

# FONDO HISPÁNICO DE LINGÜÍSTICA Y FILOLOGÍA

Vol. 39

Colección dirigida por

Juan Pedro Sánchez Méndez & María Teresa García-Godoy



Lausanne - Berlin - Bruxelles - Chennai - New York - Oxford

Miguel Calderon Campos  
e Inmaculada González Sopeña (eds.)

*Scripta manent.* Historia del español,  
documentación archivística y  
humanidades digitales

Diacronía del español y documentación histórica



Lausanne - Berlin - Bruxelles - Chennai - New York - Oxford

Bibliographic Information published by the Deutsche Nationalbibliothek  
The Deutsche Nationalbibliothek lists this publication in the Deutsche Nationalbibliografie;  
detailed bibliographic data is available online at <http://dnb.d-nb.de>.



Este libro ha sido financiado por FEDER / Junta de Andalucía - Consejería de Transformación Económica, Industria, Conocimiento y Universidades / Proyecto A-HUM-116-UGR20 (ALEA oriental-XVIII)

ISSN 1663-2648 • ISBN 978-3-0343-4727-3 (Print)  
E-ISBN 978-3-0343-4742-6 (E-PDF) • E-ISBN 978-3-0343-4743-3 (EPUB)  
• DOI 10.3726/b20916

© Miguel Calderon Campos e Inmaculada González Sopeña (eds.) 2023  
Verlegt durch Peter Lang Group AG, Lausanne, Schweiz

[info@peterlang.com](mailto:info@peterlang.com) <http://www.peterlang.com/>



Open Access: Dieses Werk ist lizenziert unter einer Creative Commons Namensnennung 4.0 Internationalen Lizenz (CC-BY 4.0) Weitere Informationen: <https://creativecommons.org/licenses/by/4.0/>

# Índice

PRESENTACIÓN .....	9
PERIODIZACIÓN E HISTORIA DE LA LENGUA ESPAÑOLA EN EL CORPUS CHARTA	
<i>Cristina Buenafuentes de la Mata</i>	
Morfología léxica histórica y lingüística de corpus: rivalidad afijal a la luz del corpus CHARTA .....	15
<i>Carlos Sánchez Lancis</i>	
Nuevas aportaciones de la lingüística de corpus a la periodización de la historia del español: el corpus CHARTA .....	39
PERIODO MEDIEVAL	
<i>Francisco Alonso de la Torre Gutiérrez</i>	
Una aproximación a los grupos paratácticos en el notariado asturiano bajomedieval (siglos XIII–XV): análisis sintáctico-discursivo .....	67
<i>Marina Gomila Albal</i>	
La interacción del registro y la variación geográfica en la difusión de los pronombres <i>nosotros</i> , <i>vosotros</i> en castellano medieval .....	95
<i>Daniel Jiménez Sánchez</i>	
La historia ecdótica del <i>Libro del Cauallero Zifar</i> : planteamiento para su edición digital .....	123
<i>Vicente J. Marçet Rodríguez</i>	
Los diptongos en el leonés occidental: la documentación medieval de San Andrés de Espinareda (El Bierzo) .....	141

# Nuevas aportaciones de la lingüística de corpus a la periodización de la historia del español: el corpus CHARTA\*

*New contributions of corpus linguistics to the periodization of the history of Spanish: the CHARTA corpus*

CARLOS SÁNCHEZ LANCIS

*Universidad Autónoma de Barcelona*

ORCID iD: <https://orcid.org/0000-0001-9111-1563>

**Resumen:** El presente trabajo sobre la periodización del español contrasta, a partir de la lingüística de corpus (cf. Rojo 2021), los resultados obtenidos en Granvik y Sánchez Lancis (2018), en donde se ha llevado a cabo un análisis cuantitativo, denominado agrupamiento jerárquico (cf. Gries y Hilpert 2008, Gries 2009), de una muestra significativa del *Corpus diacrónico del español* (CORDE) de la Real Academia Española, con los datos extraídos, mediante el mismo procedimiento de estudio, del corpus CHARTA (*Corpus Hispánico y Americano en la Red: Textos Antiguos*). Consideramos que la composición documental, no literaria, de CHARTA permite establecer una periodización en principio mucho más cercana a la realidad histórica que la que ofrecen otros corpus en los que predomina el componente literario. Además, desde otro punto de vista, no es posible concebir una periodización fidedigna de una lengua sin tener en cuenta también todo el componente documental que posee. Para ello, se estudia una serie de fenómenos gramaticales, tanto morfológicos —pronombres y determinantes, desinencias y tiempos verbales— como sintácticos —la construcción artículo + posesivo + sustantivo, los casos de interpolación entre el clítico y el verbo, el orden de los clíticos—, con el fin de que sean los datos lingüísticos, y no los extralingüísticos, los que determinen la periodización de la lengua española. Los resultados obtenidos no solo corroboran algunas observaciones establecidas sobre la periodización del español, sino que también aportan nuevos datos y nos hacen

---

\* La presente investigación ha sido parcialmente financiada con una ayuda del MICINN y FEDER (PID2021-123617NB-C41) y de la CIRIT del Comissionat per Universitats i Recerca de la Generalitat de Catalunya (2021 SGR 00787).

cuestionarnos, hasta cierto punto, la validez de las etapas tradicionales denominadas español medieval, clásico y moderno.

**Palabras clave:** Historia de la lengua española, periodización, Lingüística de corpus, corpus CHARTA, agrupamiento jerárquico

**Abstract:** This paper addresses the periodization of Spanish. Based on corpus linguistics (cf. Rojo 2021), it contrasts the results obtained from a quantitative analysis named *Hierarchical Clustering* (cf. Gries and Hilpert 2008, Gries 2009) carried out by Granvik and Sánchez Lancis (2018), wherein this method was used to consider a significant sample of the Diachronic Corpus of Spanish (CORDE) of the Royal Spanish Academy. The data for this new analysis were extracted from the CHARTA corpus (Corpus Hispánico y Americano en la Red: Textos Antiguos) using the same study procedure. We believe that the non-literary documentary composition of CHARTA allows us to establish a periodization that, in principle, is much closer to historical reality than the predominantly literary one provided by other corpora. Moreover, establishing a reliable periodization of a language does not seem possible without also considering all the documentary components it possesses. For this reason, we studied a series of grammatical phenomena including both morphological—pronouns and determiners, desinences and verb tenses—and syntactic issues—the construction article + possessive + noun, the cases of interpolation between the clitic and the verb, the order of the clitics—, to ensure that it is the linguistic data instead of the extralinguistic ones which determine the periodization of the Spanish language. The results of our analysis not only corroborate some already established observations on the periodization of Spanish but also provide new data which might call into question the validity of the traditional stages known as Medieval, Classical and Modern Spanish.

**Keywords:** History of Spanish, Periodization, Corpus Linguistics, CHARTA corpus, Hierarchical Clustering

## 1. INTRODUCCIÓN

Uno de los problemas pendientes de la historia de la lengua española, pero todavía no solucionado definitivamente, es su periodización mediante criterios básicamente de tipo lingüístico. Ciertamente, no resulta una tarea sencilla determinar lapsos de tiempo en los que se divide el eje temporal de una lengua desde un punto de vista histórico. Si bien los diversos manuales, entre los que destacan el clásico de Lapesa (1981) o los más recientes de Cano Aguilar (2005) y Torrens Álvarez (2018), establecen diversas periodizaciones —castellano o romance primitivo / español arcaico / época de los orígenes (siglos IX-XI), castellano medieval (siglos

xii-xiv), español preclásico (1474–1525), español del/de los Siglo(s) de Oro (siglos XVI-XVII) / la lengua en la España de los Austrias (1474–1700), español moderno (siglo XVIII-español actual), etc.—, en la mayoría de estas todavía siguen pesando más los aspectos externos que los internos a la propia lengua, sobre todo por lo que respecta a la lengua literaria, la cual continúa teniendo un peso fundamental en la delimitación de estos lapsos temporales. De todas formas, ya anteriormente Menéndez Pidal (1942) al menos había cuestionado la división de la historia de la lengua española por siglos, es decir, por lapsos de tiempo de cien años exactos, como así ha aparecido frecuentemente en los manuales al uso —p. ej., se habla del castellano del siglo XIII, el Siglo o los Siglos de Oro (siglos XVI y XVII), la lengua del XVIII, etc.—, por lo que propugnaba segmentaciones temporales más restringidas o precisas. Por todo ello, la cuestión de la periodización ha sido objeto de un número destacado de trabajos en los últimos treinta años<sup>1</sup>, sobre todo a partir del artículo de Eberenz (1991), primer autor en replantear claramente el concepto de periodización con criterios exclusivamente lingüísticos, denominados “parámetros esenciales”, para establecer los siguientes períodos de la lengua española: “fase antigua de estabilidad” (1200–1450), “etapa media de transformación” (1450–1650) y “fase moderna de estabilidad” (1650–actualidad). Estas etapas fueron matizadas por su complejidad casi veinte años más tarde por el propio lingüista, que además acaba concluyendo que “la periodización supone establecer una cronología absoluta de los cambios, tarea facilitada hoy en día por la disponibilidad de grandes corpus digitalizados.” (Eberenz 2009: 196).

