseologismos. Si examinamos los nombres de los diccionarios existentes entre ambos idiomas, tendremos la constancia de que todavía queda un terreno por cultivar muy amplio; lo que nos llevó a enfatizar la necesidad de la confección de un diccionario electrónico para abarcar la gran diversidad de *UFs* existentes entre las dos lenguas, dependiendo de su estructura y denominación.

Hemos de señalar que el concepto del diccionario electrónico (*DE*)¹ —que ha hecho su aparición en el mundo de la lingüística— es mucho más de lo que se podría pensar, dado que está relacionado con varias aplicaciones en el campo de la lingüística aplicada para el procesamiento automático de textos. Es decir, no debe entenderse como una sencilla lista alfabética de palabras o expresiones que podría aplicarse al análisis de textos. El *DE* es una base de datos (*BD*)² que contiene una

¹ El *DE* —término acuñado por Bernard Quemada (1987: 229-242) y se refiere al diccionario y todo lo relacionado con él; por lo tanto, se puede hablar sobre *diccionarística teórica* y *diccionarística práctica* que experimenta cambios radicales en comparación con los métodos tradicionales— es una obra dinámica que ofrece múltiples posibilidades de consulta. Este tipo de diccionarios tiene tres componentes principales: una *BD* lexicográfica; uno o varios motores de búsqueda y una o varias interfaces (Fuertes-Olivera y Bergenholtz, 2018: 5).

² El concepto de *BD* no significa corpus, ya que hay quienes equiparan ambos conceptos. El corpus, según Vargas Sierra (2002: 524), es “una colección de textos del lenguaje natural, seleccionados para caracterizar el estado de una lengua o una variedad de la misma”. Así que cuando hablamos de corpus, nos referimos a textos recopilados de acuerdo con ciertos criterios con un propósito específico.

enorme cantidad de informaciones recogidas y gestionadas por la informática (Mogorrón Huerta, 2004: 388) conforme a las necesidades programadas y deseadas. Esta *BD* puede ser rica en formatos, contenidos y herramientas de consulta.

Para la creación de un diccionario, Knowles (citado en Rafel i Fontanals, 2011: 557) señala tres situaciones bien diferenciadas:

- 1) Llevar a cabo una edición ampliada de un diccionario existente.
- 2) Elaboración de un nuevo diccionario de diseño basado en material lexicográfico existente preparado para diccionarios anteriores.
- 3) Creación de un diccionario basado en nuevos enfoques lexicográficos.

Nuestro diccionario propuesto corresponde con la tercera situación, ya que no vamos a utilizar imágenes ni textos completamente directos de una obra existente para crearlo, sino es un diccionario en línea donde vamos a empezar desde el punto CERO, por lo que nuestra propuesta sería la primera que se propone en este terreno.

Consideramos que nuestro diccionario propuesto será digital y también electrónico. Los *DEs* corresponden al grupo de los diccionarios digitalizados. En tal sentido, hemos de distinguir acerca del hecho de que cualquier documento digital es electrónico. Esto quiere decir que no todo documento electrónico tiene que ser digital. Por lo que, queremos dejar patente la diferencia entre ambos términos, dado que existe un error conceptual inherente del par *digital/electrónico*. Mientras el término *electrónico* se refiere a un material en el que se almacena información, el *digital* está designando el formato de codificación de la informática (digital/analógico).

2. Las nuevas tecnologías y la lexicografía digital

En la estela de estas notas preliminares, el avance tecnológico desde la década de los ochenta del siglo pasado —aunque hay quienes señalan que la “aplicación de los ordenadores a la lexicografía se remonta a los años 60” (Rafel i Fontanals, 2011: 557)— ha abierto el camino para la aparición de *DEs* y, a partir de entonces, se desarrollaron diversos trabajos pioneros. Por lo que, no hay duda en que la lexicografía digital o electrónica ha ganado importancia durante los últimos años, prueba de ello se puede ver en “the increasing number of online dictionaries and publications focusing on this field” (Müller-Spitzer y Möhrs, 2010: 199).

Hablando de la transformación del formato papel al electrónico, queda evidente que la confección de diccionarios ha pasado por varias fases de desarrollo. En relación con ello, la primera forma de publicación electrónica en el ambiente de

la lexicografía digital fue los diccionarios en CD-ROM. Ejemplo de ello, la Real Academia Española publicó su primer *Diccionario* en soporte magnético en 1995. Se trataba de versión electrónica del diccionario en papel, considerado una forma alternativa del diccionario impreso.

Igualmente, la Real Academia de España y ASALE anunciaron en 2017 que la primera actualización de la 23ª edición del *DRAE* fue digital; con el fin de ofrecer un verdadero *DE*, con correspondientes cambios que se suponen a todos los niveles lexicográficos (Lončar y Valero, 2020: 10). La existencia de dichos diccionarios demuestra la relación entre la lexicografía y la informática en aras de la difusión del conocimiento de la aplicación de nuevas tecnologías.

De acuerdo con Ochoa Anadón (1992: 129), se puede utilizar la informática en el campo de la lexicografía para llevar a cabo dos tareas principales. La primera es de índole auxiliar donde se fomenta el quehacer de confeccionar el diccionario con el fin de agilizar las labores de control interno del material que se maneja en redacción. La segunda se trata de la redacción para obtener datos utilizando la información almacenada en la memoria. Esto aprueba la estrecha relación entre la lingüística y la informática en el trabajo de los lexicógrafos. En este sentido, Arntz y Picht (1995: 219) declaran que el tratamiento electrónico de los datos se utiliza en creciente medida. Esto contribuye a lo que podríamos llamar *renovación de la lingüística*, donde la lexicografía ocupa un lugar especial entre los trabajos lingüísticos.

