



Universitat
de les Illes Balears

TESIS DOCTORAL
2022

**LOS PRONOMBRES *NOS(OTROS)* Y
VOS(OTROS) EN CASTELLANO MEDIEVAL:
VARIACIÓN Y CAMBIO EN EL ESPACIO
GEOGRÁFICO PENINSULAR**

Marina Gomila Albal



Universitat
de les Illes Balears

TESIS DOCTORAL
2022

Programa de Doctorado en Filología y Filosofía

**LOS PRONOMBRES *NOS(OTROS)* Y
VOS(OTROS) EN CASTELLANO MEDIEVAL:
VARIACIÓN Y CAMBIO EN EL ESPACIO
GEOGRÁFICO PENINSULAR**

Marina Gomila Albal

Director: Andrés Enrique Arias
Tutor: Andrés Enrique Arias

Doctora por la Universitat de les Illes Balears



Universitat
de les Illes Balears

Dr. Andrés Enrique Arias, de la Universitat de les Illes Balears

DECLARO:

Que la tesis doctoral que lleva por título *Los pronombres nos(otros) y vos(otros) en castellano medieval: variación y cambio en el espacio geográfico peninsular*, presentada por Marina Gomila Albal para la obtención del título de doctor, ha sido dirigida bajo mi supervisión y cumple con los requisitos necesarios para optar al título de Doctor Internacional.

Y para que quede constancia de ello firmo este documento.

Firma

Palma de Mallorca, 14 de marzo de 2022

Lista de publicaciones derivadas de la tesis

- Enrique-Arias, Andrés, Patricia Ribas Marí, y Marina Gomila Albal. En evaluación. «Algunos trasvases lingüísticos entre Castilla y el oriente peninsular en la baja Edad Media». En *La lengua española en las fuentes documentales del siglo XIII al XIX. Estudios sobre el corpus CODEA*, editado por B. Almeida Cabrejas y P. Sánchez-Prieto Borja.
- Gomila Albal, Marina. 2016. «Sobre el origen y la difusión geográfica de las formas nosotros y vosotros en castellano». *Iberoromania* 83:103-25.
- Gomila Albal, Marina. 2018. «Variación diacrónica y diatópica de con + pronombre personal de 1a y 2a persona del plural». *Bulletin of Hispanic Studies* 95(8):801-23.
- Gomila Albal, Marina. 2020. «La variación estilística en una situación de contacto de lenguas: un estudio exploratorio». *Estudios de Lingüística del Español* 42:263-281.
- Gomila Albal, Marina. 2021. «Variación y cambio de los pronombres personales nós y vós junto con los modificadores otros, todos, mismos y ambos». Pp. 97-114 en *Variación diatópica y morfosintaxis en la historia del español (Anexo 82 de Verba)*, editado por B. Garrido Martín, M. del C. Moral del Hoyo, y M. Raab. Santiago de Compostela: Servizo de Publicacións e Intercambio Científico.
- Ribas Marí, Patricia y Marina Gomila Albal. En prensa. «La variación pronominal en los romanceamientos bíblicos del XV: hacia una caracterización dialectal del corpus Biblia medieval». En *Traducción bíblica e historia de las lenguas iberorrománicas*, editado por A. Enrique-Arias. Berlín: De Gruyter.

If you were to ask practising historical linguists why they study change in language, they would give you lots of different reasons, but certainly included in their answers would be that it is fun, exciting and intellectually engaging, that it involves some of the hottest topics in linguistics, and that it has important contributions to make to linguistic theory and to the understanding of human nature.

Lyle Campbell

A Coco, a es padrí y a s'avía.

Agradecimientos

Escribo estas líneas sin terminar de asumir que de verdad esta etapa esté terminando. En este momento un poco irreal me vienen a la mente los nombres de aquellas personas sin las que dudo que hubiera podido terminar esta tesis, o siquiera empezarla.

En primer lugar, quiero dar las gracias a mi director, Andrés Enrique Arias, por descubrirme el mundo de la investigación. Todo lo aprendido estos últimos años, todas las experiencias vividas y todas las personas maravillosas a las que he conocido en este tiempo han sido posibles gracias a que un día a finales de junio de 2014 me ofreció hacer mi trabajo de fin de grado sobre historia de la lengua. Casi ocho años después, me siento tremendamente afortunada por haber podido aprender de él, y espero poder seguir haciéndolo.

Debo dar las gracias también a Ruth Miguel, Laura Camargo, Assum Rost y Toni Bernat, que fueron una parte esencial de mi formación durante el grado y con los que he podido compartir departamento y proyectos estos años.

Asimismo, estoy muy agradecida a las instituciones que me han dado su apoyo durante el doctorado: a la Universitat de les Illes Balears, que me concedió el contrato predoctoral con el que empecé la tesis en enero de 2017; y al Ministerio de Ciencia, Innovación y Universidades, que me otorgó en septiembre de 2018 el contrato FPU con el que he terminado el doctorado (FPU17/00087). También quiero agradecer los comentarios y observaciones de los dos evaluadores anónimos que han revisado esta tesis antes de su depósito definitivo.

Además, esta tesis no sería como es sin mi estancia en el Romanisches Seminar de la Universität Zürich, donde Johannes Kabatek, Carlota de Benito, Cristina Bleorțu y Marisa Gago me acogieron como a una más. Thanks also to Irene and Carl, who made me feel at home. Esos tres meses me cambiaron como investigadora y como persona.

No puedo dejar de mencionar en estas líneas a los amigos y aijhleros que he tenido la fortuna de conocer durante estos años y que ahora ocupan un lugar privilegiado en mi corazón. Podría nombrar aquí a muchísimas personas y recordar experiencias tan intensas como la de Gijón, Nador o Murcia, pero me limitaré a algunos nombres: Blanca, que me acompañó en la construcción de los cimientos de la tesis; Patricia, a la que siento mi hermana académica; y Ana, con la que he compartido penas y alegrías. Ellas se han convertido en una parte fundamental de mi vida académica y personal.

También debo dar las gracias a mis padres, a mi abuela y al resto de mi familia, que, aunque quizá no entendían del todo lo que estaba haciendo, no han dejado de apoyarme. Especialmente debo darle las gracias a Laurita, que es la que más me ha escuchado hablar de todo lo que me entusiasma y, aunque sea mi hermana pequeña, es para mí un modelo de fortaleza y de la amistad más genuina. Finalmente, gracias a Juan por ser mi apoyo incondicional, por esperarme y por aguantar los peores momentos de estos últimos años con paciencia y cariño.

Resumen

Esta tesis doctoral tiene como objetivo reexaminar los procesos de cambio que afectan al paradigma de los pronombres tónicos de primera y segunda persona plural en castellano medieval (es decir, *nós*, *vós*, *nosotros*, *vosotros*, *connusco*, *convusco*) desde el punto de vista de su evolución diacrónica y diatópica. Por tanto, mi propósito principal es trazar el origen y la difusión de cada una de las variantes lingüísticas en el tiempo y en el espacio.

En la elaboración del corpus de esta investigación he utilizado textos de tipología textual variada mediante la selección rigurosa de fuentes. Para ello, he prestado especial atención a la calidad filológica de los textos y he evitado aquellos transmitidos en copias tardías o aquellos para los que no se conoce la fecha precisa de composición o su origen geográfico. El establecimiento de una sólida base empírica, que me ha permitido trazar el origen cronológico y geográfico de los cambios lingüísticos, ha hecho posible comprobar nuevas hipótesis y examinar en detalle los patrones de difusión de las ya mencionadas variantes morfosintácticas a través de la península ibérica.

En mi análisis, examino el reemplazo de los pronombres de plural *nós*, *vós* por las formas compuestas *nosotros*, *vosotros*. Además, amplío el análisis del reemplazo de estos pronombres de plural con la incorporación de otros refuerzos pronominales enfáticos, como *todos*, *mismos* y *ambos*. Finalmente, he analizado las variantes pronominales tras la preposición *con* (*connusco*, *con nós*, *con nosotros*, y *convusco*, *con vós*, *con vosotros*).

Los resultados de esta tesis demuestran la complejidad de la interacción de un amplio abanico de factores en la difusión de las variantes lingüísticas. Esto solo puede desvelarse mediante la utilización de un corpus con coordenadas cronológicas y geográficas precisas, y con suficiente variedad de registro como para acceder, aunque sea indirectamente, al factor estilístico.

La variación geográfica y cronológica ha resultado determinante en el análisis de los tres procesos de cambio aquí presentados. La expansión de *nosotros*, *vosotros* se inicia en el oriente peninsular, fruto del contacto de lenguas con el catalán, y las nuevas formas acceden a los textos castellanos como variantes de prestigio. Asimismo, el cartografiado exhaustivo de las variantes pronominales reforzadas revela la relevancia de dos formas que habían pasado generalmente desapercibidas, *nós todos* y *vós todos*, vinculadas a las áreas occidentales. En último lugar, aunque *con nós* y *con vós* no se originan en el oriente peninsular, la ausencia de variantes en *-usco* en los documentos aragoneses lleva a pensar que el uso de las formas analíticas (*con* + pronombre tónico) pudo extenderse tempranamente por el este de la península ibérica debido al contacto con las variantes *ab nós* y *ab vós* del catalán.

En definitiva, esta investigación subraya las posibilidades alcanzadas mediante el examen cuidadoso de datos de registros variados, así como la necesidad de incorporar el análisis de la variación geográfica para poder alcanzar una comprensión completa de la difusión de innovaciones lingüísticas en situaciones de contacto.

Resum

Aquesta tesi doctoral té com a objectiu reexaminar els processos de canvi que afecten el paradigma dels pronoms tònicos de primera i segona persona plural en castellà medieval (és a dir, *nós*, *vós*, *nosotros*, *vosotros*, *connusco*, *convusco*) amb especial atenció a la seva evolució diacrònica i diatòpica. Per tant, el meu propòsit principal és traçar l'origen i la difusió d'aquestes variants lingüístiques en el temps i l'espai.

Per l'elaboració el corpus d'aquesta investigació he utilitzat textos de tipologia textual variada mitjançant la selecció rigorosa de fonts. Per això, he prestat especial atenció a la qualitat filològica dels textos i he evitat aquells transmesos en còpies tardanes o aquells per als quals no es coneix la data precisa de composició o el seu origen geogràfic. L'establiment d'una sòlida base empírica, que m'ha permès traçar l'origen cronològic i geogràfic dels canvis lingüístics, ha fet possible comprovar noves hipòtesis i examinar en detall els patrons de difusió de les variants morfosintàctiques ja esmentades a través de la península ibèrica.

En la meua anàlisi, examino la substitució dels pronoms de plural *nós*, *vós* per les formes compostes *nosotros* 'nosaltres', *vosotros* 'vosaltres'. A més, amplio l'anàlisi de la substitució d'aquests pronoms de plural amb la incorporació d'altres reforços pronominals emfàtics, com *todos* 'tots', *misimos* 'mateixos' i *ambos* 'ambdós'. Finalment, he analitzat les variants pronominals després de la preposició *con* 'amb' (*connusco*, *con nós*, *con nosotros*, i *convusco*, *con vós*, *con vosotros*).

Els resultats d'aquesta tesi demostren la complexitat de la interacció d'un ampli ventall de factors en la difusió de les variants lingüístiques. Això només es pot revelar mitjançant la utilització d'un corpus amb coordenades cronològiques i geogràfiques precises, i amb prou varietat de registre per accedir, encara que sigui indirectament, al factor estilístic.

La variació geogràfica i cronològica ha resultat determinant en l'anàlisi dels tres processos de canvi presentats aquí. L'expansió de *nosotros*, *vosotros* s'inicia a l'orient peninsular, fruit del contacte de llengües amb el català, i les noves formes accedeixen als textos castellans com a variants de prestigi. Així mateix, el cartografiat exhaustiu de les variants pronominals reforçades revela la rellevància de dues formes que havien passat generalment desapercibudes, *nós todos* i *vós todos* vinculades a les àrees occidentals. En darrer lloc, encara que *con vós* i *con vosotros* no s'originen a l'orient peninsular, l'absència de variants en *-usco* en els documents aragonesos fa pensar que l'ús de les formes analítiques (*con* + pronom tònic) va poder estendre's aviat per l'est de la península ibèrica a causa del contacte amb les variants *ab nós* i *ab vós* del català.

En definitiva, aquesta investigació subratlla les possibilitats aconseguides mitjançant l'examen atent de dades de registres variats, així com la necessitat d'incorporar l'anàlisi de la variació geogràfica per poder assolir una comprensió completa de la difusió d'innovacions lingüístiques en situacions de contacte.

Abstract

The present study aims to re-examine the processes of change that affect the paradigm of stressed pronouns of the first- and second-person plural in Old Spanish (i.e. *nós*, *vós*, *nosotros*, *vosotros*, *connusco*, *convusco*) with particular focus on the diachronic and diatopic aspects of their historical evolution. Thus, my main purpose has been to trace the origin and diffusion in time and space of each of the relevant linguistic variants.

In collecting the corpus for this study, I have gathered texts of varied typology through the rigorous and careful selection of sources. In doing so I have paid special attention to the philological quality of the texts, avoiding those transmitted in late copies, or the ones for which there is no information on date of composition or geographical origin. The establishment of a solid empirical base, which enabled me to trace the chronological and geographical origin of the changes, allowed me to test new hypotheses and to examine in detail the diffusion patterns of the aforementioned morphosyntactic variants throughout the Iberian Peninsula. The use of a wide corpus of varied textual typologies has allowed me to conduct a thorough re-examination of aspects analyzed in previous studies.

In my analysis, I examine the replacement of the plural pronouns *nós* ‘we’, *vós* ‘you-PL’ by the compound forms *nosotros*, *vosotros*. In addition, I have extended the analysis of the replacement of these plural pronouns with the incorporation of other emphatic pronominal reinforcements, such as *todos* ‘all’, *misimos* ‘selves’, and *ambos* ‘both’. Finally, I have analyzed the pronominal variants after the preposition *con* ‘with’ (*connusco*, *con nós*, *con nosotros* ‘with us’, and *convusco*, *con vós*, *con vosotros* ‘with you-PL’).

The data presented in this dissertation demonstrates the complex interaction of a wide range of factors in the diffusion of linguistic variants. This can only be revealed through the use of a corpus with precise chronological and geographical coordinates and with enough register variety to access, even indirectly, the stylistic factor.

Geographical and chronological variation has been crucial in all three processes of change presented here. The expansion of *nosotros*, *vosotros* originates in the east of the peninsula, as a result of language contact with Catalan, and the new forms are later adopted by Castilian writers as prestigious variants. Likewise, the exhaustive mapping of the reinforced pronominal variants reveals the relevance of two forms that had been previously overlooked, *nós todos* and *vós todos*, associated to the western regions. These forms, due to their brief vitality, have generally gone unnoticed in previous studies. Lastly, although *con nós* and *con vós* do not originate in the eastern regions, the absence of *-usco* variants in the Aragonese documents leads us to assume that the use of *con* + pronoun type forms could have spread earlier throughout the eastern peninsula due to contact with the variants *ab nós* and *ab vós* of Catalan.

Ultimately, this research underscores the possibilities afforded by the careful examination of data from a variety of registers, and the need to include analysis of geographic variation, in order to better understand the diffusion of linguistic innovations in language-contact situations.

Índice de contenidos

Introducción.....	23
1. Presentación.....	23
2. Objetivo principal y metodología.....	24
3. Hipótesis y resultados esperables.....	25
4. Estructura.....	26
Capítulo 1: Marco teórico.....	29
1. Presentación.....	29
2. La sociolingüística variacionista.....	29
2.1. Fases del cambio lingüístico.....	30
2.2. Factores de difusión del cambio lingüístico.....	33
3. La dialectología en el estudio de la historia de la lengua española.....	35
3.1. Panorama lingüístico y contacto de lenguas en la Iberia medieval.....	36
3.2. Modelos de difusión espacial.....	38
4. Tipos de texto, géneros y tradiciones discursivas.....	40
5. La teoría de la gramaticalización.....	41
5.1. La gramaticalización como macrocambio.....	42
5.2. Unidireccionalidad.....	44
5.3. Economía y expresividad.....	44
5.4. Frecuencia de uso y contextos.....	45
5.5. Causas y mecanismos de gramaticalización.....	46
5.6. Gramaticalización secundaria.....	48
6. Tipología y universales lingüísticos.....	49
6.1. La expresión de las relaciones de marca y el paradigma pronominal.....	49
6.2. Jerarquías o escalas implicativas y el paradigma pronominal.....	50
6.3. Distinciones gramaticales en el paradigma pronominal.....	52
7. Conclusión.....	54
Capítulo 2: Corpus y metodología.....	55
1. Presentación.....	55
2. Textos y testimonios, originales y copias.....	56
3. Control de la procedencia dialectal y variación diatópica.....	57
4. Criterios de selección.....	58
4.1. Límites temporales y periodización.....	60
4.2. Tipos de textos.....	60
4.3. Límites geográficos y categorización del <i>continuum</i> dialectal.....	61
4.4. Accesibilidad de los textos y soporte electrónico.....	62

5.	Descripción del corpus	64
5.1.	Corpus de textos documentales: CODEA+ 2015	64
5.2.	Corpus de textos literarios: fuentes y nómina.....	65
5.3.	Corpus documental auxiliar.....	85
5.4.	Representatividad y equilibrio del corpus	88
6.	Extracción de datos.....	89
7.	Herramientas de análisis y visualización de datos.....	91
Capítulo 3: La aparición de los pronombres reforzados <i>nosotros</i> y <i>vosotros</i>		93
1.	Presentación.....	93
2.	Estado de la cuestión	94
2.1.	Causas de la aparición de las formas reforzadas	94
2.2.	Cronología	95
2.3.	Orden de aparición.....	95
2.4.	Contexto sintáctico	96
2.5.	Presencia de las variantes reforzadas en otras variedades	97
2.6.	Recapitulación: preguntas de investigación.....	98
3.	Datos y metodología específica	98
4.	Análisis	102
4.1.	Primeras ocurrencias.....	102
4.2.	Cronología	104
4.3.	Orden de aparición.....	106
4.4.	Contexto sintáctico	108
4.5.	Origen y difusión de las variantes en el espacio geográfico.....	112
4.6.	Variación textual y de registro.....	118
5.	Resumen	121
Capítulo 4: Los refuerzos pronominales enfáticos		123
1.	Presentación.....	123
2.	Estado de la cuestión	124
2.1.	Puntos clave del refuerzo con <i>otros</i>	124
2.2.	El uso de pronombres reforzados en otras lenguas.....	124
2.3.	<i>Nós</i> y <i>vós</i> con modificadores distintos a <i>otros</i> en castellano.....	125
2.4.	Recapitulación: preguntas de investigación.....	125
3.	Datos y metodología específica	126
4.	Análisis	128
4.1.	Primeras ocurrencias.....	128
4.2.	Cronología	128
4.3.	Primera y segunda persona	133

4.4. Distribución geográfica	134
5. Resumen	139
Capítulo 5: El retroceso de <i>connusco</i> y <i>convusco</i>	141
1. Presentación	141
2. Estado de la cuestión	142
2.1. Origen de las formas	142
2.2. Cronología	143
2.3. Causas del cambio	145
2.4. Presencia de <i>connusco</i> y <i>convusco</i> en otras variedades romances	146
2.5. Recapitulación: preguntas de investigación.....	146
3. Datos y metodología específica	147
4. Análisis	149
4.1. Primeras ocurrencias.....	149
4.2. Cronología	151
4.3. Primera y segunda persona	153
4.4. Valor contrastivo y contexto sintáctico	155
4.5. Distribución geográfica	157
5. Resumen	161
Summary and conclusions	163
Referencias bibliográficas	167
Índice de figuras	177
Índice de tablas	179
Índice de mapas	181

Introducción

1. Presentación

El paradigma de los pronombres personales tónicos de las lenguas y dialectos iberorromances sufre una serie de cambios durante la época medieval. Los pronombres tónicos de primera y segunda persona plural *nós* y *vós*, de origen latino, fueron reemplazados en el castellano medieval por las formas reforzadas *nosotros*, *vosotros*. En este mismo período, las formas *connusco* y *convusco* medievales empezaron a ser reemplazadas en primer lugar por *con nós*, *con vós* y, posteriormente, por las formas actuales *con nosotros*, *con vosotros*.

Los reemplazos de los pronombres *nós*, *vós* y *connusco*, *convusco* aparecen nombrados con frecuencia en los manuales de historia de la lengua (Alvar y Pottier 1983:122-25; Eberenz 2000:58-84; Lausberg 1973:162-64; Penny 2006:160-61; Torrens Álvarez 2007:89-90; entre otros) y han recibido una atención considerable en estudios específicos. No obstante, la mayoría de las investigaciones que tratan estos procesos de cambio que afectan a los pronombres de primera y segunda persona del plural se han basado en el análisis de textos literarios. En el estudio de la historia de la lengua en general se han privilegiado como fuentes aquellos textos pertenecientes al canon literario, es decir, una nómina de textos reducida en la que el mérito literario era el criterio de selección (Pons Rodríguez 2006). Así, se ha tendido a analizar la lengua de los textos más conocidos con una muestra de uno o dos testimonios por siglo y sin tener en cuenta su calidad filológica o interés lingüístico. El uso de esta nómina limitada de textos literarios acarrea algunos problemas: la falta de información sobre la fecha precisa de producción, la compleja transmisión textual de muchas de estas obras o la dificultad para controlar su procedencia dialectal.

El uso de un corpus de textos sin una procedencia dialectal determinada imposibilita que se preste atención al origen y la difusión de dichas variantes pronominales en el espacio geográfico. Solo Fernández-Ordóñez (2011:78) ha profundizado más sobre la dimensión diatópica en el reemplazo de *nós*, *vós* por *nosotros*, *vosotros*, y ha situado el foco del cambio en la antigua Corona de Aragón. Esta autora destaca, además, la influencia directa del catalán y del occitano. En cuanto a otras formas pronominales propias del castellano medieval, como *connusco*, *convusco*; u otras variantes de los pronombres reforzados, como *nós todos* y *vós todos*, no tenemos información sobre su distribución dialectal durante la Edad Media. En algunos casos, la presencia de las variantes pronominales en algunas de las variedades iberorromances actuales puede darnos pista sobre su recorrido histórico. Por ejemplo, hoy en día se conservan las formas *nós* y *vós* en algunas variedades occidentales, como el bable occidental, el leonés, el mirandés o el gallego-portugués (Alvar y Pottier 1983:123; Fernández-Ordóñez 2011:78; Gili Gaya 1946:116). Asimismo, *connosco* y *convosco* se mantienen todavía hoy en gallego-portugués, aunque son formas en retroceso (Álvarez Blanco 2004). El retroceso de *nós*, *vós* y *connusco*, *convusco*, formas habituales en el castellano del siglo XIII, hacia el extremo occidental, donde se conservan hoy en día, deja patente la relevancia de la adscripción geográfica de las variantes que intervienen en sus respectivos procesos de cambio durante el período medieval.

