EL CORPUS *OCCOR*: UNA HERRAMIENTA PARA EL ESTUDIO LÉXICO DE TEXTOS EN ESPAÑOL SOBRE CHINA (1850-1939)*

THE *OCCOR* CORPUS: A TOOL FOR THE LEXICAL STUDY OF SPANISH TEXTS ABOUT CHINA (1850-1939)

Matteo De Beni
(Università di Verona)
matteo.debeni@univr.it

Elisa Sartor
(Università di Verona)
elisa.sartor@univr.it

Resumen

El objeto del presente artículo es presentar las características de diseño del corpus *OccOr*, una herramienta que se está llevando a cabo en el Departamento de Lenguas y Literaturas Extranjeras de la Universidad de Verona (Italia) en el marco de un proyecto ministerial dedicado al desarrollo de las humanidades digitales aplicadas a los ámbitos lingüístico-filológicos. *OccOr* recoge textos en español y en francés, redactados de 1850 a 1939 y dedicados a Lejano Oriente, con especial referencia a China. En el trabajo, que se centra en los documentos en lengua española, se exponen algunas de las posibles explotaciones del corpus en cuanto al léxico se refiere y se analiza una muestra de voces extractadas de la herramienta en cuestión. El examen de los términos seleccionados se desarrolla gracias a obras lexicográficas y a corpus de referencia tanto del español como de otras lenguas occidentales.

Palabras clave: corpus *OccOr*, lengua española, lexicografía y lexicología, neologismos, xenismos, oriente y occidente.

Abstract

This paper aims at presenting the design features of the *OccOr* corpus, a tool that is being developed in the Department of Foreign Languages and Literatures at University of Verona (Italy) within the framework of a ministerial project dedicated to the Digital Humanities applied to the linguistic and philological areas. *OccOr* is a corpus of Spanish and French texts written between 1850 and 1939 and concerned with the Far East, especially China. This article focuses on the Spanish language texts and explores the potential use of the corpus for lexical analysis by examining some terms extracted from *OccOr*. The analysis is carried out by using lexicographical resources and reference corpora both for Spanish language and other Western languages.

Keywords: *OccOr* corpus, Spanish language, lexicography and lexicology, neologisms, foreign words, East and West.

* Este trabajo es fruto del proyecto *Le Digital Humanities applicate alle lingue e letterature straniere* (2018-2022), financiado por el Ministerio italiano de Educación, Universidad e Investigación en el marco de una convocatoria para Departamentos de Excelencia (https://dh.dlls.univr.it/it/). Dentro de dicho proyecto, Matteo De Beni es el coordinador del área de investigación lingüística, además de coordinador del propio corpus *OccOr*, al que se dedican las páginas siguientes. Aunque el presente ensayo se ha elaborado de forma conjunta, a efectos académicos, Matteo De Beni es responsable de los apartados 1. y 2. y Elisa Sartor de los apartados 3. y 4.

DOI: https://doi.org/10.17561/rilex.4.3.6540

1. INTRODUCCIÓN

El proyecto *OccOr* –cuyo título procede de *Occidente* y *Oriente* por acronimia– pretende poner a disposición de la comunidad científica un corpus para el estudio de las relaciones lingüístico-culturales entre los países hispanófonos y Francia, por un lado, y Oriente por el otro, con particular referencia a China. La elaboración de la herramienta en cuestión, en la que colaboran investigadores de Lengua española y de Lengua francesa, se está llevando a cabo en el Departamento de Lenguas y Literaturas Extranjeras de la Universidad de Verona (Italia) en el marco de un proyecto ministerial ("Proyecto de Excelencia") dedicado al desarrollo de las humanidades digitales en los ámbitos lingüístico-filológicos: *Le Digital Humanities applicate alle lingue e letterature straniere*. Actualmente, bien para el español bien para el francés, nos encontramos en la fase de compilación del corpus. Cabe señalar que pretendemos que este, una vez completado, sea consultable como banco de datos por parte de usuarios externos a través de una interfaz.

El corpus reúne textos en español y en francés dedicados a los países de Lejano Oriente –con particular atención a China[1]– del periodo comprendido entre 1850 y 1939, esto es, hasta el estallido de la Segunda Guerra Mundial. Se trata, como explicaremos a continuación, de un corte cronológico muy significativo para las relaciones culturales, políticas y comerciales entre las dos macroáreas geográficas en cuestión.

La herramienta permitirá múltiples posibilidades para el estudio de fenómenos derivados del contacto lingüístico y cultural, también en función de la situación comunicativa. Servirá, por ejemplo, para detectar neologismos y observar la oscilación y evolución de formas lingüísticas en el tiempo y

[1] El enfoque hacia textos de Lejano Oriente, y especialmente de China, es coherente con los objetivos científicos del Proyecto de Excelencia *Le Digital Humanities applicate alle lingue e letterature straniere*, en el cual *OccOr* se enmarca. Para el área lingüística, dicho proyecto prevé tres líneas de investigación destacadas, a saber, los lenguajes de especialidad, las lenguas minoritarias del norte de Italia y las relaciones lingüístico-culturales entre Occidente y Oriente, con especial atención a China.

en géneros textuales distintos. *OccOr* hará posible, por lo tanto, examinar aspectos neológicos y de contacto entre lenguas y, por ende, estudiar la percepción de Oriente en los países occidentales aludidos. Así pues, se podrá emplear en investigaciones sobre historia del léxico, análisis del discurso e, incluso, para el estudio de fenómenos de interculturalidad que tengan interés para las áreas lingüístico-culturales consideradas[2].

En el presente trabajo, nos centraremos en el corpus de lengua española (*OccOr-es*). Expondremos, en primer lugar, los criterios que hemos tenido en cuenta a la hora de diseñarlo, que han venido determinados por las necesidades y los objetivos del proyecto, a la luz de los parámetros y consideraciones de Torruella Casañas (2017) para la construcción de corpus diacrónicos. En segundo lugar, con el propósito de ofrecer un ejemplo de utilización de *OccOr*, presentaremos una muestra de voces de interés. Entre las distintas tipologías de resultados léxicos y semánticos que nuestro corpus permitirá detectar caben: la presencia de xenismos; la neología efímera –que demuestra lo precaria que puede resultar, en ocasiones, la aclimatación de una voz en una lengua–; las oscilaciones gráficas como espejo de la vacilación frente a la representación escrita de unas palabras en otra lengua; o la retrodatación –entendiendo este fenómeno como el hallazgo de voces en fechas más tempranas con respecto a las sugeridas por la lexicografía y por los corpus de referencia del español–.

En esta ocasión, ofreceremos una muestra de las posibilidades de explotación de *OccOr* a partir de la presencia de glosas metalingüísticas que acompañan unidades léxicas de raigambre oriental. Dichas acotaciones están enfocadas a aclarar significantes cuyo valor semántico resultaría, de por sí, seguramente opaco o borroso para al lector hispano según el redactor del texto; de ahí que dichas glosas resulten fructíferas a la hora de detectar fenómenos léxicos de distinta índole documentados en los textos de nuestro corpus, entre ellos la acuñación de nuevas voces para designar significados

[2] En cuanto a esto último, señalamos que *OccOr* ya se ha utilizado para un estudio sobre la interculturalidad basado en el corpus de francés (Bonadonna y Zollo, en prensa).

inéditos en español y la inserción de xenismos. Para desarrollar nuestro análisis hemos empleado obras lexicográficas y corpus de referencia del español, que nos han permitido realizar el seguimiento de las voces estudiadas en los textos en castellano. Asimismo, con el propósito de evaluar de manera más amplia las andanzas de las unidades léxicas en cuestión en Occidente y de conseguir datos que se puedan contrastar con los encontrados en español, hemos consultado diccionarios y corpus de otras lenguas, en concreto el francés, el portugués y el inglés –todos ellos idiomas presentes en Oriente por razones coloniales y comerciales en el periodo acotado– y también, para completar el grupo de las lenguas románicas de mayor calado, el italiano.