Nuestro trabajo se enmarca en la *e-lexicografía* que se trata de una rama de la lexicografía dedicada a desarrollar teorías que pueden guiar el diseño y construcción de diccionarios en línea. En este aspecto, Internet es una tecnología que representa más que un nuevo medio lexicográfico dentro del campo de la *e-lexicografía*, donde hay que prestar atención a la relación de la tecnología con diferentes elementos como los usuarios del diccionario, los desarrollos lexicográficos y los costos. Por eso, vivimos en una fase donde Internet ha engendrado oportunidades de desarrollo lexicográfico que nunca podríamos imaginar para un formato en papel.

De hecho, el reflejo de las nuevas tecnologías en el desarrollo de la lexicografía se ha transformado en la integración de múltiples elementos como el procesamiento de textos, sonidos, fotografías y vídeos en un solo documento. Pero la diferencia más importante es la gestión de *BD*, ya que refleja una herramienta muy útil en el trabajo lexicográfico, donde se almacena gran cantidad de información en corpus compuesto de elementos del repertorio léxico que contiene significados, combinaciones sintácticas y semánticas, ejemplos, usos, etc. Huelga decir que la *BD*

o puede conseguirse mediante una creación nueva para llevar a cabo un diccionario *ex nihilo* o a través de recuperación de datos existentes.

3. Ventajas de un diccionario electrónico bilingüe

Podemos confirmar que, sin la aparición y el perfeccionamiento de la informática, era inimaginable que fuera posible realizar los procesos que caracterizan los *DEs*. En este apartado intentamos, además de resaltar los rasgos y finalidades de proponer un *DE* bilingüe, responder a la pregunta ¿por qué un *DE* tiene ventajas que otro en papel?

Antes de tratar las diferencias entre ambos tipos, podemos referirnos a los rasgos que hay en común, ya que encierran al menos tres características importantes. En primer lugar, los dos tipos han sido diseñados para cumplir funciones lexicográficas y otras comunicativas. Además, contienen datos lexicográficos para respaldar las funciones del diccionario. Se añade a ello que las estructuras lexicográficas vinculan los datos para apoyar y cumplir dichas funciones (Gouws, 2009: 27).

Sin embargo, la posibilidad de consultar diccionarios en medios electrónicos tiene ventajas indiscutibles. Este tipo de diccionarios (Porto Dapena, 2002: 42) es ágil de información a través del uso de algoritmos desarrollados para buscar *BDs* textuales, así como ocupa menos espacio y accesibilidad. Asimismo, las posibilidades de almacenamiento de la información son variadas y flexibles en un *DE*, por eso; la ventaja de este tipo de diccionarios es que cambia los modelos de acceso a la información, por lo que los usuarios no se pierden en la información organizada por las reglas de lematización. De igual modo, los *DEs* tienen la consecuencia de que el número de usuarios está creciendo. Este incremento, cada vez mayor, se debe lógicamente a la existencia de diccionarios en red y el uso habitual de nuevas tecnologías.

Otro rasgo del *DE* radica en el modelo de organización de la información léxica, ya que el diccionario impreso la organiza en entradas: *macroestructura*: con un esquema jerárquico, y *microestructura*: en una lista de entradas léxicas, sin demasiadas relaciones explícitas entre ellas. Además, en el diccionario en soporte informático, se puede realizar una búsqueda tanto alfabética como semántica de una determinada *UF*. Otra característica es que el almacenamiento organizado en la *BD* permite una mejor gestión de la información.

Conjuntamente, el *DE* es mucho más práctico que otro en papel. El primero permite un acceso más fácil a la información desde cualquier parte del mundo y en

cualquier momento, mientras que el diccionario impreso implica mantenerlo en las bibliotecas, consultarlo manualmente y, sobre todo, alfabéticamente. Por eso, una gran ventaja del diccionario propuesto sería su versión electrónica, por lo que posibilita una búsqueda rápida; permite la ampliación de la *BD* con locuciones y ejemplos que podrían añadir contenidos de todo lo que el lexicógrafo considere importante.

Se añade a ello el criterio de actualización, en el sentido amplio de la palabra, ya que se supone que el diccionario en papel se actualiza cada diez años, “dado que es imposible borrar, reescribir y añadir material a ellos” (Luque Durán, 1996: 97). Llamáramos este proceso en el *DE* como *fiabilidad sincrónica*, donde puede actualizarse en tiempo real y sin costes adicionales en la mayoría de los casos. Estamos de acuerdo en que el *DE* “allow constant additions to the basis of the dictionary, and the lemma list must similarly be open to new words and phrases” (Hilmar Jónsson, 2009: 260). Del mismo modo, Renau y Battaner Arias (2012: 152) afirman que con Internet dicha actualización se produce de forma rápida y, en consecuencia, útil para el usuario.

El rasgo culminante del *DE* es el hipertexto. Es una ventaja llamativa dado que como recurso técnico ofrece comodidad y rapidez. La flexibilidad de los saltos hipertextuales organiza el proceso de lematización de la *UF*, ya que —conforme a Fraile Vicente (2007: 225)— permite el acceso a ella por cualquiera de sus palabras que la conforman. Esto no impide que el hipertexto a veces tiene problemas como el retraso en la exhibición de material referenciado que puede limitar en última instancia su utilidad del hipertexto. Para más información técnica cf. Conklin (1987: 38 y ss.).