El análisis del origen y la difusión de las variantes lingüísticas en el espacio requiere que se disponga de textos de diferentes procedencias dialectales. Esto no siempre es sencillo: por un lado, puede que no se tenga información sobre el origen geográfico del texto; por otro, aunque se conozca, la intermediación de la copia manuscrita puede haber desvirtuado las características de un original que no se ha conservado (Enrique-Arias 2018:1555). Si desconocemos la adscripción dialectal de la muestra de textos que

utilizamos como base para nuestro análisis, la interpretación de los resultados va a ser necesariamente parcial y «podemos acabar atribuyendo a otros factores diferencias que en realidad se deben a la variación dialectal» (Enrique-Arias 2012:101). Asimismo, durante los últimos años, se ha producido una revitalización de los estudios de variación geográfica en el marco de la historia de la lengua española a partir, principalmente, del estudio de Fernández-Ordóñez (2011), quien traza algunos patrones geolingüísticos que contribuyen a explicar la conformación del español. El estudio de Fernández-Ordóñez subraya la necesidad de incorporar la perspectiva diatópica, de reconstruir el foco geográfico de las innovaciones lingüísticas y las etapas de su difusión en el espacio, para así alcanzar un conocimiento íntegro de la historia de la lengua.

2. Objetivo principal y metodología

En esta investigación me propongo reexaminar los procesos de cambio que afectan al paradigma de los pronombres tónicos de primera y segunda persona plural en el castellano medieval con especial atención a las perspectivas diacrónica y diatópica. De este modo, mi objetivo principal es trazar el origen y la difusión de cada una de las variantes lingüísticas en el tiempo y en el espacio.

Investigar el recorrido histórico y geográfico de cada una de las variantes pronominales exige una metodología específica. Durante la recopilación del corpus de esta tesis, he contado no solo con textos literarios, tal y como se había hecho en estudios previos, sino también con textos documentales. Aunque en muchos casos los textos literarios pueden ser problemáticos por la dificultad para situarlos en una fecha y localización geográfica concretas, y por su compleja transmisión textual, esto no significa que haya que renunciar a la utilización de este tipo de textos, sino que su selección y análisis deben ser cuidadosos. Por tanto, me he basado en la medida de lo posible en testimonios fiables o copias cercanas al original con data cronológica y geográfica conocidas. Asimismo, he utilizado textos de tipo documental. Este tipo de fuente apenas ha sido utilizado para los fenómenos objeto de estudio de esta tesis¹. Sin embargo, esta tipología textual tiene la ventaja de contar con una fecha y un origen geográfico definidos que incluso permiten el cartografiado de los datos.

Estos últimos años, se han desarrollado estándares filológicos de selección y presentación de textos que han sido la base de corpus diacrónicos fiables y de calidad. En la última década se han desarrollado corpus de textos documentales que permiten controlar tanto la data cronológica como la geográfica, y que, además, han sido seleccionados, transcritos y editados mediante criterios filológicos rigurosos. El principal ejemplo de estos corpus es el *Corpus de documentos españoles anteriores a 1800* (CODEA+ 2015), que es la base del corpus documental de esta tesis. Los avances en lingüística de corpus de las últimas décadas han puesto a nuestra disposición nuevos recursos que han cambiado la forma de acceder a las muestras de habla del pasado. La disponibilidad de estas herramientas permite acceder a ingentes colecciones de textos

¹ En el Capítulo 2, § 4 detallo las fuentes utilizadas en los estudios previos. En algunos trabajos sobre las variantes *conusco*, *convusco* se añaden textos documentales como complemento de un corpus principal de textos literarios (García 1997; Líbano Zumalacárregui 1988; Rini 1990). La adición de textos documentales concretamente en aquellos estudios que tratan las variantes en *-usco* probablemente se deba a la dificultad de encontrar ejemplos de estas formas y la necesidad de ampliar el corpus. En estudios más recientes se empieza a utilizar el CORDE o el *Corpus del español* para el análisis de *nos(otros)*, *vos(otros)* (Calderón Campos 2010; De Jonge y Nieuwenhuijsen 2009) pero, al menos aparentemente, no se filtran o seleccionan de ninguna manera los documentos utilizados, con los problemas que eso puede conllevar (Rodríguez Molina y Octavio de Toledo y Huerta 2017b). Además, en ninguno de estos trabajos se incluye la perspectiva diatópica en el sentido de hacer una selección de textos por distintas zonas geográficas de la península.

antiguos, lo que permite visitar fenómenos de cambio lingüístico con nuevos datos y perspectivas.

En definitiva, la presente investigación utiliza una síntesis de diferentes tradiciones y metodologías que unen el conocimiento filológico tradicional con los nuevos avances en lingüística histórica. La selección rigurosa de fuentes atendiendo a la calidad filológica y al control de la procedencia dialectal permite formar una base de datos robusta sobre la que comprobar nuevas hipótesis. Asimismo, el método de recolección y análisis de datos se nutre de las técnicas desarrolladas por la lingüística de corpus y la lingüística variacionista, y da un lugar destacado a la aplicación de métodos informáticos para la visualización y cartografiado de datos.

De este modo, en las páginas que siguen me propongo contribuir al estudio de la historia del español y al desarrollo de la dialectología histórica mediante el avance en el conocimiento de los patrones de origen y difusión de las innovaciones morfosintácticas dentro de la península ibérica.

3. Hipótesis y resultados esperables

Una parte importante del análisis que se realizará en las siguientes páginas se basará en comprobar las explicaciones de estudios previos a la luz de los nuevos datos. Dichas propuestas hacen referencia a aspectos como la cronología relativa de la evolución de cada una de las variantes, y la influencia de variables lingüísticas como la de persona o el contexto sintáctico de los procesos estudiados. Estas cuestiones se presentarán en detalle en los estados de la cuestión correspondientes (§ 2 de los Capítulos 3, 4 y 5). Las hipótesis principales de esta tesis, que presento a continuación, pretenden explicar el origen y la difusión de las variantes en el espacio geográfico.

Tal y como se ha comentado en § 1, la distribución actual de las variantes pronominales en las variedades iberorromances y la conservación de formas heredadas del latín como *nós*, *vós*, *connosco*, *convosco* en el occidente peninsular sugieren una difusión de las formas surgidas durante la baja Edad Media como *nosotros*, *vosotros* desde el oriente de la península ibérica. Por tanto, el foco oriental de estas variantes pronominales parece integrarse en el grupo de innovaciones morfosintácticas que se incorporan al castellano desde las variedades orientales. Estudios recientes muestran que ciertos fenómenos que afectan a la morfología del castellano llegan a esta lengua desde la zona oriental de la península ibérica. Fernández-Ordóñez (2011:76-81) describe algunos ejemplos de estos procesos: la pérdida de la concordancia en los tiempos compuestos, que se inicia en Navarra; la aparición de los tiempos compuestos, que llegaron a la Corona de Aragón a partir del contacto con el galorromance y se extendieron desde allí por el resto de la península; o el orden de los pronombres átonos, que llegó de nuevo desde el francés y el italiano.

En cuanto al reemplazo de *nós*, *vós* por *nosotros*, *vosotros* (Capítulo 3), Fernández-Ordóñez (2011:78) sitúa el foco de difusión de las nuevas formas en la antigua Corona de Aragón, por influencia del catalán y el occitano. En esta investigación comprobaré el recorrido de estas variantes a través del *continuum* septentrional en las diferentes variedades iberorromances, incluyendo el catalán y el gallego.

En relación con el reemplazo de *nós*, *vós*, en este estudio incorporo el análisis de otras variantes pronominales reforzadas que pudieron competir con *nosotros*, *vosotros* antes de que se gramaticalizaran en castellano, como *vós todos*, *vós mismos* o *vós ambos* (Capítulo 4). Así como *nosotros* y *vosotros* llegan al castellano desde el catalán y el occitano, y desde ahí se extienden hacia el oeste por el norte peninsular, es probable que los refuerzos pronominales *ambos*, *mismos* y *todos* se conserven más tiempo donde

nosotros y *vosotros* tarden más en llegar, es decir, en la mitad occidental de la península ibérica.

El último fenómeno que trataré es el reemplazo de *connusco*, *convusco* por *con nós*, *con vós* y, posteriormente, por *con nosotros*, *con vosotros* (Capítulo 5). Las variantes medievales en *-usco* están en la actualidad vinculadas al occidente peninsular por la conservación de *connosco*, *convosco* en gallego y portugués, por lo que es esperable que estas variantes fueran retrocediendo progresivamente hacia el oeste, empujadas por el aumento del uso de las variantes analíticas tipo *con nós* o *con nosotros*. En cualquier caso, aunque la difusión de las variantes innovadoras se produzca desde el oriente en el reemplazo de *connusco*, *convusco* del mismo modo que en el de *nós*, *vós* (Capítulo 3), es necesario comprobar si las vías de acceso de las nuevas formas y los desencadenantes de su aparición son los mismos en ambos procesos de cambio. Normalmente, para el estudio de la pérdida de las formas en *-usco* se han aplicado los mismos parámetros que para la pérdida de *nós*, *vós*. Es decir, se han considerado fenómenos muy estrechamente relacionados, ya que en ambos casos el resultado final del reemplazo son las formas que han llegado hasta la actualidad, (*con*) *nosotros*, (*con*) *vosotros*. Por tanto, se han intentado encajar las variables que explican el reemplazo de *nós*, *vós* para el de *connusco*, *convusco*, como el orden de aparición y los diferentes matices semánticos de las variantes. Mi hipótesis es que estos dos fenómenos son más independientes de lo que puede parecer a simple vista, por lo que desenmarañarlos y analizarlos de forma independiente puede servir para aclarar las causas de cada proceso de cambio.

La investigación de estas tres cuestiones a partir de fuentes con data cronológica y geográfica, en las que además se controla la tipología textual de los documentos, llevará a una comprensión completa de los procesos de cambio que afectan a los pronombres personales tónicos de primera y segunda persona de plural. La unión de las perspectivas diacrónica y diatópica permitirá la reconstrucción exhaustiva de la difusión de cada una de las variantes pronominales a través del tiempo y del espacio. A su vez, los resultados de este estudio suponen una aportación a la labor de constituir el «mapa variacional» de la lengua medieval; es decir, contribuyen a la identificación de los fenómenos sujetos a variación en cada época y área geográfica en la evolución del español.

4. Estructura

La tesis está estructurada como sigue. En el Capítulo 1: *Marco teórico* presento las distintas bases teóricas y metodológicas que han demostrado ser más eficientes para el estudio de la lengua en diacronía: la sociolingüística variacionista, la dialectología, las tradiciones discursivas, la teoría de la gramaticalización y la tipología y los universales lingüísticos. El punto de unión de estos enfoques es la percepción de la variación y el cambio como algo inherente a cualquier sistema lingüístico. Por tanto, estas disciplinas entienden la lengua como un sistema heterogéneo y variable por naturaleza, y buscan identificar y describir los mecanismos y las innovaciones que se encuentran tras el cambio lingüístico.

En el Capítulo 2: *Corpus y metodología*, comento los principales obstáculos que me he encontrado para establecer el corpus de estudio: la distinción entre texto y testimonio, original y copia (§ 2), y el control de la procedencia dialectal de los textos (§ 3). Asimismo, establezco los criterios de selección y clasificación de los textos de mi corpus, presento las fuentes de las que los he extraído y concreto la nómina de obras y testimonios (§ 5). Cierro el capítulo con algunas cuestiones prácticas sobre la extracción de datos y sobre las herramientas de análisis y visualización utilizadas en todos los capítulos.

Los tres capítulos restantes contienen el análisis de los procesos de cambio lingüístico objeto de estudio de esta tesis. En el Capítulo 3 examino el del origen y el

avance de las variantes pronominales reforzadas *nosotros* y *vosotros*, que reemplazan a las formas simples de plural *nós* y *vós* durante la baja Edad Media. Alrededor de este reemplazo gravitan los fenómenos de los capítulos siguientes: la presencia de otras variantes de los pronombres con refuerzos como *todos*, *mismos* y *ambos* (Capítulo 4); y el retroceso de las formas pronominales *connusco*, *convusco*, que fueron reemplazados por *con nós*, *con vós* y, posteriormente, por *con nosotros*, *con vosotros* (Capítulo 5).

Dentro de cada uno de estos capítulos repito la misma estructura básica. Inicio con una presentación del proceso de cambio lingüístico objeto de estudio y la relevancia de su análisis (§ 1) y continúo con un estado de la cuestión (§ 2), al final del cual planteo las preguntas de investigación de ese capítulo. En los tres casos dedico una sección a los datos y metodología específica (§ 3), en la que detallo el corpus utilizado y las particularidades del proceso de extracción de ejemplos. La parte principal la ocupa el análisis (§ 4), cuyas subsecciones dependen del fenómeno en cuestión. Finalmente, cada capítulo se cierra con un resumen en el que destaco las conclusiones más relevantes (§ 5).

Una vez cerrados los tres capítulos de análisis, presento el resumen final y las conclusiones (*Summary and conclusions*).

Capítulo 1: Marco teórico

1. Presentación

En esta tesis pretendo analizar la evolución de varios fenómenos lingüísticos que afectan al paradigma de los pronombres personales tónicos principalmente desde el punto de vista del recorrido histórico y geográfico de las variantes. Para llevar a cabo este objetivo con éxito, es necesario aprovechar las corrientes teóricas y metodológicas que han demostrado ser más eficientes para el estudio de la lengua en diacronía.

Así pues, en las páginas que siguen repasaré brevemente las cuestiones más relevantes de las siguientes perspectivas que abordan el estudio del cambio lingüístico: en primer lugar, la sociolingüística variacionista (§ 2) aporta las herramientas necesarias para explicar las fases del cambio y para identificar los factores que influyen en la difusión y adopción de variantes innovadoras. En segundo lugar, la perspectiva diatópica de la dialectología (§ 3) es imprescindible para entender el desarrollo histórico de variantes lingüísticas en el espacio peninsular. En tercer lugar, para entender de forma completa la trayectoria de las variantes innovadoras es preciso tener en cuenta la dimensión textual (§ 4), debido a que los textos son los principales transmisores de las muestras de habla del pasado. En cuarto lugar, la teoría de la gramaticalización (§ 5) resulta indispensable para comprender la evolución que experimenta una forma o construcción gramatical durante el proceso del cambio lingüístico. Y, finalmente, el estudio de la tipología y los universales lingüísticos (§ 6) permiten dar explicaciones de tipo funcional sobre el cambio lingüístico, así como delimitar los procesos de cambio posibles. El punto de unión de estos enfoques es la percepción de la variación y el cambio como algo inherente a cualquier sistema lingüístico. Se entiende la lengua como un sistema heterogéneo y variable por naturaleza, y se busca identificar y describir los mecanismos y las innovaciones que se encuentran tras el cambio lingüístico. Por tanto, en este capítulo pretendo repasar de forma sintética las perspectivas teóricas y metodológicas citadas, que he considerado imprescindibles para afrontar el estudio del cambio lingüístico en diacronía¹.

Por otro lado, cualquier investigación sobre el habla de tiempos pretéritos tiene una limitación clara en cuanto a la obtención de dichas muestras de habla, ya que la única fuente directa de la que disponemos para acceder a la lengua antigua es la documentación escrita que ha llegado hasta la actualidad. En este sentido, en lo que respecta a este trabajo se plantean otros problemas, como el de la intermediación de la copia manuscrita y el control de la procedencia dialectal. En el Capítulo 2 abordaremos cómo los últimos desarrollos en la lingüística de corpus aplicada al estudio histórico del cambio tienen en cuenta estas limitaciones y tratan de paliarlas.

2. La sociolingüística variacionista

La esencia de la sociolingüística variacionista se sustenta principalmente en el concepto de lengua como sistema ordenadamente heterogéneo. El trabajo fundacional de Weinreich, Labov y Herzog (1968:100) plantea la noción de *orderly heterogeneity* y sienta las bases del estudio de la lengua desde la perspectiva de la variación y el cambio. La variación es una característica intrínseca de la lengua, pero no es libre, no es aleatoria o caótica, sino que sigue un orden, unos patrones, y está guiada por factores internos y

¹ El marco teórico que presento en este capítulo está reducido y condensado. Para una visión más exhaustiva véanse los tratamientos generales de Campbell (1999), Joseph y Janda (2003), Penny (2004), Hock y Joseph (2009), Hernández-Campoy y Conde-Silvestre (2012), y otros.

externos. El análisis de la variación lingüística pretende identificar cuáles son los patrones que ordenan el sistema, explicar cómo se produce el cambio y qué factores intervienen.

La sociolingüística se aproxima al estudio de la lengua como algo que existe necesariamente en un contexto social y que depende del hablante que la usa, de dónde la usa y por qué. Por tanto, busca encontrar un equilibrio entre los aspectos sociales y los lingüísticos. Es decir, pretende integrar tanto los factores externos del cambio —edad, clase, género, geografía— como los internos y estructurales². Así, tal y como destacan Weinreich, Labov y Herzog (1968:188): «Explanations of language which are confined to one or the other aspect —linguistic or social— no matter how well constructed, will fail to account for the rich body of regularities that can be observed in empirical studies of language behaviour».

Weinreich, Labov y Herzog (1968) plantean cinco principios o problemas que cualquier teoría del cambio lingüístico debería tener en cuenta: restricción (*constraints problem*), transición (*transition problem*), inserción (*embedding problem*), evaluación (*evaluation problem*) y actualización (*actuation problem*). Silva-Corvalán y Enrique-Arias (2017:269) los resumen así:

- a) La *restricción* consiste en determinar qué tipos de cambios y de condiciones para la ocurrencia del cambio son universalmente posibles.
- b) La *transición* corresponde a la ruta entre dos etapas de un cambio y constituye un problema lingüístico interno.
- c) La *inserción* tiene que ver con la identificación tanto de la matriz social como la matriz lingüística en la que se verifica el cambio. Es en la resolución de este problema donde el concepto de variable lingüística y los estudios de variación encuentran su más valiosa aplicación.
- d) La *evaluación* consiste en identificar las actitudes subjetivas de los miembros de la comunidad hacia el cambio en marcha.
- e) La *actualización*, finalmente, tiene que ver con la identificación de los factores tanto lingüísticos como sociales que motivan el cambio. La complejidad de los factores que pueden inducir la evolución lingüística hace que este sea el problema más difícil de resolver.

Estos cinco principios son totalmente aplicables al estudio de la lengua en diacronía. La sociolingüística variacionista no solo tiene como objetivo estudiar fenómenos en un momento sincrónico, sino que tiene en cuenta el «estado de la lengua en relación a su desarrollo histórico» (Silva-Corvalán y Enrique-Arias 2017:262) y ha desarrollado herramientas vitales para la descripción de las fases del cambio y para la identificación de los factores que influyen en la difusión y adopción de variantes innovadoras.

2.1. Fases del cambio lingüístico

El cambio lingüístico tiene lugar cuando una variante reemplaza a otra dentro de una comunidad de habla. A diferencia de los generativistas, que consideran que el cambio se produce de forma abrupta en el paso de una generación a otra (Lightfoot 1991), la sociolingüística entiende el cambio como un proceso gradual. De este modo, las variantes en competición coexisten en una gran parte del proceso, que puede durar desde décadas a siglos.

² En el estudio de la historia de la lengua en muchos casos no se pueden tener en cuenta todos los factores externos que suelen considerarse en sociolingüística. El acceso al contexto social de los hablantes está limitado por la información que ha llegado a la actualidad a través de la documentación escrita. En esta investigación en particular no podré considerar variables extralingüísticas como el género y la edad. Aunque de forma limitada, tendré acceso al registro a través de textos de tipologías diferentes —literarios y documentales— y, dentro de estos últimos, de diferentes ámbitos de emisión (*cf.* § 4).

En sus inicios, la sociolingüística solo aborda cuestiones fonéticas. En el momento de aplicar el modelo sociolingüístico al estudio del cambio sintáctico, aparece el problema de cómo identificar aquellas variantes lingüísticas que tienen el mismo significado, cuáles son las formas que «dicen la misma cosa» y por tanto compiten por el mismo contexto. Definir una variable sintáctica o morfosintáctica puede ser problemático por distintos motivos: las variables sintácticas tienen menos variantes que las fonológicas, sus contextos de ocurrencia son más reducidos y difíciles de identificar, y, especialmente, no es fácil ver si el uso de una estructura u otra depende de la voluntad de expresar matices semánticos distintos o si en realidad es posible elegir libremente entre opciones que comunican lo mismo (Silva-Corvalán y Enrique-Arias 2017:153). La solución metodológica que resuelve la cuestión del significado es el control de aquellos contextos que se solapan, aquellos en los que las variantes son intercambiables³.

Una vez establecidas las variantes en competición, el proceso de cambio lingüístico suele dividirse en tres fases sucesivas: innovación, difusión y adopción.

La primera fase del cambio lingüístico es la **innovación**, que consiste en la creación de una nueva variante en una comunidad de habla. Es importante no confundir esta fase con el cambio lingüístico en sí mismo. Se producen usos innovadores de la lengua constantemente, pero eso no implica necesariamente la difusión o adopción de esos usos por parte de una comunidad de habla: «La innovación no basta para conferir carta de naturaleza a un cambio, pero sin esta no hay cambio posible» (Rodríguez Molina 2010:520).

No siempre podemos identificar de forma exacta la innovación en un cambio lingüístico. La innovación es un acto individual de un hablante que tiene lugar en un lugar y contexto determinados, por lo que para identificarla deberíamos localizar el primer uso de la variante nueva. Esto resulta especialmente complejo en el estudio de la historia de la lengua, ya que únicamente contamos con documentaciones escritas. Salvo casos excepcionales, solo podremos determinar la primera documentación en nuestro corpus y, al estar formado por fuentes escritas, esa primera aparición será inevitablemente posterior al primer uso en el habla. Todos estos hechos requieren tratar con cautela las primeras documentaciones: la aparición de una variante en un documento únicamente nos demuestra su existencia en ese momento.

De hecho, y paradójicamente, los lingüistas históricos solo pueden certificar la existencia de una innovación una vez esta ha alcanzado una cierta difusión en la comunidad de habla y ha sido adoptada por un número suficiente de hablantes, pues de otro modo jamás habría pasado a la lengua escrita (Rodríguez Molina 2010:525).