2. CRITERIOS DE DISEÑO

Para la creación del corpus *OccOr* hemos tenido en cuenta distintos criterios y variables, compartidos entre las dos lenguas involucradas, español y francés[3].

OccOr no tiene entre sus objetivos el de llegar a almacenar un número de textos predeterminado, sino que pretende alcanzar, por lo que a la lengua española se refiere, como mínimo 1 millón de *tokens*[4]. De todos modos, se trata de un corpus abierto, esto es, se podrá continuar alimentando tras la finalización oficial del proyecto (diciembre de 2022), acrecentando de esta manera la posibilidad de obtener términos para su examen histórico, así como muestras textuales para su observación desde los puntos de vista del análisis del discurso y del estudio de las relaciones interculturales.

[3] Los ejemplos y cuestiones tratados en estas páginas presentan casos referidos al corpus de español (*OccOr-es*) y su elaboración.

[4] Por *tokens* nos referimos al número total de las palabras que componen los textos del corpus, mientras que con *types* indicamos el número de palabras distintas en el mismo corpus. En la actualidad (septiembre de 2021) *OccOr-es* contiene 63 documentos, con un total de 503.000 *tokens* aproximadamente, mientras que los *types* superan los 42.000. Cabe destacar que estos datos se refieren a los textos almacenados en nuestra base de datos que son de dominio público o bien de los que disponemos de una autorización para su uso. Desde luego, el programa de gestión de corpus permite computar el número de *tokens* por tipología textual o por subdominio, entre otras opciones, lo que resulta muy útil a la hora de compilar un corpus equilibrado.

Los documentos recogidos son textos escritos, disponibles en bibliotecas y archivos digitales u otros repositorios electrónicos en .pdf o en otros formatos[5], que convertimos a .txt para su almacenamiento en una plataforma desarrollada por los técnicos informáticos del Proyecto de Excelencia. El análisis de *OccOr* se llevará a cabo a través de KonText[6] (Machálek, 2014), una herramienta de gestión de corpus y extracción terminológica que servirá de base para el diseño de la interfaz a través de la cual los usuarios podrán consultar el corpus cuando se ponga a disposición de forma pública (véase Fig. 1).

Hits: **21** | i.p.m.: **41.77** (related to the whole corpus) | ARF: **7.78** | Result is sorted

Line selection: simple

doc#5	depósito de reliquias . De Dagobah los europeos hemos hecha	Pagoda	, nombre genérico de tados los templos indios y chinos
doc#5	chinos ; sin embargo , William Milne opina que la palabra	Pagoda	se deriva de la india butkuda ó de la persa
doc#5	estas etimologías ; pero añadió : « Tened entendido que	pagoda	es nombre propio y exclusivo del templo búdico y no
doc#5	; el templo de Confucio no se parece á la	pagoda	búdica ; y son de distinta construccion y estilo las
doc#5	los marinos próximos á embarcarse van á orar en una	pagoda	erigida en honor de la ballena hace algunos siglos ,
doc#5	canales atestados de lorchas y juncos , y su famosa	pagoda	. El calor nos fundia , pues aunque la distancia
doc#5	conto en telas , y nos dirigimos á la gran	pagoda	, monumento notable , no por lo grandioso , sino
doc#5	de Madame Buda . Quedamos enterados . Inmediata á la	pagoda	hay una bonzería , cuyos estrechos cláustros , hoy desiertos
doc#5	la primera velada , una jóven bonza entra en la	pagoda	con su rosario en la mano ... " A la
doc#5	sus alrededores : " los bateleros se reunen en la	pagoda	de Yuan miao ; la vida es fácil para todo
doc#5	duró dos años (1858 á 1860) . La	pagoda	de Hai-Huang-Tzeu es un conjunto de pequeños templos , cuyas
doc#5	pasaba visitando y analizando las magnificencias chinas : no hay	pagoda	, fortaleza ó palacio , escultura , pintura , música
doc#5	bonzo peregrino , llamado San-iang , que predica en la	pagoda	Putusa , y madrugué para oír la explicacion de los
doc#17	malas lenguas , buvo participación clandestina . Hubiéramos visitado la	pagoda	real embaldosada de plata fundida , que encierra tesoro enorme
doc#19	al mundo para algo más importante que asistir á la	pagoda	del Buda de esmeralda y aprender los bailes rituales ,
doc#31	que un europeo no concibe una vista china sin su	pagoda	. El viajero , por todas partes , en las
doc#31	fotografía de la llamada Yin Shan Ta , es decir ,	pagoda	de la montaña de plata , que existe á espaldas
doc#31	les concedan buen viaje y pingües ganancias . La palabra	pagoda	no es china , pero los escritores europeos la han
doc#34	antes de atravesar la puerta principal , se divisa la	pagoda	, y á su derecha se ve la catedral católica
doc#56	un altar parecido al monumento de Jueves Santo ; la	Pagoda	contiene en gran desórden y desaseo muchos ídolos monstruosos ,
doc#56	de malayos y cambises , y los chinos tienen su	pagoda	. Los chinos de Hong-kong se visten con más limpieza

Figura 1. Muestra de la búsqueda de concordancias en el corpus a través de KonText

Los textos tienen a menudo una extensión dispar entre sí, como es esperable en una compilación que acoge documentos pertenecientes a distintos géneros, y, en su mayoría, se trata de textos íntegros. Un caso prototípico de excepción con respecto a la integridad lo constituyen las obras de tema misceláneo, de las que consideramos para su inclusión en el corpus solo los

[5] No obstante, no excluimos la posibilidad de pedir la digitalización de algunos documentos para su inclusión en *OccOr*, sobre todo si estos pertenecen a géneros textuales que, hacia el final del proyecto, resulten poco representados con respecto al conjunto del corpus.

[6] Se trata de un programa de acceso abierto que ha sido oportunamente adaptado por el equipo informático del Proyecto de Excelencia para cumplir con las finalidades de nuestro trabajo. Durante la fase de elaboración del corpus, el acceso a *OccOr* a través de KonText está permitido solo a los miembros del proyecto.

capítulos y secciones relativas a China y al Lejano Oriente. Además, cabe destacar que hemos creado un protocolo de limpieza de textos para prescindir, antes de incorporar los documentos al corpus, de aquellas partes y elementos que no aporten aspectos de interés coherentes con los objetivos del proyecto *OccOr* o que, incluso, puedan dificultar la extracción de datos a partir de documentos convertidos en otro formato (.txt), distinto del original; en concreto, eliminamos: portadas, datos del autor e información editorial; logotipos; índices; encabezados y pies de página; numeración de los apartados; figuras, tablas, esquemas, cuadros de texto y gráficos, así como sus títulos o pies; referencias bibliográficas; firmas (p. ej. los nombres de los países o entidades que suscriban un acuerdo internacional); resúmenes en lenguas distintas del español; también prescindimos de las notas al pie de página, con sus números dentro del cuerpo textual, cuando estas incluyan referencias bibliográficas u otros datos que no sean de interés léxico, semántico o cultural provechoso para los propósitos de *OccOr*[7]. Estos criterios de eliminación, de todos modos, son indicativos, pues es posible que encontremos excepciones que tengan valor para nuestro proyecto.

