

DOCUMENTACIÓN NOTARIAL EXTREMEÑA  
DEL SIGLO XVII EN ORALIA DIACRÓNICA  
DEL ESPAÑOL (ODE):  
EL LÉXICO DE LA VIDA COTIDIANA A TRAVÉS  
DE INVENTARIOS DE BIENES PACENSES

Inmaculada González Sopena

Departamento de Lengua Española, Facultad de Filosofía y Letras, Universidad de Granada, Campus de Cartuja, Calle del Prof. Clavera, s/n, 18011, Granada, España  
isopena@ugr.es

**Notarial documentation from Extremadura in the 17th century in Oralia diacrónica del español (ODE): the lexicon of daily life through inventories of goods from Badajoz**

**Abstract:** The objective of this study focuses on, on the one hand, developing the methodology used in the elaboration of the corpus Oralia diacrónica del español (ODE) and the subcorpus CORTENEX S. XVII (Corpus de Textos Notariales Extremeños S. XVII), which is integrated into it, and, on the other hand, analysing the lexicon of daily life that is present in the inventories of goods from Badajoz. The methodology implemented in the preparation of ODE and CORTENEX is based on the use of new technological tools: 1) the transcription of documents in the XML language (eXtensible Markup Language), and 2) the digital presentation of documentation on the TEITOK platform, in which linguistic processing tasks (normalization, lemmatization, and morphosyntactic annotation) are performed. The result translates into an open access resource that allows the viewing of a semi-paleographic edition with the facsimiles and a digitized critical edition of the selected texts. In addition, this work focuses on the historical-linguistic analysis of two terms present in the notarial documentation from Extremadura and linked to the daily life of the 17<sup>th</sup> century in relation to certain household items and the materials of which they are composed – *alioj* and *juera*. Finally, this study offers new data on the variety of Spanish developed in Extremadura from a diachronic and diatopic perspective through the historical-linguistic analysis of two cases, but, also, it develops one of the most pioneering methodologies in the elaboration of a linguistic corpus based on the application of digital tools that allows philological and computational linguistic interests to be unified.

**Keywords:** linguistic corpus; history of the lexicon; varieties of Spanish; notarial documentation; XML

**Resumen:** El objetivo de este trabajo reside, por un lado, en desarrollar la metodología llevada a cabo en la confección del corpus Oralia diacrónica del español (ODE) y del subcorpus CORTENEX (Corpus de textos notariales extremeños, S. XVII) en él integrado; y, por otro lado, en analizar de forma general el léxico de la vida cotidiana que se halla presente en los inventarios de bienes pacenses. La metodología llevada a cabo en la confección de ODE y CORTENEX se cimienta en el uso de nuevas herramientas tecnológicas: 1) la transcripción de los documentos en lenguaje marcado XML (eXtensible Mark Language) y 2) la presentación digital de la documentación en la plataforma TEITOK, en la que se realizan las tareas de procesamiento lingüístico (normalización, lematización y anotación morfosintáctica). El resultado se traduce en un recurso de acceso libre que permite visualizar una edición semipaleográfica con los facsímiles y una edición crítica digitalizada de los textos seleccionados. Con todo, este trabajo se centra en el análisis histórico-lingüístico de dos voces presentes en la documentación notarial extremeña, vinculadas a la vida cotidiana del siglo XVII en lo referente a determinados enseres domésticos y a los materiales que los componen, tales como *alioj* o *juera*. Finalmente, este estudio ofrece nuevos datos sobre la variedad de español desarrollada en Extremadura en perspectiva diacrónica y diatópica a través del análisis histórico-lingüístico de dos casos, pero, además, se desarrolla una de las metodologías más pioneras actualmente en la elaboración de corpus lingüísticos basada en la aplicación de herramientas digitales que permiten unificar intereses filológicos y de lingüística computacional.

**Palabras clave:** corpus lingüísticos; historia del léxico; variedades del español; documentación notarial; XML

## 1. Introducción

Este trabajo persigue el objetivo de desarrollar la metodología llevada a cabo en la confección del corpus Oralia diacrónica del español (ODE) y del subcorpus CORTENEX (Corpus de textos notariales extremeños, S. XVII) en él integrado; y, además, analizar de forma general el léxico de la vida cotidiana que se halla presente en los inventarios de bienes pacenses, aportando el estudio de dos casos concretos.

La variedad de español que se desarrolla en los territorios de Extremadura ha merecido escasa atención entre la comunidad científica y precisa de estudios acerca de los aspectos lingüísticos que allí prosperan en perspectiva diacrónica:<sup>1</sup>

Seguramente los viejos tópicos perduran porque a veces resultan ciertos. Así sucede con el de Extremadura como región abandonada, al menos en lo referente al estudio de su lengua [...] De las muchas carencias que todavía presenta el estudio lingüístico de Extremadura, una de ellas es la de su diacronía (Ariza 1985: 7).

Por el contrario, pueden consultarse bastantes monografías de corte dialectal sobre el habla de algunas localidades pacenses y cacereñas en el siglo XX (Viudas Camarasa 1987, Fernández de Molina 2014).<sup>2</sup> En lo que respecta a recursos lingüísticos sobre

<sup>1</sup> Contamos con estudios marginales como los de Ariza (1985), Marcos Álvarez (1992), entre algunos otros, en lo relativo a estudios estrictamente diacrónicos y apoyados en documentación archivística vinculada geográficamente a la región extremeña.

<sup>2</sup> Las escasas monografías que existen dedicadas a esta variedad lingüística de español se han centrado en el análisis de la conciencia lingüística, así como en el establecimiento de los rasgos fonético-fonológicos, morfosintácticos y léxicos que se han obtenido a raíz de estudios dialectales (Viudas Camarasa 1987, García Mouton 1996, Montero Curiel 2006).

esta variedad, también se elaboró la *Cartografía lingüística extremeña*, desarrollada por González Salgado (2003). Este recurso resulta adecuado para acercarse al habla dialectal actual de Extremadura, pero es insuficiente para constatar la evolución histórica de los fenómenos lingüísticos que allí se desarrollan a lo largo de su historia. De otro lado, se carece, además, de cualquier tipo de corpus especializado de documentación manuscrita antigua extremeña de distintos tipos y de acceso libre que pueda servir de base para estudios histórico-lingüísticos.

Bien conocida es la tradicional división de la comunidad extremeña en dos áreas lingüísticas: una denominada como la Extremadura leonesa y, otra, como la Extremadura castellana (Viudas Camarasa 1987: 52). Esta división se sustenta en los distintos tiempos históricos que ocupó su reconquista: desde mediados del siglo XII hasta 1212 se produjo la ocupación territorial cristiana hasta el río Tajo y se repartieron zonas entre Castilla y León; a partir de 1212 se produjo una extensión de León por toda Extremadura; en el año 1230 Castilla y León se unifican en un solo reino. Poco a poco, la influencia del área meridional de la Península empezó a notarse en el sur de Extremadura. Finalmente, desde el establecimiento de la corte en Madrid en 1561, la influencia lingüística castellana comenzó a sentirse en las antiguas áreas leonesa y castellana, mientras que, en el sur, la actual provincia de Badajoz quedó bajo el influjo meridional. Todos estos hechos históricos configuraron una variedad lingüística en Extremadura que, dentro de la dialectología, se ha denominado *área de hablas de tránsito* (Zamora Vicente 1974), dado que confluyen rasgos lingüísticos leoneses, castellanos y meridionales (la presencia del diminutivo *-ino*, la aspiración de *f-* inicial latina, léxico específico de cada área, etc.).

Por ello, la documentación notarial extremeña del siglo XVII que aquí se presenta permitirá obtener avances en el estado de la cuestión que presenta el estudio diacrónico de la variedad de español que se configura en Extremadura. Este tipo de textos resultan muy valiosos para plantear estudios de historia del léxico, dado que dejan constancia de voces fuera del registro general, términos no documentados, variantes diatópicas, préstamos, leonesismos, etc. (Morala 2014) y pueden servir de base para el estudio de otros fenómenos lingüísticos morfosintácticos y fonéticos, como la presencia de confusión de sonidos sibilantes, modificación de *-l-* agrupada, neutralización *r/l*, entre otros.