Cuando una innovación empieza a propagarse en una comunidad de habla, nos encontramos con la segunda fase del cambio lingüístico: la **difusión**. Mientras que la innovación generalmente tiene lugar en el individuo, la difusión involucra a la comunidad, es decir, es un proceso social que depende de la interacción entre los hablantes.

La difusión se produce de forma lenta y gradual a medida que la nueva variante va ocupando contextos de uso. Esta fase es en realidad la que nos puede dar más información sobre los factores que influyen en el cambio lingüístico: parámetros intralingüísticos, sociales, estilísticos y geográficos.

La fase de **adopción** marca la culminación del cambio. Para que podamos considerar que este ha tenido lugar, la variante innovadora debe ser adoptada por todos los miembros de una comunidad.

³ Para más detalles sobre la cuestión del significado en la variación sintáctica cf. Silva-Corvalán y Enrique-Arias (2017:152 y ss.).

La difusión de los cambios lingüísticos exitosos a través del tiempo toma normalmente la forma de una **curva en S**. La propagación de una variante lingüística no es lineal o uniforme, es decir, no avanza siempre al mismo ritmo. Habitualmente, el cambio empieza lento, luego se acelera y, en su última etapa, vuelve a ralentizarse hasta que la variante innovadora llega al último rincón de los contextos posibles. Si proyectamos la frecuencia de uso de la forma innovadora (eje Y) a lo largo del tiempo (eje X), el resultado es un gráfico similar al de la Figura 1. Es decir, la línea que representa el avance de la forma innovadora tendría forma de S o de función logística⁴.

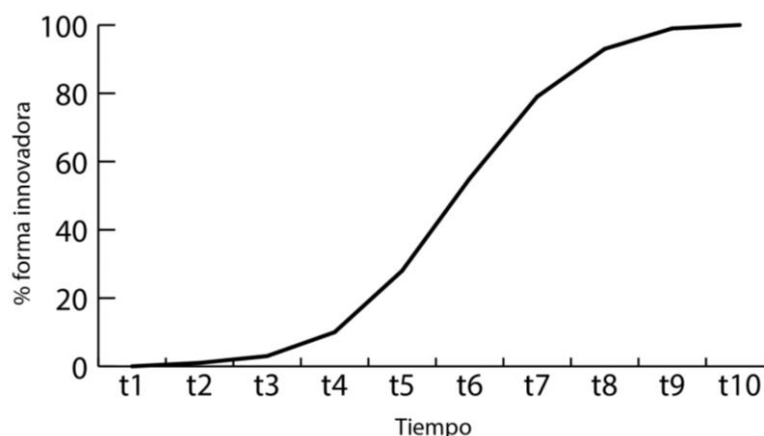


Figura 1. Representación idealizada de una curva en S. Adaptada de Denison (2003:53)

El proceso de cambio habitualmente tiene un inicio lento, ya que la variante innovadora aparece solo de forma esporádica con valores muy marcados. Posteriormente, hay una fase intermedia que se corresponde con la difusión del cambio. En esta etapa, el cambio avanza rápidamente a medida que la variante innovadora va propagándose a nuevos contextos. Finalmente, cuando la forma nueva ya es la mayoritaria, el cambio se ralentiza y se prolonga hasta llegar a reemplazar a la forma en recesión de todos sus contextos. Este proceso puede prologarse más o menos dependiendo del fenómeno.

To our knowledge there are no clearly documented cases of a change going toward completion that follows either a simple linear trajectory or an exponential curve (either slow start with rapid completion and no tapering off, or an immediate rapid increase followed by a slow completion rate) (Blythe y Croft 2012:280)

La descripción del cambio lingüístico completo como una curva en S no es una idea reciente. Aparece mencionada en el trabajo fundacional de Weinreich, Labov y Herzog (1968) y se alude a este patrón como una herramienta descriptiva útil en gran parte de la bibliografía sobre cambio lingüístico. Algunos ejemplos, que no pretenden ser exhaustivos, son Labov (1994:65, 2001), Chambers y Trudgill (2004:163), Denison (2003:58), Hernández-Campoy y Conde-Silvestre (2012), Nevalainen y Raumolin-Brunberg (2016:55), Silva-Corvalán y Enrique-Arias (2017:276) o Del Barrio de la Rosa (2018:48).

Según Nevalainen y Raumolin-Brunberg (2016:54-55), el patrón de la curva en S puede dividirse en 5 fases dependiendo de la frecuencia de aparición de la forma innovadora: una fase inicial (*incipient*) hasta el 15 %; una segunda fase en la que el

⁴ Se han propuesto diferentes funciones matemáticas para representar el modelo de la curva en S: «the Gauss cumulative function (Aldrich & Nelson 1984, Piotrovskaja & Piotrovskij 1974), the Lorentz cumulative function, and the logistic function» (Postma 2010:271). Un breve resumen de estas propuestas puede encontrarse en Denison (2003:58).

cambio empieza a acelerar (*new and vigorous*), entre el 16 y el 35 %; una fase intermedia (*mid-range*) entre el 36 y el 65 %, en la que el cambio se difunde más rápidamente; una cuarta fase en la que el avance empieza a ralentizarse (*nearing completion*), entre el 66 y el 85 %; y una fase final (*completed*), en la que la propagación vuelve a ser lenta, a partir del 86 %. Estas fases pueden ser útiles como herramienta descriptiva y para identificar en qué momento se encuentra un cambio en progreso.

Según Labov (1994:65-66), la causa detrás de este patrón de difusión estaría en la frecuencia de contacto entre los hablantes de las distintas variantes. Al principio del cambio, cuando la forma nueva aparece solo de forma esporádica, los hablantes estarían expuestos a esta variante de forma más infrecuente. A medida que se iría extendiendo, el aumento sería exponencial. Sin embargo, una vez la forma nueva fuera la mayoritaria, el cambio volvería a ralentizarse hasta poder llegar a todos los contextos posibles.

No obstante, no todos los cambios son exitosos. Los cambios incompletos o fallidos siguen un patrón diferente que ha recibido el nombre de **curva en campana**. En muchos casos, las innovaciones no se difunden lo suficiente como para llegar a generalizarse, y es posible que retrocedan hasta desaparecer. Son cambios que empiezan siguiendo la evolución de la curva en S, en los que, sin embargo, el avance de la forma innovadora se ve interrumpido y empieza a declinar. Así, el cambio se revierte, no se consume y traza una curva en forma de campana, que será más o menos acentuada dependiendo de lo avanzado que estuviera el cambio antes de revertirse: «Let us, therefore, define a failed change as a change that is composed of two factors, a rising S-curve multiplied by a falling S-curve, whose decline has an activation parameter equal to its rise» (Postma 2010:285).

Podemos encontrar algunos ejemplos de cambios fallidos en inglés en Postma (2010) y en Nevalainen (2015). También se han descrito algunos cambios que siguen este patrón en español, como el caso de los demostrativos reforzados tipo *aqueste* en Aragón (Enrique-Arias 2018; Enrique-Arias, Ribas Marí, y Gomila Albal En evaluación) o el de las formaciones adverbiales en *-mientras* (Del Barrio de la Rosa 2016, 2018).

En definitiva, para que haya cambio lingüístico, necesariamente debe haber al menos dos variantes en competición⁵. Esta situación de variación puede tener su resolución en el reemplazo de la variante antigua por la nueva o en el retroceso de la forma innovadora. Esto normalmente se refleja en dos patrones de difusión muy frecuentes: la curva en S o la curva en campana⁶.

2.2. Factores de difusión del cambio lingüístico

Una de las cuestiones más discutidas en el estudio del cambio lingüístico desde cualquier perspectiva es el establecimiento de las causas del proceso. Los factores que influyen en el cambio lingüístico son variados y complejos. Conocer estos factores no solo es relevante para determinar el detonante de la variación, sino también para entender las circunstancias que influyen en la difusión o adopción del cambio lingüístico, y, además, en su velocidad:

[...] de qué depende que con ciertos cambios se produzca una inversión de la marcación en un lapso relativamente breve en perspectiva diacrónica, dos o tres siglos [...], en otros casos tarde un milenio o más el cumplimiento de la inversión de la marcación [...] y en

⁵ Normalmente, la descripción de estos modelos de difusión se hace con dos variantes. En ocasiones, estudios sobre la curva en S como el de Nevalainen (2015) presentan casos de variación entre tres o más variantes, pero, hasta donde sé, no parece haber ningún modelo que trate específicamente este tipo de situaciones. Denison (2003:66) señala: «In linguistic work, at least, I am not aware of serious attempts to model competition among multiple variants».

⁶ Estos no son los únicos patrones de difusión posibles: cf. Nevalainen (2015).

otros casos aún, la forma innovadora permanezca por siglos en un estado temprano del canal de gramaticalización y no progrese el cambio, creando un estado de variación sincrónica permanente con convivencia de la forma innovadora y conservadora (Company 2008:20)

Habitualmente, los factores del cambio lingüístico se agrupan en dos grandes tipos: los internos o lingüísticos, y los externos o extralingüísticos.

Los **factores internos** son aquellos que forman parte del sistema lingüístico propiamente dicho y tienen que ver con la estructura de la lengua. Algunos ejemplos son la posición de la variante en un contexto determinado o la frecuencia de aparición (Hernández-Campoy y Conde-Silvestre 2012:253). Una innovación lingüística se origina en contextos que favorecen su significado y de ahí se va extendiendo progresivamente a través del sistema lingüístico y va ocupando contextos menos favorables. Durante esta difusión, la forma innovadora pierde restricciones y se vuelve independiente del contexto hasta que, si el cambio lingüístico tiene éxito, se apropia del ámbito gramatical de la forma etimológica o conservadora (Company 2008:17).

En el inicio del proceso, las innovaciones surgen y se localizan en contextos muy específicos, semántica y sintácticamente afines al significado de la forma innovadora, pero paulatinamente esta forma se va extendiendo a más y más contextos (a menudo, a expensas de una forma gramatical anterior) debido a la pérdida de las restricciones lingüísticas de partida que hacían incompatible el significado innovador con otros contextos diferentes de aquel en el que se gestó la innovación (Rodríguez Molina 2010:551).

Los **factores externos**, en cambio, son extralingüísticos y relacionan la variación o el cambio lingüístico con la comunidad o la sociedad que usa esa lengua (Hernández-Campoy y Conde-Silvestre 2012:389). La sociolingüística ha prestado especial atención a variables externas como el género, la edad o el nivel educativo de los hablantes, así como a cuestiones como el contacto de lenguas o los movimientos migratorios (Silva-Corvalán y Enrique-Arias 2017).

Desde la segunda mitad del siglo XX, la sociolingüística ha demostrado la influencia de los factores sociales en la variación lingüística. Sin embargo, en el estudio de la lengua de tiempos pretéritos hay que tener en cuenta que únicamente contamos con textos escritos, por lo que no es posible distinguir variables sociales típicas como la edad o el nivel educativo del mismo modo que se hace en la sociolingüística actual.

Uno de los conceptos más relevantes para el estudio de la variación lingüística en relación con la sociedad que se puede trasladar a investigaciones diacrónicas es el de la distinción entre cambio desde arriba y cambio desde abajo. Según Labov (1994:78), cuando se habla de la dirección del cambio, se hace en referencia al nivel socioeconómico de la comunidad y, especialmente, al nivel de consciencia del cambio por los hablantes. Por un lado, los cambios desde arriba son innovaciones conscientes introducidas generalmente por la clase social dominante. Por otro lado, las innovaciones introducidas desde abajo están por debajo de la consciencia de la comunidad, aparecen en el vernáculo y son resultado de factores lingüísticos internos (Labov 1994:78; Silva-Corvalán y Enrique-Arias 2017:270-71).

Esta cuestión es interesante por su relación con el factor geográfico (§ 3). Sánchez-Prieto (2012:12) propone dos modelos de difusión por el espacio geográfico según la dirección del cambio: a) areal para los cambios desde abajo y b) registral para los cambios desde arriba:

(1) Los cambios que se producen de abajo a arriba en la escala social son sensibles al factor geográfico, pues no parece que un cambio empiece al unísono y se desarrolle con la misma cadencia y reglas internas en todos los ámbitos regionales.

(2) Los cambios que se producen de arriba a abajo, es decir, los que empiezan en las capas culturalmente altas, que se suelen difundir, al menos inicialmente, a través de la escritura, no responden al factor geográfico (Sánchez-Prieto Borja 2012:30-31).

Por tanto, en caso de que al cartografiar las variantes no haya una continuidad en la distribución geográfica y no sea posible establecer el foco de la radiación, podemos deducir que es un cambio transmitido desde arriba, por la vía escrita (Del Barrio de la Rosa 2018:44).

En cuanto al factor registral, al que la lingüística histórica puede acceder solo de forma limitada a través de diferentes tipologías textuales, Sánchez-Prieto (2012:11) propone una escala a partir de los ámbitos de emisión de los textos documentales, de modo que se relacionarían directamente con un registro más o menos formal:

Cancilleresco > Judicial > Municipal > Eclesiástico > Particular > Declaración de partes y testigos en juicio

En cuanto al factor geográfico, este es uno de los factores externos más relevantes para esta tesis. Durante los últimos años se ha producido una revitalización de los estudios de historia de la lengua española que incorporan la perspectiva diatópica. Se están desarrollando numerosas investigaciones que buscan trazar el origen y la difusión de variantes lingüísticas a través del tiempo y del espacio, con el objetivo de conocer los patrones geolingüísticos que intervienen en la conformación del español. En esta investigación me propongo trazar el recorrido histórico y geográfico de algunas variantes de los pronombres personales tónicos del castellano medieval. Por ello, más adelante dedicaré una sección de este capítulo a dar algunas pinceladas sobre dialectología y sobre la difusión de variantes lingüísticas en el espacio geográfico (§ 3). Asimismo, en la descripción del corpus será imprescindible comentar las dificultades de controlar la procedencia dialectal de los textos (Capítulo 2, § 3) y los criterios de selección que afectan a los límites geográficos y la categorización del continuum dialectal (Capítulo 2, § 4.3).

Otro de los factores imprescindibles para el estudio de la lengua en diacronía es el tipo de texto, ya que es uno de los condicionantes externos más accesibles para los historiadores de la lengua. Sin embargo, los acercamientos teóricos a la dimensión textual son diversos y es habitual encontrarse con cierta mezcla de términos e imprecisiones. En las próximas páginas dedicaré una breve sección a los tipos de texto, géneros y tradiciones discursivas (§ 4).

Para terminar, es necesario decir que, aunque desde el punto de vista teórico los factores internos y externos se presenten por separado, esta separación no siempre está clara y, además, en muchas ocasiones estos factores pueden interactuar. En el ámbito de la sociolingüística histórica, es imprescindible tener en cuenta ambos tipos de factores para poder describir de forma completa cualquier cambio lingüístico.

3. La dialectología en el estudio de la historia de la lengua española

Podemos entender la dialectología como hermana de la sociolingüística, en tanto que «ambas disciplinas estudian la lengua hablada y establecen las relaciones que se dan entre ciertos rasgos lingüísticos y ciertos grupos de individuos» (Silva-Corvalán y Enrique-Arias 2017:10). La variable geográfica es uno de los factores externos que contempla la sociolingüística, que en ocasiones ha recibido el nombre de dialectología social o dialectología urbana. Precisamente, lo que diferencia las dos disciplinas es el tipo de lugar que ocupa su atención. La sociolingüística se fija en los centros urbanos, caracterizados

por su heterogeneidad lingüística y social, para «establecer las fronteras sociales de ciertos usos lingüísticos» (2017:12). En cambio, la dialectología habitualmente se centra en el sector rural con el objetivo de localizar las formas vernáculas y tradicionales.

La dialectología, por tanto, estudia la variación lingüística de tipo geográfico. El acercamiento a esta disciplina puede ser en sincronía o en diacronía, pero ambas perspectivas están necesariamente relacionadas. Al fin y al cabo, la variación dialectal que podemos encontrar en la actualidad es un reflejo de la historia lingüística de un territorio: «Las áreas lingüísticas, debidamente combinadas con la historia conocida de estructuras y palabras, pueden indicar con claridad los focos de creación y de difusión de las innovaciones lingüísticas» (Fernández-Ordóñez 2011:24).

Desde que Ramón Menéndez Pidal asentó los cimientos de la filología hispánica el siglo pasado, hay dos ideas fundamentales que están en la base de la interpretación de la historia de la lengua española. Ambas relacionan la evolución del castellano con las variedades iberorromances vecinas, pero de forma diferente. Tradicionalmente, se ha destacado la idea del papel predominante del castellano como origen y base en la formación del español. De este modo, se entiende que durante la Reconquista el castellano se extendió por el territorio peninsular a expensas de las variedades laterales: el asturleonés y el navarroaragonés. La otra idea, que normalmente ha pasado más desapercibida pero que recientemente ha recibido más atención, es la que entiende la formación del español a partir de una base plural, como la evolución conjunta de las tres variedades romances del centro peninsular: el castellano, el asturleonés y el navarroaragonés (Fernández-Ordóñez 2011:17).

La recuperación de esta segunda idea se enmarca en la revitalización de los estudios de variación geográfica en español. Fernández-Ordóñez (2011) traza algunos patrones geolingüísticos⁷ que contribuyen a explicar la conformación del español a partir de esta base plural y subraya la necesidad de incorporar la perspectiva diatópica al estudio de la historia de la lengua para poder reconstruir el foco geográfico de las innovaciones lingüísticas y sus etapas de difusión en el espacio.

En esta investigación me propongo trazar el recorrido histórico y geográfico de algunas variantes de los pronombres personales tónicos del castellano medieval, con el objetivo de demostrar la necesidad de incorporar la perspectiva geográfica a los estudios de historia de la lengua para conocer en profundidad los patrones geolingüísticos que intervienen en la conformación del español. Por tanto, el factor geográfico es uno de los factores del cambio lingüístico más relevantes para esta tesis.

En esta sección repasaré rápidamente el panorama lingüístico de la Iberia medieval, con atención al papel del contacto de lenguas y variedades (§ 3.1). Asimismo, presentaré brevemente algunos modelos de difusión espacial (§ 3.2).

3.1. Panorama lingüístico y contacto de lenguas en la Iberia medieval

Dos o más lenguas conviven en el mismo espacio geográfico y en la misma comunidad por circunstancias históricas como conquistas, colonizaciones o procesos migratorios. El contacto de lenguas y el bilingüismo que se producen en estas situaciones pueden provocar el cambio lingüístico, es decir, el desarrollo de innovaciones y su posterior

⁷ «En su amplio y documentado trabajo sobre la cuestión, Fernández-Ordóñez (2011) señala varios elementos morfosintácticos del español que son el resultado de innovaciones adoptadas de las variedades laterales del iberorromance (2011: 70 y ss.): la desaparición de *y* locativo y *ende*, y el desarrollo de la forma *alguien* y del plural *quienes* son elementos que el castellano adopta de las variedades occidentales; al mismo tiempo, la colocación de los clíticos abriendo cláusula, la gramaticalización de los tiempos compuestos y la creación y generalización de las formas *vosotros* y *nosotros* serían desarrollos morfosintácticos de procedencia oriental» (Enrique-Arias et al. En evaluación).

difusión y adopción; o la aceleración de cambios que ya están en marcha (Silva-Corvalán 1994:7, 119, 130; Silva-Corvalán y Enrique-Arias 2017:299; Trudgill 2011:15).

El contacto de lenguas como favorecedor del cambio ha sido estudiado en profundidad. Sin embargo, lo que ha recibido menos atención es la situación contraria: la inhibición del cambio. En algunas ocasiones, el contacto de lenguas puede favorecer el mantenimiento de rasgos que están retrocediendo en variedades que no se ven afectadas por ese contacto. Esto es lo que ocurre, por ejemplo, en zonas bilingües catalán-castellano. La presencia del catalán puede hacer que el castellano de zonas bilingües conserve determinadas formas o estructuras debido a la presencia de un paralelo en aquella lengua, mientras que en zonas monolingües estas están en retroceso o han desaparecido. Esto es lo que ocurre según Enrique-Arias (2010c, 2014, 2019b) en algunos fenómenos del castellano de Mallorca como la ralentización del yeísmo, la retención del futuro morfológico frente al avance del futuro perifrástico, el mantenimiento de la perífrasis de obligación *haber de* con infinitivo y otros. Del mismo modo, muchos de estos fenómenos son comunes con otras zonas con bilingüismo catalán-castellano (Blas Arroyo 2004, 2007).

En cualquier caso, tanto el cambio como la inhibición de este tienen consecuencias directas en la formación de las variedades. Lo que es relevante para esta tesis es ver qué procesos tienen lugar cuando dos o más lenguas o dialectos comparten espacio geográfico por un tiempo prolongado, es decir, qué efectos tiene el contacto de lenguas en la evolución de una lengua; y qué factores son los que favorecen el cambio inducido por contacto.

En esta tesis me voy a centrar en la variación geográfica del norte de la península ibérica, en el *continuum* dialectal septentrional, que se inserta a su vez en el *continuum* romance más amplio que se extiende por gran parte de Europa (Penny 2004:17). Este concepto se basa en la idea de que podemos recorrer una gran superficie de territorio sin encontrar diferencias lingüísticas abruptas:

El habla de cada localidad difiere en alguno o en varios de sus rasgos de la de cada localidad vecina, pero sin perjudicar gravemente la comprensión mutua. A medida que se atraviesa un área geográfica se van acumulando pequeñas diferencias sucesivas [...]. Cuanto mayor sea la distancia que se viaje, mayor es el número total de diferencias entre el habla de la localidad en la que se esté y la del punto de partida, y tal acumulación de diferencias causa su correspondiente grado creciente de intercomprensión mutua, hasta el punto de que el habla de un pescador gallego apenas será comprendida, si es que no es entendida en absoluto, por un pescador de la costa de Cataluña (Penny 2004:16-17)

La configuración del panorama lingüístico peninsular está determinada por la existencia del *continuum* dialectal septentrional, así como por los movimientos de población que tuvieron lugar durante la conquista y posterior repoblación de las tierras de al-Andalus. El avance progresivo de la conquista formó franjas dispuestas de norte a sur de acuerdo con el movimiento de la población. Estos ejes verticales separaron el gallego-portugués y el catalán de la zona central, así como las variedades romances centrales: el asturleonés, el castellano y el navarroaragonés (Fernández-Ordóñez 2011:25).