En cuanto a los criterios de búsqueda y selección de los documentos, un aspecto fundamental para *OccOr* es su dimensión histórica y cronológica. De hecho, los documentos recolectados deben haberse redactado en el periodo de tiempo comprendido entre 1850 y 1939: un corte cronológico amplio que se justifica por distintas razones, en particular por ser la época en la que se produce la apertura de China al mundo occidental, además de acontecimientos fundamentales –no siempre positivos– para el rumbo de las relaciones entre China y los países de Occidente[8]. De hecho, entre 1856 y 1860 tiene

[7] Conservamos, en cambio, las notas al pie de página que contengan información de interés, pero las emplazamos al final del documento a modo de lista para que no interrumpan la coherencia textual.

[8] Sobra destacar que, antes del periodo estudiado a través de nuestro corpus, no faltan los textos de interés, como demuestran, entre otros, Martínez Robles (2002), Sartor y Dal Maso (2017) y Sartor (2019).

lugar la segunda guerra del Opio; a caballo entre los dos siglos (1899-1901), en cambio, se produce la rebelión bóxer, que tuvo una gran eco en Europa; finalmente la Convención de Pekín, firmada en 1860, abre las puertas a las concesiones occidentales en China hasta 1939, que es el jalón que hemos elegido para cerrar nuestro trabajo; sobra destacar que es el año en que el comienzo de la Segunda Guerra Mundial sacude el mundo entero, desde Occidente hasta Oriente. Las concesiones occidentales llevaron consigo la apertura de los puertos chinos, la llegada al país oriental de exploradores, eruditos, arqueólogos y más personal especializado desde Occidente, además de la redacción de libros, artículos y relaciones[9]. Asimismo, por lo que concierne al mundo hispanohablante, un aspecto histórico de interés sobre cuyas consecuencias léxicas *OccOr-es* podrá esperablemente arrojar más luz es el fenómeno de las migraciones chinas a distintos países hispanoamericanos durante el periodo estudiado, como Perú y Cuba (al respecto, véase, entre otros, Rodríguez Pastor, 1989 y Jiménez Rojas, 2020).

En cuanto a los metadatos utilizados para clasificar los textos, figuran, en primer lugar, los datos de edición o de redacción del texto: obviamente su título, autor, lugar y fecha de edición o redacción. Al respecto, cabe recalcar que abarcamos también textos no publicados, como por ejemplo cartas manuscritas o informes diplomáticos, y también impresos de índole institucional, que no han pasado por los normales procesos editoriales, como los tratados internacionales. Además, para cada texto contemplamos la indicación de si se trata de una traducción o no. En cuanto a este último metadato, no descartamos las traducciones al español debido a dos razones fundamentalmente: a) las traducciones son una de las vías de incorporación de neología

[9] Obviamente, el papel de España y Francia en el proceso de penetración e intervención occidental en China no ha pasado desapercibido por parte de los historiadores. Entre varios otros, pueden verse, por lo que a Francia se refiere, Perrais (1997) y Césari y Varaschin (2003); por lo que atañe al rol de España, Rodicio García (1989) y Martínez Robles (2007). Más en concreto, sobre el papel de las misiones religiosas, véase Collineau (1997) y la ya clásica monografía de González (1955).

a una lengua, como se ha puesto en evidencia en múltiples ocasiones[10]; b) nuestro corpus tiene entre sus propósitos el de sufragar el estudio de las ideas sobre Oriente en los países hispánicos en la época referida; ahora bien, como es evidente, las traducciones, a la par de los textos originales, contribuyen a la circulación de información, ideas, estereotipos, etc. Por estos motivos, en el diseño de *OccOr* consideramos la distinción entre texto original y texto traducido para que los usuarios, a la hora de seleccionar los documentos en la herramienta de consulta del corpus, puedan tener en cuenta este parámetro.

Para la clasificación de los textos, consideramos tanto un eje tipológico como un eje temático. En cuanto al primero, abarcamos los siguientes géneros textuales, que fueron los principales vehículos escritos de transmisión de noticias, datos e informaciones sobre China y los países de Lejano Oriente en Occidente: prensa, textos institucionales, género epistolar y relaciones de viaje. Además, hemos incluido la opción "Otros" para dar cabida a textos de variada índole, que no pertenecen a las categorías aludidas, pero que no queremos desestimar por su interés para nuestros propósitos, *e. g.* informes, textos académicos y memorias. Naturalmente, como en cualquier propuesta taxonómica, puede darse algún caso dudoso, debido, por ejemplo, al solapamiento entre dos categorías en relación con un documento concreto o al hecho de que una categoría resulte más porosa que otra. Al respecto, un aspecto problemático que ha surgido a la hora de clasificar los documentos según su naturaleza textual está relacionado con la ambivalencia de una de las categorías señaladas: en efecto, una de las cuatro tipologías textuales

[10] Este fenómeno es especialmente relevante en los ámbitos especializados, aunque no es exclusivo de ellos. Recordemos que muchos de los textos especializados de los siglos pasados han entrado en el español como traducciones de otras lenguas, aunque no siempre el idioma del texto que se traducía era aquel en el que se había redactado el texto original; por ejemplo, es sabido que a menudo el francés sirvió, sobre todo en los siglos XVIII y XIX, como lengua de mediación. Gómez de Enterría (1999) analiza algunos casos de traducciones especializadas del francés y sus consecuencias léxicas. Sobre el relieve de la traducción para la historia del español de las ciencias, véase, entre otros, De Beni (2016).

antes aludidas, concretamente "Prensa", se distingue de las demás (textos institucionales, género epistolar y relaciones de viaje) por ser no solo una forma textual, sino también un medio de publicación. Por lo tanto, se ha dado el caso de algún documento que, en principio, pertenecía a la categoría "Prensa" por haberse publicado en un periódico o una revista, pero que, al mismo tiempo, exhibía rasgos propios de otro género textual: es el caso de las relaciones de viaje que se publicaron en los diarios en la época acotada y de las que tenemos varias muestras en *OccOr-es*, por ejemplo, "Relacion de un viaje al Japon", publicado en *La Ilustración* (11 de diciembre de 1852) y "Un viaje a China", extraído de *La Verdad* (4-5 de marzo de 1886). En estos casos, nos hemos decantado por clasificarlos en la categoría "Relaciones de viaje" ya que, con vistas a las investigaciones lingüísticas que se llevarán a cabo, resulta más útil poder recuperar los textos a través del filtro correspondiente a este último metadato. Con todo, la información acerca del medio de publicación no se pierde, puesto que todos los documentos se registran con el título y los datos de edición correspondientes, acompañados –en este caso concreto– por la fecha en formato día, mes y año, en conformidad con el estándar elegido para los textos aparecidos en la prensa.

En cuanto al eje temático, en la clasificación tenemos en cuenta los siguientes ámbitos: ciencias; artes y vida social; economía; política; religión y ritos. Creemos que esta clasificación refleja los principales temas de interés y que permite organizar la mayoría de los textos desde el punto de vista del asunto del que tratan[11]. Cabe destacar, además, que el dominio "Política" incluye dos subcampos especiales, que sirven para clasificar –y posteriormente recuperar a través de la herramienta de consulta– aquellos textos que se ocupan de temas asociados con la China del periodo a caballo entre el siglo XIX y el XX: "Rebelión bóxer" y "Concesiones de países occidentales". Esta elección,

[11] También en este caso, clasificamos como "Otros" los textos cuyo tema no se adscriben a ninguno de los ámbitos principales señalados.

así pues, permite destacar dos aspectos estrechamente vinculados con China y que tuvieron una enorme repercusión en Occidente.