Desde el punto de vista metodológico, este estudio se cimenta sobre los datos que aporta el Corpus de Textos Notariales Extremeños (CORTENEX S. XVII), un subcorpus especializado integrado a Oralia diacrónica del español (ODE). Sobre esta base, se realiza un estudio histórico de dos voces: *juera* y *alioj*.

Con todo, este trabajo se divide en las siguientes secciones para el tratamiento y estudio lingüístico de documentos notariales extremeños del siglo XVII: 1) una introducción al estado de la cuestión sobre el estudio de la variedad de español extremeña; 2) el análisis de las bases metodológicas de CORTENEX; 3) el análisis de la documentación notarial extremeña seleccionada en su contexto histórico; 4) un estudio histórico-lingüístico de dos voces presentes en los textos pacenses; y 5) las conclusiones generales a todo el estudio realizado.

## 2. Bases metodológicas del corpus CORTENEX S. XVII

Las fuentes documentales de carácter notarial han demostrado ser un gran recurso para la historia del léxico español (CorLexIn 2014, Puche Lorenzo 2018, Pérez Toral 2021, Morala 2016), pues en ellas se lista una gran variedad de objetos, muebles, enseres, prendas, telas e instrumentos diversos vinculados a múltiples aspectos de la vida cotidiana del pasado. Al mismo tiempo, este tipo de textos presenta, por lo general, una datación y una localización precisas y suelen presentar la misma estructura en su elaboración, lo cual facilita mucho los estudios comparativos y se asemejan a las actuales encuestas dialectales que se realizan como base de atlas lingüísticos (Morala 2012: 202). Los inventarios de bienes, asimismo, se asientan como un tipo textual que permite rastrear fenómenos propios de la oralidad, en este caso, del pasado. Ello es así dado que «suponen el traslado al papel de las declaraciones orales de tasadores [...] por lo que no es de extrañar que se cuele en los documentos rasgos lingüísticos del vernáculo de los escribanos» (Calderón y Vaamonde 2020: 169).

Teniendo en cuenta lo anterior, CORTENEX S. XVII se ha confeccionado a partir de inventarios de bienes, testamentos, particiones, cartas de dote, etc., de la primera mitad del siglo XVII vinculados a la comunidad de Extremadura. En su diseño, se ha incluido documentación de Cáceres<sup>3</sup> y de Badajoz. Por el corte cronológico y geográfico, así como por los tipos textuales contemplados, se puede caracterizar este corpus como especializado o, siguiendo la denominación de la geografía lingüística, de «pequeño dominio» (Calderón Campos 2022: 318). Este tipo de banco de datos ofrece grandes ventajas para plantear estudios de dialectología histórica, datos que, a su vez, son casi imposibles de rastrear en los grandes corpus de referencia del español, sobre todo, en lo relativo a fenómenos de variación lingüística dialectal.

La metodología sobre la que se ha elaborado CORTENEX S. XVII se basa en el uso y explotación de nuevas herramientas tecnológicas. Concretamente, este corpus se integra a Oralia diacrónica del español (ODE) como subcorpus de control y especializado en un área geográfica, a saber, Extremadura. Se aloja en la plataforma TEITOK (Janssen 2016) que ODE ha implementado para el procesamiento y consulta de documentación andaluza desde el siglo XVI hasta el siglo XIX. De este modo, en ODE confluyen intereses vinculados a la filología, por presentar ediciones digitales rigurosas de documentación archivística y, además, aspectos propios de la lingüística computacional, por realizar tareas estrictas de procesamiento lingüístico, como la anotación morfosintáctica y la lematización (Calderón y Vaamonde 2020: 169). Debido a su carácter de subcorpus en ODE, las fases en la transcripción y tratamiento de los textos de CORTENEX son las mismas: 1) tokenización, 2) normalización, 3) lematización y 4) anotación morfosintáctica.

En primer lugar, la transcripción de los textos notariales seleccionados se realiza en lenguaje marcado XML (eXtensible Markup Language), siguiendo los estándares de codificación para la digitalización de textos propuestos por el consorcio TEI (*Text Encoding Initiative*, 2016) en lo relativo a documentación manuscrita antigua. Cada

<sup>3</sup> De momento, CORTENEX solo incluye la documentación notarial de Badajoz, pero está previsto añadir documentación notarial de la provincia de Cáceres del mismo corte cronológico.

documento XML-TEI se divide en dos partes: 1) cabecera y 2) cuerpo. En la cabecera se añade toda la información metatextual correspondiente al título, la datación, los folios o el proyecto. El cuerpo se traduce en la transcripción de los textos seleccionados, haciendo uso de una serie de marcas angulares denominadas *etiquetas* (<etiqueta>). No obstante, existe una serie de etiquetas recurrentes y de uso consensuado en la elaboración del corpus ODE y CORTENEX, como la marcación de los añadidos y su posición, las tachaduras, el inicio de página y de línea o las conjeturas (Calderón Campos 2019: 25). Los criterios de transcripción se basan en mantener las grafías originales del documento, así como en separar las palabras y actualizar el uso de las mayúsculas y la puntuación según la norma vigente actual (Arrabal 2020: 69).

```

BA160117035.xml X
TEI text body p
207 <pb n="380v" facs="IMG_1492.JPG"/>
208 <p>
209
210 <lb/> Una caja de cuchillos dorada
211 <lb/> tasada en treynta y vn
212 <lb/> rreales.
213 <lb/> çinco asaderos tasados en
214 <lb/> çinco rreales.
215 <lb/> Vn plato y dos salseras de estano
216 <lb/> en syete rreales.
217 <lb/> Vna tasa y un salero de plata,
218 <lb/> que peso entr||entre anbas pieças du<lb/>çientos y veynte y seys rreales.
219 <lb/> Treze mill rreales en dinero de
220 <lb/> contado.
221 <lb/> Que en todos los dhs||dichos bi<add place="above">s</add>||bienes, |h|nerero,
222 <lb/> y prescas, suma y monta los dhs||dichos
223 <lb/> diez y seys mill y quinientos rreales,
224 <lb/> los quales en las dhas||dichas pieças, al<lb/>fayas y prescas y dinero req
225 <lb/> presençia del escriuano y t<add place="above">o</add>||testigos d||de esta
226 <lb/> de la qual paga y entrega yo, el ps<add place="above">te</add>||presente es
227 <lb/> doy fe que se hizo en mi presençia y de los
228 <lb/> t<add place="above">o</add>||testigos d||de esta carta en la forma en ella
...
    
```

Figura 1. Ejemplo de transcripción CORTENEX

En la Figura 1 se puede observar un ejemplo de transcripción de una carta de dote de Badajoz en XML-TEI. En segundo lugar, las transcripciones en formato XML-TEI, junto con las imágenes facsimilares del documento en formato JPG se importan a la plataforma TEITOK para su procesamiento lingüístico. Una vez que los documentos y sus imágenes están alojados en la plataforma, se procede a realizar los pasos fundamentales en su procesamiento lingüístico. Como se ha apuntado más arriba, primero se tokeniza el texto, así, cada palabra queda representada entre *tokens* sobre los que se irá ampliando, a base de etiquetas y atributos, toda la información relacionada con la forma transcrita, expandida y normalizada. Tras esto, se realiza la normalización semiautomática de los documentos, si bien siempre se requiere de una revisión manual.<sup>4</sup>

Una vez que se tiene por completo la forma normalizada de los documentos, la siguiente fase es anotar y lematizar cada palabra. Esto se hace forma semiautomática gracias a la herramienta *Neotag* (Janseen 2012), a través de la cual se añaden nuevos

<sup>4</sup> Dado que esta herramienta semiautomática toma necesariamente como referencia, a la hora de normalizar, un corpus de entrenamiento, es fundamental revisar manualmente las formas que no hayan sido correctamente normalizadas según la norma actual del español.