Tal y como hemos mencionado más arriba, la convivencia de todas estas variedades en el espacio geográfico peninsular provocó el trasvase constante de formas y estructuras lingüísticas en distintas direcciones. Habitualmente se destaca la difusión del castellano a expensas de las variedades iberorromances laterales, el asturleonés y el navarroaragonés. No obstante, la influencia lingüística no fue solo del castellano hacia las variedades vecinas, sino que estas también contribuyeron a lo que hoy conocemos como español. Por tanto, podemos entender el español como una *koiné* resultado de la evolución conjunta de castellano, navarroaragonés y leonés: «la variedad que conocemos

como español estándar ha surgido de una serie de mezclas dialectales, y ha sufrido repetidamente nuevas formaciones dialectales o *koineizaciones*, desde al menos el siglo IX» (Penny 2004:87).

Una de las consecuencias principales del contacto de lenguas y dialectos mutuamente inteligibles como el que se dio en la Iberia medieval es la acomodación. Este proceso tiene lugar cuando los hablantes adaptan su habla a la de la persona con la que conversa. Durante la acomodación un hablante puede eliminar rasgos que identifica como particulares de su variedad a favor de formas y estructuras más generales. Así, la acomodación puede dar lugar a que un rasgo que aparece debido a la adaptación provisional de un hablante en una situación de contacto puede propagarse a través de toda una comunidad de habla (Penny 2004:74-75).

Otros mecanismos que intervienen en el contacto prolongado entre diferentes variedades son la nivelación de diferencias lingüísticas o la simplificación de los sistemas. Esto lleva a que se conserven un menor número de variantes y a que las soluciones que prevalezcan sean las más simples aunque no sean mayoritarias (Penny 2004:78, 87). Trudgill (2011:33) defiende que los procesos de simplificación tienen lugar más frecuentemente en las situaciones de contacto en las que participan hablantes adultos. En cambio, en situaciones de contacto a largo plazo, en periodos de hasta cientos de años, se produce la adquisición bilingüe en la infancia y tienen lugar procesos de complejificación.

Los factores sociales que determinan la convergencia de rasgos lingüísticos son, por ejemplo, la duración y la intensidad del contacto, así como el tamaño y la densidad de la población (Trudgill 2011). Si el contacto es intenso y se produce a través de diferentes generaciones, será esperable que se produzcan fenómenos lingüísticos fruto del contacto de lenguas. También es relevante el estatus de la lengua para los hablantes: normalmente, será la variedad mayoritaria la que influya sobre la minoritaria. Esto ocurre porque es más probable que los hablantes de la lengua minoritaria se vuelvan bilingües y adopten rasgos de la otra lengua (Thomason 2001:66). Esto está estrechamente relacionado con la actitud respecto a la otra lengua: si una variedad es prestigiosa, es más probable que se transfieran rasgos; mientras que, si la actitud es negativa, esta será un impedimento para la transferencia.

La velocidad de los cambios en situaciones de contacto también depende del nivel lingüístico que se vea afectado: los elementos fonéticos y léxicos se prestan con mayor facilidad que los gramaticales. Precisamente, se considera que los pronombres personales son uno de los elementos más difíciles de transferir. Forman parte de una clase cerrada a medio camino entre lo léxico y lo gramatical y constituyen una pieza esencial de cualquier sistema lingüístico (Thomason y Everett 2001). Sin embargo, esto no significa que sea imposible que un pronombre pase de una lengua a otra. Uno de los motivos por los que se puede tomar un pronombre de otra lengua es porque se perciba un hueco en el paradigma pronominal o haya una necesidad de reestructuración (2001:305).

Finalmente, también es importante tener en cuenta la cercanía tipológica entre las lenguas de contacto. Cuanto más cercanas sean, más probable es que ocurran fenómenos lingüísticos provocados por el contacto de lenguas. Este es el caso de las lenguas y variedades iberorromances, que, a excepción del Euskera, son muy cercanas tipológicamente; y, en concreto, del *continuum* dialectal septentrional que va desde el oeste con el gallego al este con el catalán.

3.2. Modelos de difusión espacial

Los dos modelos de difusión espacial más frecuentes son el modelo de ondas y el modelo jerárquico o en cascada.

Según el **modelo de ondas**, las innovaciones se extienden de forma gradual por el espacio geográfico en forma de ondas concéntricas. Para ilustrar este modelo se suele utilizar la analogía de las ondas que forma un guijarro al caer en un estanque. El punto en el que la piedra entra en contacto con el agua es el foco del cambio a partir del que se difunde la innovación en todas las direcciones (Campbell 1999:187; Chambers y Trudgill 2004:166). Por tanto, según esta teoría, las innovaciones se difundirían de forma continua a través del espacio. Además, el uso de la forma innovadora estaría más arraigado en el foco del cambio, de modo que tendría una frecuencia de uso más alta y aparecería en más contextos. En consecuencia, a medida que nos alejamos del centro del cambio se irían reduciendo la frecuencia y los contextos en los que aparece. Por tanto, «podemos asociar mayor frecuencia y mayor número de contextos de un cambio determinado a una mayor antigüedad en un territorio dado e ir desbrozando el camino seguido por la difusión siguiendo en sentido inverso su recorrido espacial» (De Benito Moreno 2015:76-77).

Según el **modelo jerárquico o en cascada**, las innovaciones se difunden desde los núcleos urbanos. En lugar de basarse únicamente en la proximidad geográfica como el modelo de ondas, el modelo jerárquico incorpora el factor social. Así, las formas nuevas se propagan de un centro urbano a otro a través de las vías de comunicación, de modo que pueden tardar en llegar a poblaciones más pequeñas que, aunque pueden estar cerca del foco del cambio, están peor comunicadas (Hernández-Campoy y Conde-Silvestre 2012:455).

Ambas teorías parten de la existencia de áreas focales, desde donde se difunden las innovaciones; áreas residuales, donde tarda más en llegar el cambio (si llega) y, por tanto, se conservan las formas anteriores; y las áreas de transición, donde se alternan ambas variantes (Campbell 1999:192).

Según Chambers y Trudgill (2004:166), estos dos modelos pueden combinarse en uno solo. Para ello, modifican la analogía de la piedra en el estanque y matizan que, en lugar de caer, la piedra es lanzada y salta por la superficie del agua. Así, se crean distintos puntos focales, normalmente ciudades, desde donde el cambio se propaga de forma continua en el espacio geográfico.

A esto podríamos añadir que las piedras no solo se lanzan en una dirección. Es decir, la presión lingüística puede irradiar desde varios centros de prestigio. Esto puede provocar que en ocasiones la continuidad en la difusión de una variante en el espacio geográfico puede verse cortada coincidiendo con fronteras políticas, que marcan el límite de la influencia de un modelo de prestigio. De este modo, habría lugares en los que se podría distinguir claramente la transición entre variedades diferentes:

la extensión de rasgos a modo de ondas desde un centro de prestigio concreto puede detenerse en una frontera política, más allá de la cual el centro de prestigio no ofrece ningún influjo, ya que los que viven al otro lado de la frontera están sujetos a presiones lingüísticas procedentes de otra dirección (Penny 2004:57).

Además, la dirección y el área de la influencia del prestigio pueden variar con el paso del tiempo. Un ejemplo claro de esto es lo que sucede con las formas reforzadas de los demostrativos tipo *aqueste* en Aragón, que se extienden hasta ser variantes categóricas y posteriormente retroceden hasta desaparecer. Esto se explica por la fluctuación del modelo de prestigio para Aragón. Hasta el siglo XV, Aragón se ve influida por el modelo lingüístico catalán que irradiaba desde la Cancillería Real, donde las formas largas de los demostrativos se generalizaron muy tempranamente. Posteriormente, el origen del modelo de prestigio fluctuó al castellano, lo que provocó un retroceso de los demostrativos reforzados, que ya no estaban presentes en esta variedad (Enrique-Arias et al. En evaluación).

En mis datos, puede verse cómo los pronombres *nosotros*, *vosotros* entran en castellano desde el oriente peninsular, desde el catalán. En el Mapa 2 (Capítulo 3, § 4.5) puede verse con más detalle esta difusión. No obstante, el porcentaje de aparición de las formas innovadoras no aumenta de forma totalmente gradual y continua de un área a otra, sino que hay cierta diferencia abrupta entre el área catalana y las colindantes. Esto hace pensar que la frontera política frena en cierta medida el avance de las variantes reforzadas con ALTEROS hacia el este, pero estas terminan entrando igualmente. Las variantes innovadoras arraigan primero en Navarra y Aragón, zonas que se ven influidas por el prestigio del modelo lingüístico catalán, y de ahí pasan al resto de la península. Es decir, la frontera política no llega a romper el *continuum*, pero sí puede funcionar como un tamiz que ralentiza momentáneamente la difusión de las variantes.

En definitiva, en los cambios lingüísticos que se difunden a través del espacio influyen tanto la proximidad geográfica como la proximidad mediante el comercio o las vías de comunicación, así como los lazos entre comunidades producidos por la dirección de los modelos lingüísticos de prestigio.

4. Tipos de texto, géneros y tradiciones discursivas

Otro de los factores imprescindibles para el estudio de la lengua en diacronía es el tipo de texto, ya que es uno de los condicionantes externos más accesibles para los historiadores de la lengua. Sin embargo, los acercamientos teóricos a la dimensión textual son diversos y es habitual encontrarse con cierta mezcla de términos e imprecisiones.

Es innegable que las innovaciones pueden verse favorecidas por ciertos géneros textuales, aunque no hay acuerdo en cómo se produce exactamente esta influencia. Según Company (2008:38), las diferencias entre tipos de texto se reflejan en las frecuencias de uso, no en la presencia o ausencia de las formas innovadoras. Es decir, según esta autora, el factor textual no determina el foco de un cambio lingüístico, pero sí puede marcar las vías de difusión de las innovaciones. En cambio, según Kabatek (2018:140), las innovaciones sí pueden originarse en una tradición discursiva concreta: «En algunos casos, se observa cómo los elementos gramaticales son generados en tradiciones determinadas, van de unas a otras, se expanden o quedan restringidos, se fosilizan o mueren».

Durante los últimos años ha ido tomando relevancia la noción de tradiciones discursivas (TD). Desde su aparición, es habitual la mezcla de este término con otros que hacen referencia al factor textual, como *tipo de texto*, *género* o similares. Sin embargo, es necesario señalar sus diferencias. Kabatek (2018:170) define las TD como «la repetición de un texto o de una forma textual o de una manera particular de escribir o de hablar que adquiere valor de signo propio (por lo tanto, es significable)». Además, especifica que «no se trata de un sinónimo de «tipo textual», «género», etc., sino de un concepto más amplio que incluye todo tipo de tradiciones del hablar identificables, también subgéneros o tradiciones dentro del mismo género» (2018:189).

La clasificación en TD se extrae a partir del conocimiento exhaustivo de los textos, de sus características lingüísticas, y, especialmente, de la relación que establecen con otros textos (Kabatek 2018:194). Es decir, son categorías establecidas *a posteriori*, tras el análisis de los textos y, más que una herramienta, podríamos considerarlo un objeto de estudio en sí mismo. En cambio, otras distinciones que hacen referencia al factor textual son categorizaciones hechas *a priori* a partir de la forma o el contenido. Algunos ejemplos son «textos literarios», «textos documentales», «textos científicos», «poéticos» o «jurídicos» (2018:178). Para Kabatek (2018:178-81) este tipo de categorizaciones son insuficientes, ya que dentro de un mismo género pueden confluír distintas tradiciones discursivas.

Debido al uso impreciso del término *tradición discursiva* en numerosos estudios, se ha propuesto uno alternativo, el de *tradicionalidad discursiva* (López Serena 2021; Octavio de Toledo y Huerta 2018a), que matizaría la definición expuesta unos párrafos más arriba. Este término haría referencia específicamente a la «propiedad asociada a los fenómenos, que estos van adquiriendo y perfilando conforme aumenta su repetición en determinadas constelaciones textuales, y que, por lo demás, no tiene en absoluto por qué ser lineal ni constante a lo largo del tiempo» (2018a:121). Un ejemplo de la aplicación del concepto de *tradicionalidad discursiva* es el estudio Octavio de Toledo y Huerta (2018a), en el que extrae un mapa de *tradicionalidad*⁸ a partir del análisis afinado de la distribución textual de las construcciones modales con infinitivo antepuesto en una variedad de textos.

De este modo, la *tradicionalidad discursiva*, entendida como factor interno del cambio lingüístico, sería una herramienta útil para describir con precisión las vías de difusión de una variante que se establezca en unos textos y se traspase a otros similares por su estructura o su temática.

El corpus de esta tesis se divide en dos amplias tipologías textuales: textos literarios y documentales (Capítulo 2, § 5). En los textos literarios he priorizado los textos en prosa y he evitado en la medida de lo posible los textos en verso, ya que pueden plantear problemas para distinguir cuándo las soluciones lingüísticas están determinadas por la métrica y/o la rima. En el corpus de textos documentales incluyo cartas de compraventa, cartas privadas, actas y declaraciones. La fuente principal para este tipo de textos es el *Corpus de documentos españoles anteriores a 1800* (CODEA+), que distingue entre ámbitos de emisión, que, como ya sabemos, tienen una relación directa con el nivel de formalidad; de mayor a menor: cancilleresco, judicial, municipal, eclesiástico y particular (Sánchez-Prieto Borja 2012:11).

La identificación de *tradicionalidades discursivas* a partir de los textos que forman el corpus queda fuera de los objetivos de este estudio. No obstante, durante el análisis es posible que se señalen algunas tendencias relacionadas con el factor textual.

5. La teoría de la gramaticalización

La gramaticalización es un concepto teórico fundamental para entender la evolución de muchos fenómenos de la historia de la lengua. Esta importancia queda reflejada en la gran cantidad de bibliografía dedicada a profundizar sobre la definición, las causas, los mecanismos y las consecuencias de la gramaticalización.

La teoría de la gramaticalización tiene como principal objetivo la descripción del origen y el desarrollo de las formas gramaticales. Estas formas gramaticales pueden surgir a partir de piezas léxicas o construcciones que en ciertos contextos desarrollan funciones gramaticales. Por tanto, la gramaticalización suele definirse como el proceso en el que un elemento léxico se convierte en gramatical o un elemento gramatical se vuelve todavía más gramatical (Company 2003:9; Garachana 1997:23; Heine 2003:575; Heine y Song 2011:590; Hopper y Traugott 2003; Kuteva et al. 2019:3). En el caso de los pronombres personales tónicos y sus refuerzos analizados en esta tesis estaremos, por tanto, ante fenómenos que se insertan en la segunda parte de la definición, ya que partimos de elementos que son gramaticales desde el inicio.

Así, una palabra o estructura pierde o ve debilitado su significado etimológico y pasa a tener un significado gramatical más abstracto. Es decir, es un proceso que tiene como resultado la creación de «nuevas formas o nuevos mecanismos de expresión en la

⁸ Cf. Octavio de Toledo y Huerta (2018a:116).

gramática» (Company 2008:18). Este cambio se caracteriza como un proceso largo, gradual, unidireccional e irreversible hacia la abstracción (Company 2003:9).

Como cualquier proceso de cambio lingüístico, la gramaticalización es intrínsecamente diacrónica. Su marco teórico permite describir el recorrido histórico de una forma o construcción a medida que se va volviendo más gramatical, así como establecer cuáles son los caminos posibles en el futuro (Bybee, Perkins, y Pagliuca 1994:3; Heine 2003:575, 598; Hopper y Traugott 2003:2).

Meillet acuña el término de gramaticalización (*grammaticalisation*) a principios del siglo XX y lo define como «l'attribution du caractère grammatical à un mot jadis autonome» (1912:131). Su objetivo es explicar la historia de la gramática, los antecedentes de la estructura en una lengua, a través de la historia de las formas gramaticales específicas (Hopper 1991:17). Sin embargo, la idea que subyace bajo al término es más antigua y se remonta al siglo XVIII (Garachana 1997:24 y ss.; Lehmann 2015:1).

En los años setenta, con la revitalización de los estudios sobre tipología y universales lingüísticos, se reaviva el interés por la gramaticalización. Su marco teórico permite explicar patrones evolutivos en distintas lenguas que se corresponden con procesos comunes de gramaticalización y, por tanto, dar cuenta de cambios en la tipología de las lenguas (Bybee et al. 1994:14; Garachana 1997:37). A su vez, la aplicación de la teoría de la gramaticalización a los estudios sincrónicos obliga a poner los problemas lingüísticos en un marco diacrónico.

5.1. La gramaticalización como macrocambio

El proceso de gramaticalización incluye toda una cadena de cambios lingüísticos que afectan a una pieza léxica o construcción en distintos niveles a medida que se va volviendo más gramatical:

[...] grammaticalization is usually thought of as that subset of linguistic changes whereby a lexical item or construction in certain uses takes on grammatical characteristics, or through which a grammatical item becomes more grammatical (Hopper y Traugott 2003:2)

[...] there is no one simple mechanism of change that produces grammatical meaning, but rather that there are several mechanisms or types of change. These different mechanisms that lead to semantic change and eventually grammatical meaning may be associated with different points along grammaticization paths and thus with different semantic substance (Bybee et al. 1994:24).

Grammaticalization is a process leading from lexemes to grammatical formatives. A number of semantic, syntactic and phonological processes interact in the grammaticalization of morphemes and of whole constructions. A sign is grammaticalized to the extent that it is devoid of concrete lexical meaning and takes part in obligatory grammatical rules» (Lehmann 2015:IX).

Algunos de los subprocesos que intervienen en la gramaticalización son la extensión, la generalización del significado, la decategorialización y la erosión fonética⁹.

⁹ Se han establecido distintas listas de procesos que se consideran requisitos para identificar un proceso de gramaticalización. Algunas son muy restrictivas, de modo que algunos fenómenos que normalmente se han identificado como procesos de gramaticalización se quedan fuera; y otras demasiado generales, de modo que es difícil identificar qué no es gramaticalización (Heine y Song 2011:591). Pueden encontrarse otros catálogos de parámetros en Hopper (1991) y en Lehmann (2015).

La **extensión** o **generalización contextual** consiste en el avance de la expresión lingüística a nuevos contextos y en la difusión dentro de la comunidad de habla (Heine y Song 2011:592).

La **generalización del significado**, que también ha recibido el nombre de desemantización o *semantic bleaching*, consiste en la pérdida o desgaste de los rasgos semánticos originales y en su generalización. Se pasa de un significado léxico a uno gramatical, más abstracto y, por tanto, aplicable a más contextos (Bybee 2003:605; Heine y Song 2011:592).

La **deategorialización** implica la pérdida de propiedades morfosintácticas propias de la forma original, habitualmente léxica, como los rasgos de flexión o de palabra independiente (clitización, afijación) (Heine 2003:579; Heine, Claudi, y Hünemeyer 1991:213). Dentro de este parámetro también se incluyen la **obligatorificación**, de modo que una forma que en principio es opcional se vuelve obligatoria; la **coalescencia**, el colapso de dos formas adyacentes; y la **fijación**, el establecimiento de un orden fijo en la estructura (Lehmann 2015:132).

Finalmente, otro de los parámetros que más se suele mencionar es el de **erosión** o **reducción fonética**, que se basa en la pérdida de sustancia fónica. Suele producirse en último lugar e, incluso, no se considera requisito indispensable para que se produzca la gramaticalización (Heine y Song 2011:593).

Cada uno de estos parámetros afecta a un nivel distinto de la lengua: pragmática, semántica, morfosintaxis y fonética respectivamente. Excepto el primero, el resto implican una pérdida de propiedades (Heine 2003:579; Heine y Song 2011:591; Kuteva et al. 2019:3). Todos ellos se interrelacionan en mayor o menor medida en distintos procesos de gramaticalización. Por ejemplo, se ha señalado frecuentemente el paralelismo entre el desgaste semántico y fonético (Lehmann 2015:134). A su vez, el desgaste semántico hace que la palabra o estructura pueda aplicarse a más contextos, lo que aumenta su frecuencia de uso y puede tener como consecuencia un mayor desgaste fonético o la coalescencia con formas adyacentes (Bybee et al. 1994:6-8).

Si examinamos el reemplazo de los pronombres de primera y segunda persona plural *nós* y *vós* por las variantes reforzadas con el indefinido *nosotros* y *vosotros*, podemos identificar todos los subprocesos mencionados. En primer lugar, se produce la extensión de las formas compuestas a través de los contextos y de la comunidad de habla desde el siglo XIII con las primeras apariciones hasta el siglo XVI, cuando son las formas generales. El indefinido *otros* pierde el significado de alteridad y pasa a formar parte indivisible del pronombre personal. En último lugar, podemos encontrar ejemplos de la erosión fonética de las variantes reforzadas en atlas regionales como el Atlas Lingüístico y Etnográfico de Aragón, Navarra y Rioja (ALEANR: lámina 1851, mapa 1598), como *naltros*, *nantros* o *natros*. Además, el catalán y el occitano, que gramaticalizan los pronombres reforzados de forma más temprana registran todavía más variantes que muestran esta erosión, como *nosatres*, *natres*, *nantres*, *nosaltros*, *noltros* o *naltros* en catalán; y *nosaus*, *nautres*, *naltres*, *natres* o *nantres* en occitano (Nowikow 1994:288). En cambio, otros modificadores como *todos*, *mismos* o *ambos*, que también podían acompañar a los pronombres *nós* y *vós*, no pasaron por el proceso de gramaticalización, por lo que en todo momento mantienen su significado y su independencia respecto del pronombre.

Uno de los puntos débiles que se han señalado sobre la teoría de la gramaticalización es que estos mecanismos no son exclusivos de este proceso, sino que también intervienen en otros cambios lingüísticos (Hopper 1991:32). Al mismo tiempo, no siempre se consideran imprescindibles para que un cambio lingüístico se identifique como gramaticalización. Esto lleva a algunos autores a argumentar que la

gramaticalización no es un proceso con estatus independiente: «Cases of grammaticalization are explained adequately by the other mechanisms of linguistic change, and grammaticalization explains nothing by itself but must rely on these other mechanisms and kinds of linguistic change» (Campbell 2001:158).

Sin embargo, esto no le quita valor al concepto de gramaticalización, que resulta imprescindible para agrupar ciertos cambios con características similares, para identificar patrones de cambio lingüístico recurrentes y comunes en distintas lenguas que podrían perderse en el panorama general del cambio lingüístico:

The enthusiasm for grammaticalization has brought forth a range of examples (especially the *lexical > grammatical, less grammatical > more grammatical* sort) which is important for providing a broad and useful database and for focusing attention on a particular set of changes within the broader picture of linguistic change. Work in grammaticalization has given us a very rich assembly of examples we would be unaware of had it not been done (Campbell 2001:158).