En total, y en consonancia con Gelpí (2003: 315), dichos rasgos terminarán derrotando el diccionario de papel, especialmente en lo que atañe a las restricciones de formato y posibilidades de búsqueda que posibilita el enlace hipertexto. Este tipo de diccionario requiere “una programación específica, por lo que se habla de auténticos diccionarios digitales” (Lončar y Dal Maso, 2018: 30). Sin duda, tales características (Santana Suárez, 1996: 74) se ven favorecidas por el nivel de conectividad con otras herramientas y tipos de datos. En palabras de Bosque y Barrios Rodríguez (2018: 646) “online bilingual dictionaries have become almost indispensable tools for a large number of professionals nowadays”.

4. Marco teórico-práctico de la propuesta

Cabe decir que los diccionarios español-árabe existentes carecen de criterios claros para su producción. Por lo cual, ha sido necesario desarrollar principios que

orienten y sistematicen métodos, procedimientos y recursos fraseográficos que sean flexibles y adaptables a las necesidades específicas de cada repertorio. Por eso, hemos de tener en cuenta criterios fijos para nuestra propuesta.

En efecto, la elaboración de nuestra propuesta depende de algunos factores entre ellos los criterios de usabilidad, el perfil de usuarios y los objetivos a conseguir con dicha herramienta (Roiss y Zimmermann González, 2012: 11). Además, y para la *planta* de este diccionario, en lo siguiente vamos a tratar de planificar el diccionario a través de abordar sus distintos elementos. En primer lugar, nuestra propuesta tendrá varias fases que, conforme a Bailey y Robinson (1969: 4), pueden distinguir tres pasos principales:

1. Colección de información en la que se basa el diccionario.
2. Preparación de entradas, equivalentes, selección de ejemplos ilustrativos.
3. Producción del trabajo final.

Los dos primeros pasos los podríamos denominar *fase de planificación* caracterizada por ser especulativa, teórica e intelectual donde el lexicógrafo marcará las normas y pautas que ha de seguir en la fase siguiente. Estas dos fases implican la elaboración de un plan, llamado a veces *planta*. El tercer paso se considera como realización o desarrollo del diccionario en sí. No hemos de olvidar de que la computadora, o mejor dicho la informática, afecta cada uno de estos tres pasos; dado que ya el trabajo de los lexicógrafos depende en gran medida de los recursos técnicos (Hass y Schmitz, 2010: 2).

Para el proceso de la compilación del diccionario propuesto, hemos de tener en cuenta que habrá que tratar varias fases o aspectos teóricos y metodológicos. Su planificación debe cubrir todos los aspectos de su desarrollo. Por esta razón, pensamos aplicar parámetros analíticos divididos en dos aspectos: lexicográficos y técnicos. A continuación, exponemos las particularidades y los criterios principales de estos dos aspectos del diccionario cuyo título sería: *Diccionario electrónico de fraseología bilingüe español/árabe (DEFBEA)*:

4.1.Aspectos lexicográficos

4.1.1. Diccionario pasivo

El diccionario propuesto será pasivo, es decir que parte de lengua extranjera (o L2) a lengua materna (o L1); ya que su finalidad está destinada a la comprensión de textos en una lengua extranjera (tiene la función descodificadora). Por lo que será

diccionario bilingüe unidireccional (español-árabe). Este tipo de diccionarios conserva sus ventajas, ya que —de acuerdo con Valero Fernández y Mušura (2019: 100)— tiene sus lados positivos donde “permite la posibilidad de traducir y producir activamente en la lengua meta”. Por lo dicho, concebimos nuestra propuesta como una obra de descodificación dirigida en primer lugar al usuario arabo-parlante.

4.1.2. *Locuciones*

De entre las tareas que hay que realizar es almacenar las locuciones en archivos electrónicos donde se marcan los posibles campos. Se añade a ello, la redacción, el uso de cada locución, el equivalente más adecuado y el contenido en *HTML*. Este proceso, conforme a Llanillo Gutiérrez (2008: 719), requiere planificación, así como recursos materiales y humanos. En relación con el proceso de selección de las locuciones, hemos de hacer hincapié en cuatro criterios:

- a. Criterio funcional, se refiere a diferentes categorías de los fraseologismos. Al respecto, *a priori*, las locuciones forman el corpus de nuestro diccionario propuesto, porque dicha esfera se considera como un todo. Por eso, se consignarán locuciones de tipo nominal, adjetival, adverbial, verbal, pronominal, preposicional y conjuntivo.
- b. Criterio diatópico, relacionado con las variedades geográficas, donde nos basamos en el conjunto de locuciones pertenecientes a la península española. Esto porque una misma expresión puede tener varias formas en más de un país. Ejemplo de ello la *UF cara y cruz* que se dice así en España, pero varía según otros países. Veamos que Colombia, Chile, Panamá, Perú y Venezuela se dice *cara o sello*; mientras que en Costa Rica se dice *cara o escudo*; *cara o sol* en Nicaragua; *cara o ceca* en Argentina; *escudo o corona* en Costa Rica; *sello o cruz* en Ecuador; *águila o sol* en México; *escudo o estrella* en Cuba, etc. Por lo que nos limitaremos a la variedad castellana.
- c. Criterio diafásico, se trata de las etiquetas que usan los lexicógrafos para indicar que es un elemento asociado con un cierto nivel de uso. Sin duda, la presencia de etiquetas diafásicas en los artículos lexicográficos es muy importante para el usuario, ya que permiten delimitar los distintos registros lingüísticos: formal, informal, vulgar, etc. (Boughaba, 2021: 100-103). En nuestra propuesta, proponemos la extracción de las locuciones marcadas por su inclusión a un registro formal.
- d. Criterio diatécnico, o simplemente temático, se representa a través de aquellas unidades que informan sobre a qué área pertenece la unidad léxica, es decir, si se

trata de un campo especializado. En este sentido, hemos de aclarar que, para recopilar y almacenar la información fraseológica de la *BD* de nuestra propuesta, tenemos pensado no limitarnos a la recopilación de textos literarios; porque partimos de la hipótesis de que el nuevo estilo lingüístico ha penetrado profundamente en todos los niveles lingüísticos.