atributos a cada *token*: 1) un atributo POS (Part of speech) y 2) un atributo lema. Así puede observarse en la Figura 2:

ode/corpus/BA160117035.xml

### Carta de pago de dote de Beatriz de Mesas

```

76 </p>
77 <pb n="380r" facs="IMG_1491.JPG" id="e-81"/>
78 <p id="p-3">
79 <lb id="e-82"/> <tok id="w-552" lemma="verde" pos="AQ0CS0">verde</tok> <tok id="w-553" lemma="y" pos="CC">y</tok> <tok id="w-554
80 <lb id="e-83"/> <tok id="w-560" lemma="en" pos="SPS00">enc</tok> <tok id="w-561" nform="ciento" lemma="100" pos="Z">ciento</tok>
81 <lb id="e-84"/> <tok id="w-564" lemma="y" pos="CC">y</tok> <tok id="w-565" nform="cuatro" lemma="4" pos="Z">quatro</tok> <tok id
82 <lb id="e-85"/> <tok id="w-568" lemma="2" pos="Z">Dos</tok> <tok id="w-569" lemma="tapete" pos="NCP000">tapetes</tok> <tok id="w
83 <lb id="e-88"/> <tok id="w-586" nform="un" lemma="un" pos="DI0MS0">Vn</tok> <tok id="w-587" lemma="cobertor" pos="NCMS000">cober
84 <lb id="e-89"/> <tok id="w-591" lemma="en" pos="SPS00">enc</tok> <tok id="w-592" nform="treinta" lemma="30" pos="Z">treynta</tok>
85 <lb id="e-90"/> <tok id="w-595" lemma="2" pos="Z">Dos</tok> <tok id="w-596" lemma="mesa" pos="NCFP000">mesas</tok> <tok id="w-59
86 <lb id="e-91"/> <tok id="w-599" lemma="entero" pos="AQ0FP0">enteras</tok> <tok id="w-600" nform="romanescas" lemma="romanesco" p
87 <lb id="e-92"/> <tok id="w-602" nform="cincuenta" lemma="50" pos="Z">cinquenta</tok> <tok id="w-603" nform="reales" lemma="real"
88 <lb id="e-93"/> <tok id="w-605" lemma="adarnascado" pos="AQ0FP0">adarnascadas</tok> <tok id="w-606" lemma="," pos="Fc"/> </tok> <tok
89 <lb id="e-94"/> <tok id="w-609" nform="cuarenta" lemma="40" pos="Z">quarenta</tok> <tok id="w-610" nform="reales" lemma="real" p
90 <lb id="e-95"/> <tok id="w-612" lemma="otro" pos="DI0FP0">otras</tok> <tok id="w-613" lemma="2" pos="Z">dos</tok> <tok id="w-614
91 <lb id="e-97"/> <tok id="w-619" lemma="en" pos="SPS00">enc</tok> <tok id="w-620" nform="cincuenta" lemma="50" pos="Z">cinquenta</
92 <lb id="e-98"/> <tok id="w-623" lemma="otro" pos="DI0FS0">otra</tok> <tok id="w-624" lemma="mesa" pos="NCF5000">mesas</tok> <tok
93 <lb id="e-99"/> <del rend="cancelled"><tok form="--" id="w-628" nform="lienzo" lemma="lienzo" pos="NCMS000">lienço</tok></del> <
94 <lb id="e-100"/> <tok id="w-633" nform="reales" lemma="real" pos="NCP000">rreales</tok> <tok id="w-634" lemma="," pos="Fp"/> </to
95 <lb id="e-101"/> <tok id="w-635" lemma="3" pos="Z">tres</tok> <tok id="w-636" lemma="mesa" pos="NCFP000">mesas</tok> <tok id="w
96 <lb id="e-102"/> <tok id="w-639" lemma="de" pos="SPS00">de</tok> <tok id="w-640" nform="lienzo" lemma="lienzo" pos="NCMS000">lie
97 <lb id="e-103"/> <tok id="w-645" lemma="otro" pos="DI0FS0">otra</tok> <tok id="w-646" lemma="mesa" pos="NCF5000">mesas</tok> <tok
98 <lb id="e-104"/> <tok id="w-649" nform="más" lemma="más" pos="R">mas</tok> <tok id="w-650" lemma="pequeño" pos="AQ0FS0">pequeña<
99 <lb id="e-105"/> <tok id="w-656" nform="un" lemma="un" pos="DI0MS0">Vn</tok> <tok id="w-657" lemma="almohada" pos="NCF5000">almo
100 <lb id="e-106"/> <tok id="w-660" lemma="su" pos="DP3CP0">sus</tok> <tok id="w-661" pos="NCFP000" lemma="bellota">bellotas</tok>
101 <lb id="e-107"/> <tok id="w-665" nform="reales" lemma="real" pos="NCP000">rreales</tok> <tok id="w-666" lemma="," pos="Fp"/> </to
102

```

Figura 2. Ejemplo de normalización, anotación y lematización

Las etiquetas morfosintácticas se basan en una versión simplificada de las que ofrece el grupo EAGLES.<sup>5</sup> Tras esta fase, es también necesario revisarlas y corregir manualmente aquellas que no se correspondan<sup>6</sup> con el sentido del texto. La visualización final de los documentos, gracias al empleo de todas estas herramientas, se ofrecen en una interfaz cómoda a la vista humana, como aparece en las Figuras 3 y 4, con diferentes opciones según se quiera ver la transcripción semipaleográfica o la forma normalizada:

Sepan quantos esta carta de pago y dote vieren como nos, Manuel Xuarez, escriuano del numo d esta ciudad de Badajoz, y Beatris de Mesas, mi mugr, vos d esta ciudad, con licencia y facultad y espreso consentimiento y yo, la dha Beatris de Meza, pido y demando a uos, el dh mi marido, para otorgar esta carta y la jurar con licencia y facultad y yo, el dh Manuel Xuarez, conçoedo a uos, la dh mi mugr, quan bastante de dro se requiere; y d ella vsando dezimos que por quanto Jorxe de Meza, nro yerno, y Alo Lopez y Leonor de Meza, su mugr, vos d esta ciudad y estantes al presente en ella como principales y Ju Diaz de Morales, voz d esta dha ciudad, como su abonador, se obligaron de dar y pagar



Figura 3. Visualización de la forma transcrita y el facsímil

<sup>5</sup> <<https://www.cs.upc.edu/~nlp/tools/parole-sp.html>>.

<sup>6</sup> Se producen problemas con el etiquetador semiautomático en palabras homónimas que pertenecen a categorías gramaticales diferentes o en palabras que, por la grafía con las que se han escrito, se confunden (*sera* con el valor ‘cera’ y con el valor del futuro del verbo *ser*).

Sepan cuantos esta carta de pago y dote vieren como nosotros, Manuel Juárez, escribano del número de esta ciudad de Badajoz, y Beatriz de Mesas, mi mujer, vecinos de esta ciudad, con licencia y facultad y expreso consentimiento y yo, la dicha Beatriz de Mesas, pido y demando a vos, el dicho mi marido, para otorgar esta carta y la jurar con licencia y facultad y yo, el dicho Manuel Juárez, concedo a vos, la dicha mi mujer, cuan bastante de derecho se requiere; y de ella usando decimos que por quanto Jorge de Mesas, nuestro yerno, y Alonso López y Leonor de Mesas, su mujer, vecinos de esta ciudad y estantes al presente en ella como principales y Juan Díaz de Morales, vecinos de esta dicha ciudad, como su abonador, se obligaron de dar y pagar



Figura 4. Visualización de la forma normalizada y el facsímil

Gracias a todas las fases descritas en el procesamiento lingüístico de los textos, se obtiene como resultado un corpus anotado con un motor de búsquedas CQL (Corpus Query Language), que permite realizar consultas en el corpus a nivel muy avanzado y a través de diversos criterios (el año, el lema, la etiqueta morfosintáctica, la forma transcrita, el tipo de documento, etc.). En la Figura 5 se visualiza el ejemplo de la búsqueda por lema de la voz *cedazo* en Badajoz, junto con las concordancias que se encuentran:

Búsqueda en   
 CQL:

12 resultados • ipm: 21.68 / 159.63

Texto:  Forma transcrita  Expansión  Forma expandida  Forma normalizada  Transcripción

Etiquetas:  Etiqueta POS  Lema  Notas lingüísticas

contexto	en tres reales.   Vn <b>cedaco</b> un real.   Dos tovajas	1601 Badajoz
contexto	en seys relaes.   Dos <b>cedacos</b> e un criuo e juera	1601 Badajoz
contexto	con sus bancos.   Tres <b>cedacos</b> .   Vn peso de garfios	1601 Badajoz
contexto	en tres reales.   Dos <b>cedaços</b> en quatro reales.   Vna	1601 Badajoz
contexto	.   Vn artesa y dos <b>sedacos</b> y vn   criuo, en	1635 Badajoz
contexto	hierro en quatro reales   Dos <b>sedasos</b> y una juera en ocho	1645 Badajoz
contexto	en tres ducados.       Dos <b>sedasos</b> , criuo y juera,	1633 Badajoz
contexto	en mo ducado.   Dos <b>sedasos</b> y un criuo en   seis	1601 Badajoz
contexto	treze reales que   costo   Dos <b>sedazos</b> y un criuo   y una	1629 Badajoz
contexto	en seis rreales.   Dos <b>sedaços</b> y un criuo en seis	1601 Badajoz

Figura 5. Ejemplo de búsqueda por lema de la palabra *cedazo*

### 3. La documentación notarial pacense del siglo XVII. Contextualización y primeros resultados de análisis léxico

Lingüísticamente, la explotación filológica de textos notariales extremeños antiguos ha sido escasa. Son muy pocos los estudios que existen de este carácter, como los de Marcos Álvarez (1988, 1992)<sup>7</sup> o los datos que ofrece el corpus CorLexIn,<sup>8</sup> los cuales han sido el punto de referencia para comenzar la elaboración de CORTENEX S. XVII.

La búsqueda y selección de los textos ha sido llevada a cabo a través del Archivo Histórico Provincial de Badajoz.<sup>9</sup> Para ello, se han establecido cortes cronológicos que dividen la primera mitad del siglo XVII en intervalos que oscilan entre los 5 y 10 años, con el objetivo de obtener una muestra de análisis lo suficientemente representativa para plantear estudios léxicos. Actualmente, la relación de documentos seleccionados y disponibles para su consulta en el corpus, ordenados por notarios y fechas, así como por tipo textual, es la siguiente:

<b>Diego Martín Sequera (1615-1645)</b>	Inventarios de bienes	7
	Cartas de dote	11
	Testamentos	11
<b>Diego López (1601)</b>	Inventarios de bienes	1
	Cartas de dote	10
<b>Baltasar Suárez (1624)</b>	Inventarios de bienes	3
<b>Manuel de León (1633-1635)</b>	Dotes	4
	Inventarios	5
	Testamentos	1
<b>Diego Martín Gamo (1561)</b>	Dote	1
	Inventario	
	Inventarios	2
<b>Total</b>		56

Tabla 1. Documentación CORTENEX S. XVII disponible en ODE

Dado que este estudio parte de los postulados de la lingüística externa que vincula los cambios políticos, económicos y sociales de un momento concreto a los cambios léxico-semánticos que se observan en la lengua, es necesario contextualizar toda la documentación seleccionada.<sup>10</sup> Se ha procurado tener en cuenta el criterio histórico, esto es, la división en relación con los diferentes reinados que se suceden desde

<sup>7</sup> Habría que incluir los estudios históricos realizados por Arcadio Guerra Guerra (1977, 1980), dado que aportan una descripción bastante precisa de la documentación notarial pacense del siglo XVII.

<sup>8</sup> Se suma a estos trabajos la tesis doctoral de Sánchez Sierra (2019), que recopila un volumen importante de documentación notarial extremeña.

<sup>9</sup> Este archivo cuenta con un fondo de Protocolos Notariales de 5253 legajos datados desde 1523 hasta 1917, <<https://acortar.link/nVNapn>>.

<sup>10</sup> Esta concepción de los factores externos como impulsores de cambios en la lengua ha quedado plasmada en multitud de trabajos como los de Álvarez de Miranda (1992) o García-Godoy (2012), entre muchos otros.

finales del siglo XVI hasta mediados del siglo XVII, por marcar todo el contexto económico y social que se desarrolla a continuación.

En primer lugar, de forma escueta, es preciso aportar algunos datos sobre el perfil poblacional y ocupacional de los habitantes pacenses a comienzos de esta centuria, así como ciertos datos de la historia social, política y económica de la zona en el período señalado. Por un lado, la actual comunidad extremeña pertenece a la corona castellana desde su anexión a finales del siglo XIII. Por otro lado, si ponemos el foco en Badajoz, sus villas y ciudades vivieron una época de crecimiento poblacional desde finales del siglo XVI hasta aproximadamente mediados del siglo XVII, esto supone el período cronológico que abarca parte del reinado de Felipe III (1598-1621) y el de Felipe IV (1621-1665).

En términos generales, puede decirse que, desde finales del XVI hasta mediados del XVII, la población extremeña se estima en unos 10.000 habitantes (Guerra 1977: 7). Nos encontramos ante una población eminentemente rural, dedicada a labores vinculadas al sector primario mayoritariamente (agricultura y ganadería), seguido de los sectores secundarios y terciarios (cargos administrativos y artesanos agrupados en gremios). Aunque el número de personas dedicadas al trabajo gremial era significativamente menor, destacan los dedicados a la manufactura textil, sobre todo, en la confección de paños. Mucho menor fue el número de personas dedicadas a otras profesiones o cargos administrativos, como guardas, barberos, herradores, maestros o médicos. El poder adquisitivo y las rentas familiares fueron bastante bajos, lo cual contrasta con la enorme riqueza de telas, atuendos, prendas y vestimentas que se observan en las relaciones de bienes de la primera mitad del siglo XVII (Marcos Álvarez 1992: 1162). La explicación a esto debe buscarse en el contexto social de la época. El siglo XVII se cataloga tradicionalmente como el siglo del Barroco español, época marcada por un declive político, económico y moral creciente hasta finales de siglo, pero que, al mismo tiempo «aspira a enmascarar en la grandilocuencia, la ampulosidad o la evasión, la crisis que se debate sobre el país» (Ubieto *et alii* 1983: 380).<sup>11</sup>

Es importante destacar que desde el inicio del siglo comenzó un periodo de decadencia en todos los planos, en correspondencia con el resto de Europa. En este contexto es importante no olvidar hechos históricos fundamentales como la expulsión definitiva de los moriscos (1609)<sup>12</sup> o los conflictos bélicos que se producen simultáneamente en Cataluña y Portugal (1640).<sup>13</sup> La enorme crisis política y económica se dejó ver en la creciente inflación,<sup>14</sup> en las nuevas cargas fiscales impuestas a la

<sup>11</sup> El siglo XVII se resume en la síntesis «de dos factores: la pobreza de las fábricas y la magnificencia decorativa» que se observa, sobre todo en la arquitectura (Ubieto *et alii* 1983: 408). Esto tiene su correlación con lo que se observa en el plano literario, con las corrientes conceptista y culterana.

<sup>12</sup> Este hecho produjo una merma considerable en los frutos obtenidos de la explotación agrícola, que en parte se frenaron a raíz de la introducción de nuevos cultivos como el maíz o la patata (Ubieto *et alii* 1983: 387).

<sup>13</sup> Bajo el reinado de Felipe IV comenzó la Guerra de la Restauración Portuguesa (1640), cuyas consecuencias demográficas se estiman catastróficas, sobre todo para la zona centro occidental de la actual Extremadura (Blanco Carrasco 1998).