La existencia en la actualidad de estos corpus diacrónicos representativos del español (cf. Torruella Casañas 2017, Sánchez Lancis 2009 y 2022) posibilita la aplicación de un análisis cuantitativo denominado agrupamiento jerárquico (cf. Gries 2009), con el fin de conseguir una imagen mucho más precisa de los diferentes períodos históricos de una

---

1 Además de Eberenz (1991 y 2009), bastantes autores han tratado también el tema de la periodización del español, ya sea de forma general (Marcos Marín 1992 y 1995, Cano Aguilar 1992, Ridruejo Alonso 1993, Martínez Alcalde y Quilis Merín 1996, Sánchez Lancis 1997–1998, 1999 y 2001, Echenique Elizondo y Martínez Alcalde 2000, Buenafuentes 2002a y 2002b, Granvik y Sánchez Lancis 2018, entre otros), o para precisar particularmente etapas ya delimitadas (Serradilla Castaño 2001, Melis, Flores y Bogard 2003, Girón Alconchel 2005 y 2008, Octavio de Toledo y Huerta 2007, Sánchez Lancis 2012, entre otros).

lengua<sup>2</sup>. El agrupamiento jerárquico es un método de análisis, que se emplea para identificar semejanzas y diferencias entre grupos. Gracias al cálculo matemático, es posible establecer un modelo descriptivo que se fundamenta en la semejanza numérica entre distintas categorías, de tal modo que las categorías que presentan un conjunto de datos más semejantes entre sí formarán grupos, los cuales se distanciarán de los grupos de las categorías con datos también similares entre sí, pero diferentes de los anteriores. Por consiguiente, la finalidad de este análisis es hallar, a partir de variables, las similitudes que posibilitan agrupar los datos dentro de grupos lo más homogéneos posible (cf. Cardillo 2015). Sin embargo, como en la periodización lingüística se busca agrupar categorías cronológicamente sucesivas y no solo categorías que comparten rasgos o propiedades, ha sido necesario recurrir a una variante del agrupamiento jerárquico aplicable a la periodización denominada “Variability-based Neighbor Clustering” (VNC), algoritmo desarrollado por Gries y Hilpert (cf. Gries y Hilpert 2008 y 2012, Hilpert 2012), que permite agrupar juntos períodos de tiempo sucesivos que poseen elementos lingüísticos semejantes, diferenciándose así del resto. En nuestro caso, los datos numéricos se han obtenido a partir del análisis de la frecuencia relativa por millón de palabras, es decir, su frecuencia normalizada, de una serie de fenómenos gramaticales representativos diacrónicamente por cada década que compone el corpus estudiado. El resultado del análisis de VNC es un dendrograma. De este modo, cada ramificación señala una cierta diferencia cuantitativa entre décadas contiguas, mientras que las décadas más semejantes según las cifras obtenidas se agrupan juntas, constituyendo un periodo determinado.

## 2. EL CORPUS CHARTA Y LA PERIODIZACIÓN

El *Corpus Hispánico y Americano en la Red: Textos Antiguos* (CHARTA) es un corpus diacrónico del español de textos archivísticos, formado por 13 subcorpus diferentes, compuesto por 2076 documentos con un total de

---

2 Para una explicación más detallada sobre el agrupamiento jerárquico y su aplicación a la periodización lingüística a partir de los datos gramaticales de un corpus, cf. Granvik y Sánchez Lancis (2018: 756–758).

1.346.094 formas, que van desde el siglo IX hasta el siglo XIX, si bien los textos en romance aparecen básicamente a partir del siglo XIII. Además, el corpus distingue tres parámetros: cronológico —por años: desde el 822 hasta 1871, aunque existe una gran desproporción, ya que solo el siglo XIII contiene el 41,15 % del total con 554.020 formas, el XIV el 16,17 %, el XVI el 14,54 %, el XV el 11,54 %, el XVII el 6,77 %, el XVIII el 5,91 %, el XII el 1,26 % y el XIX el 0,97 %—, geográfico —por países: España 93,51 %, Venezuela 2,01 %, México 1,70 %, Ecuador 1,17 %, Colombia, Cuba, El Salvador, Guatemala, Inglaterra, Panamá, Portugal, República Dominicana— y tipológico —cartas de compraventa y contratos 39,34 %, textos legislativos 38,17 %, otros 9,19 %, actas y declaraciones 6,68 %, testamentos e inventarios 2,08 %, certificaciones 2,08 %, cartas privadas, estatutos, informes y relaciones, notas breves, recibí—. El valor primordial de utilizar este corpus para hacer un estudio sobre la periodización del español estriba precisamente en su tipología, ya que se trata por entero de textos no literarios, frente a un corpus como, por ejemplo, el CORDE, en donde los textos literarios conforman una gran parte de este. Sin embargo, también hay que tener en cuenta de cara a los resultados la gran diferencia de tamaño entre uno y otro, pues este último ya es un corpus grande cerrado que contiene más de 34.500 documentos con casi 245 millones de formas, mientras que el primero todavía es un corpus pequeño no cerrado, muy especializado, con las consiguientes dificultades para incrementar progresivamente el número de formas. A pesar de ello, es de gran utilidad poder disponer de un corpus como CHARTA para poder observar y contrastar con otros corpus de distinta tipología la evolución de ciertos cambios lingüísticos en la historia del español y, sobre todo, su periodización.

Por ello, en el presente estudio se ha realizado un análisis de la evolución de 13 fenómenos gramaticales, tanto morfológicos como sintácticos, significativos en la diacronía del español por marcar una frontera lingüística temporal, ya sea por representar el fin de una etapa o el principio de una nueva. La finalidad de esta propuesta es que solo sean los datos gramaticales los que sirvan de base para la periodización de la lengua. Los fenómenos estudiados son los siguientes<sup>3</sup>:

---

3 La elección de los fenómenos gramaticales se ha realizado a partir de los que se estudiaron para la periodización del CORDE en Granvik y Sánchez Lancis (2018). Sin embargo, dadas las características del corpus CHARTA en función de la tipología de los textos que lo componen, no ha sido posible analizar los 27 fenómenos,

- (1) Pronombres y determinantes (8)
  - a) nos > nos otros > nosotros
  - b) vos > os ~ hos
  - c) este ~ aqueste ~ aquest > este
  - d) que / qui ~ qi ~ ki > que
  - e) so / sue ~ sua ~ su > su
  - f) quien > quienes ~ qienes ~ kienes ~ qujenes
  - g) unas
  - h) las ~ elas > las
  
- (2) Morfología verbal (2)
  - a) so > soy ~ soi
  - b) sodes > sois
  
- (3) Construcciones sintácticas (3)
  - a) la su > su
  - b) lo no ~ lo non > no lo
  - c) \*r lo he ~ \*r lo hé ~ \*r lo e ~ \*r lo é > lo \*ré

## *2.1. Pronombres y determinantes*

El reanálisis y posterior gramaticalización de la secuencia sintáctica formada por el pronombre personal de primera persona del plural “nos” seguido del adjetivo “otros” con valor de exclusividad es un cambio característico en español de finales de la Edad Media, que, según las gramáticas históricas, comenzó en el siglo XIV y acabó por generalizarse en el siglo XVI. En CHARTA se documenta la combinación analítica “nos otros” en el texto paleográfico y la forma sintética “nosotros” en el texto crítico en 94 casos en 49 documentos, de 1414 a 1871, pero no antes del

---

ya que no se habrían obtenido datos cuantitativos de todos ellos, por lo que se han seleccionado los 13 más representativos. Por otra parte, hay que tener también en cuenta que el estudio de estos fenómenos, tanto en el CORDE como en CHARTA, presenta una serie de limitaciones, ya sea por el tamaño del primero o al no estar ambos corpus etiquetados morfosintácticamente, por lo que, por ejemplo, el análisis de la construcción “artículo + posesivo + sustantivo” o de la interpolación entre el clítico y el verbo solo se ha podido realizar en una de sus diferentes combinaciones posibles.

siglo xv —27,66 % (26/10<sup>4</sup>) siglo xv, 59,57 % (56/29) siglo xvi, 4,26 % (4/4) siglo xvii, 1,06 % (1/1) siglo xviii, 7,45 % (7/5) siglo xix—, aunque también se constata su mayor difusión en el siglo xvi. Su primera documentación se puede observar en el siguiente ejemplo:

- (4) Sepan cuantos esta carta vieren cómo **nós** e la abadesa, e frairas e convento del monasterio de Santa Clara de Gerca, que es cerca la cibdat de Astorga, estando juntas al nuestro cabillo por son de campana tañida segund que lo avemos de uso e de costumbre de nos ayuntar, otorgamos e conocemos por esta carta que arrendamos a vós Domingo Pascoal, e Antona Janes vuestra muger e a vuestros hijos Diego, e Alfonso, e Catalina, e Aldonça e María, moradores en Val de Rey toda la heredat, e casas, e huertos e todas las otras cosas que **nosotros** por nombre del dicho monasterio avemos en término de Val de Rey so campanas de la eglesta de Santa María que fue e fincó de María Alfonso, fija que fue de Alfonso Joánez e de María Alfonso de Val de Rey, fraira que fue del dicho monasterio por en todos los días de vuestra vida e de cadaúno de **nós** porque avedes a dar e pagar al dicho monasterio e a **nosotros** por nombre d'él e al que lo ovier de recabdar por **nós** de cada año por ello en renta dos moyos de buen pan, metad trigo e metad centeno, limpio, e seco e medido por la <...> derecha; (CODEA-0432, 1414, España, León, CHARTA).