En el diccionario propuesto, ofrecemos el tipo de equivalente para cada locución. Para ello, contamos con la tripartición tradicional (Lončar y Dal Maso, 2018: 31) que consta de tres niveles de equivalencia: total, parcial y nula. Merece la pena mencionar al respecto tres parámetros que, de acuerdo con Mellado Blanco (2015: 165-169), establecen el equivalente más adecuado: la extensión y estructura semántica, el componente connotativo-pragmático y la combinatoria sintagmática.

4.1.3. *Usuarios*

Podemos confirmar que la sociedad contemporánea está inmersa en la era de la informatización, donde el uso de las tecnologías y de las redes informáticas permite a las personas estar constantemente actualizadas y comunicadas con todos los fenómenos que ocurren en los diversos extremos de la información. Por eso, y en palabras de Thuillier (2014: 460), “una de las principales preocupaciones ha de ser la de lograr una mayor adaptación a las necesidades reales de los usuarios”, ya que la insistencia en dichas necesidades manifiesta una característica común de la investigación de diccionarios en línea.

Hemos de tener en cuenta los que utilizan la obra lexicográfica: los nuevos usuarios de los *DEs*, o sea la *comunidad digital*, que tienen sus propias necesidades. Así que tenemos diccionario digital y usuario digital. Un diccionario digital debería permitirle al usuario acceder a la información seleccionada atendiendo al perfil que le caracterice (Gelpí, 2004: recurso electrónico).

Sin lugar a duda, los diferentes tipos de información lexicográfica se determinan en función de las necesidades de los usuarios y de las situaciones que el diccionario pretende cubrir (Nomdedeu Rull, 2018: 293). Podemos decir que cada vez más hay usuarios que hagan consultas a diccionarios en línea, cada vez menos son los que utilicen los diccionarios en papel o en otros soportes. Por eso, y en palabras de Granger (2012: 16), “dictionaries have become quite complex tools which require specialized training”.

4.2. Aspectos técnicos

4.2.1. Programario de base de datos

A través de la base informática se gestiona la información lingüística recopilada. En esta fase, se ilustrarán recursos necesarios para exportar la *BD* a una página web (Ilie, 2014: 157). En este sentido, hemos de hacer entender que el programa aplicado para la parte informática no vaya a realizar maravillas por mucho que se quiera a la hora de gestionar la información lexicográfica, sino que su parte se trata de, según Mogorrón Huerta (2004: 388), “analizar la información que la base de datos lingüística haya recopilado”. Son dos procesos paralelos, por lo que el lingüista debe saber lo que necesita el informático, y el informático a su vez lo que desea conseguir el lingüista. Por ende, el diseño de un *DE* requiere la selección de un programa informático que permita la transferencia de datos a la red.

Al respecto, la representación de información del diccionario se realiza a través del hipertexto que puede entenderse como el proceso de organizar los datos a través de una red de nudos conectados por enlaces. De acuerdo con Persico (1992: 101), los sistemas de hipertexto permiten explorar los conocimientos de manera no lineal interactiva, presentando los datos de forma de unidades a través de enlaces entre una unidad y otra, de modo que cada usuario pueda moverse en todas las maneras posibles. Del mismo modo, los sistemas del hipertexto pueden transformar el concepto del texto en algo que se puede explorar según las necesidades del usuario.

Así que, el proceso de búsqueda puede comenzar en distintos nudos, según el tipo de información buscada, según la configuración del diccionario o incluso según los caprichos del usuario. Pensamos que el fundamento técnico del diccionario propuesto, la página web, se compone de varias secciones como, por ejemplo: *Presentación, Normas de uso, Acerca del diccionario, Fuentes, Participación de usuarios*, etc.

Para emprender el diseño técnico de la *BD* de la propuesta del diccionario, hemos hecho más de una búsqueda para contar con un *programario* (software) conveniente y de fácil manejo. Podemos decir que para ello va a servir el *software* de *MySQL*. La función principal de *MySQL* es administrar mediante una *BD*, por eso se llama administrador o motor de *BD*. Hemos de señalar que este programario no es el único motor de *BD* que existe. Hay otros como el DB2, MARIA DB, SQL SERVER, POSTGRESQL, ORACLE, etc. De acuerdo con Ilie (2014: 157), *MySQL* en Internet es mucho más adecuado para administrar una *BD*.

Además de que no es costoso, *MySQL* se caracteriza por su gran flexibilidad además de que está adaptable a diversas versiones de *Windows*, así como de otros sistemas operativos; por lo que opinamos que es el más apropiado para nuestra propuesta. También, este programario no es un proyecto típico de código abierto, lo que ofrece una combinación única de la libertad asociada con el código abierto y la confianza del *software* compatible.

4.2.2. *Interfaz*

Tenemos pensado que la versión automática del diccionario está equipada con una interfaz avanzada, flexible e independiente de la *BD*. De acuerdo con la interfaz desarrollada, a la que acceden los usuarios, se permite al usuario consultar las diferentes acepciones de una palabra dada (Read y Bárcena Madera, 1999: 432). Esta interfaz está escrita en Java, que permite el acceso al diccionario desde la computadora o Internet a través de la *World Wide Web* (WWW).