También forman parte de los metadatos de nuestro corpus un campo dedicado a las palabras clave, que contribuyen a la descripción del texto, y los datos relacionados con el acceso al texto mismo: la fuente utilizada, la fecha de consulta y los derechos de autor. Por último, como campo útil para el trabajo de recopilación y estudio del corpus, hemos previsto un apartado para eventuales notas, en el que poder incluir información adicional sobre el texto o sobre su valor cultural.

3. MUESTRA DE VOCES

Con vistas a ofrecer una demostración de las potenciales investigaciones léxicas que *OccOr* permite efectuar, a través de KonText hemos realizado un primer análisis de los documentos archivados en la plataforma que almacena dicho corpus. Para ello, se han empleado los 63 textos que conformaban *OccOr-es* hasta la fecha (septiembre de 2021), previamente convertidos en formato .txt según el protocolo de tratamiento de textos descrito *supra* (*cfr*. apartado 2.).

Partiendo de la idea de que los textos recopilados pueden mencionar –o incluso describir– realidades extralingüísticas ajenas a la lengua y la cultura hispánicas, hemos buscado glosas explicativas con el propósito de identificar, gracias a ellas, extranjerismos y neologismos que puedan haber surgido a raíz del contacto entre la cultura de Oriente y la lengua de acogida. Dichas glosas recurren a menudo a *verba dicendi*, a verbos que sirven para designar –*llamar*(*se*), *denominar*(*se*), etc.– o a otros elementos explicativos (por ejemplo, *es decir*), o bien se enmarcan en el discurso entre elementos ortográficos con función parentética.

De ahí que, empleando la función de KonText que permite extraer las concordancias, hayamos realizado una búsqueda a través de marcadores discursivos para identificar glosas textuales, por ejemplo, *llam**, *denomin** o el símbolo ortográfico del paréntesis. A continuación, ofrecemos un botón de muestra de los 332 resultados que ha devuelto la búsqueda de *llam**[12]:

(1) Las ventanas ocupan todo un lienzo de la habitacion y están cubiertas de papel, cristal, láminas de mica, ó una especie de concha **llam**ada *vidrio de la China* (placuna). (doc#38) [1864]

(2) Los elegantes trasforman en verdadera americana la chaqueta, y se la ponen con su correspondiente camisa almidonada y corbata, ésto combinado con el sarong, **llam**ado en el país *kampon*, puesto de la manera ya descrita. (doc#10) [1915]

La búsqueda del paréntesis, en cambio, ha producido 470 ocurrencias[13], de entre las que nos conformamos con copiar dos ejemplos:

(3) Entre las entradas figura la contribución territorial por 47.338,520 yens, y la contribución sobre el *sake* **(**vino de arroz), industria anexa, por 55.465,767. (doc#35) [1900]

(4) …desechos de seda y de capullos *Soya* **(**salsa aromática del Japon) Azufre Tè. (doc#37) [1872]

En los ejemplos reproducidos destacan las voces *vidrio de la China*, *kampon*, *sake* y *soya*. Con el propósito de ilustrar las posibilidades que ofrece *OccOr* para aportar datos sobre la historia del léxico del español, someteremos a una mirada escrutadora los primeros dos términos mencionados, sin duda los menos conocidos entre los lectores hispanófonos.

El caso de *vidrio de la China* es especialmente representativo, puesto que, al mencionar una realidad extralingüística nueva o desconocida, esto es, como se lee en la cita, la "especie de concha" con la que se fabrican las hojas de las ventanas en el país oriental, el autor del texto –publicado en 1864[14]– emplea

[12] A diferencia de otras herramientas de gestión de corpus, como Sketch Engine, la búsqueda por asterisco en KonText se efectúa a través de la opción "Word part" y el fragmento textual correspondiente, en este caso, *llam*. En los ejemplos, evidenciamos con negrita el elemento objeto de la búsqueda (marcador discursivo, elemento ortotipográfico) y con cursiva el término que lo acompaña y que, por ende, puede tener un interés para los estudios históricos del léxico. A excepción de estos resaltes gráficos, los fragmentos se reproducen tal y como aparecen en los documentos almacenados.

[13] Sobra destacar que, en la mayoría de los contextos devueltos por KonText, el paréntesis no delimita una glosa metalingüística de interés. Es preciso, así pues, aquilatar los resultados y seleccionarlos a partir de una lectura sistemática de los fragmentos. Con todo, hemos podido comprobar que se trata de una búsqueda provechosa.

[14] Desde luego, somos conscientes de que esta lexía podría aparecer en textos publicados en fechas anteriores, sin embargo, el término *a quo* del corpus *OccOr* es 1850. Esta fecha, como ya se ha adelantado, se ha elegido por razones extralingüísticas y representa uno de los criterios de diseño compartidos por las áreas de Lengua española y Lengua francesa involucradas en el proyecto.

una dúplice modalidad: en primer lugar, elige una denominación basada en la función que dicho material desempeña (la del vidrio) con un modificador toponímico; en segundo lugar, añade entre paréntesis[15] una denominación alternativa del molusco cuyas conchas sirven para la producción de las láminas translucidas aludidas. En este caso, dicha denominación coincide con el nombre científico del género del animal, según la taxonomía de Linneo.

En efecto, únicamente gracias a esta segunda estrategia descriptiva, el lector contemporáneo puede entender la referencia textual, puesto que la lexía *vidrio de la China* no se recoge en ninguna de las obras lexicográficas disponibles en el *Nuevo tesoro lexicográfico de la lengua española* (*NTLLE*) ni en ninguno de los corpus académicos que hemos consultado, tanto los históricos –el *Corpus Diacrónico del Español* (*CORDE*) y el *Corpus Diacrónico y Diatópico del Español de América* (*CORDIAM*)– como los destinados a representar el español contemporáneo –el *Corpus de Referencia del Español Actual* (*CREA*) y el *Corpus del Español del Siglo XXI* (*CORPES XXI*)–[16]. La búsqueda en el *Diccionario histórico de la lengua española* (*DHLE*), obra todavía

[15] De ahí que la lexía *vidrio de la China* –que, a efectos de este estudio, se ha extraído realizando una búsqueda con el fragmento *llam**– se pueda encontrar también efectuando una consulta de las concordancias a través del paréntesis como elemento para enmarcar una glosa explicativa o, como en este caso, un sinónimo: "una especie de concha llamada vidrio de la China (placuna)".

[16] Una búsqueda en el *CORDE* realizada a partir del nombre de una región conocida por la producción de estas conchas, esto es, la provincia filipina de Cápiz, ha permitido recabar algún dato más. Entre los resultados que nos devuelve la base de datos mencionada, figura el contexto que copiamos a continuación, del que se desprende cierta variación terminológica –que evidenciamos en cursiva– relacionada con las conchas de *placuna* y su empleo artesanal: "Por la solidez y duración de que disfruta, suelen aplicarla los carpinteros á la fabricación de *barrillas de Cápiz*, (*así llamamos las vidrieras que se usan en Filipinas*); fórmanse de unos listones de madera fuerte, cuyo grosor no excede de un dedo; se ensamblan entre sí formando cuadraditos iguales del grandor de la palma de la mano, y en ellos se ponen unas conchas del mar, á modo de grueso talco, que llaman en algunas islas *tipaq*, y en Manila *capis*. Estas son las que sirven en todas partes de vidrieras en las puertas y ventanas; permiten pasar á través de sí bastante cantidad de luz, y adornan los edificios y balcones, cuando los hiere el sol, por ser la materia una especie de nácar ó madreperla. Llámanse estas conchas en Bohol, *lampiron*, en Leyte y Sámar *tipay*" (Delgado, Juan José, *Historia general sacro-profana, política y natural de las islas del Poniente llamadas Filipinas*, Imp. de El Eco de Filipinas de D. Juan Atayde, Manila, 1892 [1754]; *apud CORDE*).

in fieri, no ha devuelto resultados y tampoco la consulta del *Fichero general* de la Real Academia Española ha producido ninguna ocurrencia de este término. En esta última herramienta, sin embargo, hemos encontrado dos compuestos formados por el mismo procedimiento morfológico, *vidrio de Jena* (también conocido como *vidrio óptico*) y *vidrio de Cuba* (es decir, la planta *Sesuvium portulacastrum*), que se suman a los términos pluriverbales con la misma estructura recogidos únicamente por el diccionario de la editorial Gaspar y Roig (1853-1855, *apud NTLLE*), como subentradas del lema *vidrio*:

> Vidrio de Islandia: obsidiana de Irlanda [*sic*].
>
> Vidrio de Moscovia: hojas de mica que llegan a tener hasta media vara de diámetro y se usan en Rusia como cristales, en las ventanas.