La forma pronominal “nos otros” > “nosotros” era empleada para la función de sujeto —“**nosotros** [...] avemos”— o como término de preposición —“avedes a dar e pagar al dicho monasterio e a **nosotros**”—, como así se puede comprobar en el ejemplo anterior. Sin embargo, la forma etimológica simple “nos” convive con la anterior con valor mayestático referida a una persona en singular —“**nós** e la abadesa, e frairas e convento [...] otorgamos e conocemos”—, aunque también puede aparecer sin este valor como plural en los mismos contextos preposicionales —“de cadaúno de **nós**”; “lo ovier de recabdar por **nós**”—. Este hecho, a falta de otros datos, podría indicar tal vez que se trataría de un cambio que se difundió más tardíamente en los textos no literarios, como así demostraría este contrato de arrendamiento. De ahí su importancia al tratarse de un cambio que afecta a la morfología pronominal personal.

La evolución fónica del clítico pronominal de segunda persona del plural, tanto acusativo como dativo, “vos” > “os” marca también un punto temporal en la evolución del español. Según las gramáticas históricas, la

---

4 En x/y, el primer valor indica el número de casos y el segundo el de documentos en los que se halla una determinada forma o conjunto de formas lingüísticas, mientras que el tanto por ciento que le precede hace referencia a su frecuencia relativa.

forma “os” no presenta sus primeros ejemplos hasta el siglo XIV y acaba difundiéndose en el siglo XVI. Sin embargo, en CHARTA, con 126 casos en 40 documentos, se halla, junto a su variante gráfica “hos”, ya desde 1284 hasta 1735, aunque muy esporádicamente tanto en el XIII como en el XV y ya de manera frecuente en el XVI —1,58 % (1/1) siglo XIII, 0/0 siglo XIV, 0/0 siglo XV, 71,43 % (90/32) siglo XVI, 25,40 % (32/6) siglo XVII, 1,59 % (2/1) siglo XVIII, 0/0 siglo XIX—. Su primera documentación en el corpus es la siguiente:

- (5) E mando al escrivano público que me embíe <dar> con este mio omne un estru-  
mento público de cual dia <...> **os** fue mostrada esta carta (CODCAR-0519,  
1284, España, Jaén, CHARTA).

Por su parte, en el caso del pronombre demostrativo masculino singular de primer grado, interesa constatar el momento en que la forma característica medieval y en gran parte también clásica “aqueste” es sustituida completamente por su variante “este”. En el corpus estudiado, “aqueste” se encuentra en 17 casos en 16 documentos, desde 1233 hasta 1697, por lo que, curiosamente, hasta a partir de principios del siglo XVIII no dejaríamos de hallar esta forma —50,00 % (10/9) siglo XIII, 20,00 % (4/3) siglo XIV, 0/0 siglo XV, 20,00 % (4/4) siglo XVI, 10,00 % (2/2) siglo XVII, 0/0 siglo XVIII, 0/0 siglo XIX—, como se puede observar en los siguientes ejemplos:

- (6) a. Ye yo don Elo con mies fiyos, ye yo doña Sancha con mies fiyos **aquesta**  
carta roboramos ye confirmamos, ye **aqueste** signo en ela mandemos fazer.  
(CODEA-0416, 1233, España, Zamora, CHARTA).  
b. [...] renusciación non vala salvo si la dicha ley no renuscia e yo así lo renus-  
cio, pero sea entendido y entiéndase lo susodicho que sea sin perjuicio de  
cualquier derecho que la dicha cibdad e yo tengamos en la propiedad de la  
dicha plaça e **aqueste** derecho quede a salvo. (CODEA-0202, 1531, España,  
Toledo, CHARTA).  
c. y, si la persona a quien se comete esta numeración no es de muy ajustada con-  
siensia, fuera de los salarios que lixitimamente lleba, puede aprobecharse en  
gran daño de Vuestra Magestad y encomenderos, reserbando o no entrando a  
tributar ni numerar a los que se lo pagaren, aunque **aqueste** interés no puede  
ser muy cuantioso por ser dádibus de indios que tienen poco. (DCEP-0903,  
1697, Ecuador, Quito, CHARTA).

El pronombre relativo “qui” es una forma característica del español medieval para remitir a personas, mientras que la variante “que” es empleada, normalmente durante esa misma época, para hacer referencia a elementos no humanos. En este caso, para la periodización importa el momento en

que dejan de coexistir ambas formas y “que” pasa a ser el único pronombre relativo. En CHARTA “qui” se halla en 762 casos en 390 documentos, desde 1200 a 1529 —66,66 % (508/296) siglo XIII, 18,50 % (141/70) siglo XIV, 14,04 % (107/20) siglo XV, 0,78 % (6/4) siglo XVI, 0/0 siglo XVII, 0/0 siglo XVIII, 0/0 siglo XIX—, por lo que presenta una destacada frecuencia de uso durante todo el español medieval, desapareciendo, como era esperable, a principios del siglo XVI, como se puede ver en los siguientes ejemplos:

- (7) a. *Notum sit omnibus hominibus tam presentibus quam futuris quod ego Día Gómez empeñé un solar en Vuelna in loco pernominato ubi dicitur Sancta Marina a frater Didacus con suas pertinencias por XX moravedís, en concejo de Orzales; le dio por mañero a Álvar Gonzálvez qui:l metió por en el solar <...> él en les le entró por fiador que:l faga coientur de año e de día.* (CODEA-0216, 1202, España, Cantabria, CHARTA).
- b. Signo de mí Joán Campi, notario público de la ciudat de Çaragoça, **qui** a las sobredichas cosas juntamente con los testimonios de la carta de arriba nombrados presente fue e en parte según fuero escrebí e lo otro escribir fize, con <...>, [...]. (CODEA-0798, 1529, España, Zaragoza, CHARTA).

Por lo que respecta a los pronombres posesivos, las gramáticas históricas explican la pérdida de la distinción de género, por ejemplo, en el tercer grado —masculino “so(s)” / femenino “sue(s)” ~ “sua(s)” ~ “su(s)”— en posición antepuesta al sustantivo, con la función de adjetivo o más bien de determinante, ya durante el español medieval. En este caso la forma originariamente femenina “su”, fruto de la apócope de la vocal, será la que quedará como variante única para ambos géneros en dicha secuencia. En CHARTA se atestiguan 6081 casos en 1243 documentos del posesivo “su”, desde 1162 a 1871 —0,06 % (4/2) siglo XII, 23,63 % (1437/506) siglo XIII, 12,84 % (781/197) siglo XIV, 12,76 % (776/119) siglo XV, 18,89 % (1149/173) siglo XVI, 13,59 % (827/118) siglo XVII, 15,49 % (942/91) siglo XVIII, 2,71 % (165/37) siglo XIX—, en los que se puede observar a partir del siglo XIII en adelante, y no antes, un alto índice de frecuencia. A continuación, se recoge su primera documentación seguido de un sustantivo masculino ya en el siglo XII, aunque sea de forma un tanto esporádica, y en el mismo fragmento que la forma antigua “so”:

- (8) D'este fecho son testigos: don Alfonso López e don Joán Alfonso, **so** fijo; e don Sant Garciez de Salcedo, e don Diego Garciez de Alfaro, e don Ruy Díaz de Velascor, e don Joán Alfón Carriello, e don Martín de Aivar, e don Gonçalo Gómez de Agüero, e don Ferrant García Çatón, e don fray Domingo de Medina,

doctor de los predigadores; e fray Pedro, su compañero, e don Martín de Estella con conseyo de Cañas. (CODEA-0405, 1162, España, CHARTA).