<sup>14</sup> Se sucedieron numerosas bancarrotas (1627, 1647, 1656, 1664) (Ubieto *et alii* 1983: 383).

población, en la reducción de comercio con América y en la pérdida de otros territorios del imperio español, fruto de la pésima gestión económica que se hizo más patente bajo el reinado de Felipe IV (Marcos Álvarez 1988).<sup>15</sup> Estos hechos se ven, a su vez, enmascarados en la excesiva decoración en las artes o en la enorme grandilocuencia en las letras, que tanto caracterizan al Barroco español (Blanco Carrasco 1998: 96).<sup>16</sup>

Teniendo en cuenta todo lo anterior, el tipo de documentación notarial seleccionada para este estudio proporciona una idea bastante aproximada de cómo vestían los extremeños, de su nivel económico, de su alimentación, de ciertas tradiciones jurídicas y religiosas, así como del nivel de industrialización de la ciudad (Guerra 1980: 446). La riqueza y variedad léxica que se atestigua en los primeros análisis del CORTENEX S. XVII se reparte por multitud de campos léxico-semánticos de la vida cotidiana. En lo que se refiere a los tipos de telas, de hilos y de otros elementos empleados en la confección de la indumentaria de la época, el corpus documenta en reiteradas ocasiones voces como *tafetán, naval, paño, anascote, esterlín, jerguilla, jerga, felpa, bayeta, burato, estopa, esterilla, fustán, cotonía, cambray, seda, hilo de pita, hilo de acarreto, piel de rata, brocatel, cochinilla, lana, grana, tabí, damasco, encaje, perpetuán, raso, terciopelo, vaqueta, galón, bramante, espolín, gorgorán, tirela, esparragón, bocací, bombasí, camadillo, gurbión, sarga, lona*, además de todo tipo de lienzos (nuevos, viejos, caseros, randados, leonados, de laberinto, con borlas, con flecos, etc.).

Muy variada es también la indumentaria propiamente dicha, así como ropa de cama: aparecen *guadamecés, paños* de muchos colores (pardos, carmesíes, azules, amarillos, colorados, noguerados, de granadilla, negros, blancos, verdes, verdegay, largueados), *jubones, frisas, ferreruelos, fajas, basquiñas, medias, faldellines, valonas, hábitos, cobertores, colchas, pasamanos, antepuertas, antecamas, sobre colchas, sobremesas, sábanas, vestidos, camisas, camisonos, chapines, mantas, mantillas, conteras, capotes, cintas, tranzaderas, mitanes, toallas, manteles* (alemaniscos, romanescos, reales). Además, tanto las telas como las prendas son de muy diversa procedencia, tanto nacional como internacional (de Moscovia, de Portugal, de Francia, de Holanda, de la India, de la China, de Toledo, de Segovia, de Palencia, de Valencia). Se han podido documentar numerosos objetos de decoración, adornos y joyas, así como armas: *sortijas, pendientes, zarcillos, gargantillas* (de plata y de oro, pero también de plata y oro falso), con *esmeraldas, con diamantes, con perlas, de aljófar, de nácar, rosarios, cadenas, cuadros, hechuras* de diversos personajes religiosos, *cruces* de perlas, de cristal, *peras* engastadas en oro, *peinadores, rascamoños, porcelanas, galones, cálices, arcabuces, alabardas y espadas*.

Entre los muebles y objetos de casa se documentan, asimismo, *mesas* de distintos materiales (de madera, de ébano, de bordo, de nogal), *escritorios, bufetes, bufetillos, contadores, mostradores, sillas imperiales, sillas de respaldo, escabelillos, arcas* (de madera,

<sup>15</sup> El reinado de Felipe IV estuvo marcado por una pésima situación en la hacienda pública, fruto de las cargas fiscales mencionadas más el coste de mantenimiento de los territorios que aún formaban parte de la Corona española (Ubieto *et alii* 1983, Marcos Álvarez 1988), además de las secuelas que todavía dejaba la peste en ese momento.

<sup>16</sup> Además, el azote de enfermedades como la peste todavía continuaba causando estragos en Europa (Marcos Álvarez 1988: 663).

marfil, encoradas) *medias camas, camas enteras, camas de campo, camas con rieles, cortinas y alamares, acericos, almohadas, cabeceros, cojines, alfombras, catalufas, colchones, traveseros, espejos.*

Igual de variados son los enseres de cocina: *bacías de latón, fuentes de peltre, de azófar, tinajas (grandes, pequeñas, medianas), vasijas, tembladeras, potes, jarros, mesas de manteles, fruteros, cofres, artesas, almiércoles, salvillas, morteros de mármol o alioj, picheles, peroles, pesos, morillos, asadores, trébedes, espumaderas, vinagreras, sartenes, bernegales, cestas, escudillas, canastillas, loza, palmatorias, almofías, taburetes, bancos, banquillos, estrados, saleros, cuchillos, cucharas, tenedores (de hierro, de plata), tenazas, candeleros, candiles, badiles, ralladores, tachos, braseros, calentadores, barreños, alcuzas, cazos, cazos de cabo (de cobre, de hierro, de peltre), balanzas, cepillos, alquitaras, etc. También se codifican vocablos relativos a las actividades agrícolas y ganaderas, tales como *cedazo, criba, zaranda, ceranda, juera, harnero, timones de arado, raederas, bueyes, mulas, yeguas, potros, añacas, cerdos, caballos, jumentos, rozas, viñas, olivos, costales de cebada, de centeno, de harina, de trigo, entre otros cultivos, como el cáñamo, el azúcar o el tabaco.* Otro léxico de interés es el que se vincula con los pesos y monedas de la época, por ejemplo, aparecen sistemáticamente *reales, ducados, maravedíes, monedas de vellón, tostones o tarjas, así como medidas de múltiples tipos: resmas de papel, treintanarios de misas, gruesas de botones, pies, fanegas, entre otras.**

Como puede observarse, la documentación notarial seleccionada incluye muchos aspectos relacionados con el campo, la agricultura, el tipo de cultivos, así como de joyas, tipos de telas, vestimenta, objetos y muebles de cocina. De esta forma se proporciona una visión bastante precisa sobre la vida cotidiana del Badajoz del siglo XVII, a modo de atlas lingüístico diacrónico, de pequeño dominio. Así, el volumen documental aportado permite abordar múltiples tipos de estudios de historia del léxico español: análisis de casos de préstamos de diferentes lenguas, de extremeñismos, de occidentalismos, andalucismos, estudios de competencia y sustitución léxicas o estudios comparativos con los datos que arroja ODE para Andalucía con el fin de establecer áreas dialectales de convergencia y de divergencia, por presentar documentos de la misma tipología textual.

Ejemplos de ello se encuentran en términos léxicos propios de Extremadura como *juera*, occidentalismos como *almofía*, se constatan casos de competencia léxica entre voces de diferentes orígenes etimológicos como *mármol* y *alioj*, *mortero* y *almirez*, así como de voces del mismo origen árabe como *azófar* y *latón*. Se constatan casos de andalucismos como *acarreto* (hilo de acarreto) o el uso de palabras tradicionalmente catalogadas como propias de Andalucía occidental como *entenado/a*, actualmente reemplazadas por *hijastro/a*.

#### 4. Estudio de casos

El análisis que se plantea a continuación incluye el establecimiento de la hipótesis etimológica de los términos seleccionados, los ejemplos encontrados en el corpus, el tratamiento lexicográfico que han recibido, así como un estudio comparativo con los datos que aportan algunos corpus de referencia, junto a otros corpus de tipo

especializado, con el objetivo de poder proporcionar información acerca de si tales voces han tenido un área léxica más amplia durante ese período; además de observar las convergencias y divergencias dialectales del léxico seleccionado con otras áreas hispanohablantes. Finalmente, el análisis se cierra con el uso actual que puedan tener, teniendo en cuenta otros recursos digitales de corte sincrónico.