La búsqueda de *placuna*, en cambio, ha sido algo más fructífera, aunque el empleo de la palabra está supeditado a la clasificación científica linneana: esto es, los contextos encontrados no sugieren un empleo de la voz como nombre común, ya que siempre hacen referencia a la denominación del taxón. La consulta de los recursos lexicográficos incluidos en el *NTLLE* devuelve un único resultado, puesto que el término se lematiza solo en el ya mencionado diccionario de la editorial Gaspar y Roig con la siguiente definición: "S. f. Zool. (lámina): género de moluscos conchíferos monomiarios, que atendiendo a su organización ofrecen especies intermedias entre los anomiarios y los ostáaceos"; además, el *Fichero general* de la RAE conserva una cédula que reproduce una definición, aun incompleta, procedente del *Diccionario Enciclopédico UTEHA* (1953): "Zool. Género de moluscos bivalvos marinos perteneciente a la familia de los anomíidos. La *P. sella* se encuentra en los océanos Pacífico e Índico. La *P. placenta*, tiene forma laminar, es trans-" [*sic*]. Cabe observar que ambas definiciones presentan la marca diatécnica relativa al ámbito de la zoología y no hacen mención de la explotación de la concha en cuestión para usos fabriles.

La palabra *placuna* ha dejado algún rastro en el *CORDE*: en este banco de datos hemos encontrado dos casos en un solo documento; se trata, en

concreto, de un texto del ámbito zoológico en el que se menciona también el uso de las conchas de este molusco en la fabricación de láminas translucidas para las ventanas en China:

> En la misma familia se incluye el género Placuna que tiene especies de gran tamaño, conchas irregulares y bello aspecto nacarado, superior al de las anomias de nuestros mares. En Cantón, muchas de las más antiguas y suntuosas mansiones se iluminan con la delicada luz que pasa a través de las semitransparentes conchas de la Placuna placenta, colocadas en ventanas y vidrieras a modo de nacarados cristales. (Rioja Lo-Bianco, Enrique, *Los moluscos [Historia natural. Vida de los animales, de las plantas y de la tierra. Zoología]*, Instituto Gallach de Librería y Ediciones, Barcelona, 1926; *apud CORDE*)[17]

La voz *placuna* en la enciclopedia *Britannica* remite a las denominaciones comunes *windowpane oyster* y *windowpane shell*, lo cual confirma que en la cultura anglosajona también se había difundido el conocimiento del uso por parte de los chinos de las láminas fabricadas con las conchas de este molusco en la construcción de ventanas. El *Oxford English Dictionary*, por su parte, recoge *window oyster* (*s.v. oyster*), con marca diacrónica de uso *obsolete* y remisión a *window shell*. El primer ejemplo citado en la entrada *window oyster* incluye justamente el gentilicio *Chinese*, mientras que el segundo es menos relevante para nuestro estudio:

> 1815. E. J. Burrow. *Elements of conchology*. 197 [Genus] Anomia... [Species] Placenta. Chinese window Oyster.
>
> 1907. *Mining Journal*. 27 July 114/2 Provision has been made for the protection of pearl oysters, including the 'window' oysters, in the territorial waters of the colony.

La definición del *OED* de *window shell* (s.v. *window*) es la siguiente:

> [A]ny of the bivalve molluscs constituting the family Placunidae, found in the Indian Ocean and adjoining seas, esp. *Placuna placenta*, having a thin flat translucent

[17] En cuanto al Instituto Gallach, apuntamos de paso que forma parte de una historia editorial interesantísima. De hecho, los manuales Soler-Gallach son un ejemplo señero de divulgación dedicada a los ámbitos técnicos y científicos. Dichos textos se editaron primero por Manuel Soler, luego por José Gallach y finalmente por Calpe. A la postre, en 1924 se inauguró una segunda etapa con el nombre de Instituto Gallach de Librería y Ediciones. Sobre el relieve lexicográfico de los manuales en cuestión, véase Moreno Villanueva (2017).

shell much used in decorative craftwork (also called *capiz*); (also) any of several similar bivalves of the genus *Anomia* (family Anomiidae).

1801. *Catalogue Museum M. de Calonne*. 96 The glassy placuna, or Chinese window shell.

1887. *Guide to Shell & Starfish Galleries* (Brit. Mus. Nat. Hist.). 46 The *Placunidæ*, sometimes called Window-shells and Saddle-Oysters, are very flat pearly shells with a remarkable hinge.

2015. Seilacher & A. Gishlick. *Morphodynamics* xv. 244/1 Best known is the 'window shell' *Placuna placenta* because Philippines artisans use it for making lampshades of various kinds.

Esta entrada resulta eminentemente interesante por distintas razones: en primer lugar, por la confirmación de la denominación de origen geográfico *capiz*, a pesar de que a este término no se le asigne en el *OED* una entrada independiente; en segundo lugar, por la ocurrencia, en el primer ejemplo citado, de la lexía formada a través del gentilicio *Chinese*, que se configura como variante léxica de *Chinese window Oyster* (s.v. *oyster*); y, en tercer lugar, por la fecha reciente del tercer ejemplo citado, que, de hecho, constituye un testimonio de uso actual del vocablo *window shell*, a estas alturas ya sin la indicación de su procedencia, puesto que, de todas formas, su posterior elaboración en productos de artesanía se realiza, según el autor del texto mencionado, en las Filipinas. Finalmente, es reseñable que en uno de los contextos citados se consigne un sinónimo que remite directamente al uso artesanal de la concha para la producción de un equivalente del vidrio: "The *glassy placuna*, or Chinese window shell" (cursiva nuestra).