En el caso del pronombre relativo, tanto masculino como femenino, “quien”, interesa la datación de su forma plural “quienes”, ya que, según las gramáticas históricas, esta no se desarrolla hasta los siglos XVI-XVII, durante el español clásico, pues anteriormente tampoco presentaba distinción de número. En CHARTA se halla el pronombre relativo plural “quienes” en 19 casos en 17 documentos, desde 1426 hasta 1840 —5,26 % (1/1) siglo xv, 5,26 % (1/1) siglo xvi, 10,53 % (2/2) siglo xvii, 57,89 % (11/9) siglo xviii, 21,05 % (4/4) siglo xix—, por lo que ya es posible datar su aparición, aunque sea de forma muy ocasional, a partir del siglo xv, como se puede observar en el siguiente ejemplo:

- (9) Sea manifiesto a todos omes que yo Domingo Rodrigo, carnicero, vezino de Ruvielos, aldea de la ciudat de Teruel, non seducto ni engañado, mas con franco coraçon e agradable voluntat, de mi cierta ciencia vendo e de present livro o quasi livro e en corporal posesión e tenencia pongo de día e non de noche, segunt fvero de la dita ciudat quiere e manda, a vós Joán Çapata, vezino del dito lugar de Ruvielos, qui sodes present e a los vuestros o a **quienes** vós querredes por siempre, es a saber, una viña e tierra yerma que yo é sitiada en el Donatue, término del dito lugar de Ruvielos, que afronta con viña de mí, dito Domingo Rodrigo, vendedor quí desús, e con viña de Pero Marco, e con viña de Estevan, cabrero mayor, e con viña de Estevan, cabrero menor. (CODEA-0796, 1426, España, Teruel, CHARTA).

En relación con los determinantes, resulta también muy significativa la evolución por gramaticalización del adjetivo numeral “un” al valor de artículo indeterminado de manera progresiva durante la época medieval. Para nuestro estudio, por razones operativas de búsqueda en el corpus, con el fin de evitar la presencia de formas gramaticalmente ambiguas, se ha seleccionado el determinante femenino plural “unas”, el cual se encuentra en 137 casos en 111 documentos, desde 1179 hasta 1815 —0,72 % (1/1) siglo xii, 25,54 % (35/30) siglo xiii, 17,51 % (24/20) siglo xiv, 15,32 % (21/18) siglo xv, 18,24 % (25/19) siglo xvi, 11,67 % (16/12) siglo xvii, 10,21 % (14/10) siglo xviii, 0,72 % (1/1) siglo xix—. En este caso, importa su frecuencia relativa de empleo en los diferentes siglos. Su primera documentación en CHARTA es la siguiente:

- (10) Ego doña María de Castañeda vendo vobis doña Eulalia una tierra que avía en carrera del Campo e tomé en precio XV morabetinos e so pagada de precio e de

robra e nichil remansit pro dare e por sennas fue la robra **unas** luvas. (CODEA-0222, 1179, España, Palencia, CHARTA).

Por último, dentro de los pronombres y determinantes estudiados, como complemento del caso anterior aunque por razones diferentes, es interesante además tener en cuenta el artículo determinado, en nuestro estudio a partir de la forma “las”, la cual aparece en CHARTA en 10028 casos en 1483 documentos, desde 1174 hasta 1871 —0,11 % (12/4) siglo XII, 39,00 % (3911/736) siglo XIII, 14,79 % (1484/234) siglo XIV, 13,08 % (1312/129) siglo XV, 19,45 % (1951/175) siglo XVI, 7,54 % (757/101) siglo XVII, 5,28 % (530/81) siglo XVIII, 0,70 % (71/23) siglo XIX—. La evolución de su frecuencia de empleo desde su origen pronominal latino a su nuevo valor como determinante durante el español medieval deviene un dato muy significativo en el establecimiento de la periodización, pues no tuvo, según las gramáticas históricas, una difusión general desde un principio en todos los sintagmas nominales, dependiendo a veces de la función sintáctica que estos desempeñaban. Una muestra de su uso ya desde el siglo XII se recoge en el siguiente ejemplo:

- (11) *In Dei nomine, amen. Notum sit tam presentibus quam futuris quod concilio de Confortos e el abat de Santa María de Aguilar ovieron tal vida en uno que la eglesia de Santo Micael e la eglesia de Santa María es una colación e los hijos de la eglesia facen sennas raciones, e los fraires sacan como uno d'elos, e si por aventura no oviere ý filio de la villa que canten los fraires amas **las** eclesiás e tomen todo; e otrosí si los fraires no podieren servir **las** eclesiás e hijos oviere ý de la eclesia sirvan **las** eclesiás e tomen todo. E sobre esto cumpliendo los hijos de **las** eclesiás e los fraires ningún omne non aya poder de meter capellano de fora.* (CODEA-0207, 1174, España, Palencia, CHARTA).

## 2.2. Morfología verbal

Por lo que respecta a la morfología verbal, existen dos aspectos de gran repercusión en la historia del español. El primero de ellos es la evolución, por razones en las que todavía hoy no se han puesto completamente de acuerdo las gramáticas históricas, de la primera persona del singular del presente de indicativo de los verbos “ser”, “estar”, “dar” e “ir” en el paso del español medieval al español clásico. En nuestro estudio, nos centramos en la evolución final de la forma “so” > “soy/soi”, contexto

en donde CHARTA recoge de esta última 108 casos en 82 documentos, desde 1244 hasta 1853 —5,56 % (6/5) siglo XIII, 3,70 % (4/3) siglo XIV, 9,26 % (10/10) siglo XV, 56,48 % (61/41) siglo XVI, 13,89 % (15/13) siglo XVII, 7,41 % (8/7) siglo XVIII, 3,70 % (4/3) siglo XIX—. Si bien se explica mayoritariamente este cambio como un proceso característico del paso de la lengua medieval a la clásica, resulta llamativo hallar ejemplos en el corpus analizado ya en el siglo XIII, e incluso en los siglos posteriores de la Edad Media, aunque sea de forma muy minoritaria, lo cual no impide que su mayor frecuencia de empleo sea a partir del siglo XVI, como era de esperar. Su primera documentación en el corpus aparece en el siguiente ejemplo:

- (12) [...] *et ego Dominus Petrus Servandi, de mandato Munio Felices, soy* fiador e manero de defender el cabildo con esta heredade si nul omne fuesse de sus filios o de sos parentes la quisés demandar o de otro omne que les quisés contrariar sobr'ela. (CODEA-0509, 1244, España, Salamanca, CHARTA).

El segundo aspecto relacionado con la morfología verbal, también de gran incidencia temporal, es la evolución de la desinencia general de número y persona de la segunda persona del plural “des” > “is”, según las gramáticas históricas ya a partir del siglo XV en las formas llanas y en el XVII en las esdrújulas, aunque solo se tendrá en cuenta la primera. Así, en el corpus, la forma verbal “sodes”, que marca el final de una etapa lingüística, se halla en 71 casos en 61 documentos, desde 1224 hasta 1524 —43,66 % (31/31) siglo XIII, 35,21 % (25/17) siglo XIV, 16,90 % (12/10) siglo XV, 4,23 % (3/3) siglo XVI, 0/0 siglo XVII, 0/0 siglo XVIII, 0/0 siglo XIX—. Como se puede constatar por los datos anteriores, la desinencia “des” en las formas llanas presenta una gran vitalidad en CHARTA durante los siglos XIII y XIV, comienza a disminuir su frecuencia en el XV y acaba desapareciendo finalmente a principios del XVI. Los ejemplos de su primera y última documentación en el corpus son los siguientes:

- (13) a. *Notum sit presentibus e futuris quod ego Micael, abas Sante Marie de Aguiar, una cum conventu eiusdem loci, de bona voluntate damus racione ad vos Ferrant González e a vuestra mugier doña Sancha Roíz porque sodes* nuestros familiares e que la ayades en toda vuestra vida tal racione cada año: (CODEA-0239, 1224, España, Palencia, CHARTA).  
b. Yo don Alonso Pérez de Guzmán, duque de Medinasidonia, conde de Niebla, marqués de Caçaça, señor de la noble cibdad de Gibraltar, hago saber a vos el mi contador mayor, e a vos el mi recabrador mayor, e a vuestros

lugares tenientes, e a vós los mis armadores, e escrivanos, e mayordomos e otros cualesquier fatores de las mis almadravas de Conil que agora **sodes** e **seredes** de aquí adelante que mi voluntad es de hazer limosna al señor prior, e frailes e convento del monesterio de mi señora santa María de Guadalupe de cadaún año para siempre jamás de quinze dozenas de atunes para provissión del dicho monesterio, [...]. (CODEA-0101, 1524, España, Huelva, CHARTA).