#### 4.1. Un caso de extremeñismo: *juera*

La peculiaridad que ofrece el vocablo *juera* reside en el hecho de que está marcado en el DLE (2014, s.v.) con la etiqueta geográfica de *extremeñismo* y se define como ‘Harnero espeso de esparto para limpiar o ahechar el trigo’. Etimológicamente, *juera* se constituye como voz derivada de *joyo* y se documenta por ver primera en 1555 (DCECH, s.v. *joyo*). A su vez, la voz está tomada del portugués *joeira*, que además cuenta con otra forma análoga en el gallego *xueira* (DCECH).<sup>17</sup>

No obstante, la búsqueda de esta voz en los corpus de referencia diacrónicos para el español, como el CORDE o el CNDHE, no aporta ningún ejemplo.<sup>18</sup> La consulta en otros corpus especializados de corte diacrónico otorga más luz sobre esta palabra. Este es el caso del corpus CODEA+ 2015 en documentos localizados en Badajoz exclusivamente y solo en el siglo XVII:

- a) 1637: Dos zedazos para amasijo, una fuera
- b) 1677: Un cedaço, en tres reales, y una juera, otros tres reales
- c) 1694: Un zedaço y una juera

El CorLexIn documenta el término en varias localidades pacenses con diversas variantes ortográficas, también únicamente durante el siglo XVII (Morala 2016: 323):

- d) 1638: Vna artesa con sus xueras y cedazos
- e) 1639: Una juera y un zedazo
- f) 1652: Un zedaço y una huera y criba
- g) 1656: Vna escoba y una huera

No se documenta en otras provincias andaluzas en el corpus ODE. La documentación notarial de Badajoz atestigua en reiteradas ocasiones la voz *juera* y, además, diferencia léxicamente, de forma muy minuciosa, todos los lemas que se corresponden con el significado genérico de ‘criba’, esto es, se codifican palabras casi-sinónimas como *criba*, *cedazo*, *zaranda* (*ceranda*) y *harnero*. Con todo, existen diferencias entre todas ellas, bien en el tamaño, bien en el material del que están hechas esas cribas o en el espesor de las cuadrículas de la malla que conforma el tejido de su red.<sup>19</sup>

CORTENEX proporciona las siguientes concordancias en la primera mitad del siglo XVII:

---

<sup>17</sup> Así se observa en el *Diccionario de Galego*, s.v. <<https://digalego.xunta.gal/es/termo/65400/xoeira>>.

<sup>18</sup> Todas las concordancias que arrojan estos corpus codifican la confusión *juera* por *fuera* (tanto del verbo *ser* como del verbo *ir*).

<sup>19</sup> Por ejemplo, se aporta una pequeña descripción en un diccionario de agricultura del siglo XVIII: «Las cribas con el fondo de espartos poco distantes entre sí, llamadas más comúnmente jueras, solo se emplean para separar del trigo el polvo y las semillas menudas» (Álvarez Guerra 1797: 120).

contexto	en quatro reales.   Vna <b>fuera</b> en <sup>un</sup> real.   Vn	1601 Badajoz
contexto	siete reales y medio   Vna <b>juera</b> y un çedaço, en	1615 Badajoz
contexto	dosel [...]   un criuo y una <b>juera</b> , tasado todo   en diez	1629 Badajoz
contexto	y un criuo   y una <b>juera</b> , en seis   reales   Todos	1629 Badajoz
contexto	y un criuo y una <b>juera</b>   en onze reales que costo	1629 Badajoz
contexto	reales   Dos sedasos y una <b>juera</b> en ocho reales   Una caldera	1645 Badajoz
contexto	Dos sedasos, criuo y <b>juera</b> , en ocho rreales.	1633 Badajoz
contexto	cedacos e un criuo e <b>juera</b> en   doze reales.   Un	1601 Badajoz
contexto	.   Un rregadelo y vna <b>juera</b> en dos rs.   Un	1601 Badajoz
contexto	aposeno quatro tozinos   enteros   Dos <b>jueras</b> y un <b>cribo</b> nuevas   Media	1645 Badajoz

Figura 6. Concordancias de CORTENEX para juera

Esta documentación completa a la proporcionada en otros corpus y, al tener esta palabra sus correspondientes términos en gallego y en portugués, sería plausible la hipótesis apuntada por Corominas acerca de que el término se trate de un lusismo que ha pervivido en la zona centro-sur de Badajoz, en la línea fronteriza con Portugal. Aun así, es llamativo que solo se documente durante el siglo XVII y no se tenga constancia de la palabra a través de ejemplos documentales de siglos posteriores hasta los siglos XX y XXI. En contraste, en gallego es posible rastrear el término en Martín Sarmiento hacia 1745 en una crónica sobre el viaje que realizó a Galicia, en la que señala que *joeira* «significa especie de cribo, cedazo o *peneira*», además de apuntar el verbo que designa la acción de separar el trigo del grano: *joear* (DdD, s.v.).

En el siglo XX, estudios dialectales plasmados en algún vocabulario específico para Extremadura codifican *juera* dentro de los instrumentos de labranza (Santos Coco 1940: 85). También, la *Cartografía lingüística extremeña* (2003) registra dos casos en el mapa lingüístico dedicado al concepto ‘criba’ en puntos geográficamente localizados en Badajoz y junto a Portugal, concretamente, en las localidades de Cheles y Usagre (mapa 63, BA 304 y 600), con aspiración de la velar y con la vocal protética *a* (*juera* y *ajuera*). El ALEA registra *juera* con aspiración de la velar en un único punto del norte de Huelva, justo en la frontera con la provincia de Badajoz (ALEA, mapa 70, H 100) en la Sierra de Aracena. Precisamente, para Andalucía, el TLHA (2000 s.v. *juera*) recoge el término a través de Venceslada y del ALEA definido como ‘cedazo para cerner harina’ y ‘criba para cereales’, además de apuntar que existe la variante *huera*, fruto de la aspiración antes señalada.

Lexicográficamente, el registro de esta palabra tiene su origen en la lexicografía académica,<sup>20</sup> que codifica este término como *extremeñismo* desde 1780 (NTLLE, s.v.), definido, asimismo, como ‘especie de harnero’. Con esa misma definición se mantiene

<sup>20</sup> No hay rastro de esta voz en diccionarios del Siglo de Oro, ni en diccionarios dedicados al español medieval.

hasta la actualidad, si bien se ha incluido la característica de que el material del que está hecho es el esparto.<sup>21</sup> Esto mismo se observa en la lexicografía extraacadémica que recoge el NTLLE (Núñez de Taboada 1825, Salvá 1846, Domínguez 1853, Zerolo 1895<sup>22</sup>).

Tras los datos aportados, podemos concluir que, a pesar de que actualmente, de forma dialectal, *juera* se localiza en algún punto andaluz, la voz tiene un claro arraigo en Extremadura bajo diversas variantes fonéticas y ortográficas. Se han podido rastrear la siguientes: *juera*, *huera*, *xuera*, *fuera*, *ahuera*. No obstante, falta constatar más documentación de este sustantivo en siglos posteriores al XVII.

Actualmente, salvo las obras dialectológicas dedicadas a la variedad de español de Extremadura, el término *juera* no figura en diccionarios del español actual extra-académicos. Tampoco ha tenido expansión por otras zonas hispanohablantes americanas en el siglo XVII (CORDIAM).

#### 4.2. Un caso de competencia léxica: *alioj* y *mármol*

El arabismo *alioj* es definido como ‘mármol’ y se etiqueta como voz desusada (DLE 2014, s.v.). Se documenta desde 1495 en Nebrija (Alonso 1986, DCECH, CNDHE). Etimológicamente, procede el árabe hispánico *yašb* a través del arameo y del acadio (Corriente 1999, s.v.).

En este caso, el CORTENEX registra *alioj* en diversos textos a lo largo de toda la primera mitad del siglo XVII, casi siempre como material del que están hechos los morteros de piedra. Sin embargo, no se consigna el término homólogo de origen latino *mármol* en la documentación pacense, aunque sí se constatan ejemplos en otras provincias estudiadas en ODE en textos del mismo tipo, pero del siglo XVIII (Granada, Almería, Cádiz).<sup>23</sup> Tampoco se documentan ejemplos de *alioj* en el CorleXin.