En cuanto a la lexicografía francesa, la búsqueda en las distintas ediciones del *Dictionnaire de l'Académie Française* [1694, 1762, 1798, 1835, 1932-5] no ha producido resultados reseñables. *Le trésor de la langue française*, por su parte, en la definición de *aronde* incluye una cita textual sacada del tomo I de *Leçons d'anatomie comparée* de George Cuvier (1805), en la que también aparece el término *placune*: "En effet, les huîtres, les placunes, les pélérines, les arondes, etc. n'ont point de dents du tout à leur charnière" (s.v. *aronde*). Es en el *Dictionnaire de la langue française* de Littré (1872-1877, s.v. *vitre*) donde se

halla una referencia significativa, aunque escueta y poco circunstanciada: "Vitre chinoise, coquille bivalve"[18]. En la definición del lexicógrafo galo se mencionan secuencialmente, en primer lugar, el producto ('vidrio chino', equivalente francés de *vidrio de la China*) y, a continuación, la 'concha bivalva' del que el primero se produce, pero sin dilucidar la relación entre estos elementos. Además, se ha encontrado una definición mucho más amplia y pormenorizada en el *Dictionnaire pittoresque d'histoire naturelle et des phénomènes de la nature*, un diccionario enciclopédico coordinado por Félix-Édouard Guérin y publicado en París entre 1833 y 1839 (*s. v. placune*)[19]. Los autores del artículo enciclopédico nos informan de que la *placune* es una "Coquille bivalve appartenant à la famille des Ostracés de M. de Lamarck" y que una de las tres especies conocidas de este molusco es la *placune vitrée*, de la que ofrecen la siguiente descripción:

> La PLACUNE VITRÉE, P. placenta, Lam., est aussi très grande, mais n'atteint pas tout-à-fait la taille de la première. Cette coquille a quelque fois sept pouces de diamètre. Son aplatissement et aussi sa transparence la font employer dans certains pays comme les vitres, et elle est connue vulgairement sous le nom de Vitre chinoise. On trouve cette coquille en très grande abondance dans la mer Rouge et dans l'océan Indien. C'est celle que nous avons représentée, pl. 570, fig. 1.

Reproducimos a continuación (Fig. 2) un detalle de la lámina aludida en la cita, en la que la figura marcada con 1 se refiere a la voz *placune vitrée*:

Desde luego, la lexía *placune vitrée* es de especial interés para nuestra investigación, puesto que se compone de una parte de la denominación científica de la concha bivalva y de un adjetivo que alude a su potencial explotación en la manufactura de láminas translucidas: se trata, así pues, de una acuñación diferente de las soluciones realizadas en lengua española.

---

[18] La lexía francesa *vitre chinoise* pudo inducir, en ocasiones, la creación del compuesto *vidrio chinesco* en español, como muestra la traducción al castellano, realizada por José Garriga y Baucís, de *Tableau élémentaire de l'histoire naturelle des animaux* de Georges Cuvier, publicada en 1834. En efecto, el tratado en lengua francesa (1798) describe "*La vitre chinoise (Pl. placenta)*" (*p. 419*), mientras que la versión española propone "*El Vidrio chinesco (Placuna placenta)*" (*t. II, p. 52*). Para profundizar en la traducción española del tratado de Cuvier, véase De Beni, 2019.

[19] Los autores quieren expresar su gratitud a la Dr.ª Silvia Domenica Zollo por haber señalado esta interesante referencia lexicográfica.

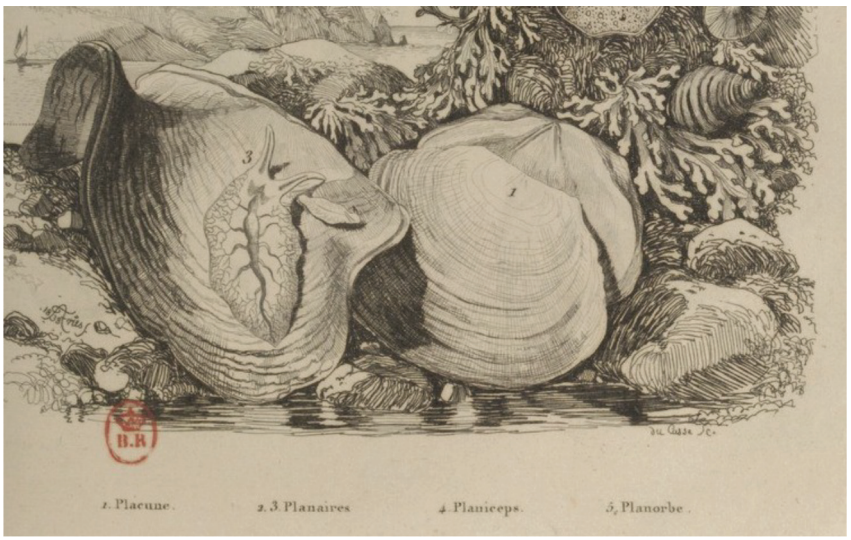

Figura 2. Detalle inferior de la lámina 570 del Dictionnaire pittoresque d'histoire naturelle et des phénomènes de la nature, *t. 8.*

En resumen, las muestras léxicas detectadas en francés y en inglés, además de las huellas encontradas en la lexicografía y los corpus españoles[20], indican que existieron, en lenguas de países que entraron en contacto con Oriente y la cultura china, denominaciones distintas desde el punto de vista de su formación que designaban tanto 'placuna como molusco bivalvo' como, por metonimia, la concha de este último y, finalmente, el producto obtenido de la elaboración de las valvas. Dichas denominaciones pluriverbales circularon en particular en el siglo XIX, mientras que su uso fue menguando en el siglo XX. Por lo que se refiere al compuesto español *vidrio de la China*, lo consideramos un neologismo efímero, surgido para denominar

[20] En cuanto a la lexicografía italiana, hemos consultado las distintas ediciones del diccionario de la Academia de la Crusca (1612, 1623, 1691, 1729-1738, 1863-1923), el diccionario Tommaseo-Bellini (1861-1874) y el diccionario histórico de Battaglia (1961-2004), pero ninguno de ellos consigna *placuna* o una lexía equivalente a *vidrio de la China*. Asimismo, por lo que atañe a la lengua portuguesa, la búsqueda de las voces en cuestión no ha devuelto ningún resultado ni en el diccionario *Michaelis* ni en el *Corpus de Referência do Português Contemporâneo* del Centro de Linguística da Universidade de Lisboa, que, aunque reúne sobre todo textos de 1970 a 2006, también incluye documentos a partir de mediados del siglo XIX.

una realidad extralingüística antes desconocida y luego caído en desuso con el paso del tiempo.

En cuanto al segundo ejemplo que nos hemos propuesto estudiar, *kampon*, hemos podido comprobar que esta voz substancialmente no ha dejado rastro en el caudal léxico del español; sin embargo, el término del que *kampon* pretende ser un equivalente en el ejemplo citado de *OccOr*, es decir, *sarong* ("sarong, llamado en el país kampon"), se ha revelado de considerable interés histórico-lexicográfico por tratarse de un xenismo de origen malayo nunca lematizado en los diccionarios académicos[21] y no académicos reunidos en el *NTLLE*, pese a tener cierta difusión panhispánica en la actualidad. De hecho, una búsqueda en el *CREA* devuelve 8 casos en 5 documentos redactados entre 1976 y 2004 en España, Perú, México y Argentina, mientras que el *CORPES XXI* nos consigna 11 casos hallados en 8 documentos publicados en España, Republica Dominicana, Perú, Chile y Nicaragua entre 2004 y 2010. En cuanto a la atestiguación más antigua, el *CORDE* recoge un ejemplo de uso del año 1973 salido de la pluma de Pablo Neruda en su autobiografía *Confieso que he vivido*. Ahora bien, el dato de interés propiciado por nuestra investigación estriba en el hecho de que *OccOr* nos brinda un ejemplo de uso que antecede en varias décadas el único testimonio registrado en el *CORDE*, ya que el documento en el que aparece la voz *sarong* se publicó en Manila en 1915. *OccOr*, por ende, nos ofrece un caso de retrodatación con respecto a las obras de referencia del español, lo cual permite observar que este

[21] Pese a no haberse todavía incluido en el *Diccionario histórico de la lengua española* (*DHLE*), a la voz *sarong* corresponden dos cédulas en el *Fichero general* de la RAE, lo cual indica que los académicos en algún momento debieron de tomar en consideración este término: la primera cédula solo lleva una indicación abreviada que remite al *Diccionario técnico de la música* de Felipe Pedrell de 1894 (1894 PEDRELL DMús. MA [¿?]-118); la segunda, en cambio, reproduce un fragmento del segundo tomo de *La vuelta al mundo, de un novelista* de Vicente Blasco Ibáñez, que se editó en tres volúmenes entre 1924 y 1925, "Los hombres que desempeñan oficios prescinden del pañal llamado sarong y usan calzoncillos blancos" (subrayado en el texto), acompañado por la notación 7-A 378.

xenismo estaba presente en la lengua española, aunque de forma puntual, ya en la segunda década del siglo XX[22].