### 2.3. *Construcciones sintácticas*

Por último, en relación con la sintaxis, hay tres construcciones muy características de una determinada etapa temporal. La primera de ellas es la secuencia formada por “artículo determinado + posesivo + sustantivo”, distintiva del español medieval, pero que se desgramaticaliza en el español clásico. En CHARTA se ha analizado la combinación “la su”, presente en 131 casos en 102 documentos, de 1255 a 1560 —30,53 % (40/33) siglo XIII, 38,93 % (51/35) siglo XIV, 19,85 % (26/21) siglo XV, 10,69 % (14/13) siglo XVI, 0/0 siglo XVII, 0/0 siglo XVIII, 0/0 siglo XIX—. Como se puede observar, esta construcción posee una gran vitalidad en el corpus durante todo el español medieval, sobre todo en el siglo XIV, e incluso mantiene también una cierta presencia durante la primera mitad del siglo XVI, aunque significativamente no se encuentra hasta la segunda mitad del siglo XIII:

- (14) a. E pues que los reyes d'este señor e d'este rey avemos el nombre e d'él tomamos el poder de fazer justicia en la tierra e todas las onras e todos los bienes d'él nacen e d'él vienen e él quiso guardar **los nuestros derechos**, sin que él es señor sobre todo e puede fazer como él quisiere en todo, por el amor que él nos muestra en guardar **los nuestros derechos**, grand razón es e gran derecho que nós le amemos e que-l temamos e que guardemos **la su onra** e **los sus derechos**, e mayormientre el diezmo que él señaladamiento guardó e retovo para sí, por mostrar que él es señor de todo e d'él e por él vienen todos los bienes. (CODCAR-0141, 1255, España, Burgos, CHARTA).  
b. y que mandará que así les sea guardado y cumplido y que persona alguna no les baya ni pase contra lo susodicho ni quiera cosa alguna ni parte d'ello agora ni en ningún tiempo ni por alguna manera so pena de **la su merced** y de las penas en los dichos privilegios e cartas contenidas, todo lo cual vuestra magestad como rey y señor d'estos reinos a suplicación de los procuradores de cortes que están presentes jura y promete y otrosí confirma y dize. (CODEA-0201, 1560, España, Toledo, CHARTA).

Así, en (14a), el primer caso de este tipo en CHARTA, de principios de la segunda mitad del siglo XIII, proliferan diferentes posibilidades de esta combinación —“los nuestros derechos”, “la su onra e los sus derechos”— en un mismo contexto. Sin embargo, en (14b), último ejemplo hallado en el corpus, solo aparece un único caso de esta construcción en este fragmento al tratarse de un texto de finales del segundo tercio del siglo XVI, cuando la secuencia ya se encuentra en declive.

La segunda construcción sintáctica, incluso más característica del español medieval que la anterior, es la que permite la interpolación de elementos lingüísticos, normalmente sustantivos o adverbios, entre el pronombre personal clítico y el verbo en forma personal. Una de las combinaciones más habituales, analizada en el presente corpus, es la que intercala el adverbio de negación “no(n)”. Así, en CHARTA, a pesar del tipo de textos que lo componen, la secuencia “lo non/no” seguida de un verbo en forma personal se encuentra en 183 casos —172 “lo non” y 11 “lo no”— en 144 documentos, desde 1231 hasta 1524 —59,56 % (109/86) siglo XIII, 23,49 % (43/36) siglo XIV, 13,66 % (25/17) siglo XV, 3,27 % (6/5) siglo XVI, 0/0 siglo XVII, 0/0 siglo XVIII, 0/0 siglo XIX—. Como se puede apreciar por los datos anteriores, se trata de una construcción claramente medieval, de gran vigor principalmente en el siglo XIII, cuya presencia va disminuyendo progresivamente a lo largo de toda esta época, hasta acabar por desgramaticalizarse en el segundo cuarto del siglo XVI. Se recogen a continuación el primer y el último ejemplo hallados:

- (15) a. Mando que todos los que vinieren a mercado a Sant Fagund que den el portadgo assí como dize en la carta del rey don Alfonso, mio avuelo; e si alguna cosa quisieren dezir por que **lo non devén dar**, mando que vengan ante mí. E si non, mando a los merinos que los peindren fata que den so portadgo, o que vengan ante mí mostrar por qué **lo non devén dar**. (CODCAR-0007, 1231, España, Valladolid, CHARTA).
- b. E los unos nin los otros no hagades nin hagan ende ál so pena de la mi merced e de diez mil maravedís a cualquier que contra ello fuere e **lo non guardare** nin cumpliere para la mi cámara. (CODEA-0101, 1524, España, Huelva, CHARTA).

Finalmente, la tercera y última construcción sintáctica seleccionada, y en cierto punto también relacionada con la anterior, es la secuencia formada por la intercalación del pronombre personal clítico en la forma perifrástica

del futuro imperfecto de indicativo, entre el infinitivo y el auxiliar. Se trata también de una combinación más propia de la época medieval, como se comprueba en el corpus estudiado con “\*r lo é”, la cual aparece solo en 4 casos en 4 documentos, todos en textos legislativos, en 1268, 1278, 1300 y 1301 —50,00 % (2/2) siglo XIII, 50,00 % (2/2) siglo XIV, 0/0 siglo XV, 0/0 siglo XVI, 0/0 siglo XVII, 0/0 siglo XVIII, 0/0 siglo XIX—, como se puede apreciar en los siguientes ejemplos:

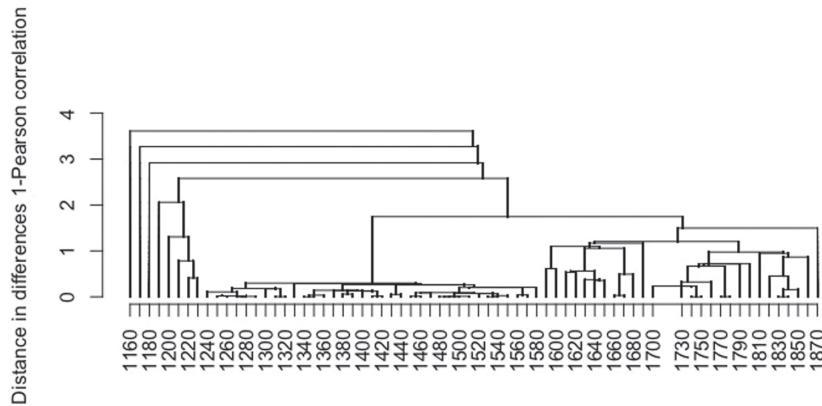
- (16) a. E si ál fiziessen, non gelos cosintades e embiátmelo dezir e yo **escarmen-**  
**tar lo é e porné** otro en so logar. (CODCAR-0290, 1268, España, Cádiz,  
CHARTA).
- b. Onde vos mando, luego vista esta mi carta, que l dedes sobre bonos fiadores  
e, si dar non lo quisiéredes, embiadme la pesquisa que fue fecha contra él e la  
razón porque l prisiestes, escripta e sellada con vuestros seallos, e yo **veer**  
**lo é e mandaré** y lo que toviere por bien. (CODCAR-0407, 1278, España,  
Valladolid, CHARTA).
- c. E aquello que montare en los pechos que a mí ovieren a dar d'aquí adelante,  
que los descuenten de la cabeza de la villa o del lugar do ellos fueren vecinos  
e moradores e que lo cuenten en la cabeza de los pechos que a mí ovieren  
a dar; yo **recebir lo é** en cuenta. (CODCAR-0700, 1300, España, Burgos,  
CHARTA).
- d. Otrossí tengo por bien que si cartas mías desaforadas algunos mostraren que  
sean contra los privilegios e cartas que an los concejos de los reyes onde yo  
vengo e de mí que les yo confirmé, que las tomen los alcalles del logar o los  
merinos e que non usen d'ellas e que me embíen mostrar el traslado d'ellas  
en cómo dizan que son contra sus privilegios, e yo **librar lo é** como tovier  
por bien e fallar que es derecho. (CODCAR-0704, 1301, España, Burgos,  
CHARTA).

Así como la presencia del clítico pronominal parece condicionar la preferencia por la forma verbal perifrásica analítica y su interpolación en todos estos casos —“escarmentar lo é”, “veer lo é”, “recebir lo é”, “librar lo é”—, su ausencia parece conllevar la selección de la forma sintética, incluso coordinada con la anterior —“porné”, “mandaré”—. Por otra parte, llama la atención los pocos casos que se han encontrado de esta construcción en el corpus, todos concentrados en la segunda mitad del siglo XIII y el inicio del siglo XIV, cuando se trata de una secuencia con una mayor frecuencia en otro tipo de corpus que recogen textos medievales. Por tanto, es posible suponer que el género textual que conforma CHARTA tal vez no favorezca una mayor presencia.