4.2.3. *Muestra lexicográfica*

Una vez comprobada la escasez de la lexicografía español-árabe en general y sobre todo la carencia de un *DE* fraseológico propiamente dicho, planteamos una muestra del diccionario propuesto según los criterios anteriormente señalados. En esta propuesta, y por limitaciones de espacio del que disponemos, presentamos tres *UFs* —como ejemplo para mostrar los tres tipos de equivalencia— del diccionario que proponemos llevarlo a cabo en un medio plazo. Los ítems de la tabla siguiente son los que vamos a poner en el programario de base de datos *MySQL*; para que luego en el futuro próximo nos dediquemos a su exportación a una página web.

Unidad fraseológica	echar aceite al fuego	agujero negro	no amueles
Categoría	Locución verbal	Locución nominal	<i>fórmula oracional</i>
Definición	Echar leña al fuego	Causa de un gran gasto o pérdida económica.	Se usa para expresar rechazo, asombro o incredulidad ante lo que se acaba de
Ejemplo	Deberían percatarse del carácter complejo y delicado de estos últimos, para que al tratar de ayudarnos a encontrar soluciones eviten arrojarnos aceite al fuego.	La financiación de los partidos, el agujero negro de la democracia.	- El Boni me he dicho que es un ingeniero de su fábrica. - ¿Un ingeniero? No amueles.
Equivalente	صب الزيت على النار	سبب أزمة	لا أصدق
Tipo de equivalencia	Total	Parcial	Nula
Traducción del ejemplo	ينبغي أن يكونوا على دراية بالطبيعة المعقدة والحساسية لهذا الأمر، حتى أنهم في محاولتهم مساعدتنا في إيجاد الحلول يتجنبون صب الزيت على النار.	تمويل الأحزاب هو سبب أزمة الديمقراطية	- أخبرني بوني أنه مهندس من مصنعه. - مهندس؟ لا أصدق.

Elaboración propia de los ítems de la propuesta del diccionario en una tabla

Para poder realizar esta propuesta, ha sido necesario instalar en el ordenador el programario que hemos seleccionado para la elaboración de *BDs*.