5.2. Unidireccionalidad

Generalmente, se describe el proceso de gramaticalización como un proceso unidireccional que avanza de menos gramatical a más gramatical (Hopper y Traugott 2003:16). Salvo contadas excepciones, las formas gramaticales no se convierten en léxicas. Sin embargo, los cambios lingüísticos en los que una forma léxica pasa a ser gramatical son muy frecuentes y están bien documentados (2003:99).

The lexical items that become grammaticalized must first be semantically general and serve commonly needed discourse functions. They then become syntactically fixed (they become constructions), and may eventually amalgamate morphologically, say, as stem and affix. The basic assumption is that there is a relationship between two stages A and B, such that A occurs before B, but not viceversa. This is what is meant by unidirectionality (Hopper y Traugott 2003:100)

Este principio, que muchos consideran fundamental para definir la gramaticalización, no está exento de controversia. Para Campbell, la unidireccionalidad no sirve como criterio para identificar la gramaticalización, «since there are significant counterexamples to the unidirectionality claim» (2001:157). Según Hopper y Traugott (2003:99) y Kuteva *et al.* (2019:6), estos ejemplos son esporádicos frente a las abundantes evidencias que apoyan el principio en diferentes lenguas.

Lehmann analiza ejemplos que aparentemente van en contra del principio de unidireccionalidad, es decir, que ejemplificarían casos de degramaticalización, y termina descartándolos: «We may therefore conclude this discusión with the observation that no cogent examples of degrammaticalization have been found» (Lehmann 2015:21).

5.3. Economía y expresividad

En el cambio lingüístico y, en concreto, en el proceso de gramaticalización intervienen dos tendencias opuestas en constante competencia: la economía y la expresividad. Por un lado, encontramos muchos fenómenos, especialmente los de tipo fonético, que se deben a una tendencia a la simplificación, al desgaste, a la eliminación de sonidos, rasgos o palabras. Estos procesos forman parte de las estrategias de los hablantes que tienen el objetivo de facilitar la producción. Sin embargo, en otros muchos casos encontramos una tendencia contraria a la expresividad y al énfasis. Estas estrategias basadas en la redundancia se centran en el oyente y tienen como fin asegurar la comunicación. Entre estas dos tendencias o motivaciones se encuentra el uso lingüístico (Company 2003:12-13; Lehmann 2015:4).

Normalmente, se suele dar un lugar privilegiado a la tendencia a la simplificación y la economía en el cambio lingüístico. Sin embargo, en el fenómeno principal de esta tesis, el paso de *nós* y *vós* a *nosotros* y *vosotros*, encontramos un ejemplo de cambio antieconómico, un cambio en el que la tendencia a la transparencia y la expresividad es la que vence y rompe esa tensión dinámica. En cualquier caso, una vez que las formas *nosotros* y *vosotros* se han gramaticalizado, pueden verse afectadas por la erosión fonética, por lo que veríamos una tendencia a la economía. Esto es lo que sucede en distintas realizaciones sujetas a variación dialectal como los ejemplos ya citados *naltros*, *nantros* o *natros*, que se documentan en el ALEANR (lámina 1851, mapa 1598).

El encadenamiento de las tendencias a la economía y la expresividad produce frecuentemente cambios cíclicos. Uno de los ciclos más conocidos es el *Ciclo de Jespersen*, utilizado para explicar la evolución de la expresión de la negación en distintas lenguas. Este ciclo consiste en el debilitamiento y el refuerzo sucesivos de las partículas de negación, como ha ocurrido en francés: *je ne dis* > *je ne dis pas* > *je dis pas*.

Van Gelderen (2013) repasa algunos ciclos conocidos y hace especial énfasis en la existencia de distintos tipos de ciclos dependiendo del alcance del cambio dentro del sistema lingüístico. También destaca que, al contrario de lo que podamos pensar en un primer momento, los cambios cíclicos no son circulares, sino unidireccionales. Es decir, no se repiten de forma idéntica las estructuras de etapas anteriores. Por tanto, los ciclos pueden caracterizarse, más que como círculos, como espirales.

5.4. Frecuencia de uso y contextos

La frecuencia de uso de una variante es una herramienta metodológica imprescindible para el estudio empírico del cambio lingüístico. Además, como ya hemos visto más arriba, es uno de los parámetros concomitantes del proceso de gramaticalización (Hopper y Traugott 2003:126-27). Sin embargo, una mayor frecuencia de uso no es solo un resultado de la gramaticalización, sino también una fuerza activa que impulsa el proceso (Bybee 2003:602).

Se suele distinguir entre dos tipos de frecuencia: la frecuencia de tipo (*type frequency*) y la frecuencia de caso (*token frequency*). La frecuencia de tipo está relacionada con el grado de productividad de un elemento lingüístico o una construcción, mientras que la frecuencia de caso hace referencia al número de veces que aparece una variante particular en el discurso y en cómo cambia a lo largo del tiempo (Hopper y Traugott 2003:127).

La frecuencia de caso (*token frequency*) tiene efectos que a primera vista pueden parecer contradictorios. Un mayor uso contribuye a una mayor erosión fonética de las formas (*reduction effect*) y, al mismo tiempo, tiene como consecuencia una mayor retención de formas irregulares (*conservation effect*): «Forms that are isolated in a morphological paradigm will tend to conform to the paradigm unless they are especially frequent» (Hopper y Traugott 2003:128).

La alta frecuencia fija el uso, lo rutiniza, otorga apoyo paradigmático y crea estabilidad en el sistema [...], pero también los signos más frecuentes son los que están expuestos a mayor erosión, y son los que más pronto experimentan cambios, ya que la mucha frecuencia genera alomorfia y una morfofonémica compleja. Por su parte, la baja frecuencia suele conllevar equivocaciones en el uso de las formas y es un factor determinante para que esas formas sean sometidas a nivelaciones paradigmáticas (Comany 2003:43-44).

Por tanto, desde el punto de vista teórico, es importante tener en cuenta el concepto de frecuencia de uso, «puesto que incide en el diseño de las gramáticas y puesto que puede

predecir, en principio, la estabilidad o inestabilidad diacrónica de las formas conservadora e innovadora» (Company 2008:32).

Normalmente, los cambios de frecuencia en el cambio lingüístico son lentos y graduales. Del mismo modo, la gramaticalización avanza paulatinamente a través de contextos lingüísticos y extralingüísticos. Company (2003:33) explica el avance a través de estos contextos mediante una jerarquía: las formas innovadoras progresan desde los márgenes, más inestables, más susceptibles al cambio; hacia el centro, más estable, más resistente al cambio.

Las formas innovadoras aparecen inicialmente en contextos muy específicos de acuerdo con el valor que expresan. En las primeras apariciones, estos contextos son muy marcados y menos frecuentes. Sin embargo, con el tiempo, la forma innovadora puede ir invadiendo los ámbitos de uso de la forma antigua. De este modo, al menos en el caso de los cambios exitosos, el significado de la forma innovadora se generaliza y se va volviendo menos marcado hasta que finalmente pierde sus restricciones¹⁰. Así, «el avance de la gramaticalización supone una progresiva independización del contexto» (Company 2003:26).

Esto es precisamente lo que ocurre con las variantes reforzadas de los pronombres, *nosotros* y *vosotros*. Como explicaré en más detalle en la sección correspondiente, estas variantes surgen para expresar un contraste y para desambiguar el número del referente. Por este motivo, el modificador *otros* se extiende primero en la segunda persona, *vós*, que por definición marca un contraste con la primera persona del emisor del mensaje. Del mismo modo, el uso de *vós* con referente singular como forma de tratamiento es mayor que el de *nós* como plural mayestático, por lo que la necesidad de añadir una marca de plural es más frecuente.

5.5. Causas y mecanismos de gramaticalización¹¹

Las causas de la gramaticalización son múltiples y pueden tener su origen en todos los niveles de la lengua. Podemos destacar dos que muchas veces tienen lugar a la vez: la ambigüedad y la mala integración paradigmática (Company 2003:41).

Uno de los requisitos indispensables para que se produzcan reanálisis o reinterpretaciones y, por tanto, se lleve a cabo la gramaticalización, es que el signo sea ambiguo u opaco (Company 2003:42). Otra de las cuestiones que suele marcarse frecuentemente como causa del cambio lingüístico y de la gramaticalización es la mala integración paradigmática. Se suele hablar de que un elemento se integra mal en su paradigma cuando es formalmente distinto al resto de elementos que lo forman. Para regularizar o nivelar el paradigma puede surgir una variante que, por presión analógica, sea más parecida formalmente al resto y, por tanto, se integre mejor en el paradigma.

Tanto el argumento de la falta de transparencia como el de la unidad del paradigma se han utilizado para explicar el reemplazo de *connusco*, *convusco* por *con nós*, *con vós*. Las variantes etimológicas, tras los cambios sufridos desde las formas latinas NOBISCUM y VOBISCUM, no se integran bien en el paradigma pronominal y resultan extrañas para los hablantes: la posposición CUM ha pasado a tener formas diferentes en el singular y el plural, *-go* y *-co* respectivamente; se ha vuelto a añadir CUM como preposición; y los morfemas pronominales han quedado en interior de palabra. El resultado es una forma morfológicamente poco transparente, que formalmente no parece un plural y que se diferencia del resto de pronombres personales tónicos. Por todos estos motivos, por la

¹⁰ Cf. § 6.1 sobre la expresión de las relaciones de marca y la inversión de marca.

¹¹ Por supuesto, no son los únicos mecanismos que pueden intervenir en la gramaticalización. Para una descripción completa, cf. Rodríguez Molina (2010:84 y ss.).

presión del resto del paradigma y, especialmente, por las formas tónicas que se utilizan tras el resto de preposiciones, surgen *con nós* y *con vós*.

El reanálisis y la analogía son mecanismos relevantes para la descripción del cambio lingüístico en general, especialmente en el cambio morfosintáctico, y, además, tienen un papel clave en el proceso de gramaticalización¹².

El reanálisis se produce cuando el oyente interpreta la estructura de una palabra de forma distinta a como lo hace el hablante (Hopper y Traugott 2003:50). Así, «supone una manipulación conceptual y/o formal de las formas lingüísticas por parte del oyente» (Company 2003:45). Suele originarse por la falta de transparencia del signo lingüístico. De este modo, cuando una secuencia puede interpretarse según dos estructuras sintácticas distintas, se «intenta recuperar una relación biunívoca entre el valor del significado de una forma o expresión y su manifestación formal» (2003:45).

Las modificaciones producidas por el reanálisis afectan a la interpretación del significado, pero no a la forma, al menos no en un principio (Hopper y Traugott 2003:39). El reanálisis solo resulta evidente una vez que se produce una modificación estructural que lo revela, lo que se entiende como la actualización del reanálisis: «the reanalysis itself is covert until some recognizable modification in the forms reveals it» (2003:50).

En muchos casos, junto al reanálisis tiene lugar la analogía. Los hablantes, al detectar similitudes formales o semánticas entre distintas palabras, pueden alterarlas para que sean más parecidas. Esto suele suceder especialmente dentro de paradigmas, en los que la evolución de una o varias palabras no puede explicarse únicamente por su evolución fonética, sino que es necesario recurrir a la presión analógica. En palabras de Torrens Álvarez, «cuando la forma de una palabra hace que resulte anómala dentro de su clase, es probable que actúen sobre ella los mecanismos de la analogía para lograr que encaje mejor» (2007:81).

Reanalysis essentially involves linear, syntagmatic, often local, reorganization and rule change. It is not directly observable. On the other hand, analogy essentially involves paradigmatic organization, change in surface collocations, and in patterns of use. Analogy makes the unobservable changes of reanalysis observable (Hopper y Traugott 2003:68)

Un ejemplo habitual de reanálisis y analogía es el de los pronombres personales con la adposición *cum* pospuesta: *mecum*, *tecum* y *secum*. En las formas singulares, la oclusiva sonoriza al estar entre vocales. Esto, junto con la evolución habitual de *-um* en *-o*, da lugar a una terminación *-go* que no resulta transparente para los hablantes. Por otro lado, las formas plurales *noscum* y *voscum*, como ya hemos mencionado, presentan una evolución distinta en *-co*, ya que la oclusiva no se encuentra entre vocales. La opacidad de estas terminaciones, especialmente en las formas singulares, provoca que «se reinterpreté la adposición como parte de la base pronominal» (Company 2003:45) y se

¹² Hay cierta controversia en cuanto a la relación entre estos mecanismos y el proceso de gramaticalización: «it [grammaticalization] has been described by some as a process that necessarily involves reanalysis, while others insist that the two notions should be separated» (Heine 2003:592). Tanto Hopper y Traugott como Campbell son de la opinión de que hay una relación de inclusión, en el sentido de que la gramaticalización implica reanálisis, aunque no todos los casos de reanálisis implican gramaticalización. Otros autores como Heine y Haspelmath, en cambio, opinan que son dos fenómenos independientes. El distinto uso del término de reanálisis ha llevado a ciertos malentendidos, por lo que Heine (2003:593) propone evitar su utilización al hablar de gramaticalización. Algo similar ocurre con la analogía. Según Hopper y Traugott (2003:61-62), la analogía es un requisito indispensable para que haya gramaticalización, mientras que Heine (2003:593) opina que no es un concepto útil para entender este proceso: «analogy is a ubiquitous strategy that can be invoked for many different phenomena, to the extent that its use is sometimes not very helpful for describing and/or understanding the specifics of grammaticalization».

añada de nuevo la adposición (actualización del reanálisis), esta vez ante el pronombre: *conmigo*, *contigo*, *consigo*. Este cambio se extiende por analogía a las formas plurales *connusco*, *convusco*.

5.6. Gramaticalización secundaria

En el inicio de esta sección hemos definido la gramaticalización como el proceso en el que un elemento léxico se convierte en gramatical o un elemento gramatical se vuelve todavía más gramatical (Company 2003:9; Garachana 1997:23; Heine 2003:575; Heine y Song 2011:590; Hopper y Traugott 2003; Kuteva et al. 2019:3). Esta definición bipartita distingue entre dos posibles orígenes de la gramaticalización: un elemento léxico o uno ya gramatical. En ocasiones, a partir de esta distinción se han diferenciado dos tipos de gramaticalización, la primaria y la secundaria respectivamente.

Ha habido multitud de acercamientos a la gramaticalización secundaria, lo que ha llevado a diferentes definiciones. Generalmente, en su sentido más amplio, se utiliza la gramaticalización secundaria para identificar aquellos cambios lingüísticos que tienen un *input* gramatical que se vuelve todavía más gramatical (Breban 2014:492-93). Otros usos de los términos de gramaticalización primaria y secundaria entienden estos procesos como sucesivos. De este modo, la gramaticalización secundaria se entendería como una fase posterior dentro del proceso de gramaticalización consecuencia del aumento de frecuencia y la rutinización; o incluso como una segunda gramaticalización a partir de un elemento ya gramaticalizado (Breban 2014:493).

Un ejemplo de gramaticalización secundaria similar a los que se presentarán en esta tesis es el que identifica Octavio de Toledo y Huerta (2018b) en el caso de las variantes *esotro* y *estotro*, formadas por los demostrativos *este / ese* y el indefinido *otro*. Para argumentar que los demostrativos compuestos se corresponden con un proceso de gramaticalización secundaria, Octavio de Toledo y Huerta (2018b) hace referencia, en primer lugar, a la definición de este término en el sentido amplio. Así, los dos elementos de partida son de tipo funcional, gramatical. Además, hace especial énfasis en dos conceptos: por un lado, en la coalescencia avanzada que tiene lugar entre los componentes, que da lugar a una amalgama; por otro lado, en la aportación de un nuevo valor gramatical, de un nuevo contraste, al paradigma en el que se inserta.

La coalescencia de un indefinido de alteridad con un determinante o pronombre para formar una nueva pieza gramatical de carácter pronominal constituye, pues, una ruta evolutiva (un *canal* de gramaticalización: *cfr.* Lehmann 2002:22) no inhabitual en las lenguas del grupo y de la familia del español (Octavio de Toledo y Huerta 2018b:325)

En el caso de los procesos de gramaticalización objeto de estudio de esta investigación, todos ellos parten de elementos que ya son gramaticales, por lo que se trataría de un proceso de gramaticalización secundaria en un sentido amplio. Es decir, se trataría de procesos en los que un elemento gramatical se vuelve más gramatical.

En definitiva, la teoría de la gramaticalización y los subprocesos que incluye son imprescindibles para explicar los fenómenos de cambio lingüístico analizados en esta tesis, que tienen como resultado la aparición de nuevos elementos gramaticales. La generalización contextual y del significado, la obligatorificación, la coalescencia, la fijación, la erosión fonética, el ciclo de economía y expresividad, la frecuencia de uso, la mala integración paradigmática, la analogía y el reanálisis afectan en mayor o menor medida a cada uno de los procesos de cambio examinados.

6. Tipología y universales lingüísticos

La tipología lingüística se acerca al estudio de la lengua especialmente desde un enfoque comparativo y tiene como objetivo la extracción de universales. Entiende la variación y cambio como propiedades inherentes a todo sistema lingüístico y, a su vez, contempla la existencia de la arbitrariedad. Es decir, entiende que no todo se puede explicar a través de universales lingüísticos, sino que hay rasgos particulares de cada lengua.

A typologist uses an inductive method of analysis, by constructing a sample of the world's languages and seeking language universals via cross-linguistic generalizations. [...] it is only through exploring linguistic diversity that one is able to discover the limits to variation, that is, the universals of human language (Croft 2003:282).

Este enfoque, además de permitir desarrollar generalizaciones sobre la lengua objeto de estudio, implica un paso posterior: la evaluación de esa generalización en el contexto del lenguaje como un todo (2003:7).

Habitualmente, se identifican tres tipos de universales: los absolutos, los implicativos y las tendencias. Los **universales absolutos** se dan en todas las lenguas naturales y no tienen excepciones. Algunos ejemplos de universales absolutos son «todas las lenguas tienen vocales y consonantes» o «todas las lenguas distinguen nombres y verbos» (Ruiz Antón 1998:14). Los **universales implicativos** relacionan la aparición de una característica con otra, del tipo «si A entonces B». Es decir, una propiedad solo puede estar presente si otra ya lo está. El ejemplo que pone Comrie (1989:17) para este tipo es el siguiente: si una lengua tiene pronombres reflexivos de 1ª y 2ª persona, también los tiene para 3ª persona. Las **tendencias**, en cambio, sí tienen excepciones. Tienen una gran frecuencia, pero no son absolutas. Podemos encontrar tendencias no implicativas, como el hecho de que casi todas las lenguas tengan consonantes nasales (Comrie 1989:19); y tendencias implicativas, como «las lenguas SOV tienden a tener posposiciones» (Ruiz Antón 1998:14).

En principio, los conceptos de tipología y universales pueden parecer contrarios. Mientras que los universales buscan encontrar aquellas propiedades comunes a todas las lenguas, la tipología busca encontrar los hechos diferenciales. Sin embargo, ambos enfoques están interrelacionados y en muchos casos son inseparables el uno del otro (Comrie 1989:33-34).

Ambas disciplinas pueden ser útiles para el estudio de la lengua en diacronía. La tipología y los universales aportan herramientas imprescindibles para entender los mecanismos y las motivaciones del cambio sintáctico, como la expresión de las relaciones de marca o las jerarquías implicativas. Nos permiten dar explicaciones de tipo funcional sobre el cambio lingüístico y delimitar las innovaciones posibles, así como las direcciones posibles de los procesos de cambio.

En esta sección me centraré en repasar los rasgos tipológicos que afectan al paradigma de los pronombres personales tónicos, especialmente a las formas de plural.

6.1. La expresión de las relaciones de marca y el paradigma pronominal

Desde la aparición del concepto de marca en la Escuela de Praga, este se ha ido desarrollando y usando en diferentes enfoques. En un sentido amplio, se dice que una forma marcada tiene un significado más específico, aplicable únicamente a una serie de contextos restringidos; mientras que una forma no marcada tiene un significado más general, menos restringido. En tipología, se utiliza para hacer referencia a cómo se codifica la función gramatical en la forma (Croft 2003:87-88).

La oposición *marcado* frente a *no marcado* dentro de un paradigma puede reflejarse en la estructura de las variantes lingüísticas: cuando una variante es menos marcada que

otra, no puede tener una mayor realización superficial que la forma marcada (Croft 2003:88-89). Este principio suele ejemplificarse con la distinción de número. Tipológicamente, las formas singulares no marcadas y las plurales marcadas. En muchas lenguas, el singular no se expresa mediante un morfema explícito, mientras que el plural sí. Otras lenguas que sí usan un morfema explícito para el singular, también lo usan para el plural. Por tanto, esto se puede resumir con el siguiente universal implicacional: si una lengua usa un morfema explícito para el singular, entonces también usa uno para el plural (Croft 2003:89; Greenberg 1966a:32).

Por ejemplo, los pronombres personales de 1ª y 2ª persona plural *nós* y *vós*, que en origen son formas no marcadas, se oponen a *nosotros* y *vosotros*, que surgen con un valor enfático, marcado. De este modo, la forma marcada tiene más sustancia material que la forma no marcada.

Las relaciones de marca también se pueden manifestar en las diferencias flexivas dentro de un paradigma. De este modo, los elementos no marcados dentro de un paradigma generalmente tendrán una flexión más rica¹³. Por ejemplo, en el caso de los pronombres personales, el singular, categoría menos marcada que el plural, tiene más distinciones morfológicas de género. Esto se corresponde con el universal 37 de Greenberg (1966b:113), que indica que «a language never has more gender categories in nonsingular numbers than in the singular». Así, en los pronombres personales de 3ª p. tenemos tres distinciones para el singular (*él, ella, ello*) y dos para el plural (*ellos, ellas*). No obstante, en castellano podemos encontrar una excepción a esta regla, ya que la 1ª y la 2ª persona del plural sí distinguen género (*nosotros -as, vosotros -as*), mientras que en singular no hay ninguna distinción (*yo, tú*). Esto se debe probablemente a la evolución de estos pronombres personales tónicos. Las formas simples del castellano medieval *nós* y *vós* no distinguían género, pero al añadir el modificador *otros* se abre la posibilidad de diferenciar entre género masculino y femenino. En cualquier caso, habría que ver si en el futuro esta distinción se pierde y los pronombres terminan volviendo a adecuarse al esquema de la expresión de las relaciones de marca. Al fin y al cabo, las variantes correspondientes en catalán, *nosaltres* y *vosaltres*, no distinguen género.