### 3. EL CORPUS CHARTA Y EL AGRUPAMIENTO JERÁRQUICO

Para obtener los datos necesarios para el análisis cuantitativo con el fin de poder realizar su agrupamiento jerárquico, se ha extraído la frecuencia relativa por millón de palabras por década de cada uno de los fenómenos gramaticales anteriores, principalmente desde el año 1200 al estar los documentos redactados claramente en romance, o antes si los datos hallados así lo justificaban<sup>5</sup>, hasta 1871, fecha del último texto incluido en el corpus. A los resultados conseguidos se les ha aplicado el algoritmo de VNC para poder dibujar un dendrograma que permita tener una representación aproximada de su periodización. A continuación, se analizan los cuatro árboles obtenidos a partir del análisis de estos fenómenos gramaticales en CHARTA tanto por separado como en conjunto, comparando el resultado con el del CORDE, para poder observar las posibles similitudes y diferencias.

En primer lugar, la combinación de los datos extraídos del estudio de los ocho pronombres y determinantes seleccionados ha permitido obtener el Gráfico 1:



**Gráfico 1** Periodización de CHARTA según las frecuencias relativas de los pronombres y los determinantes

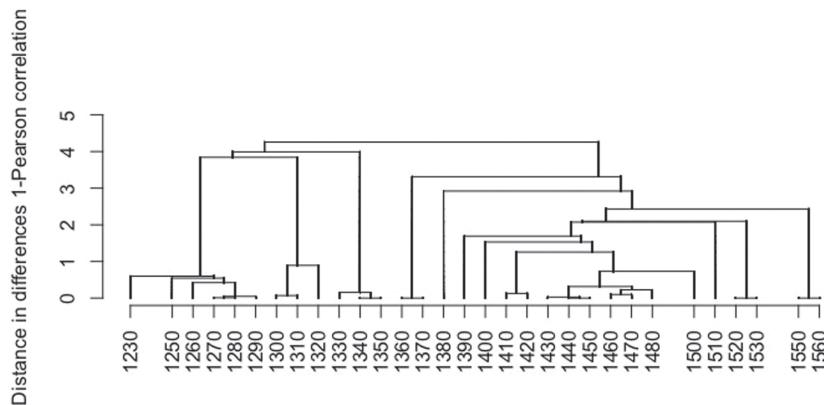
5 Así sucede con el pronombre posesivo “su”, el artículo determinado “las” y el artículo indeterminado “unas”, que presentan ejemplos ya desde 1162, 1174 y 1179, respectivamente.

El grupo de pronombres y determinantes es el más numeroso y completo de nuestro análisis, y el que más datos cuantitativos arroja, por lo que los resultados conseguidos tienen una especial relevancia para el establecimiento de períodos históricos. Si se observa detenidamente el dendrograma del Gráfico 1, a pesar de que no todas las décadas contienen datos o incluso un número suficiente de ellos, es fácil comprobar la existencia de una etapa muy bien definida y cohesionada, que iría básicamente desde el año 1240 hasta el año 1550/1580 —podría alargarse hasta 1580, pero no parece tan claro—. Este periodo sería el que representaría las características propias de lo que hasta ahora se ha denominado español medieval. Sin embargo, no sería un bloque único, sino que podría tener tres posibles subperiodos: a) de 1240 a 1320, primera etapa del español medieval, en la que encontraríamos simultáneamente sus formas más características, como “aqueste” y “qui”, junto a “os”, “su”, “unas” y “las”; b) de 1340 a 1420, en la que se produciría el desarrollo de la frecuencia de uso de las anteriores; y c) de 1450 a 1550, en la que, además de hallar los últimos casos de “qui”, tendríamos los primeros de “quienes”. Aparte de este periodo, es posible establecer otro, de 1590 a 1860, que se podría dividir a su vez en dos más, de 1590 a 1690 y de 1700 a 1860. El primero de estos dos subperiodos se caracterizaría por presentar los últimos casos de “aqueste”, mientras que el segundo englobaría ya exclusivamente las formas modernas.

En el caso del CORDE, los datos de los pronombres y los determinantes (cf. Granvik y Sánchez Lancis 2018: 760) permitían establecer una primera etapa anterior, de 1000 a 1130, caracterizada por la baja frecuencia de empleo de “las” y “su”, por la ausencia de “quienes” y “nosotros”, y el uso esporádico de “este” y “os”; y una segunda etapa, a partir de 1420, con la aparición de “quienes” junto con la disminución de la frecuencia de “qui”, desde 1470 con un destacado aumento de empleo de “os” y desde 1550 de “nosotros”. Entre medio, de 1140 a 1410 no habría cambios destacables, por lo que también tendría una homogeneidad como etapa histórica.

Si comparamos los resultados del análisis de ambos corpus por lo que respecta a los pronombres y los determinantes estudiados, ciertas coincidencias nos permitirían establecer, en principio y con gran cautela, una frontera temporal en el año 1420, considerado como el fin de las características definitorias de la etapa medieval o el inicio de un nuevo periodo lingüístico, no necesariamente diferente en su totalidad del anterior.

Por lo que respecta a las construcciones sintácticas analizadas, los datos conseguidos han permitido establecer el siguiente dendrograma:

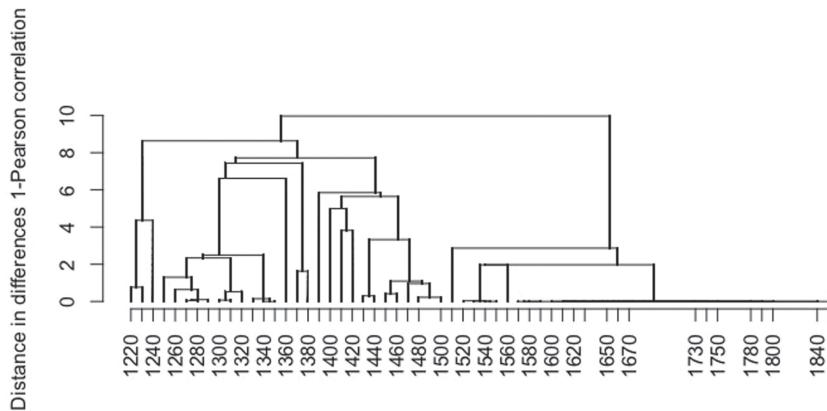


**Gráfico 2** Periodización de CHARTA según las frecuencias relativas de las construcciones sintácticas

Al ser las dos primeras construcciones sintácticas definitorias del español medieval, es lógico que el resultado obtenido en el árbol muestre un periodo que comprenda casi exclusivamente esta etapa. Sin embargo, resulta interesante constatar que, exceptuando la intercalación del clítico en las formas de futuro perifrásicas, de muy escasa presencia en CHARTA, el resto mantiene en parte su presencia más allá del siglo xv, por lo que podríamos establecer un periodo, a partir de los datos exclusivamente sintácticos, desde 1230 hasta 1530/1560, que coincidiría bastante con el que obteníamos antes con los pronombres y los determinantes, de 1240 a 1550 aproximadamente. En otras palabras, los datos sintácticos permitirían avalar los resultados anteriores.

En el caso del CORDE (cf. Granvik y Sánchez Lancis 2018: 762), si bien el número de construcciones analizadas es mayor, su análisis posibilita establecer también un árbol muy similar al de CHARTA, en el que se establecen dos períodos principales, con bastante variación. La primera etapa va de 1170 a 1520, en la que confluyen los casos de interpolación de la negación entre el clítico y el verbo “lo no(n)” y el empleo del artículo ante posesivo + sustantivo “la su”. La segunda etapa incluiría el resto, de 1530 a 1970, al desgramaticalizarse las secuencias anteriores.

En relación con el análisis de la morfología verbal, dado el escaso número de elementos hallados, ha sido necesario combinar los datos conseguidos con los de las construcciones sintácticas para obtener un dendrograma suficientemente significativo, que presentamos en el Gráfico 3:



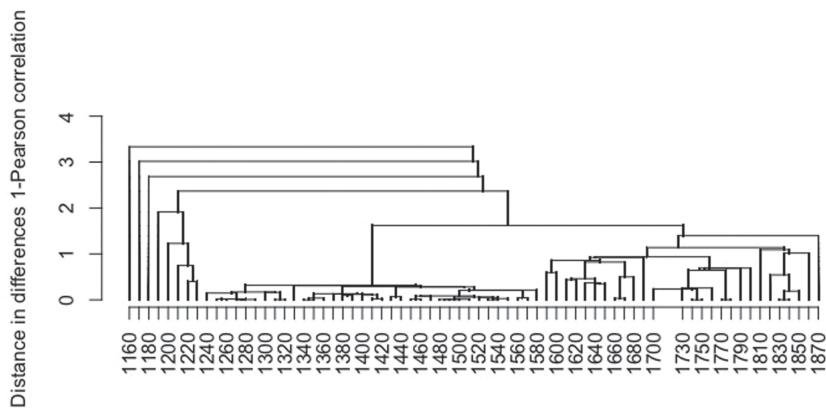
**Gráfico 3** Periodización de CHARTA según las frecuencias relativas de las formas verbales y las construcciones sintácticas

En este caso, podemos establecer también a partir de los datos de CHARTA dos grandes etapas históricas. La primera va desde 1220 hasta aproximadamente 1500, en la que los rasgos medievales característicos tienen una gran presencia, como, por ejemplo, la desinencia general de número y persona de la segunda persona del plural “sodes”, además de la existencia de las construcciones sintácticas anteriores. La segunda comprende desde 1520 hasta 1850, en la que ya prácticamente no perviven las formas medievales y, por ejemplo, predomina la primera persona del singular del presente de indicativo “soy”. De nuevo, el año 1520 volvería a marcar una frontera temporal en la periodización del español.