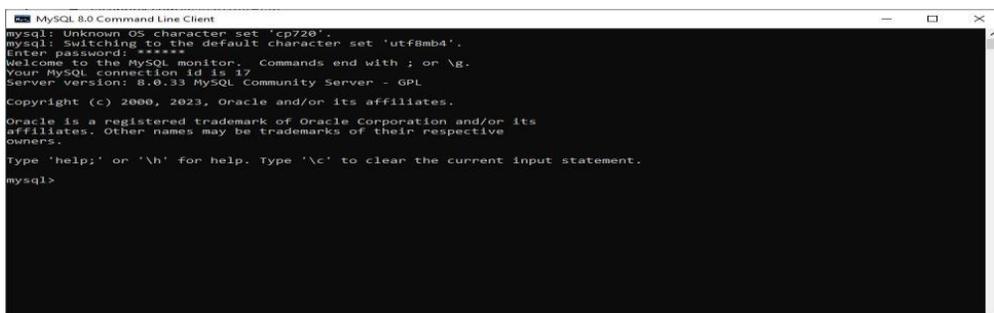


Figura de la instalación de *MySQL*

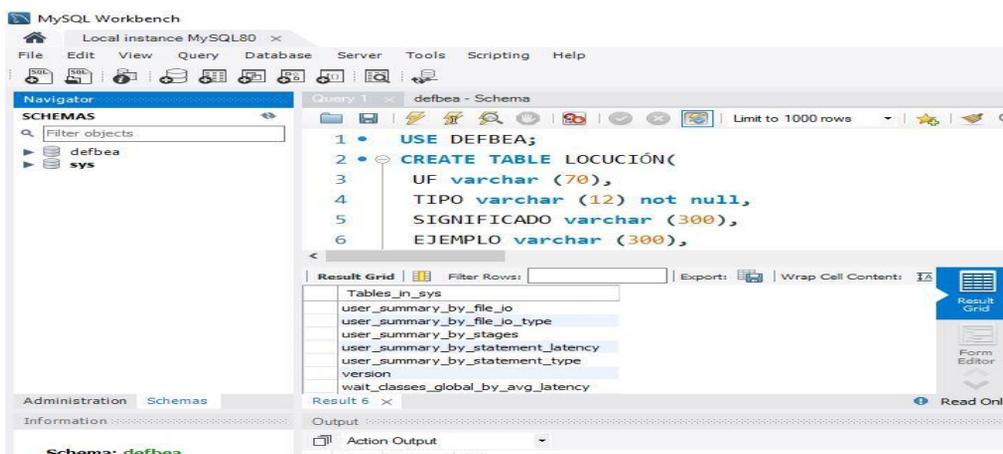


Figura de la creación de una tabla dentro de *MySQL*

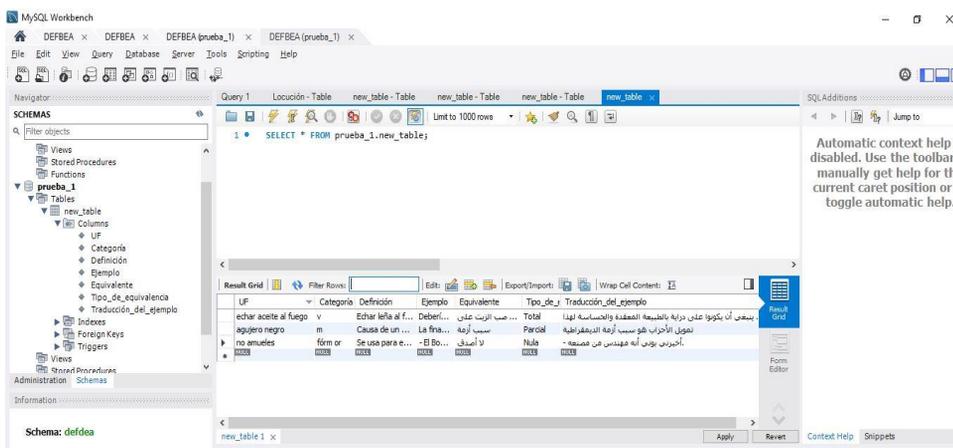


Figura de la inserción de los ítems y datos en la tabla anterior en *MySQL*

5. Conclusiones

La principal conclusión esgrimida es que, desde finales del siglo xx, la disciplina de la lexicografía sufrió un cambio de soporte material que modificó sustancialmente la estructura de los diccionarios y transformó radicalmente las formas en que se construyen. La misión de los lexicógrafos ya no es la de poner a disposición de los usuarios un producto impreso, ordenado alfabéticamente, etc.; ahora es el trabajo con la conformación de *BD*, dotada de programas y herramientas para la adecuada consulta de los nuevos destinatarios.

Se ha establecido que es posible la creación de un *DE* de fraseología español-árabe. Las herramientas informáticas utilizadas para este fin son versátiles y corresponden con el contenido. Por eso, y en comparación con la situación de la lexicografía español-árabe y si tomamos en consideración los diccionarios existentes, nos hallamos ante una situación paradójica con vistas a la relación entre ambos idiomas y que ahora más que nunca requiere confeccionar un diccionario de fraseologismos.

Dado que el desarrollo de este tipo de herramientas es obviamente difícil, hemos decidido resumir a través de este artículo los avances de la propuesta de creación de bases para un nuevo *DE* español-árabe de fraseología. El objetivo de esta propuesta ha sido, por un lado, contribuir al desarrollo de la fraseología digital, cuyo primer fin es resolver la brecha entre los equivalentes interlingüísticos del español y del árabe y; por otro lado, proporcionar a los usuarios una herramienta fácil y de libre acceso; capaz de promover la comprensión y el uso de la fraseología.

A lo largo del presente trabajo, hemos pretendido resaltar la relación entre la lexicografía y la informática y, a la vez, valorar la naturaleza de esta relación. Por lo que, hemos dejado claro la diferencia entre los diccionarios en papel y los de formato electrónico que permiten acceder a la información por distintas formas. El *DE* nos ayuda a entenderlo como una obra de encuentro entre tradición y modernidad: lexicografía e informática. En general, los diccionarios en línea representan una nueva obra lexicográfica con dimensiones de usos muy diferentes y tienen la particularidad de ser obras vivas que se actualizan constantemente. De hecho, la posibilidad de actualización inmediata de los contenidos, y sin coste adicional, ofrece una ventaja que no debería desaprovechar un diccionario que pretenda ser sincrónico.

Asimismo, una de las características de nuestro diccionario propuesto radica en que revela uno de los medios informáticos para familiarizar al usuario con los recursos multimedia relacionados con la lexicografía. Además, el diccionario

propuesto sería de interés para los que estén interesados en el uso de diccionarios fraseológicos en ambas lenguas y, por consiguiente, para el aprendizaje del español como lengua extranjera. Hemos de señalar que los diccionarios disponibles en línea tienen el rasgo de recoger su corpus de forma sistemática y unificada y, a la vez, presentar un tratamiento adecuado en cambio de los diccionarios en papel que no lo han podido lograr.

Llegados al término de este artículo, el modelo que hemos hecho en esta propuesta podría ser mejorable, dado que la elaboración de un diccionario se trata de un quehacer laborioso que, a menudo, requiere el trabajo de un equipo integrado por informáticos y lexicógrafos a la vez. Por eso, las limitaciones que podrían aparecer en esta propuesta deben ser tratadas en el futuro. Consideramos importante seguir trabajando en este campo y aspiramos a que esta propuesta sea uno de los intentos serios que pone a disposición de diccionarios en línea y contribuciones de investigación de una *BD* en la que se pueden buscar *UFs* español-árabe.

Referencias

- ARNTZ, R. y PICTH, H. (1995). *Introducción a la terminología*. Fundación Germán Sánchez Ruipérez, Madrid: Pirámide.
- AZNÁREZ MAULEÓN, M. y SANTAZILIA SALVADOR, E. (2016). Un acercamiento a la conceptualización de algunas emociones en el patrimonio fraseológico del euskera y del castellano. *Huarte de San Juan. Filología y Didáctica de la Lengua*, 16, pp. 13-33.
- BAILEY, R. W., ROBINSON, J. L. (1969). The Computer in Lexicography. En *Paper presented to the section on Computer Research in Language and Literature of the Midwest Modern Language Association*, October 24, pp. 1-22. <https://files.eric.ed.gov/fulltext/ED037701.pdf>
- BARTON, D. y LEE, C. (2013). *Language online. Investigating digital texts and practices*. Londres: Routledge.
- BOSQUE, I. y BARRIOS RODRÍGUEZ, M. A. (2018). Spanish lexicography in the Internet era. En Pedro Antonio Fuertes-Olivera (ed.), *The Routledge Handbook of Lexicography*, cap. 40 (pp. 636-660). Londres: Routledge.
- BOUGHABA, M. (2021). Las marcas diafásicas y diastráticas de las locuciones en los diccionarios generales y fraseológico. *Revista de Estudos da Linguagem*, xxix(1), pp. 99-119.
- CAÑETE GONZÁLEZ, P. (2006). Formación de unidades terminológicas especializadas para la elaboración de un diccionario multilingüe en línea. *RLA. Revista de lingüística teórica y aplicada*, 44(1), pp. 81-94.