Por lo que concierne a la lengua portuguesa, otro idioma presente en Oriente por razones comerciales y coloniales, la búsqueda de la forma *sarong* en el *Corpus de Referência do Português Contemporâneo* devuelve dos ocurrencias de la variedad peninsular –de las cuales una está desprovista de fecha y la otra se remonta a 1993– y una de la grafía empleada en la variedad brasileña, *sarongue*, con fecha de 1994, lo que atestigua un uso limitado de este término en la actualidad. En cuanto a la variedad brasileña, el *Dicionário Brasileiro da Língua Portuguesa* (*Michaelis*), como era de esperar, lematiza la voz con la grafía *sarongue*; de todos modos, la etimología y la definición que proporciona confirman los resultados encontrados en las demás lenguas.

En lo que se refiere a la lengua italiana, en cambio, cabe mencionar que la voz *sarong* está lematizada en el diccionario histórico de Salvatore Battaglia (1961-2004); su registro lexicográfico se apuntala en un único contexto del año 1908, que se debe a la novela *La riconquista del Mompracem* del escritor veronés Emilio Salgari: "Una barca […] montata da un personaggio importante, a giudicarlo dalla ricchezza del suo sarong e dalla mole del suo turbante, […] abbordò Fyacht". No es una casualidad que se trate de un ejemplo salido de la pluma de un autor especializado en novelas de aventuras, muchas de ellas, como la citada, de ambientación oriental: un género sin duda poroso hacia la inserción de xenismos y elementos culturales ajenos a la cultura de recepción.

[22] A este propósito, cabe recordar una vez más que nuestro corpus está formado por textos publicados o redactados a partir de 1850 y, por lo tanto, no se excluye el empleo de la voz *sarong* en testimonios que se remontan a fechas anteriores, no solo originales sino también traducidos, como es el caso de la versión castellana de *Historia de América* (tomo XV de *El mundo. Historia de todos los pueblos*) de Hippolyte Belloc que Juan Cortada realizó a partir del francés en 1845. De hecho, la obra ofrece una ocurrencia de la palabra: "El vestido de esas gentes es muy sencillo como que en la mayor parte de las islas los hombres no llevan mas que una especie de blusa ó saco abierto llamado *sarong*, y en la cabeza una pañoleta de algodon puesta con bastante gracia" (1845: tomo II, p. 366).

Este hallazgo está en línea con las evidencias de la circulación de *sarong* en las demás lenguas de los países europeos que habían establecido contactos comerciales con Oriente, como el Imperio Británico, cuya presencia en el archipiélago malayo desde finales del siglo XVIII puede explicar por qué este término aparece en textos redactados en inglés en fechas tan tempranas como 1834, como demuestra el siguiente contexto procedente de *Wanderings New S. Wales* de G. Bennett: "He was attired in a dirty sarong around his waist, and a loose baju or jacket" (*Oxford English Dictionary*, s.v. *sarong*). En cuanto al francés, el *Dictionnaire historique de la langue française* (2010) de Alain Rey registra el año 1836 –muy próximo a la primera atestiguación en inglés– como fecha de incorporación de este préstamo malayo, que define como "un vêtement drapé, porté par les hommes et par les femmes"[23].

Desde luego, según se vaya ampliando *OccOr* –corpus que, cabe recordar, todavía está en fase de compilación– y se incorporen más textos, es posible que se encuentren otros ejemplos de uso del término *sarong* entre 1850 y 1939 aún no reseñados, en particular en documentos redactados en la segunda mitad del siglo XIX, aseverando así, quizás, un uso de este vocablo en textos en español anteriores al año 1915.

4. CONCLUSIONES

Este ensayo nos ha brindado la oportunidad de presentar los criterios de diseño del corpus *OccOr* en lengua española y de realizar un primer estudio de caso para demostrar su potencial utilidad. Para ello, hemos efectuado una búsqueda en KonText –la herramienta de gestión y análisis de corpus desarrollada en el seno del Proyecto de Excelencia *Le Digital Humanities applicate alle lingue e letterature straniere*– de elementos orientados a insertar glosas explicativas y formas sinónimas o equivalentes –concretamente, *llam**

[23] *Le trésor de la langue française* confirma la fecha de aparición de *sarong* citando *L'Univers. Océanie ou Cinquième partie du monde*, publicado por Grégoire Louis Domeny de Rienzi en 1836, y agrega ejemplos sacados de documentos de 1844 y de 1933. El *Dictionnaire de l'Académie Française*, en cambio, no deja constancia de esta voz.

y el signo de paréntesis– con vistas a encontrar material léxico relevante para nuestra investigación. Efectivamente, hemos documentado ocurrencias de términos con escasa o ninguna presencia en la lexicografía y los corpus de referencia de la lengua española, como *vidrio de la China* y *sarong*, de los que hemos ofrecido una muestra de análisis tomando en consideración sus equivalentes léxicos en francés, inglés, italiano y portugués.

Nuestras pesquisas nos han permitido detectar las pautas de implantación y difusión de estas voces, que no siempre son homogéneas entre las distintas lenguas europeas. En el caso concreto de *vidrio de la China*, por ejemplo, hemos puesto de relieve el carácter efímero de esta creación neológica en español, contrariamente a sus equivalentes en inglés, que aparecen en documentos redactados a comienzos del siglo XIX y llegan, con las debidas adaptaciones, hasta al siglo XXI. Además, hemos dejado constancia de la variación denominativa que atañe a la incorporación de un mismo referente en diferentes idiomas: al lado de una lexía, como la española, construida sobre la base de un topónimo (*vidrio de la China*), figuran otras aún descriptivas (*Chinese window shell*) o bien términos que radican en la taxonomía linneana (*placune vitrée*).

El presente estudio también ha puesto en evidencia como el corpus *OccOr* permite sacar a luz préstamos ocasionales, como es el caso del xenismo *kampon*, nunca consignados en las obras lexicográficas ni en los corpus de referencia del español. Asimismo, el estudio de la voz *sarong* ha arrojado algo de luz sobre su circulación en las distintas lenguas que entraron en contacto con la cultura malaya a partir del siglo XIX: por un lado, hemos señalado su incorporación temprana al inglés y al francés y, por otro, hemos podido adelantar la fecha de la primera ocurrencia de este extranjerismo en español, puesto que *OccOr* ha devuelto una ocurrencia que antecede en algunas décadas el único caso registrado en *CORDE*.

Aun tratándose de una investigación basada en un número limitado de documentos, ya que en la actualidad el corpus *OccOr* en lengua española está todavía en fase de compilación, los ejemplos que hemos encontrado –y su posterior análisis– indican una perspectiva de investigación prometedora

en el ámbito de los estudios históricos del léxico, especialmente en relación con la circulación de neologismos surgidos a raíz del contacto entre el mundo hispanófono y las culturas del Lejano Oriente.