Por su parte, los datos del CORDE (cf. Granvik y Sánchez Lancis 2018: 761–762) en el análisis de las formas verbales permiten distinguir cuatro períodos: de 1150 a 1580, de gran variación; de 1590 a 1670, con el aumento de la frecuencia de la forma “soy” y la práctica desaparición de la forma “sodes”; de 1690 a 1750, en la que dicha desinencia general evoluciona también definitivamente en las palabras esdrújulas; y de 1760 a 1970, en la que se encuentra el sistema verbal actual. Consideramos que

la gran disparidad de tamaño y tipología textual existente entre ambos corpus, así como el menor número de formas verbales estudiadas en CHARTA, ofrecen esta diferencia de periodización de la morfología verbal, aspecto que se tendrá en cuenta para estudios posteriores.

Finalmente, se han englobado los datos proporcionados por los 13 fenómenos gramaticales analizados en un solo dendrograma, como se puede ver en el Gráfico 4:



**Gráfico 4** Árbol general de CHARTA basado en las frecuencias relativas de los 13 fenómenos gramaticales

Como se puede apreciar en este gráfico, se reafirman en gran parte las diferentes etapas históricas definidas antes en el corpus, debido sobre todo al peso en el estudio tanto de los pronombres y los determinantes como de las construcciones sintácticas, los cuales permiten establecer, junto con el análisis de las formas verbales, una periodización. De este modo, obtendríamos dos grandes etapas: una primera periodización, que va desde 1240 hasta 1550/1580, con sus posibles divisiones internas —1240–1320, 1340–1420 y 1450–1550—, que caracterizaría un modelo lingüístico específico y se correspondería, hasta cierto punto, con el denominado español medieval, en donde podría observarse una subetapa final de transición; y una segunda periodización, desde 1590 hasta 1870, también con sus divisiones —1590–1680, 1700–1800 y 1810–1860—, en la que se va perfilando progresivamente el modelo actual, desde una ya clara distinción con la etapa anterior.

Si comparamos estos datos con los obtenidos en el CORDE (cf. Granvik y Sánchez Lancis 2018: 759), a pesar de las diferencias existentes tanto entre los corpus como entre los análisis, salvando las distancias, se observa la repetición de algunas fechas, lo que en principio da cierta validez a los resultados conseguidos tanto en uno como en otro. Así, en el corpus académico se establece en primer lugar un primer periodo de orígenes, de 1000 a 1180, y un extenso segundo periodo, de 1190 a 1970, que se subdividiría en dos etapas, de 1190 a 1410 y de 1420 hasta finales del siglo xx.

#### 4. CONCLUSIÓN

El presente trabajo ha demostrado, en primer lugar, la gran utilidad de la aplicación de la lingüística de corpus al estudio de la periodización de la historia del español. Gracias a la existencia en la actualidad de una mayor variedad de corpus históricos representativos es más fácil aproximarse al pasado de la lengua española. Entre ellos, el *Corpus Hispánico y Americano en la Red: Textos Antiguos* (CHARTA), a pesar de su reducido tamaño, pero gracias a su especialización en cuanto a su género textual, permite, por un lado, obtener una visión del cambio lingüístico desde los textos no literarios y, por otro, convertirse en un corpus de control de los resultados aportados por otros corpus mucho más grandes, como, en el caso que nos ocupa, el *Corpus diacrónico del español* (CORDE).

En segundo lugar, se ha constatado que el estudio de fenómenos exclusivamente gramaticales posibilita el establecimiento de etapas históricas mucho más cercanas a la realidad de la época, en las que el concepto de siglo, como se ha demostrado, ya no tiene sentido, pues la evolución de los distintos fenómenos sugiere la existencia de una sucesión continua de diferentes períodos, en los que de manera progresiva se van abandonando unos elementos para ser sustituidos por otros. De esta forma, es posible establecer en la historia del español grandes etapas que a su vez permiten ser subdivididas en otras de menor duración.

En tercer y último lugar, se ha comprobado también la validez de la aplicación del método de análisis cuantitativo denominado agrupamiento jerárquico para el establecimiento de la periodización, ya que nos permite obtener una visión global más precisa de la evolución de una lengua,

que, en definitiva, no es más que la representación del conjunto de textos compilados en un determinado corpus, en este caso CHARTA. En el momento en que CHARTA disponga de un conjunto más equilibrado y mucho mayor de textos para el análisis de un número más elevado y variado de fenómenos gramaticales, podremos avanzar mucho más en el estudio de la periodización de la historia del español.

## Bibliografía

### CORPUS

- [CHARTA]: *Corpus Hispánico y Americano en la Red: Textos Antiguos*. Disponible en: <http://www.corpuscharta.es>. Fecha de consulta: mayo de 2022.
- [CORDE] Real Academia Española: *Corpus diacrónico del español*. Disponible en: <http://corpus.rae.es/cordenet.html>. Fecha de consulta: mayo de 2022.

### REFERENCIAS BIBLIOGRÁFICAS

- BUENAFUENTES, Cristina (2002a): “Cambio gramatical y periodización en español”, *Res Diachronicae*, 1, 96–105.
- BUENAFUENTES, Cristina (2002b): “Sintaxis y periodización en español”, *Res Diachronicae*, 1, 393–398.
- CANO AGUILAR, Rafael (1992): “La sintaxis española en la época del Descubrimiento”, en José Antonio Bartol Hernández, Juan Felipe García Santos y Javier de Santiago Guervós (eds.), *Estudios filológicos en homenaje a Eugenio de Bustos Tovar*, Salamanca: Universidad de Salamanca, 183–197.
- CANO AGUILAR, Rafael (coord.) (2005): *Historia de la lengua española*, Barcelona: Ariel, 2<sup>a</sup> ed.

- CARDILLO, Marcelo (2015): *Análisis de agrupamiento jerárquico*, material de enseñanza publicado en Academia.edu <<https://conicet.academia.edu/MarceloCardillo/Teaching-Documents>> [julio de 2015].
- EBERENZ, Rolf (1991): “*Castellano antiguo y español moderno*: reflexiones sobre la periodización en la historia de la lengua”, *Revista de Filología Española*, 71, 79–106.
- EBERENZ, Rolf (2009): “La periodización de la historia morfosintáctica del español: propuestas y aportaciones recientes”, *Cahiers d'études hispaniques médiévales*, 32, 181–201.
- ECHENIQUE ELIZONDO, María Teresa y María José MARTÍNEZ ALCALDE (2000): *Diacronía y gramática histórica de la lengua española*, Valencia: Tirant lo Blanch.
- GIRÓN ALCONCHEL, José Luis (2005): “Cambios gramaticales en los Siglos de Oro”, en Rafael Cano Aguilar (coord.), *Historia de la lengua española*, Barcelona: Ariel, 859–893, 2<sup>a</sup> ed. actualizada.
- GIRÓN ALCONCHEL, José Luis (2008): “La lengua de un embajador y un marino del siglo XVIII: ¿español moderno ya, o todavía clásico?”, en Concepción Company Company y José G. Moreno de Alba (eds.), *Actas del VII Congreso Internacional de Historia de la Lengua Española (Mérida (Yucatán), 4–8 de septiembre de 2006)*, vol. II, Madrid: Arco/Libros, 2243–2253.
- GRANVIK, Anton y Carlos SÁNCHEZ LANCIS (2018): “Un acercamiento cuantitativo a la periodización en la historia del español”, en M<sup>a</sup> Luisa Arnal *et al.* (eds.), *Actas del X Congreso Internacional de Historia de la Lengua Española*, vol. I, Zaragoza: Institución “Fernando el Católico”, 751–766.
- GRIES, Stefan Thomas (2009): *Statistics for Linguistics with R*, Berlin/Boston: De Gruyter Mouton.
- GRIES, Stefan Thomas y Martin HILPERT (2008): “The identification of stages in diachronic data: variability-based neighbor clustering”, *Corpora*, 3/1, 59–81.
- GRIES, Stefan Thomas y Martin HILPERT (2012): “Variability-based neighbor clustering: A bottom-up approach to periodization in historical linguistics”, en Terttu Nevalainen y Elizabeth Closs Traugott (eds.), *The Oxford Handbook oh the History of English*, Oxford: Oxford University Press, 134–144.
- HILPERT, Martin (2012): “Diachronic collostructional analysis: how to use it and how to deal with confounding factors”, en Kathryn