Otra cuestión relevante sobre las relaciones de marca es la versatilidad o *distributional potential* (Croft 2003:95), que se refleja en la frecuencia de uso. Las formas no marcadas, al tener un significado más general, aplicable a más contextos sintácticos, son más frecuentes. Además, al ser más habituales para los hablantes, las categorías menos marcadas también son más resistentes a ciertos cambios de tipo morfológico (Elvira 1998:58).

El concepto de frecuencia de uso es especialmente relevante, ya que puede provocar cambios en las relaciones de marca. Cuando una forma marcada aumenta en frecuencia y se generaliza, pierde el valor marcado. Al mismo tiempo, la forma a la que reemplaza, en origen no marcada, ve reducidos sus contextos de uso y termina convirtiéndose en una forma marcada. Es decir, el cambio en la frecuencia de uso produce una *inversión de marca*.

6.2. Jerarquías o escalas implicativas y el paradigma pronominal

Como hemos comentado más arriba, los universales implicativos relacionan la aparición de una característica con otra. Es decir, una propiedad solo puede estar presente si otra ya lo está. Además, es posible encontrar varios universales implicativos relacionados entre sí, de modo que formen lo que se llama jerarquía o escala gramatical.

¹³ Esto es lo que Croft (2003:95) denomina como *inflectional potential* (§ 4.1.2 *Behavioral potential*).

A grammatical hierarchy is an ordered ranking of grammatical types which constrains possible language types (or at least the relative cross-linguistic frequency of actual language types) with respect to one or more grammatical phenomenon (Croft 2003:123)

Estas escalas juegan un papel fundamental en la explicación del avance o retroceso de determinados rasgos lingüísticos (Comrie 1989:188-89). Algunas de las más conocidas son la escala de animacidad y la de definitud, que en ocasiones se presentan en una sola escala combinada, la jerarquía de animacidad extendida (Croft 2003:130; Dixon 1979:85):

Pronombres de primera y segunda persona < pronombres de tercera persona < nombres propios < nombres comunes con referente humano < nombres comunes con referente animado no humano < nombres comunes con referente inanimado.

Hay multitud de rasgos tipológicos que se ven afectados por el parámetro de animacidad. Uno de ellos es la distinción de número. Si comparamos el uso de la marca de plural en distintas lenguas, veremos que esta puede extenderse más o menos por la jerarquía. Como mínimo, deberemos encontrar distinción de número en los pronombres de primera y segunda persona, y, si una lengua marca el plural en los nombres comunes con referente animado, también marcará todos los elementos previos en la escala.

Las escalas implicacionales que hacen referencia a los pronombres personales suelen agrupar la primera y la segunda persona en un solo escalón de la jerarquía. Sin embargo, uno de los temas que suele tenerse en cuenta en el estudio de *nosotros*, *vosotros* y el resto de las variantes con las que se relacionan es el orden de aparición. Por tanto, sería interesante determinar si existe algún tipo de relación jerárquica entre la primera y la segunda persona.

Según Croft (2003:161), podemos distinguir diferentes jerarquías en la categoría de persona dependiendo del parámetro que tengamos en cuenta: la animacidad, el número, el género, la cortesía o las formas verbales de imperativo. Así, se distinguen un total de cinco posibles patrones tipológicos:

- | | |
|-------------------------|-------------|
| a) Animacidad extendida | 1, 2 < 3 |
| b) Número | 1 < 2 < 3 |
| c) Género | 3 ? < 2 < 1 |
| d) Cortesía | 2 < 3 < 1 |
| e) Imperativo | 2 < 1, 3 |

Por tanto, parece que la percepción de qué persona es más o menos marcada depende mucho del contexto, de la saliencia perceptual o cultural.

Otra escala que puede ser relevante para los fenómenos que afectan al paradigma pronominal es la jerarquía de relaciones gramaticales o *GR hierarchy* (Croft 2003:142 y ss.): Sujeto < objeto directo < objeto indirecto < oblicuo¹⁴. Según esta escala, determinados fenómenos o rasgos sintácticos pueden estar sujetos a las relaciones que se establecen entre el verbo y sus argumentos. Un ejemplo es la distribución de las marcas de caso, que respalda esta jerarquía. De este modo, si una posición sintáctica de la jerarquía no recibe marca de caso (Croft 2003:146), los elementos superiores tampoco la recibirán. Asimismo, cuanto más a la derecha de la escala se sitúen las posiciones

¹⁴ La jerarquía que presenta Croft en realidad es la siguiente: *subject < (direct) object < oblique*, y más abajo indica que incluye el dativo dentro del oblicuo. En las páginas siguientes comenta la jerarquía de accesibilidad de la frase nominal (*NP accessibility hierarchy*) de Keenan y Comrie (1977) sobre la posibilidad de relativizar sintagmas nominales dependiendo del rol del participante, que se basa en la jerarquía de relaciones gramaticales: *subject < direct object < indirect object < oblique*. Esta es la que he presentado aquí.

sintácticas, más marcadas serán. Es decir, que la posición de sujeto es la menos marcada, y la de oblicuo la más marcada. A partir de esto, volviendo a cuestiones sobre las relaciones de marca, podemos deducir que el sujeto, como contexto sintáctico menos marcado, será la posición más resistente al cambio. En cambio, el contexto más marcado e inestable sería el oblicuo.

Los pronombres personales tónicos pueden aparecer principalmente en dos contextos sintácticos: en posición de sujeto y en término de preposición. De estas dos categorías, la menos marcada es la de sujeto. Por tanto, lo esperable sería que el cambio avanzara antes tras preposición. Así, es posible que los pronombres reforzados, con un valor más marcado, enfático, se manifiesten en primer lugar en posiciones más bajas de la jerarquía.

Podemos aplicar un razonamiento similar al hecho de que, ante la necesidad de desambiguar el referente, se marquen y refuercen los pronombres plurales y no los singulares. Las formas simples *–nós* o *vós–* pueden utilizarse tanto con referente singular y como plural y ya hemos mencionado más arriba que el plural es más marcado que el singular. Por tanto, si se debe reforzar morfológicamente uno de los dos, lo esperable es que se haga con el plural.

6.3. Distinciones gramaticales en el paradigma pronominal

a) Distinción de género

Tipológicamente, la distinción de género es característica de los pronombres personales de tercera persona antes que los de segunda o primera. De acuerdo con el universal 44 de Greenberg (1966b:70), «if a language has gender distinctions in the first person, it always has gender distinctions in the second or third person or in both».

A su vez, la distinción de género también es más propia del singular que del plural, tal y como expresa el universal 45 de Greenberg: «If there are any gender distinctions in the plural of the pronoun, there are some gender distinctions in the singular also» (1966b:70).

En el castellano estándar actual se distingue género en los pronombres personales de tercera persona del singular (*él, ella*) y en todos los del plural (*nosotros, nosotras, vosotros, vosotras, ellos, ellas*)¹⁵. En catalán solo se distingue género en la tercera persona a pesar de que también tiene los pronombres reforzados *nosaltres* y *vosaltres*. En cambio, otras lenguas como el inglés o el alemán solo hacen esta distinción en la tercera persona del singular, donde además incluyen una forma neutra (en inglés *he, she, it*; en alemán *er, sie, es*) (Ruiz Antón 1998:39).

b) Distinción de número

Los pronombres personales suelen tener distinción de número en la mayoría de las lenguas. Como hemos visto más arriba, los pronombres personales son los elementos que están más arriba en la escala de animacidad. Por este motivo, si una lengua no tiene la distinción de número en los pronombres personales, tampoco la tendrá en otros elementos (Daniel 2013).

Cysouw (2001:66-67) critica el concepto de pluralidad aplicado a los pronombres personales debido a que no sirve para explicar semánticamente las diferencias entre las formas *yo-nosotros, tú-vosotros*. Generalmente, los pronombres de plural no son simplemente varios ‘yo’ o varios ‘tú’, sino que su significado prototípico es asociativo;

¹⁵«Instances of gender in the first and second person non-singular are found in Europe (Spanish, Lithuanian, Slovene), New Guinea (Baniata, Dumu) and Australia (Anindilyakwa, Ndjébbana, Nunggubuyu)» (Siewierska 2013).

es decir, ‘yo y otros’ o ‘tú y otros’. Hablar de pronombres de grupo en lugar de plurales permitiría explicar mejor estas diferencias.

A causa de las diferencias semánticas entre los pronombres singulares y los plurales o de grupo, no es tan habitual encontrar sistemas pronominales que marquen el plural del mismo modo que los sustantivos:

In those few cases where a nominal marker is found in the pronominal paradigm, it is either found in part of the paradigm only, or it is functionally empty (ie the singular and plural forms are still different if the plural marker would be removed), or it is only optionally used (Cysouw 2001:67).

Según Daniel (2013), el español se caracteriza por tener pronombres personales con distintas raíces que distinguen persona y número, y, además, tener un afijo de plural específico (*-otros*). Asimismo, según este autor, el español es uno de los casos particulares en los que encontramos dos marcas de plural en un solo pronombre, que además están invertidas: *no-s-otros*, *vo-s-otros*. Sin embargo, es cuestionable que hoy en día la gente distinga esta /s/ en interior de palabra como una marca de plural.

En los paradigmas pronominales puede tener lugar lo que Cysouw (2001:181) denomina homofonía horizontal (*horizontal homophony*), que consiste en el solapamiento de las categorías singular y no singular de alguna de las tres personas, como es el caso de *you* en inglés. En castellano medieval ocurre algo similar, cuando *vós* y *nós*, pronombres de primera y segunda persona del plural, pueden utilizarse también como formas de tratamiento de cortesía en singular.

c) Distinción de inclusión o exclusión del oyente

Algunas lenguas tienen un pronombre específico para explicitar cuándo se incluye o se excluye la persona a la que se dirige el hablante. La distinción entre pronombres inclusivos y exclusivos es extraña para las lenguas de Europa, y el español no es ninguna excepción (Cysouw 2013). Sin embargo, creo que es necesario dedicarle una pequeña sección, ya que la distinción entre estos valores ha sido una de las justificaciones para la aparición de las formas reforzadas *nosotros* y *vosotros* durante la Edad Media.

Según Cysouw (2001:69), la distinción inclusivo-exclusivo concierne únicamente a la primera persona del plural. En ningún caso se hace referencia a la segunda o tercera. Esta distinción permite hacer referencia a alguno de los tres valores que en lenguas como el castellano o el inglés se agrupan tras los pronombres *nosotros* o *we* (2001:70-73):

- a) Inclusivo mínimo: la primera persona plural incluye al destinatario del mensaje y a nadie más (yo + tú)
- b) Inclusivo aumentado: la primera persona plural incluye al destinatario del mensaje y a otras personas (yo + tú + otros)
- c) Exclusivo: la primera persona del plural no incluye al destinatario del mensaje (yo + otro(s))

En numerosas ocasiones se ha explicado que, en el momento en que surgen las variantes reforzadas *nosotros* y *vosotros*, estas aportarían un valor exclusivo frente a *nós* y *vós*. Sin embargo, no considero que el uso de este término sea adecuado en este contexto, ya que esta distinción no es aplicable a la segunda persona del plural que, a la luz de los datos, parece ser la que lidera el cambio. Otros autores en cambio han planteado la aparición de las formas reforzadas no tanto con la idea de exclusividad, sino de énfasis y contraste, que me parece más adecuada.

7. Conclusión

En resumen, en este capítulo he repasado conceptos básicos y esenciales de la sociolingüística variacionista, la dialectología, las tradiciones discursivas, la teoría de la gramaticalización y la tipología y los universales lingüísticos. Todas estas perspectivas teóricas, junto con los avances metodológicos de la lingüística de corpus y el análisis de datos lingüísticos en diacronía han permitido grandes avances en el estudio del cambio lingüístico y constituyen el método idóneo para afrontar los problemas que me planteo en esta tesis.

Capítulo 2: Corpus y metodología

1. Presentación

Uno de los mayores inconvenientes para el estudio de la variación lingüística desde un punto de vista histórico es la dificultad para obtener muestras de habla real. La única fuente directa de la que disponemos para acceder a la lengua antigua es la documentación escrita que ha llegado hasta la actualidad, lo que conlleva claras limitaciones. Esta cuestión ha sido abordada en multitud de ocasiones (Dworkin 2011:156; Fernández-Ordóñez 2001:389-90; Kabatek 2013:9; Oesterreicher 2011:307; Pons Rodríguez 2006:74; Rodríguez Molina 2010:602; Rodríguez Molina y Octavio de Toledo y Huerta 2017b:5). En muchas de ellas se hace referencia a la famosa cita de Labov (1994:11), que describe la lingüística histórica «as the art of making the best use of bad data».

Parafraseando a Labov (1972, 1994): los historiadores de la lengua no tienen control sobre los textos de sus fuentes, que han llegado hasta nuestros días por accidente, por casualidad. Los datos históricos son por naturaleza fragmentarios y no podemos saber hasta qué punto reflejan el vernáculo de los hablantes.

El corpus de textos medievales que se han conservado no solo es inevitablemente finito, sino que también plantea otros problemas: el de la intermediación de la copia manuscrita y el control de la procedencia dialectal. En las dos últimas décadas se ha hecho especial énfasis en la compleja transmisión textual de muchas obras medievales que solían utilizarse como base para estudios de lingüística histórica y en la necesidad de distinguir entre manuscritos originales y copias.

Es necesario recordar aquí que, aunque cada testimonio manuscrito efectivamente constituye un texto, con sus peculiaridades y su personalidad, normalmente no conserva el «texto», es decir, lo que el autor escribió. En la transmisión de textos romances, los copistas suelen intervenir activamente modificando su modelo: la gama de intervenciones puede oscilar desde el grado mínimo, la reformulación gráfica y fonética; hasta el máximo, que implica la creación de una nueva obra a través de la refundición (Fernández-Ordóñez 2001:391).

Esta circunstancia provoca que la rigurosidad al seleccionar los textos que integran el corpus sea un pilar fundamental en cualquier investigación en lingüística histórica. Así, el investigador no puede limitarse a utilizar las ediciones más autorizadas de los textos que desea incluir en su corpus, sino que debe plantearse «el problema de su representatividad a la hora de establecer generalizaciones» (Fernández-Ordóñez 2001:397). Esto ha llevado en los últimos años a la necesidad de revisar recursos que se habían utilizado hasta el momento. Una de las muestras de esta revisión de recursos es el artículo de Rodríguez Molina y Octavio de Toledo y Huerta (2017b) con la herramienta del *Cordemáforo*. En este artículo, los autores examinan cada uno de los textos literarios medievales incluidos en el *Corpus diacrónico del español* (CORDE) y asignan un código de colores a dependiendo de su fiabilidad y representatividad como testimonio.

Otro de los pilares fundamentales de la confección de un corpus, en especial para los objetivos de esta tesis, es el control de la procedencia dialectal. Si no se tiene en cuenta el origen geográfico de un texto, se corre el peligro de generalizar rasgos lingüísticos propios de una zona como rasgos de una época. Además, tal y como ya he mencionado con anterioridad y como desarrollaré en las próximas páginas, en este estudio es indispensable el punto de vista diatópico. En palabras de Fernández-Ordóñez (2001:464) «la historia de la gramática necesita, pues, construirse sobre un corpus seguro en cuanto a la localización espacio-temporal de sus integrantes para poder evitar falsas generalizaciones».

En este capítulo mi objetivo principal es describir los criterios de selección y clasificación de los textos de mi corpus (§ 4), así como presentar las fuentes de las que los he extraído y concretar la nómina de obras y testimonios (§ 5). Sin embargo, antes de esto voy a desarrollar los factores que ya he presentado en esta introducción y que han determinado los principales obstáculos que me he encontrado para establecer el corpus de estudio: la distinción entre texto y testimonio, original y copia (§ 2), y el control de la procedencia dialectal de los textos (§ 3).

2. Textos y testimonios, originales y copias

Cualquier estudio basado en un corpus lingüístico de textos medievales debe tener en cuenta la compleja transmisión textual de algunas de las obras literarias de este periodo y los problemas de la intermediación de la copia manuscrita.

En primer lugar, debemos distinguir entre texto y testimonio. Cuando hablamos de ‘el texto’ frente al testimonio, nos referimos al texto concreto que escribió el autor, mientras que los testimonios son cada una de las copias conservadas. Cada uno de estos testimonios «aun cuando no representen fielmente el texto, se encuentran, debido a su materialidad, firmemente anclados en una realidad lingüística y temporal concreta» (Rodríguez Molina 2010:613). No obstante, los estudios lingüísticos monográficos han tendido a no distinguir texto y testimonio, de modo que se ha privilegiado ‘el texto’, «un concepto ideal, escurridizo por naturaleza» (2010:613), que no suele poder adscribirse a una fecha y un lugar exactos. Así, frecuentemente se han utilizado testimonios que transmiten copias tardías o ediciones críticas en lugar de manuscritos concretos. El peligro, además, está en generalizar los datos extraídos a la lengua del texto o a su periodo cronológico. De este modo, en las últimas décadas se ha insistido en que, para la investigación en lingüística, se debe priorizar el uso de testimonios frente al de textos.

Sin embargo, no todos los testimonios son igual de válidos. Durante el proceso de copia, además de introducir errores, el copista adapta en mayor o menor medida la lengua a su propia variedad e introduce modificaciones que pueden ir desde lo más superficial hasta el cambio de lengua en las traducciones (Fernández-Ordóñez 2012:51-52). Asimismo, es necesario destacar que el proceso de transmisión afecta en diferente medida a los distintos niveles lingüísticos. El nivel en el que suele haber más modificaciones es el gráfico-fonético, seguido por el morfosintáctico y, finalmente, el léxico (2012:58). Así, en cada copia se van superponiendo las variedades de cada uno de los copistas. En consecuencia, en la mayoría de las ocasiones no es posible distinguir los rasgos de los copistas de los del texto original, problema que se acentúa cuanto mayor es la distancia entre original y testimonio, y cuantas más copias se han intercalado entre ambos. En las copias tardías se van entrecruzando variedades de zonas geográficas y épocas distintas hasta que dejan de corresponderse a un estado real de lengua (Rodríguez Molina y Octavio de Toledo y Huerta 2017b:47).

Sabiendo esto, el corpus ideal de cualquier estudio lingüístico estaría basado en manuscritos de autor. Según Fernández-Ordóñez (2009), se entiende por manuscrito de autor no solo aquel borrador o copia autógrafa del autor, sino también los apógrafos, copias del texto que han sido autorizadas por el autor o por su entorno. En todos estos casos hay cercanía cronológica y una vinculación estrecha entre texto y copia (2009:92). Así, se debe evitar en la medida de lo posible la utilización de copias tardías, ya que esto puede falsear los resultados del estudio (Rodríguez Molina y Octavio de Toledo y Huerta 2017b:47). No obstante, esto no siempre es posible:

Si la supervivencia de manuscritos medievales es rara circunstancia, el que a nosotros hayan llegado códices que podamos juzgar de mano del autor, o supervisados o instigados

por él, o incluso acometidos en su entorno cercano, es portentosa fortuna que aún más raramente acontece (Fernández-Ordóñez 2009:117).

Fernández-Ordóñez (2006:1790-91) propone el medio siglo como límite convencional para considerar que una copia es contemporánea al original. Su propuesta se basa en el estudio de un conjunto de cambios en tiempo aparente de Chambers (2013). De sus resultados se extrae que el período mínimo de difusión del cambio lingüístico es de aproximadamente medio siglo¹. Por tanto, podemos considerar copias tardías aquellas que hayan sido compuestas más de cincuenta años después que el original.

El límite de 50 años es, claro, arbitrario, pero es preciso fijar el corte en algún punto y, dada la menor movilidad geográfica y social de los hombres medievales y sus menores posibilidades de interacción en diferentes entornos lingüísticos respecto de la época actual, parece sensato mantener esta cifra. No obstante, en cada caso particular hay que considerar la posible distancia entre el estado lingüístico que creemos original y el estado lingüístico que efectivamente recoge un códice en concreto [...] (Rodríguez Molina 2010:616).

Para el corpus de esta tesis he priorizado los testimonios originales o más próximos al original. En casos puntuales, he decidido incluir algunos testimonios que se consideran copias tardías. Un ejemplo es el ms. O del *Libro de Alexandre*, códice de principios del XIV compuesto más de medio siglo después de la fecha original de producción del texto (1240-1250). Otro caso es el ms. S del *Libro de buen amor*, copiado por Paradinas hacia 1415, mientras la fecha original de producción se sitúa a finales de la primera mitad del XIV. La inclusión de estas copias tardías en el corpus se debe a que en ambos ejemplos nos encontramos con testimonios que por sus características lingüísticas resultan imprescindibles para describir el habla del occidente peninsular. De cualquier modo, durante el análisis habrá que tener en cuenta las particularidades de cada texto.

3. Control de la procedencia dialectal y variación diatópica

El segundo pilar esencial en la confección del corpus de esta tesis ha sido el control de la procedencia dialectal de los textos. Como ya hemos mencionado en el Capítulo 1 (§ 3), durante los últimos años, se ha producido una revitalización de los estudios de variación geográfica en el marco de la historia de la lengua española a partir, principalmente, del estudio de Fernández-Ordóñez (2011), quien traza algunos patrones geolingüísticos que contribuyen a explicar la conformación del español.

En su trabajo señala varios elementos morfosintácticos del español que son el resultado de innovaciones adoptadas de las variedades laterales del iberorromance (2011: 70 y ss.): la desaparición de *y* locativo y *ende*, y el desarrollo de la forma *alguien* y del plural *quienes* son fenómenos de cambio lingüístico que llegan al castellano desde las variedades occidentales; al mismo tiempo, la colocación de los clíticos abriendo cláusula, la gramaticalización de los tiempos compuestos y la creación y generalización de las formas *vosotros* y *nosotros* serían desarrollos morfosintácticos de procedencia oriental. Así, Fernández-Ordóñez insiste en la necesidad de reconstruir el foco geográfico de las innovaciones y las etapas de su difusión en el espacio a través de estudios empíricos que confronten el castellano con el resto de variedades peninsulares (2011: n. 87). En

¹ «parece sensato admitir como fuente de datos de la lengua de una época aquellas copias que sean contemporáneas de la composición del texto. El problema está en determinar qué entendemos por contemporáneas. Atendiendo a las constataciones que hoy nos proporciona la sociolingüística sobre el período mínimo de difusión de los cambios estudiados en tiempo aparente, me atrevería a proponer a copias que no hayan rebasado la frontera del medio siglo entre la composición de la obra y la transcripción del códice» (Fernández-Ordóñez 2006:1790-91).

definitiva, no podemos entender el castellano como una variedad lingüística uniforme, sino como un complejo dialectal de límites difusos. Es necesario tratar las lenguas y dialectos de la península ibérica como un *continuum* lingüístico, ya que solo así se aportará una visión de conjunto que nos permitirá comprender mejor los factores que intervienen en el cambio lingüístico.