Se puede observar una incipiente competencia léxica entre un arabismo y un término de origen latino, cuyo devenir ha dado como resultado el éxito del segundo tal y como ha sucedido en otros casos (González Sopena 2021: 87). De hecho, la lexicografía académica marca *alioj/aliox* como voz anticuada desde la edición de su diccionario de 1770/1832, respectivamente, dando como definición genérica para esta entrada la de *mármol*. Además, CODEA + 2015 documenta un caso pacense en el que sí aparecen las dos voces:

h) 1608: Cosas semaxantes de alioj y marmol

De esta forma, parece que ambos términos se usaron como sinónimos en el siglo XVII. Anteriormente, *alioj* ha estado presente en otras obras de los Siglos de Oro elaboradas por extranjeros y nacionales, como se recoge en el NTTLE, s.v. (Nebrija 1492, Alcalá 1505, Oudin 1607, Rosal 1611, Franciosini 1620). El arabismo *alioj* no figura en los diccionarios dedicados al español actual fuera del ámbito de la Real Academia. Se puede confirmar, en este caso, el relevo léxico de dos términos

<sup>21</sup> Algunas ediciones del DLE han señalado que la etimología es de origen incierto, otras han propuesto la base latina *\*ioiarium*.

<sup>22</sup> Es el primero en apuntar que se trata de una voz que procede del portugués *joiera*.

<sup>23</sup> «Yttem, vna pileta de mármol pa agua bendita» (1766, Almería, ODE).

que codifican el mismo referente en favor de la palabra de origen latino *mármol*, tal y como se muestra en la búsqueda de los corpus diacrónicos: en el CNDHE o en el CORDE, se recuperan más de 4500 concordancias para *mármol*, mientras que los resultados son nulos para *alioj*. Lo mismo sucede en la misma búsqueda en el CREA. No se hallan ejemplos en otros corpus correspondientes a otras zonas de habla hispana en el período de estudio (CORDIAM). Existe un caso más en el CdE (NOW) que recoge un ejemplo a través de una web colombiana dedicada a la lengua española en 2017:

- i) 2017: Con 5 letras están alioj, que es una piedra arcillosa y calcificada con una textura consistente o resistente, mejor dicho, un mármol.

Este ejemplo resulta esclarecedor dado que, aparte de codificar el término, explica qué es y lo compara con un tipo de mármol, si bien refleja un uso lingüístico. Con todo, es necesario recopilar más documentos que puedan dar constancia de la equiparación del arabismo y de la voz de origen latino al mismo nivel semántico.

contexto	rs   Un mortero blanco de <b>aliej</b> os en quatro rreales   Un candelero	1601	Badajoz
contexto	Un mortero de piedra de <b>alio</b> x   blanco y dos candeleros de	1615	Badajoz
contexto	mortero blanco de piedra de   <b>alio</b> x en seis reales   [...] platos y	1628	Badajoz
contexto	mortero blanco   de piedra de <b>alio</b> x   en cinco reales   Otro mortero	1629	Badajoz
contexto	Un mortero de piedra de <b>alio</b> x con   su mano de madera	1629	Badajoz
contexto	mortero blanco de piedra de <b>alio</b> x   con su mano de palo	1630	Badajoz
contexto	mortero blanco de piedra de <b>alio</b> x con su mano   Una silla	1645	Badajoz
contexto	mortero blanco de piedra de <b>alio</b> x   Dos caços de açofar con	1645	Badajoz
contexto	mortero blanco de piedra de <b>alio</b> x   tasado en quatro reales   Un	1645	Badajoz
contexto	mortero blanco de piedra   de <b>alio</b> x con su mano de   palo	1646	Badajoz
contexto	.   Vn mortero blanco de <b>alio</b> x en quatro   reales.   Los	1601	Badajoz
contexto	mortero blanco de piedra de <b>alio</b> x   Una dozana de platos y	1643	Badajoz
contexto	mortero blanco de piedra de <b>alio</b> x   tassado en seis res   Una	1645	Badajoz
contexto	rreales.   Un mortero de <b>alio</b> x en quatro rs.   Loza	1601	Badajoz
contexto	.   Un mortero blanco de <b>alio</b> x   en quatro rreales.   Unas	1601	Badajoz

Figura 7. Concordancias de alioj en CORTENEX S. XVII

## 5. Conclusiones

En este trabajo se han podido llevar a cabo los dos objetivos planteados. Esto es, por un lado, describir la metodología llevada a cabo en la confección del CORTENEX S. XVII gracias a la implementación de nuevas tecnologías en el ámbito de las

humanidades digitales. Ello se consigue a través de transcribir documentación antigua en XML y del uso de la plataforma TEITOK que aloja al corpus ODE actualmente, al que el corpus descrito se suscribe como corpus de control.

Además, la tipología textual de las relaciones de bienes ha resultado ser especialmente fructífera para ahondar en la historia del léxico español de una zona geográfica bastante desconocida en términos diacrónicos en muchos campos léxico-semánticos. Se ha podido contextualizar la documentación pacense y constatar los primeros resultados.

Por último, se ha puesto el foco en dos palabras presentes en el corpus elaborado en lo referente a la documentación pacense de la primera mitad del siglo XVII: la voz *juera* se constata como extremeñismo y casi no tiene vitalidad actualmente, salvo en localidades muy concretas de Badajoz cercanas a Portugal. Con respecto a esta palabra, sigue siendo necesario encontrar documentación posterior al siglo XVII para poder rellenar el hueco documental existente hasta el siglo XX. El arabismo *alioj* ha tenido su huella documental para venir a designar lo mismo que *mármol*, si bien, actualmente, es una voz desusada y se ha impuesto en el uso la voz de origen latino frente al arabismo. Al igual que en el caso de *juera*, no se encuentran muchos ejemplos textuales no lexicográficos después del siglo XVII y tampoco en otras zonas geográficas según las bases de datos consultadas.

## Agradecimientos

El presente artículo ha sido realizado dentro del marco del proyecto «Atlas Lingüístico y Etnográfico de Andalucía, S. XVIII. Patrimonio documental y humanidades digitales» (Proyectos I+D+i Junta de Andalucía-FEDER, P18-FR-695) y del Proyecto «Hispane Testium Depositiones», FFI 2017-83400-P (MINEO/AEI/ FEDER, UE).

### Referencias bibliográficas

- ALEA = ALVAR, Manuel (1961-1973), *Atlas lingüístico y etnográfico de Andalucía (ALEA)*, Granada: Universidad de Granada-CSIC.
- ALONSO, Martín (1986), *Diccionario medieval español. Desde las Glosas Emilianenses y Silenses (S. X hasta el S.XV)*, Salamanca: Universidad Politécnica de Salamanca.
- ÁLVAREZ DE MIRANDA, Pedro (1992), *Palabras e ideas: el léxico de la ilustración temprana en España (1680-1760)*, Madrid: Anejos del BRAE.
- ÁLVAREZ GUERRA, Juan (1797), *Diccionario universal de agricultura (Abate Rozier)*, Madrid: Imprenta Real, Tomo I.
- ARIZA VIGUERA, Manuel (1985), «Dos estudios de historia lingüística de Extremadura», *Anuario de Estudios Filológicos* VIII, 7-18.
- ARRABAL RODRÍGUEZ, Pilar (2020), «Edición de un corpus digital de inventarios de bienes», *Procesamiento de Lenguaje Natural* 65, 67-74.
- BLANCO CARRASCO, José Pablo (1998), «Las ciudades extremeñas en tiempos de Felipe II: Problemas fiscales y demográficos», en MARTÍNEZ MILLÁN, J. (dir.), *Felipe II (1527-1598): Europa y la monarquía católica*, Madrid: Parteluz, vol. 2, 89-102.
- CALDERÓN CAMPOS, Miguel (2019), «La edición de corpus históricos en la plataforma TEITOK. El caso de Oralía diacrónica del español», *Chimera* 6, 21-36.