- Allan y Justyna A. Robinson (eds.), *Topics in English Linguistics: Current Methods in Historical Semantics*, Berlin: De Gruyter Mouton, 133–160.
- LAPESA, Rafael (1981): *Historia de la lengua española*, Madrid: Gredos, 9<sup>a</sup> ed.
- MARCOS MARÍN, Francisco (1992): “Spanisch: Periodisierung”, en Günther Holtus, Michel Metzeltin y Christian Schmitt (eds.), *Lexikon der Romanistischen Linguistik*, vol. VI/1, Tübingen: Max Niemeyer, 602–607.
- MARCOS MARÍN, Francisco (1995): “La periodización”, en M<sup>a</sup> Antonia Martín Zorraquino y Túa Blesa (coords.), *Homenaje a Félix Monge. Estudios de lingüística hispánica*, Madrid: Gredos, 325–333.
- MARTÍNEZ ALCALDE, M<sup>a</sup> José y Mercedes QUILIS MERÍN (1996): “Nuevas observaciones sobre la periodización en la historia de la lengua española”, en Alegría Alonso González, Ladislao Castro Ramos, Bertha Gutiérrez Rodilla y José Antonio Pascual Rodríguez (eds.), *Actas del III Congreso Internacional de Historia de la Lengua Española (Salamanca, 22–27 de noviembre de 1993)*, vol. I, Madrid: Arco/Libros, 873–886.
- MELIS, Chantal, Marcela FLORES y Sergio BOGARD (2003): “La historia del español. Propuesta de un tercer período evolutivo”, *Nueva Revista de Filología Hispánica*, 51, 1–56.
- MENÉNDEZ PIDAL, Ramón (1942): “El lenguaje del siglo XVI”, en Ramón Menéndez Pidal, *La lengua de Cristóbal Colón, el estilo de Santa Teresa y otros estudios sobre el siglo XVI*, Madrid: Espasa-Calpe, 47–84.
- OCTAVIO DE TOLEDO Y HUERTA, Álvaro S. (2007): “Un rasgo sintáctico del primer español moderno (ca. 1675–1825): las relaciones interoracionales con *ínterin (que)*”, en Marta Fernández Alcaide y Araceli López Serena (coords.), *Cuatrocientos años de la lengua del Quijote: estudios de historiografía e historia de la lengua española. Actas del V Congreso Nacional de la Asociación de Jóvenes Investigadores de Historiografía e Historia de la Lengua Española*, Sevilla: Universidad de Sevilla, 421–442.
- RIDRUEJO ALONSO, Emilio (1993): “¿Un reajuste sintáctico en el español de los siglos XV y XVI?”, en Ralph Penny (ed.), *Actas del Primer Congreso Anglo-Hispano*, vol. I, Madrid: Castalia, 49–60.

- Rojo, Guillermo (2021): *Introducción a la lingüística de corpus*, London/New York: Routledge.
- SÁNCHEZ LANCIS, Carlos (1997–1998): “Cambio morfológico y periodización en español preclásico”, *Estudi General*, 17–18, 173–199.
- SÁNCHEZ LANCIS, Carlos (1999): “Sintaxis histórica, informática y periodización del español”, en José Manuel Blecua, Gloria Clavería, Carlos Sánchez y Joan Torruella (eds.), *Filología e Informática. Nuevas tecnologías en los estudios filológicos*, Lleida/Bellaterra: Milenio/Seminario de Filología e Informática (Universidad Autónoma de Barcelona), 415–446.
- SÁNCHEZ LANCIS, Carlos (2001): “Historia de la lengua, gramática histórica y periodización en español”, *Estudi General*, 21, 395–412.
- SÁNCHEZ LANCIS, Carlos (2009): “Corpus diacrónicos y periodización del español”, *Cahiers d'études hispaniques médiévaux*, 32, 159–180.
- SÁNCHEZ LANCIS, Carlos (2012): “Periodización y cambio gramatical: el siglo XVIII, ¿frontera temporal del español?”, en M<sup>a</sup> Teresa García-Godoy (ed.), *El español del siglo XVIII. Cambios diacrónicos en el primer español moderno*, Bern/Berlin/Bruxelles/Frankfurt am Main/New York/Oxford/Wien: Peter Lang, 21–51.
- SÁNCHEZ LANCIS, Carlos (2022): “Corpus diacrónicos del español de España”, en Giovanni Parodi, Pascual Cantos Gómez y Chad Howe (eds.), *Lingüística de corpus en español / The Routledge Handbook of Spanish Corpus Linguistics*, London/New York: Routledge, 33–44.
- SERRADILLA CASTAÑO, Ana (2001): “Frecuencia de preposiciones en el régimen de construcciones transitivas: hacia una periodización lingüística de la Edad Media”, *Cahiers de linguistique et de civilisation hispaniques médiévaux*, 24, 143–165.
- TORRENS ÁLVAREZ, M<sup>a</sup> Jesús (2018): *Evolución e historia de la lengua española*, Madrid: Arco/Libros, 2<sup>a</sup> ed. actualizada.
- TORRUELLA CASAÑAS, Joan (2017): *Lingüística de corpus: génesis y bases metodológicas de los corpus (históricos) para la investigación en lingüística*, Frankfurt am Main: Peter Lang.

## NÓMINA DE EVALUADORES

Belén Almeida Cabrejas, Universidad de Alcalá de Henares

Beatriz Arias Álvarez, Universidad Nacional Autónoma de México

Florencio del Barrio de la Rosa, Universidad de Venecia

Teresa Bastardín Candón, Universidad de Cádiz

Matteo de Beni, Università degli Studi di Verona

Ivo Buzek, Universidad Masaryk

Víctor Caballero Gómez, Universidad Alcalá de Henares

Jesús Camacho Niño, Universidad de Jaén

Elena Carmona Yanes, Universidad de Sevilla

Francisco Carriscondo Esquivel, Universidad de Málaga

Emili Casanova Herrero, Universidad de Valencia

Viorica Codita Petrea, Universidad de Neuchâtel

Antonio Corredor Aveledo, Universität Zürich

Francisco Díaz Montesinos, Universidad de Málaga

Rocío Díaz Moreno, Universidad de Alcalá de Henares

Pedro Dono López, Universidade do Minho

Rolf Eberenz, Universidad de Lausanne

Cristina Egido Fernández, Universidad de León

Marta Fernández Alcaide, Universidad de Sevilla

Mar Garachana Camarero, Universidad de Barcelona

Amalia García Pedraza, Universidad de Granada

Patricia Giménez Eguíbar, Western Oregon University

María Nieves González Herrero, Universidad de Salamanca

Clara Grande López, Centro de Estudios de la RAE (DHLE)

Martin Hummel, Universidad de Graz

Javier Jiménez López de Eguileta, Universidad de Cádiz

Alejandro Junquera Martínez, Universidad de León

Víctor Lara Bermejo, Universidad de Cádiz

Magdalena León Gómez, Università degli studi dell'Aquila

Leyre Martín Aizpuru, Universidad de Sevilla

Aurelia Martín Casares, Universidad de Málaga

Lorena de Mattéis, Universidad Nacional del Sur

Francisca Medina Morales, Universidad de Granada

Javier Medina López, Universidad de La Laguna

Idanely Mora Peralta, Universidad Nacional Autónoma de México

María del Carmen Moral del Hoyo, Universidad de Cantabria

José Ramón Morala Rodríguez, Universidad de León

Maksim Ósipov, Universidad de Sevilla

Rafael Padrón Fernández, Universidad de la Laguna

Antonio Peláez Rovira, Universidad de Granada

Hermógenes Perdiguero Villareal, Universidad de Burgos

Manuel Pérez Saldanya, Universidad de Valencia

Marta Pérez Toral, Universidad de Oviedo

Ricardo Pichel Góterrez, Universidad de Alcalá

David Porcel Bueno, Universidad de Granada

Miguel Ángel Puche Lorenzo, Universidad de Murcia

Diego Antonio Reinaldos Miñarro, Universidad de Granada

Marcela Rivadeneira Valenzuela, Universidad de Arturo Prat

Sara Robles Ávila, Universidad de Málaga

Antonio Rafael Rubio Flores, Universidad de Granada  
Engracia Rubio Perea, Universidad de Málaga  
Julián Sancha Vázquez, Universidad de Cádiz  
Marta Torres Fernández, Universidad de Jaén  
Joan Torruella Casañas, Universidad Autónoma de Barcelona  
Hiroto Ueda, Universidad de Tokio  
Santiago Vicente Llavata, Universidad de Valencia  
Andrzej Zielinski, Universidad Jaguelónica de Cracovia