-
- CASANOVAS CATALÁ, M. (2016). Las herramientas 2.0 en la escritura académica: buscadores y diccionarios. *Folios: Revista de la Facultad de Humanidades*, 43, pp. 77-88.
- CONKLIN, J. (1987). Hypertext: An Introduction and Survey. *Computer*, xx(9), pp. 16-41.
- CORPAS PASTOR, G. (2003). *Diez años de investigación en fraseología: análisis sintáctico-semánticos, contrastivos y traductológicos*. Madrid: Iberoamericana.
- COSERIU, E. (1981). *Lecciones de lingüística general*. Madrid: Gredos.
- FERNÁNDEZ-PAMPILLÓN CESTEROS, A. y MATESANZ DEL BARRIO, M. (2006). Los Diccionarios Electrónicos: hacia un nuevo concepto de diccionario. *Estudios de Lingüística del Español (ELIES)*, 24. <http://elies.rediris.es/elies24/pampillon.htm>
- FRAILE VICENTE, E. (2007). El tratamiento terminográfico de las expresiones idiomáticas en los diccionarios y glosarios de economía en línea en inglés y español: su utilidad para el traductor. *Sendeban: Revista de la Facultad de Traducción e Interpretación*, 18, pp. 223-261.
- FUERTES-OLIVERA, P. A. y BERGENHOLTZ, H. (2018). Diccionarios del español para la producción de textos. *RILEX. Revista sobre investigaciones léxicas*, 1(1), pp. 5-28.
- GELPÍ, C. (2003). El estado actual de la lexicografía: los nuevos diccionarios. En Antonia María Medina Guerra (coord.), *Lexicografía española* (pp. 307-332). Barcelona: Ariel.
- GELPÍ, C. (2004). Diccionarios digitales especializados por la temática: estado actual y perspectivas. *Scripta Nova. Revista Electrónica de Geografía y Ciencias Sociales*, 8(170). <http://www.ub.es/geocrit/sn/sn-170-69.htm>
- GOUWS, R. H. (2009). Reviewing printed and electronic dictionaries. Access and lexicographic functions. En Sandro Nielsen y Sven Tarp (eds.), *Lexicography in the 21st Century* (pp. 23-41). Amsterdam/Philadelphia: John Benjamins.
- GRANGER, S. (2012). Electronic lexicography from challenge to opportunity. En Sylviane Granger y Magali Paquot (eds.), *Electronic Lexicography*, 1-11 (pp. 1-16). Oxford: Oxford University Press.
- HASS, U. y SCHMITZ, U. (2010). Lexikographie im Internet 2010: Einleitung. *Lexicographica. International Annual for Lexicography*, pp. 26, 1-18. <http://elies.rediris.es/elies24/index.htm>
- HILMAR JÓNSSON, J. (2009). Lexicographic description: An onomasiological approach on the basis of phraseology. En Sandro Nielsen y Sven Tarp (eds.),

-
- Lexicography in the 21st Century* (pp. 257-280). Amsterdam/Philadelphia: John Benjamins.
- HUETE GARCÍA, Á. (2019). Hacia la conceptualización de un e-diccionario de fútbol: definición de funciones lexicográficas específicas. En Julia Sanmartín Sáez y Mereces Quilis Merín (coords.), *Retos y avances en lexicografía: los diccionarios del español en el eje de la variación lingüística* (pp. 219-230). València: Universitat de València.
- ILIE, L. (2014). Diseño y elaboración de un diccionario sanitario trilingüe en línea (Español-Rumano-Inglés). *FTISPos International Journal*, Vol. I, pp. 150-158.
- KNOWLES, F. E. (1990). The Computer in Lexicography. En Franz J. Hausmann *et al.* (eds.), *Wörterbücher. Ein internationales Handbuch zur Lexicographie. Dictionaries. An International Encyclopedia of Lexicography. Dictionnaires. Encyclopédie internationale de lexicographie* (pp. 1645-1672). Berlin: Walter de Gruyter.
- LLANILLO GUTIÉRREZ, F. F. (2008). Proyecto para un diccionario de lenguajes especiales en línea. En Dolores Azorín Fernández *et al.* (coords.), *El diccionario como puente entre las lenguas y culturas del mundo: actas del II Congreso Internacional de Lexicografía Hispánica* (pp. 715-720). Madrid: Biblioteca Virtual Miguel de Cervantes.
- LONČAR, I. y DAL MASO, E. (2018). Propuesta de elaboración lexicográfica de la equivalencia y polisemia asimétricas en un diccionario fraseológico multilingüe en línea: El caso de español, italiano y croata. En Antonio Pamies Bertrán (coord.), *Lenguaje figurado y competencia interlingüística (II): Aplicaciones lexicográficas y traductológicas* (pp. 29-52). Granada: Comares.
- LONČAR, I. y VALERO, P. (2020). Propuestas sobre la representación de la fraseología en diccionarios en línea. *Círculo de lingüística aplicada a la comunicación*, 82, pp. 9-26.
- LUQUE DURÁN, J. (1996). Tipos de diccionarios y el diccionario del futuro (Sobre las aplicaciones de la informática y de la lingüística computacional a la lexicografía), *Segundas jornadas sobre el estudio y la enseñanza del léxico* (pp. 93-102). Granada: Método ediciones.
- MARTÍNEZ EZQUERRO, A. (2002). Diccionarios "on-line" en el Aula de Lengua. *Educación y futuro: Revista de investigación aplicada y experiencias educativas*, 6, <https://dialnet.unirioja.es/servlet/articulo?codigo=2016054>
- MELLADO BLANCO, C. (2015). Parámetros específicos de equivalencia en las unidades fraseológicas (con ejemplos del español y el alemán). *Revista de filología*, 33, pp. 153-174.