Ese mismo año, en 2011, se publica además la primera versión del *Corpus de Documentos Españoles Anteriores a 1700* (CODEA), confeccionado por el Grupo de Investigación de Textos para la Historia del Español (GITHE) y dirigido por Pedro Sánchez-Prieto Borja. En ese momento, el corpus consta de alrededor de 1500 textos de tipo documental, una tipología textual que se presta especialmente al análisis de la variación dialectal, ya que en la mayoría de casos permite controlar la procedencia geográfica y la fecha de composición de los textos².

A partir de este caldo de cultivo, estos últimos años se han desarrollado numerosos estudios en los que se busca trazar el origen y la difusión de variantes lingüísticas a través del tiempo y del espacio. Algunos ejemplos son la tesis doctoral de Rodríguez Molina (2010) sobre la gramaticalización de los tiempos compuestos, el estudio de Octavio de Toledo acerca de los relacionales locativos (2016), el trabajo de Del Barrio (2016) sobre la distribución de las variantes de los adverbios en *-mente*, el análisis de Enrique-Arias (2018) acerca de la variación entre las formas *este* y *aqueste*, o de algunos trasvases lingüísticos entre Castilla y el oriente peninsular (Enrique-Arias et al. En evaluación), o el trabajo de Moral del Hoyo (2015) en el que traza la distribución del imperfecto y del futuro de subjuntivo, entre muchos otros.

Otra muestra de la vitalidad de los estudios de dialectología histórica es la reciente organización de la sección titulada «Variación diatópica y morfosintaxis en la historia del español» en el XXII Congreso de la Asociación Alemana de Hispanistas, celebrado en 2019 en Berlín. Asimismo, varias tesis recientes o en curso tienen como eje central la perspectiva de la dialectología histórica en algún aspecto del español, como la de Serrano Marín (2018), en la que estudia fenómenos que afectan a la morfología verbal del español mediante la documentación del CODEA; o la de Primerano (en preparación), que trata la gramaticalización del futuro y el condicional en las lenguas y dialectos iberorromances.

El desarrollo de la dialectología histórica y, por tanto, el avance en el conocimiento de los patrones de origen y difusión de las innovaciones en todos los niveles lingüísticos es relevante no solo en sí mismo, sino que puede servir como medio para situar espaciotemporalmente obras importantes de la literatura para las que no tenemos información explícita. En este sentido, en la actualidad hay varios proyectos en desarrollo que pretenden cubrir algunas lagunas de conocimiento: como el de Rodríguez Molina sobre la estratigrafía de manuscritos del siglo XIV (2018), el método de Kawasaki (2014) para datar cronológica y geográficamente documentos medievales españoles de forma automática o la tesis de Ribas Marí (2022) sobre la caracterización dialectal de los romanceamientos bíblicos.

4. Criterios de selección

Es bien sabido que tradicionalmente en el estudio de la historia de la lengua se han privilegiado como fuentes aquellos textos pertenecientes al canon literario, es decir, una nómina de textos limitada y con prestigio. Pons Rodríguez (2006) analiza las obras que forman el corpus de estudio de los principales manuales y de casi un centenar de estudios empíricos de historia de la lengua del siglo XV, y prueba la innegable influencia del canon literario. De este modo, los textos que han recibido más atención en la historia de la

² Además, en la versión de 2015 (CODEA+2015) se añade al corpus una herramienta de cartografiado de los resultados.

literatura, como *La Celestina* y el *Corbacho*, y que además son más accesibles porque tienen más ediciones y más recientes, son los que más se han estudiado y que han servido como modelo para la historia de la lengua.

Del mismo modo, Fernández-Ordóñez (2006) examina cómo se ha utilizado la historiografía medieval como fuente de datos lingüísticos en gramáticas históricas. Es habitual que se repitan los mismos títulos, que en ocasiones pueden ser problemáticos, como la *Primera crónica general*, que mezcla diversas versiones de diferentes siglos.

Así, generalmente, se ha tendido a analizar la lengua de los textos literarios más conocidos, con una muestra de uno o dos testimonios por siglo y sin tener en cuenta su calidad filológica o interés lingüístico. Por supuesto, de este modo no es posible adoptar una perspectiva diatópica.

Es evidente que si la historia de la lengua se restringe al estudio de la lengua literaria tal y como se ha planteado hasta el presente, a partir del análisis de una serie de cortes cronológicos por centurias en los que se examina la lengua de uno o dos testimonios (o a lo sumo tres) —digamos el *Poema de mio Cid*, los *Milagros de Nuestra Señora* de Gonzalo de Berceo, la *Estoria de España* de Alfonso X, el *Libro del conde Lucanor*, el *Libro de buen amor*, el *Corbacho* y *La Celestina*, por citar el canon más frecuente en las monografías gramaticales— difícilmente se podrá apreciar variación geográfica alguna. Y es que la percepción de la variación dialectal exige una diversidad de testimonios mayor que la mera confrontación de un testimonio con otro (Rodríguez Molina 2010:693).

El caso concreto de los estudios sobre la evolución histórica del paradigma pronominal no es muy diferente. Al margen de los manuales de historia de la lengua y las monografías, ya examinados por Pons (2006) y Fernández-Ordóñez (2006), si nos centramos en los trabajos que tratan específicamente la evolución de las variantes de los pronombres personales tónicos de plural en español, vemos que la mayoría se basan únicamente en un corpus formado por textos literarios, donde destacan algunos títulos que se repiten en mayor medida. En los diez estudios que he examinado (Calderón Campos 2010; García 1989, 1991, 1997; García et al. 1990; Gili Gaya 1946; De Jonge y Nieuwenhuijsen 2009; Líbano Zumalacárregui 1988; Rini 1990, 1999), obras como *Calila e Dimna*, *Tristán de Leonís* y *El libro del caballero Zifar* aparecen citadas en 6 de ellos; el *Libro de buen amor* en 5; y *El Corbacho* y *La fazienda de Ultramar* en 4. Solo en algunas ocasiones, el corpus de textos literarios se complementa con colecciones documentales, como es el caso de Líbano Zumalacárregui (1988), García (1997) o Rini (1990). Es llamativo que en estos tres trabajos las variantes objeto de estudio son *connusco* y *convusco* y sus correspondientes reemplazos. Es probable que, debido a que la posición sintáctica de estas variantes es muy concreta, la obtención de un número más reducido de ocurrencias haya favorecido la utilización de un corpus más amplio.

Por otro lado, hay que tener en cuenta que la mayor parte de estos trabajos fueron publicados en la década de los 90, lo que determina los recursos a disposición de los investigadores. De este modo, los más recientes ya incluyen datos del CORDE, como el de De Jonge y Nieuwenhuijsen (2009), que lo usan como corpus complementario de su corpus principal, formado por textos literarios principalmente en prosa; o el de Calderón Campos (2010), que lo utiliza como corpus principal junto con el *Corpus del español* de Mark Davies. En todo caso, no parece que en ninguno de los dos trabajos se filtren o seleccionen los textos de alguna manera, con los problemas que eso puede conllevar (Rodríguez Molina y Octavio de Toledo y Huerta 2017b). Además, en ninguno de estos trabajos se incluye la perspectiva diatópica en el sentido de hacer una selección de textos por distintas zonas geográficas de la península.

En cuanto al corpus de este trabajo, mi objetivo es utilizar textos tanto literarios como documentales y basarme en la medida de lo posible en testimonios fiables o copias

cercanas al original, y que puedan situarse cronológica y geográficamente. Con esto como punto de partida, a continuación describiré en detalle los cuatro principales criterios que he tenido en cuenta durante la recopilación del corpus: a) los límites temporales y la periodización, b) los tipos de textos, c) los límites geográficos y la categorización del *continuum* dialectal y d) la accesibilidad de los textos y el soporte electrónico.

4.1. Límites temporales y periodización

Para este trabajo me voy a centrar en textos de entre los siglos XIII y XVI. Esto se debe a la propia evolución que se ha registrado de las variantes de estudio: todas las formas innovadoras de los fenómenos que voy a analizar empiezan a documentarse a mediados del siglo XIII y se considera que están establecidas en el XVI.

En el corpus de textos literarios, formado *ad hoc* para este trabajo, he intentado representar cada siglo de forma equilibrada, pero esto no siempre ha sido posible debido a la misma disponibilidad de testimonios con localización geográfica conocida. Así, el periodo con una peor representación es el siglo XIII, con 7 testimonios; mientras que el mejor representado es el XV, con 20.

En cuanto al corpus principal de textos documentales, el CODEA+, afortunadamente el grueso de sus documentos se concentra en esos siglos.

En cuanto a la periodización, debido a la complejidad de la datación de algunos textos literarios, he optado por una división amplia por siglos. Es cierto que podría ser recomendable hacer divisiones cada 40 o 50 años. Sin embargo, al tener que dividir también en zonas geográficas, los datos se fragmentarían demasiado. En el caso de los textos documentales, que generalmente tienen una fecha más concreta, sí que he podido hacer divisiones más pequeñas para casos concretos.

4.2. Tipos de textos

Como comentábamos al inicio de este capítulo, el estudio de la historia de la lengua está limitado por los textos que se han conservado hasta la actualidad. El hecho de que solo se puedan extraer datos de textos escritos filtra irremediablemente los rasgos más cercanos a la oralidad. No obstante, dentro de lo escrito es posible de algún modo acceder a diferentes niveles de formalidad dependiendo del registro y la tipología textual. Además, es bien sabido que «existe una relación inversamente proporcional entre la variación geográfica y la variación social» (Rodríguez Molina 2010:609). Así, los textos de registros más elevados normalmente dejan penetrar rasgos dialectales en menor medida.

Como ya he expuesto en detalle más arriba, la mayoría de los trabajos de morfosintaxis histórica se han centrado en el análisis de textos literarios, que, como hemos visto, en muchas ocasiones pueden ser problemáticos por la dificultad de situarlos en una fecha y localización geográfica concretas, y por su compleja transmisión textual. No obstante, esto no significa que haya que renunciar a la utilización de este tipo de textos, sino que su selección y análisis deben ser cuidadosos:

Aunque sin renunciar a los textos literarios como manantial de datos lingüísticos, es conveniente analizar por separado los obtenidos en cada texto, y no generalizar las conclusiones a una época hasta haber concluido los análisis parciales. En fin, una reconstrucción óptima de un cambio en sus coordenadas históricas buscará estar construida sobre documentación original y localizada en el tiempo y en el espacio, localización que muy raramente ofrecen los textos literarios, conservados en copias tardías y anónimas (Fernández-Ordóñez 2001:415).

En todo caso, dentro de estos textos literarios he priorizado los textos en prosa, ya que, como bien se sabe, los textos en verso pueden plantear problemas para distinguir cuándo las soluciones lingüísticas están determinadas por la métrica y/o la rima.

Asimismo, también usaré textos documentales. Ya a principios del siglo XX, Menéndez Pidal destacaba en sus *Documentos lingüísticos* la utilidad de este tipo de textos en el estudio de la historia de la lengua, ya que generalmente tienen una fecha clara y en la mayoría de casos pueden situarse en un punto concreto de la geografía (1919:V). Por estos motivos, en esta parte del corpus incluyo cartas de compraventa, cartas privadas, actas y declaraciones, textos legislativos, etc. Tal y como detallaré más adelante, la principal fuente de textos documentales es el *Corpus de documentos españoles anteriores a 1800* (CODEA+), que incluye textos de los ámbitos cancilleresco, judicial, municipal, eclesiástico y particular, en orden de mayor a menor formalidad (Sánchez-Prieto Borja 2012:15, 2016:17; Serrano Marín 2018:82-83).

No obstante, a pesar de las ventajas que proporcionan los textos documentales en el estudio de la variación diacrónica y diatópica frente a los textos literarios, este tipo de textos tiene un inconveniente conocido. Estos documentos favorecen la aparición recurrente de ciertas fórmulas fijas, especialmente en el encabezado o el cierre³, que en muchos casos incluyen las variantes pronominales objeto de estudio. De este modo, las variantes innovadoras tardarían más en introducirse en contextos como los que vemos en (1), extraídos de documentos del siglo XVI:

- (1)
- a) a *vós* el concejo, justicia, regidores, caballeros, escuderos, oficiales e omnes buenos de la villa de Vilvao, salud e gracia (CODEA-1425, Bilbao, 1506)
 - b) a vuestro alcalde en el dicho oficio e a cadauno de *vós* a quien esta mi carta fuere mostrada, salud e gracia (CODEA-0296, Burgos, 1515)
 - c) a los que agora son como a los que serán de aquí adelante e a cadauno de *vós* a quien esta nuestra carta fuere mostrada (CODEA-0277, Valladolid, 1523)
 - d) e a cada uno e cualquier de *vós* en vuestros lugares e jurisdicciones, e a otras cualesquier personas de cualquier estado e condición (CODEA-0306, Madrid, 1552)

De este modo, no podemos esperar que los textos documentales reflejen de forma exacta la lengua hablada en un punto geográfico y cronológico, ya que «sus usos están mediatizados por la *scripta* o tradición de escritura en que se insertan» (Almeida Cabrejas, Sánchez-Prieto Borja, y Vázquez Balonga 2018:1540).

En definitiva, la combinación de textos literarios y documentales, así como la distinción de diferentes niveles de formalidad, permitirá obtener un panorama lo más completo posible de la variación estilística o diastrática.

4.3. Límites geográficos y categorización del *continuum* dialectal

En mi análisis me voy a centrar principalmente en el estudio de la mitad norte peninsular y, debido a la dificultad de localizar de forma más concreta los textos literarios, he decidido categorizar el *continuum* del centro-norte en tres grandes áreas verticales, de oeste a este: 1) León y Castilla occidental, 2) Castilla central y oriental y 3) Navarra y Aragón. En el caso de los textos documentales, para los que normalmente se conoce una localización más detallada, muchas veces a nivel de población, optaré generalmente por el cartografiado por puntos. Sin embargo, cuando haya que comparar textos literarios y documentales, utilizaré la división en áreas. De este modo, en el área de León y Castilla occidental incluiré los datos de los puntos correspondientes a las actuales provincias de Asturias, León, Zamora, Salamanca, Valladolid, Palencia y Ávila; y en Castilla central y oriental, los de Cantabria, Segovia, Madrid, Burgos, Guadalajara, Soria y La Rioja.

³ Para un análisis de este tipo de elementos formúlísticos en documentos notariales del CODEA, cf. Almeida Cabrejas (2015).

Esta clasificación no pretende representar variedades dialectales individuales o isoglosas, sino que surge de la necesidad de comparar los textos desde una perspectiva diatópica.

Además, como he expuesto hace algunas páginas, considero imprescindible incluir en el análisis todas las variedades romances peninsulares. De este modo, para los extremos oriental y occidental, que actualmente se corresponden con las zonas bilingües con el catalán o el gallego-portugués, he utilizado dos corpus específicos que contienen tanto textos literarios como documentales, el *Corpus Informatizat del Català Antic* (CICA) y el *Tesouro Medieval Informatizado da Lingua Galega* (TMILG) respectivamente (ver § 5.c). El CICA proporciona la información del dialecto al que pertenecen los textos: septentrional, central y balear —bloque oriental—; y noroccidental y valenciano —bloque occidental. Esto ha permitido separar los datos del CICA en otras dos franjas verticales, una oriental y otra occidental⁴. Del mismo modo, en muchos casos el TMILG aporta información de provincia, por lo que he agrupado Lugo y Ourense por un lado, y A Coruña y Pontevedra por otro.



Mapa 1. Propuesta de categorización del *continuum* norte peninsular en 7 franjas de norte a sur.

4.4. Accesibilidad de los textos y soporte electrónico

Los avances en lingüística de corpus de las últimas décadas han puesto a nuestra disposición nuevos recursos que han cambiado la forma de acceder a las muestras de habla del pasado. La disponibilidad de herramientas informáticas que permiten acceder a ingentes colecciones de textos antiguos y analizar grandes bases de datos han permitido el resurgir de la lingüística histórica.

En cualquier caso, el empleo de corpus diacrónicos en línea no exime de la necesidad de acercarse a los textos a través del conocimiento filológico y de analizar los datos con la cautela y la rigurosidad del método científico:

Advertimos ya de entrada que el empleo de los recursos informatizados en línea no sirve para sustituir los principios y habilidades que siempre deben acompañar al filólogo, que son el conocimiento amplio de la historia, lenguas, cultura y sociedad medievales y, como en cualquier actividad investigadora, la atención al detalle, la capacidad de evaluar el mérito relativo de diferentes hipótesis, la habilidad para valorar la validez empírica de los datos, el razonamiento lógico y, finalmente, esa virtud difícil de definir que suele estar

⁴ La ilustración de estas dos zonas en el Mapa 1 únicamente es una ayuda visual y no pretende ser totalmente fiel a las fronteras de los dialectos del catalán, que, por supuesto, van más allá de las fronteras políticas de Cataluña y se prolongan hacia el oeste en Aragón, hacia el norte en Francia y hacia el sur en Valencia.

presente en los grandes descubrimientos: la intuición afinada por la experiencia (Enrique-Arias 2019a:335-36).

Estos últimos años, se han desarrollado estándares filológicos de selección y presentación de textos que han sido la base de corpus diacrónicos fiables y de calidad. Las ediciones críticas no son suficientes para el estudio de la lengua medieval, ya que «durante el proceso [de edición] los editores descartan muchas variantes de interés lingüístico e invariablemente confieren al texto resultante de una engañosa uniformidad» (Enrique-Arias 2019a). La red CHARTA ha tenido un papel fundamental en el establecimiento del estándar de edición de los textos medievales, que propone una triple presentación que da acceso a tres niveles del texto: el facsímil digital del manuscrito, la transcripción paleográfica y el texto normalizado (Sánchez-Prieto Borja 2011). A esta presentación en tres niveles, que solo es viable en soporte electrónico, se le suman herramientas de búsqueda que faciliten la obtención y el filtrado de los resultados⁵.

Para esta tesis he utilizado principalmente corpus diacrónicos informatizados, especialmente de acceso libre en la red, como el CODEA+, *Biblia Medieval*, el *Diccionario del castellano del siglo XV en la Corona de Aragón* (DiCCA-XV), el *Corpus diacrónico del español* (CORDE) o la *Biblioteca Digital de Textos del Español Antiguo* (BiDTEA).

He priorizado aquellos corpus que permitieran acceder libremente a la transcripción paleográfica completa de los testimonios y, siempre que fuera posible, también al facsímil. Los dos corpus que cumplen con estos requisitos son el CODEA+ y *Biblia medieval*. La BiDTEA, aunque no pone a disposición del usuario los facsímiles, sí permite la descarga de las transcripciones semipaleográficas completas, lo que facilita enormemente el trabajo de extracción y codificación de los ejemplos.

Otros corpus utilizados, como el DiCCA-XV y el CORDE, no dan acceso al texto completo, sino a las concordancias y al contexto próximo de los resultados de las búsquedas. No obstante, el DiCCA-XV tiene una gran ventaja, y es que en los resultados de la búsqueda el corpus proporciona información que facilita la codificación. Al buscar la lista de contextos para cada lema, una de las columnas se corresponde con la acepción. En el caso de la búsqueda de los pronombres personales *nós* y *vós*, estas acepciones distinguen entre: 1) variante tónica con referente plural, 2) variante átona con referente plural, 3) variante tónica con referente singular, y 4) variante átona con referente singular. Por tanto, este ha sido el único corpus en el que no ha sido necesario identificar el referente de cada uno de los ejemplos en la codificación.

En algunas ocasiones, las que menos, he optado por ediciones digitalizadas de textos concretos disponibles en línea, como la edición de Pons Rodríguez (2008) de *Virtuosas e claras mugeres* o la de Garrido Martín (2016) de *La Atalaya de las corónicas*, o ediciones accesibles a través del propio autor, como la de D'Ambruso (2012) de *Crónica troyana promovida por Alfonso XI*.

Aunque el requisito del acceso electrónico y libre de los textos limita los materiales disponibles, considero que es un criterio indispensable, ya que facilita replicabilidad del estudio y la comparabilidad con trabajos que utilicen una nómina de textos igual o similar. Además, permite poner a prueba y aprovechar las herramientas que se han ido desarrollando durante la última década.

⁵ Para un análisis comparativo de las herramientas de búsqueda y el acceso a las concordancias y los textos de los corpus más importantes del español, véase De Benito Moreno (2019).

5. Descripción del corpus

Una vez que hemos repasado los criterios fundamentales de la formación del corpus, es necesario exponer las fuentes y la nómina final de textos. Así, he dividido esta sección en tres apartados: en primer lugar, describiré la fuente principal de textos documentales, el CODEA+ (§ 5.1); en segundo lugar, presentaré las fuentes de los textos literarios y listaré las obras seleccionadas (§ 5.2); y, finalmente, expondré las fuentes de las que he extraído los textos documentales que forman el corpus auxiliar (§ 5.c).

5.1. Corpus de textos documentales: CODEA+ 2015

Una de las fuentes de datos más importantes para el corpus de esta tesis y, más concretamente, para el corpus de textos documentales es el *Corpus de documentos españoles anteriores a 1800* (CODEA+), desarrollado por el Grupo de Investigación de Textos para la Historia del Español (GITHE) de la Universidad de Alcalá. Como es bien sabido, es un corpus que se encuentra disponible en línea y es de acceso libre. Además, el CODEA+ utiliza el sistema de triple presentación de la red CHARTA, por lo que es posible consultar la transcripción paleográfica de los documentos, que reproduce las opciones gráficas del original; la presentación crítica con grafías normalizadas, que eliminan la variación gráfica superflua; y las imágenes de los facsímiles.

El corpus contiene textos desde el siglo XI al XVIII, aunque el grueso de documentos y palabras se concentra entre los siglos XIII y XVI. De este modo, coincide perfectamente con el periodo cronológico objeto de estudio de esta tesis.