- CALDERÓN CAMPOS, Miguel – VAAMONDE, Gael (2020), «Oralia diacrónica del español: un corpus de la Edad Moderna», *Scriptum digital* 9, 167-189.
- CALDERÓN CAMPOS, Miguel (2022), «Andalucismos en el corpus del reino de Granada», en CORBELLA, D. – FAJARDO, A. – LANGENBACHER, J. (eds.), *Historia del léxico español y Humanidades digitales*, Berlín: Peter Lang, 319-341.
- CdE = MARK DAVIES (2012-2019), *Corpus del español (NOW)*, <<https://www.corpusdelespanol.org/now/>> [13/03/2022].
- CODEA+ 2015 = GHITE (GRUPO DE INVESTIGACIÓN TEXTOS PARA LA HISTORIA DEL ESPAÑOL), *Corpus de documentos españoles anteriores a 1800*, <<http://corpuscodea.es>> [16/02/2022].
- CORDIAM = ACADEMIA MEXICANA DE LA LENGUA, *Corpus Diacrónico y Diatópico del Español de América*, <[www.cordiam.org](http://www.cordiam.org)> [2/02/2022].
- CORLEXIN = MORALA RODRÍGUEZ, José Ramón (dir.), *Corpus léxico de inventarios*, <<https://apps2.rae.es/CORLEXIN.html>> [15/02/2022].
- CORRIENTE, Federico (1999), *Diccionario de arabismos y voces afines en iberorromance*, Madrid: Gredos.
- DCECH = COROMINAS, Joan – PASCUAL, Antonio (1980-1991), *Diccionario crítico y etimológico castellano e hispánico*, Madrid: Gredos.
- DdD = SANTAMARINA, Antón (coord.), *O Diccionario de diccionarios*, Santiago de Compostela: Instituto da Lingua Galega, <<http://ilg.usc.gal/ddd/index.php>> [13/03/2022].
- FERNÁNDEZ DE MOLINA, Elena (2014), «La investigación científica en el habla de Extremadura: monografías dialectales y estudios sobre fonética y fonología extremeña», *Anuario de Estudios Filológicos* 37, 5-20.
- GARCÍA-GODOY, María Teresa (2012), «La lengua de las primeras constituciones hispánicas: el cambio léxico-semántico», en BASTARDÍN, T. – RIVAS, M. (eds.), *Actas del IX Congreso Internacional de Historia de la Lengua Española*, Madrid/Frankfurt: Iberoamericana/Vervuert, 107-132.
- GARCÍA MOUTON, Pilar (1996), «El extremeño», en *Lenguas y dialectos de España*, Madrid: Arco/Libros, 31-34.
- GONZÁLEZ SALGADO, José Antonio (2003), *Cartografía lingüística de Extremadura*. Madrid: Universidad Complutense, <<http://www.geolectos.com/>> [16/01/2022].
- GONZÁLEZ SOPEÑA, Inmaculada (2021), «Indications of lexical-semantic changes in the Arabisms of the Kingdom of Granada (1493-1612) the lexical competition between almádena and martillo», *Onomázein* 51, 83-98.
- GUERRA GUERRA, Arcadio (1977), «Escribanos badajocenses del siglo XVII», *Revista de estudios extremeños* 33/1, 5-68.
- GUERRA GUERRA, Arcadio (1980), «Suntuaria, muebles y enseres del Badajoz del siglo XVI», *Revista de estudios extremeños* 3, 445-488.
- JANSSEN, Maarten (2012), «Neotag: a POS tagger for grammatical neologism detection», en *Proceedings of the tenth international conference on Language Resources and Evaluation, LREC 2012*, Estambul, 1-7.
- JANSSEN, Maarten (2016), «TEITOK: Text-Faithful Annotated Corpora», en *Proceedings of the Language Resources and Evaluation Conference (LREC 2016) ELRA*, Portoroz, Eslovenia, 4037-4043.
- MARCOS ÁLVAREZ, Fernando (1988), «Badajoz en 1640: anotaciones laborales», *Revista de estudios extremeños* 44/3, 661-722.

- MARCOS ÁLVAREZ, Fernando (1992), «Algunas precisiones léxicas sobre indumentaria española en el siglo XVII», en *Actas del II Congreso Internacional de Historia de la Lengua Española*, Madrid: Pabellón de España, 1161-1172.
- MONTERO CURIEL, Pilar (2006), *El extremeño*, Madrid: Arco/Libros.
- MORALA RODRÍGUEZ, José Ramón (2012), «Léxico e inventarios de bienes en los Siglos de Oro», en CLAVERÍA NADAL, G. – FREIXAS, M. – PRAT SABATER, M. – TORRUELLA, J. (eds.), *Historia del léxico: perspectivas de investigación*, Madrid/Fránkfort: Iberoamericana/Vervuert, 199-218.
- MORALA RODRÍGUEZ, José Ramón (2014), «El *CorLexIn*, Un corpus para el estudio del léxico histórico y dialectal del Siglo de Oro», *Scriptum Digital* 3, 5-28.
- MORALA RODRÍGUEZ, José Ramón (2016), «Léxico histórico: sobre cribas, cedazos y harneros», en LÓPEZ SERENA, A. (ed.), *El español a través de los tiempos*, vol 1, Sevilla: Universidad de Sevilla, 309-324.
- ODE = CALDERÓN CAMPOS, Miguel – GARCÍA-GODOY, María Teresa (2010-2019), *Oralia diacrónica del español (ODE)*, <<http://corpora.ugr.es/ode>> [16/01/2022].
- PÉREZ TORAL, Marta (2021), «Léxico dialectal documentado en textos notariales asturianos del siglo XVII», *Revista de lexicografía* 26, 85-112.
- PUCHE LORENZO, Miguel Ángel (2018), «Léxico dialectal/léxico especializado en los inventarios de bienes», en *Actas del X Congreso Internacional de Historia de la Lengua*, Zaragoza: Institución Fernando el Católico, 1655-1670.
- REAL ACADEMIA ESPAÑOLA, *Corpus diacrónico del español (CORDE)*, <[www.rae.es](http://www.rae.es)> [27/01/2022].
- REAL ACADEMIA ESPAÑOLA, *Corpus del Nuevo Diccionario Histórico del Español (CNDH)*, <[www.rae.es](http://www.rae.es)> [27/01/2022].
- REAL ACADEMIA ESPAÑOLA, *Corpus de referencia del español actual (CREA)*, <[www.rae.es](http://www.rae.es)> [27/01/2022].
- REAL ACADEMIA ESPAÑOLA. *Diccionario de la lengua española (DLE)*, <[www.rae.es](http://www.rae.es)> [27/01/2022].
- REAL ACADEMIA ESPAÑOLA (1960-1996), *Diccionario histórico de la lengua española*, <[www.rae.es](http://www.rae.es)> [2/02/2022].
- REAL ACADEMIA ESPAÑOLA, *Nuevo tesoro lexicográfico de la lengua española (NTTLE)*, <[www.rae.es](http://www.rae.es)> [2/02/2022].
- SÁNCHEZ SIERRA, Diego (2019), *Edición y estudio léxico de fuentes documentales extremeñas de los siglos XVI y XVII*, Madrid: Universidad de Alcalá de Henares.
- SANTOS COCO, Francisco (1940), «Vocabulario extremeño», *Revista del Centro de Estudios Extremeños*, 65-96.
- TEI CONSORTIUM (ed.) (2016), *TEI P5: guidelines for electronic text encoding and interchange. Text Encoding Initiative Consortium*, <<http://www.tei-c.org/release/doc/tei-p5-doc/en/Guidelines.pdf>>.
- TLHA = ALVAR EZQUERRA, Manuel (2000), *Tesoro léxico de las hablas andaluzas*, Madrid: Arco/Libros.
- UBIETO, Antonio – REGLÁ, Juan – JOVER, José María – SECO, Carlos (1983), *Introducción a la historia de España*, Barcelona: Teide (14ª edición).
- VIUDAS CAMARASA, Antonio – ARIZA VIGUERA, Manuel – SALVADOR PLANS, Antonio (1987), *El habla en Extremadura*, Junta de Extremadura: Consejería de Educación y Cultura.
- ZAMORA VICENTE, Alonso (1974), *Dialectología española*, Madrid: Gredos.