BIBLIOGRAFÍA

Academia Mexicana de la Lengua. *Corpus Diacrónico y Diatópico del Español de América* (*CORDIAM*). http://www.cordiam.org

Accademia della Crusca (1612, 1623, 1691, 1729-1738, 1923). *Vocabolario degli Accademici della Crusca*. http://www.lessicografia.it

ATILF - CNRS & Université de Lorraine (s.f.). *Le trésor de la langue Française informatisé*. http://atilf.atilf.fr/

Battaglia, S. (1961-2004). *Grande dizionario della lingua italiana*. http://www.gdli.it

Belloc, H. (1845). *Historia de América. Desde los tiempos más remotos hasta nuestros días*, vol. II, traducción de Juan Cortada, en *El mundo. Historia de todos los pueblos*, tomo XV. Imprenta de A. Brusi.

Bonadonna, M. F. & Zollo, S. D. (en prensa). Indices d'interculturalité dans le corpus *OccOr*. En M. T. Zanola & P. Puccini, *Terminologie et interculturalité. Problématiques et perspectives*. Bononia University Press.

Centro de Linguística da Universidade de Lisboa. *Corpus de Referência do Português Contemporâneo* (*CRPC*). https://cutt.ly/lRKvMNq

Césari, L. & Varaschin, D. (dirs.) (2003). *Les relations franco-chinoises au vingtième siècle et leurs antécédents*. Artois Presses Université.

Collineau, P. (1997). Les missionnaires, agents de l'influence française. Le problème politico-missionnaire au Sichuan (1898-1937). En J. Weber (dir.), *La France en Chine (1853-1943)*. Presses académiques de l'Ouest, Centre de recherches sur l'histoire du monde atlantique, Université de Nantes.

Cuvier, G. (1797-1798*). Tableau élémentaire de l'histoire naturelle des animaux*. Baudouin. https://doi.org/10.5962/bhl.title.45918

Cuvier, G. (1834). *Lecciones elementales de la historia natural de los animales, dadas por G. Cuvier en la Escuela del Panteon de París i traducidas por don José Garriga y Baucis*. Imprenta de Cabrerizo, 2 tomos.

De Beni, M. (Ed.) (2016). *Ciencias y traducción en el mundo hispánico*. Universitas Studiorum (Col. *Pliegos Hispánicos*, 3).

De Beni, M. (2019). Georges Cuvier en España. La traducción de José Garriga y Baucís de *Lecciones elementales de la historia natural de los animales*. *Quaderns de filología. Estudis lingüístics*, 24, 309-328. https://doi.org/10.7203/qf.24.16314

Dictionnaire de l'Académie Française [1694, 1762, 1798, 1835, 1932-5]. http://artfl-project.uchicago.edu/node/17/

Enciclopedya Britannica (2021). Britannica. https://www.britannica.com/

Gómez de Enterría Sánchez, J. (1999). Las traducciones del francés. Cauce para la llegada a España de la ciencia ilustrada. Los neologismos en los textos de botánica. En F. Lafarga (ed.), *La traducción en España (1750-1830). Lengua, literatura, cultura* (pp. 143-155). Edicions de la Universitat de Lleida.

González, J. M. (1955). *Historia de las Misiones dominicanas en China: 1800-1900*. Juan Bravo.

Guérin, F.-É. (dir.) (1833-1839). *Dictionnaire pittoresque d'histoire naturelle et des phénomènes de la nature*. Au bureau de souscription.

Jiménez Rojas, Y. (2020). Aproximaciones al estudio de la inmigración china en Cuba: contextos, tendencias y espacios baldíos. *Revista de la Red de Intercátedras de Historia de América Latina Contemporánea*, 7/12, 104-124.

Littré, É. (1872-1877). *Dictionnaire de la langue française*. https://www.littre.org/

Machálek, T. (2014). KonText – Corpus Query Interface. FF UK, Praha. http://kontext.korpus.cz/

Martínez Robles, D. (2002). Las fuentes españolas sobre China en los siglos XVIII y XIX. En F. J. Antón Burgos & L. Óscar Ramos Alonso (eds.), *Traspasando fronteras: el reto de Asia y el Pacífico*. Vol. 1 (pp. 383-392). Asociación Española de Estudios del Pacífico y Universidad de Valladolid.

Martínez Robles, D. (2007). *La participación española en el proceso de penetración occidental en China: 1840-1870* [tesis doctoral. Universidad Pompeu Fabra]. Tesis Doctorals en Xarxa. https://cutt.ly/5RKxDCN

Melhoramentos (2021). *Michaelis. Dicionário Brasileiro da Língua Portuguesa*. http://michaelis.uol.com.br

Moreno Villanueva, J. A. (2017). Los Manuales Gallach: materiales para la historia de la lexicografía especializada. En I. Sariego López, J. Gutiérrez Cuadrado & C. Garriga Escribano (eds.), *El diccionario en la encrucijada, de la sintaxis y la cultura al desafío digital* (pp. 647-664). Escuela Universitaria de Turismo Altamira – Asociación Española de Lexicografía Hispánica.

Oxford University Press (2021). *Oxford English Dictionary Online* (*OED*). http://www.oed.com/

Perrais, F. (1997). Français et Chinois à Shanghai (1850-1880). En J. Weber (Dir.), *La France en Chine (1853-1943)*. Presses académiques de l'Ouest, Centre de recherches sur l'histoire du monde atlantique, Université de Nantes.

Real Academia Española. *Corpus de Referencia del Español Actual* (*CREA*). https://cutt.ly/KRKcoVA

Real Academia Española. *Corpus del Español del Siglo XXI* (*CORPES XXI*). https://cutt.ly/qRKcfKZ

Real Academia Española. *Corpus Diacrónico del Español* (*CORDE*). https://cutt.ly/wRKcjjM

Real Academia Española. *Diccionario histórico de la lengua española* (*DHLE*). https://cutt.ly/fRKclXp

Real Academia Española. *Nuevo tesoro lexicográfico de la lengua española* (*NTLLE*). https://cutt.ly/ERKcc4a

Real Academia Española & Fundación Rafael Lapesa. *Fichero general*. https://cutt.ly/RRKce9N

Rey, A. (Dir.) (2010): *Dictionnaire Historique de la langue française*. 3.ª ed. Le Robert.

Rodicio García, S. (1989). España y las guerras del opio. En F. Solano, F. Rodao & L. E. Togores (eds.), *Extremo Oriente Ibérico*. Publicaciones del Instituto de Cooperación para el Desarrollo – AECI.

Rodríguez Pastor, H. (1989). *Hijos del celeste imperio en el Perú (1850-1900). Migración, agricultura, mentalidad y explotación*. Instituto de Apoyo Agrario.

Sartor, E. (2019). Historia de un fragmento textual sobre la exploración de la China. Notas léxicas. *RILEX. Revista sobre investigaciones léxicas,* 2/1, 40-55. https://doi.org/10.17561/rilex.v2.n1.3

Sartor, E. & Dal Maso, E. (2017). Fortuna y traducción de las relaciones acerca del Lejano Oriente en España y Europa. Estudio textual y léxico sobre el *Discurso de la navegacion que los Portugueses hazen à los Reinos y Provincias del Oriente, y de la noticia que se tiene de las grandezas del Reino de la China* (1577) de Bernardino de Escalante y sus versiones inglesas (1579; 1745). En F. del Barrio de la Rosa (Ed.), *Palabras Vocabulario Léxico. La lexicología aplicada a la didáctica y a la diacronía* (pp. 237-249). Venezia, Edizioni Ca' Foscari (Col. *VenPalabras. Estudios de lexicología española*).

Tommaseo, N. & Bellini, B. (1861-1874). *Dizionario della lingua italiana*. http://www.tommaseobellini.it

Torruella Casañas, J. (2017). *Lingüística de corpus: génesis y bases metodológicas de los corpus (históricos) para la investigación en lingüística*. Peter Lang. https://doi.org/10.3726/b11170