-
- MOGORRÓN HUERTA, P. (2004). Los diccionarios electrónicos fraseológicos, perspectivas para la lengua y la traducción. *ELUA: Estudios de Lingüística* 2, pp. 381-400.
- MÜLLER-SPITZER, C. y MÖHRS, C. (2010). The "Online Bibliography of Electronic Lexicography" (OBELEX). En Sylviane Granger y Magali Paquot (eds.), *Lexicography in the 21st Century: New challenges, new applications: proceedings of eLex 2009*, (pp. 199-202). Louvain-la-Neuve.
- NOMDEDEU RULL, A. (2018). El desafío del tratamiento de los datos culturales en un diccionario en línea para aprendices de español como LE/L2. *Revista de Filología de la Universidad de La Laguna*, 36, pp. 277-306.
- OCHOA ANADÓN, J. A. (1992). Observaciones sobre la aplicación de la informática a la lexicografía general grecolatina. *Estudios Humanísticos. Filología*, 14, pp. 129-142.
- ORTEGO ANTÓN, M. T. y FERNÁNDEZ NISTAL, P. (2015). Aproximación a las unidades con significado en el campo de la informática adquirido por extensión metafórica en los diccionarios generales de inglés y español. *Revista de Lingüística y Lenguas Aplicadas*, 10, pp. 44-54.
- PERSICO, D. (1992). El hipertexto y la tecnología multimedia: un paradigma para los diccionarios del futuro. *Cl & E: Comunicación, lenguaje y educación*, 13, pp. 101-107.
- PORTO DAPENA, J. Á. (2002). *Manual de técnica lexicográfica*. Madrid: Arco libros.
- QUEMADA, B. (1987). Notes sur lexicographie et dictionnaire. *Cahiers de lexicologie*, 51, pp. 229-242.
- RAFEL I FONTANALS, J. (2011). Lexicografía e informática. Aplicación a la lengua catalana, *Pirinioetako hizkuntazk: oraina eta lehena: Euskaltzaindiaren XVI* (pp. 557-575). Biltzarra: Euskaltzaindia.
- READ, T. y BÁRCENA MADERA, E. (1999). Hacia un modelo de formalización del conocimiento léxico con fines informáticos. *Epos: Revista de filología*, 15, pp. 425-434.
- RENAU, I. y BATTANER ARIAS, M. P. (2012). Respuesta a la necesidad de diccionarios en línea para el aprendizaje de español como lengua extranjera. En María Nieves Vila Rubio (ed.), *Tonto el que no lea: enseñar lengua y literatura castellanas en la era digital* (pp. 149-158). Lleida: Edicions de la Universitat de Lleida.
- ROISS, S. y ZIMMERMANN GONZÁLEZ, P. (2012). Del tratamiento de falsos amigos en un diccionario electrónico de dudas. *Estudios de Lingüística Aplicada*, 56, pp. 9-33.
- RUIZ GURILLO, L. (1997). *Aspectos de fraseología teórica española*. Valencia: Universitat de València.

-
- SANTANA SUÁREZ, O. *et al.*, (1996). Diccionarios en soportes informáticos. *Cuadernos Cervantes de la lengua española*, 11, pp. 68-73.
- SECO, M. *et al.* (2004). *Diccionario fraseológico documentado del español actual. Locuciones y modismos españoles*. Madrid: Aguilar.
- SHABAN MOHAMMAD SALEM, T. (2013). *La fraseología en español y en árabe: estudio, comparación, traducción y propuesta de un diccionario* (Tesis doctoral, Universidad Complutense de Madrid). <https://docta.ucm.es/entities/publication/3f9337e9-dd1e-4fe9-92d0-0d7633b1eb64>
- Tarp, Sven (2019): La ventana al futuro: despidiéndose de los diccionarios para abrazar la lexicografía. *Rilex. Revista sobre investigaciones léxicas*, II(2), pp. 5-36.
- THUILLIER, S. (2014). De una obra digitalizada a una obra digital: reflexiones en torno al Diccionario Griego-Español en línea. En Sagrario López Poza y Nieves Pena Sueiro (eds.), *Humanidades Digitales: desafíos, logros y perspectivas de futuro* (pp. 459-471). A Coruña: Universidade da Coruña, Editores SIELAE.
- VALERO FERNÁNDEZ, P. y MUŠURA, J. (2019). Acerca de la valoración del diccionario como herramienta para la traducción por parte de estudiantes universitarios croatas de Español: resultados de una encuesta metalexigráfica. *Rilex. Revista sobre investigaciones léxicas*, II(1), pp. 86-123.
- VARGAS SIERRA, C. (2002). Utilización de los programas de concordancias en la traducción especializada. En Comisión Europea (comp.), *El español, lengua de traducción: Actas del I congreso internacional* (pp. 522-539). Puebla (México): Esletra.
- ZUCCALÀ, D. (2015). Fraseología, traducción y fraseodidáctica: Elaboración de un instrumento en línea para el estudio de UFs aplicado a la traducción y a la fraseología en el par de lenguas italiano y español. En Miriam Seghiri *et al.* (coords.), *Nuevos horizontes en los estudios de traducción e interpretación (Comunicaciones completas)*: Conferencia AIETI7 (29 al 31 de enero de 2015 en Málaga) (pp. 264-279). Madrid: Editions Tradulex.