En su estado actual consta de 2491 textos de fuentes documentales, es decir, cartas de compraventa, cartas privadas, actas y declaraciones, textos legislativos, etc. La mayoría de estos documentos (alrededor de un 94 %) tienen una data geográfica y cronológica conocida, lo que convierte al CODEA+ en una herramienta ideal para trazar el origen y la difusión de las diferentes variantes lingüísticas en el tiempo y en el espacio. Al mismo tiempo, incluye información sobre otros parámetros, como la tipología documental, el tipo diplomático, el ámbito de emisión o la participación de la mujer. Así, el CODEA+ es una «herramienta completa para el estudio lingüístico de la documentación histórica desde diversas perspectivas» (Almeida Cabrejas et al. 2018:1549).

Respecto al alcance geográfico, el CODEA+ incluye documentos en español en un sentido amplio. Es decir, incorpora textos en castellano, pero también de las variedades leonesa, aragonesa y navarra. Además, frente a la versión de 2011, la actual añade documentos de zonas bilingües (Galicia, País Vasco, Cataluña y Valencia) con el objetivo de representar toda la geografía peninsular. Sin embargo, esta representación todavía sigue siendo minoritaria, especialmente hasta el siglo XVI. De la totalidad del corpus, 201 documentos pertenecen a estas regiones, es decir, un 8 %. Solo 27 de estos se sitúan entre los siglos XIII y XVI (2 de Cataluña, 3 de Galicia, 3 de Valencia y 19 del País Vasco). Para cubrir este hueco y conseguir un panorama completo del *continuum* romance del norte peninsular, completaré los datos del CODEA+ con los de los textos documentales del *Corpus Infortmatizat del Català Antic* (CICA) y el *Tesouro Medieval Informatizado da Lingua Galega* (TMILG).

En la Figura 2 puede verse un resumen del volumen de palabras en cada una de las zonas geográficas (cf. § 4.3) que hemos establecido para el corpus de este trabajo desde el siglo XIII al XVI en el CODEA+. En general, el corpus se muestra equilibrado. Únicamente, es necesario destacar que en la zona de Navarra y Aragón encontramos un descenso considerable en densidad de documentos y cantidad de palabras a partir 1425 aproximadamente, lo que se acentúa en el siglo XVI.

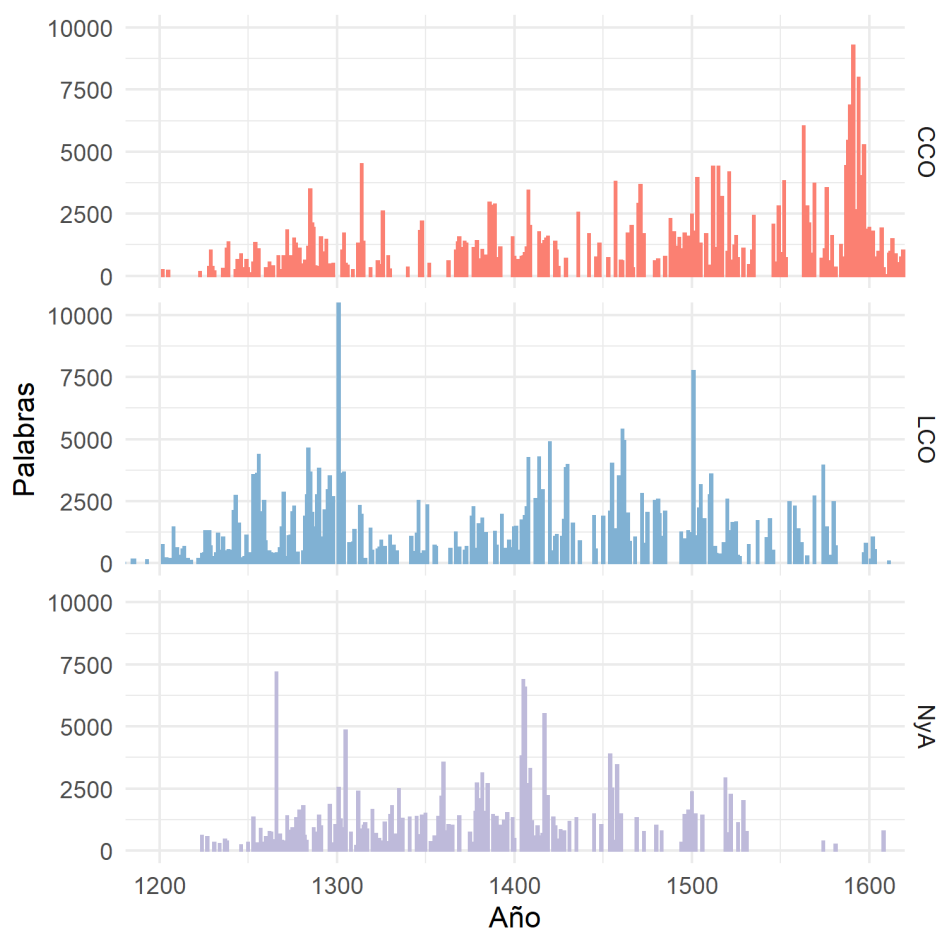


Figura 2. Volumen de palabras entre los siglos XIII-XVI en Castilla central y oriental (CCO), León y Castilla occidental (LCO) y Navarra y Aragón (NyA) en el CODEA+.

5.2. Corpus de textos literarios: fuentes y nómina

La compilación de textos literarios ha sido diferente a la de los textos documentales. Con el objetivo de cumplir con los criterios que se han presentado en § 4 para conseguir una muestra fiable, representativa y equilibrada, he hecho una selección de testimonios *ad hoc* para este trabajo. Para ello, ha sido imprescindible manejar una mayor diversidad de fuentes. Esto no significa que la nómina de textos que se va a presentar en esta sección no sea perfeccionable, ya que no es posible controlar al 100 % todas las variables —fecha, localización geográfica, tipología textual, extensión del texto— al mismo tiempo que se mantiene una muestra equilibrada en cada una de ellas.

En esta sección haré una breve descripción de las fuentes de las que he extraído los textos y, a continuación, presentaré la nómina de textos literarios.

a) Fuentes

Las principales fuentes de las que se han extraído textos literarios han sido la *Biblioteca Digital de Textos del Español Antiguo* (BiDTEA) y los textos y concordancias del *Hispanic Seminary of Medieval Studies* (HSMS), *Biblia medieval* y *Biblias hispánicas*, el *Corpus diacrónico del español* (CORDE) y el *Diccionario del castellano del siglo XV en la Corona de Aragón* (DiCCA-XV). Al mismo tiempo, también se han incorporado

algunas ediciones concretas de textos que no se han encontrado en estas fuentes y que se especificarán en § 5.2.b.

- *Biblia medieval*: 2 textos, 1.016.697 palabras
- BiDTEA: 14 textos, 3.101.730 palabras
- Textos y concordancias de Madison: 5 textos, 800.538 palabras
- CORDE: 16 textos, 3.349.524 palabras
- DiCCA-XV: 5 textos, 439.661 palabras
- Ediciones individuales: 4 textos, 894.995 palabras

La Biblioteca Digital de Textos del Español Antiguo y el Hispanic Seminary of Medieval Studies

Una de las fuentes de la que se han extraído más textos ha sido la *Biblioteca Digital de Textos del Español Antiguo* del *Hispanic Seminary of Medieval Studies*, que empezó a estar disponible en 2011 con la publicación de la *Obra en prosa de Alfonso X el sabio*. La BiDTEA nace en el seno del HSMS como heredera de la idea del *Old Spanish Archive* (OSA) y como continuación del proyecto del *Dictionary of the Old Spanish Language* y de la publicación en microficha y luego en CD-ROM de los *Textos y Concordancias* (Gago Jover 2015).

La BiDTEA es una base de datos interactiva disponible en línea y en acceso abierto que contiene transcripciones semi-paleográficas de testimonios de textos medievales, así como índices (alfabéticos, de frecuencia y alfabéticos inversos) y concordancias en formato KWIC (*key word in context*) (Gago Jover 2015:10)⁶. Los testimonios que se incluyen en cada uno de sus subcorpus han sido seleccionados con gran rigor filológico y se basan en la información recogida en PhiloBiblon⁷.

Además de la propia interfaz de la BiDTEA, la base de datos ofrece la posibilidad de descargar los textos. Al no ser esta la única fuente de mi corpus, he optado por descargarlos para unificar en la medida de lo posible la extracción de datos con el resto de fuentes.

En su versión actual⁸, la BiDTEA dispone de los siguientes sub-corpus:

- Obra en prosa de Alfonso X el sabio (20 textos)
- Textos médicos españoles (62 textos)
- Textos navarro-aragoneses (52 textos)
- Textos legales españoles (66 textos)
- Textos bíblicos españoles (19 textos)

⁶ «A pesar de las mejoras ofrecidas, *BiDTEA* no satisface los objetivos establecidos por Nitti para el *Old Spanish Archive* y es por ello que en el año 2015 Francisco Gago Jover y F. Javier Pueyo Mena, ambos colaboradores del HSMS, comenzaron a desarrollar el *Old Spanish Textual Archive* (OSTA), un corpus lingüístico accesible en línea, que permitirá efectuar búsquedas lingüísticas complejas en las más de 400 transcripciones semi-paleográficas de textos, escritos en castellano, asturiano, leonés, navarro-aragonés y aragonés, y producidos entre los siglos XIII y XVII, realizadas por los colaboradores del *Hispanic Seminary of Medieval Studies* (HSMS) desde los años 70 del pasado siglo» (Gago Jover y Pueyo Mena 2018a:167-68). El proyecto se describe en Gago Jover y Pueyo Mena (2018a, 2018b). En Gago Jover y Pueyo Mena (2018b) se presentan los aspectos técnicos de la lematización y etiquetado gramatical del corpus. En Gago Jover y Pueyo Mena (2018a) se describen los textos, códigos y transcripciones. Diferencia respecto a BiDTEA: «permitirá combinar búsquedas en todos los niveles de marcación anotados y filtrar los resultados según los campos recogidos en la meta-descripción de cada obra y de cada código» (Gago Jover y Pueyo Mena 2018b).

⁷ *PhiloBiblon* es «una base de datos bio-bibliográfica sobre textos romances escritos en la península ibérica en la Edad Media y temprano Renacimiento, de acceso gratuito a través de Internet» (Faulhaber 1997). Los principales responsables del proyecto son Charles B. Faulhaber, Arthur L-F. Askins, Harvey L. Sharrer, Gemma Avenozza y John G. May.

⁸ Fecha de consulta: 16/09/2020.

- Textos poéticos españoles (37 textos)
- Textos tempranos de Celestina (21 textos)
- Textos cronísticos españoles (49 textos)
- Textos del Lazarillo de Tormes (1554) (4 textos)
- Textos del Fuero general de Navarra (19 textos)
- Textos coloniales (11 textos)

Es necesario mencionar que hay cierto solapamiento entre los corpus citados aquí arriba, ya que hay textos que se incluyen en varios a la vez. Por ejemplo, los distintos testimonios de la *Estoria de España* y de la *General Estoria* aparecen tanto en el corpus de textos alfonsíes como en el de textos cronísticos. A su vez, las dos partes de la *Crónica de los conqueridores* están recogidas tanto en los textos cronísticos como en los navarro-aragoneses. Tal y como se verá en la nómina de textos, los tres corpus mencionados en este párrafo son los que han resultado más útiles para esta tesis.

Algunos textos, como el *Exemplario por ABC* o el *Libro de las propiedades de las cosas*, transcritos por el HSMS, al menos por el momento no están disponibles en la BiDTEA, por lo que los he tomado de los *Textos y Concordancias* publicados primero en microficha y luego en CD-ROM.

Biblia Medieval y Biblias Hispánicas

Otra de las fuentes de datos de esta tesis han sido los recursos de libre acceso en línea del proyecto *Biblia Medieval*, que incluyen transcripciones paleográficas, ediciones normalizadas e imágenes facsimilares del corpus integral de los romanceamientos bíblicos medievales. En su versión actual, el corpus recopila los testimonios conservados datados entre los siglos XIII y XVI, y los alinea en paralelo con sus fuentes hebrea y latina, lo que facilita la comparación entre cada uno de los textos. Asimismo, es necesario destacar que el corpus *Biblias Hispánicas* está lematizado y contiene etiquetado gramatical.

Como es bien sabido, las traducciones bíblicas castellanas son las más numerosas en el conjunto de los vernáculos premodernos europeos y forman una fuente rica y variada para el análisis lingüístico. Así, disponer de varias versiones de un mismo texto constituye un recurso especialmente apropiado para el estudio de la variación diacrónica (Enrique-Arias 2009).

No obstante, para muchos de estos textos no se conoce de forma precisa el contexto de composición de la obra o su procedencia dialectal. Por este motivo, he seleccionado únicamente dos de los testimonios que incluye el corpus. Ambos están ampliamente descritos y han sido localizados geográficamente: el texto conservado en el códice Escorial I.i.6 (E6) y la *Biblia de Arragel*. En el caso de E6 (siglo XIII), se trata de un códice que conserva junto con Escorial I.i.8 (E8) la traducción de la *Biblia prealfonsí*, hecha a partir del latín. En cuanto a la *Biblia de Arragel* (siglo XV), es uno de los romanceamientos que ha sido descrito con más detalle y sobre el que tenemos más información. Gracias al extenso prólogo de la obra conocemos las características del proceso romanceador y aspectos como la fecha y el lugar de composición, el origen geográfico del autor, el rabino Mosé Arragel, natural de Guadalajara; y el mecenas y destinatario de la traducción, Luis de Guzmán⁹.

Del mismo modo que con la BiDTEA, para la extracción de datos de *Biblia medieval* no he utilizado la propia interfaz del corpus, sino que he descargado los textos disponibles en la web¹⁰ para unificar el proceso con el resto de fuentes.

⁹ Para más información sobre la fecha y el origen geográfico de estos dos testimonios, véase § 5.2.b. Nómina de textos.

¹⁰ Disponibles en este enlace: <http://www.bibliamedieval.es/index.php/recursos/recursos-textos>.

El Corpus diacrónico del español y el Cordemáforo

Otra de las fuentes de la que he extraído más textos ha sido el *Corpus diacrónico del español* (CORDE). El corpus de la Real Academia Española contiene textos de todos los tipos desde los orígenes hasta 1974¹¹. Es decir, contiene textos históricos tanto documentales como literarios, en prosa o en verso, o científico-técnicos. Con 250 millones de palabras¹², es uno de los corpus diacrónicos más grandes y visibles (Rojo 2016:289). Sin embargo, durante los últimos años se han destacado algunos problemas sobre su composición que obligan a utilizarlo con cautela.

No cabe duda de la utilidad del CORDE, pero autores como Rodríguez Molina y Octavio de Toledo (2017b:5) han subrayado «la escasa fiabilidad filológica de las dataciones que ofrece y de algunos de los textos que reúne». Estos autores cuestionan la inclusión de algunos textos y ediciones, la mezcla de «materiales de muy distinta calidad» y que no se expliciten los criterios de su selección (2017b:6-9). Así, el CORDE mezcla indiscriminadamente textos y testimonios, tiene obras duplicadas, errores de catalogación, textos que aparecen como anónimos a pesar de tener autor, ediciones que modernizan lingüísticamente el texto; y no parece tener un criterio claro para la datación de las obras, ya que «no siempre sigue la fecha propuesta en la edición que utiliza» (2017b:13).

Por estos motivos, en lugar de utilizar el corpus en su totalidad, he extraído algunos textos que no he podido encontrar en otras fuentes. A su vez, he complementado la utilización del CORDE con el Cordemáforo (Rodríguez Molina y Octavio de Toledo y Huerta 2017a, 2017b).

El Cordemáforo es una valiosa herramienta que clasifica cromáticamente los textos literarios (excluyendo cancioneros) del CORDE hasta 1492 según su fiabilidad filológica. Rodríguez Molina y Octavio de Toledo (2017b:44) adjudican «un color a cada texto en función de si este es bastante fiable (verde), relativamente fiable (ámbar) o poco o nada fiable (rojo)». Además, en esta base de datos incluyen otra información útil como la fecha del testimonio, la fecha de copia, la BETAmanid o el tipo de testimonio (edición crítica o paleográfica). Esto no quiere decir que en el corpus de esta tesis solo haya incluido textos marcados en verde o que haya descartado sistemáticamente todos los que aparecen marcados en rojo, pero sí me ha servido como guía durante el proceso de selección y tendré en cuenta la clasificación para tratar los textos más problemáticos con una mayor cautela.

Por otro lado, a diferencia de la BiDTEA o de *Biblia Medieval*, el CORDE no permite descargar los textos y solo es posible acceder a un contexto limitado, por lo que la extracción de datos se ha hecho a través de su propia interfaz.

El Diccionario del castellano del siglo XV en la Corona de Aragón

El DiCCA-XV incluye textos en castellano compuestos en la Corona de Aragón durante el siglo XV, transcritos directamente desde el original por el *Grup d'Història i Contacte de Llengües* (GHCL) de la Facultat de Filologia de la Universitat de Barcelona.

¹¹ Información completa en la página web: <https://www.rae.es/recursos/banco-de-datos/corde>.

¹² Como bien exponen Rodríguez Molina y Octavio de Toledo (2017b:7), este número varía en distintas descripciones: «La cifra de 250 millones que figura en la página web de la Academia y refrenda Rojo no coincide con otras cifras más cercanas a los 300 millones que proporcionan otros investigadores (Sánchez y Domínguez 2007, Muñoz Armijo 2009). Los datos más precisos acerca de la composición del CORDE los ofrece Sánchez-Prieto (2015: 117), quien con fecha de consulta 31/03/2014 cifra en 390.922.575 el total de palabras recogidas en este corpus, de las cuales 196.106.277 corresponden a España y 37.562.461 a Hispanoamérica, a los que hay que sumar 3.041.176 procedentes de otros espacios lingüísticos.»

El corpus del diccionario está formado por textos tanto literarios como documentales. Concretamente, distinguen cuatro tipologías: a) textos jurídicos y administrativos, b) textos científicos, técnicos y doctrinales, c) textos históricos y descriptivos y d) textos novelescos y poéticos. Cada una de estas cuatro secciones forma aproximadamente un 25 % del corpus. Para esta tesis, con el objetivo de mantener la comparabilidad con el resto de textos de mi corpus, he decidido seleccionar los textos de tipo histórico y descriptivo.

Asimismo, cabe destacar que el diccionario permite extraer todos los contextos de un lema. A pesar de que el contexto accesible es más limitado que el del CORDE, es la única fuente en la que hay un etiquetado con información sobre el tipo de referente (individual o plural/colectivo) y la posición sintáctica, marcas muy relevantes para el análisis.

b) Nómima de textos

A continuación, presento la nómina de textos que forman el corpus literario de esta tesis. Como ya se ha ido mencionando en las páginas anteriores, esta lista de textos está lejos de ser perfecta y definitiva. Aunque he procurado que los textos fueran lo más fiables y representativos posibles, esto no significa que no haya algunos casos problemáticos.

En las próximas páginas, listo por siglos (XIII-XVI)¹³ cada uno de los textos literarios de mi corpus. Junto al título de cada obra adjunto una breve ficha con algunos datos que he considerado que eran relevantes. Así, incluyo la clave que utilizaré a lo largo del estudio para hacer referencia al texto, la etiqueta de la zona geográfica asignada de acuerdo con los criterios (ver § 4.3), la ubicación del testimonio seleccionado y su signatura, la fuente, el número de palabras¹⁴ y, finalmente, un apartado para observaciones y apuntes. Este último apartado no pretende ser en ningún caso una descripción exhaustiva de los testimonios, su historia textual o su lengua. Únicamente se incluyen algunas referencias que han sido la clave para escoger un testimonio o situarlo cronogeográficamente, y que pueden ser útiles para aquellos que quieran obtener más información.

Como puede verse en las fichas, hay una serie de recursos que aparecen de forma recurrente, como la base de datos PhiloBiblon (Faulhaber 1997), «que ofrece un catálogo codicológico y ecdótico exhaustivo sobre los textos medievales iberrrománicos y sus testimonios» (Rodríguez Molina 2018:103) o el *Diccionario filológico de literatura medieval española* de Alvar y Lucía Megías (2002).

Siglo XIII

Biblia Escorial I.i.6

Clave	E6
Zona geográfica	Castilla central y oriental
Testimonio	San Lorenzo de El Escorial: Real Biblioteca del Monasterio de San Lorenzo de El Escorial (RBME), I.I.6 (BETA manid 1161)
Fuente	<i>Biblia Medieval</i>
Palabras	428.697

¹³ Ante la dificultad de situar en una fecha más concreta algunos de los testimonios y de conseguir una muestra equilibrada geográficamente en cortes cronológicos más pequeños, he optado por una clasificación amplia en siglos. Queda pendiente pulir y refinar esta nómina de textos para conseguir una división, al menos, para cada medio siglo.

¹⁴ Siempre que ha sido posible, esta información se ha tomado de la fuente. El CORDE o la BiDTEA incluyen este dato en sus índices de textos. Para otros, como los de *Biblia Medieval* o los de *Textos y Concordancias* de Madison, se ha tomado el recuento de palabras que ofrece *AntConc* (Anthony 2019).

Observaciones y apuntes	<p>«el códice E6 es la biblia castellana más antigua que conservamos», «traducción de gran calidad de la Vulgata que contiene el <i>AT</i> a partir de los <i>Proverbios</i> y el <i>NT</i> en su totalidad» (Enrique-Arias 2010a:15).</p> <p>Sánchez-Prieto Borja y Torrens Álvarez (2010:33) destacan su «castellanismo lingüístico». Matute y Pato (2010:64): «en general se observa un empleo semejante al de los textos plenamente castellanos», aunque con oscilaciones.</p> <p>Copia cercana a su fecha de composición, hacia 1250 (Sánchez-Prieto Borja y Torrens Álvarez 2010:37-38).</p> <p>Códice hermano de E8, de hacia 1400 y con rasgos lingüísticos orientales (aragoneses, navarros o riojanos según diversos análisis) (Enrique-Arias 2010b:68).</p>
--------------------------------	--

General Estoria I

Clave	GE1
Zona geográfica	Castilla central y oriental
Testimonio	Madrid: Biblioteca Nacional de España, MSS/816 (BETA manid 1055)
Fuente	BiDTEA
Palabras	535.516
Observaciones y apuntes	<p>Subcorpus <i>Obra en prosa de Alfonso X el sabio; Textos cronísticos españoles</i>.</p> <p>Transcrito por: Lloyd Kasten, John Nitti, y Wilhelmina Jonxis-Henkemans.</p> <p>Corregido por: Pedro Sánchez-Prieto Borja, Rocío Díaz Moreno, Elena Trujillo Belso, y Francisco Gago Jover.</p> <p>Códice original del <i>scriptorium</i> alfonsí, c. 1270 (Alvar y Lucía Megías 2002:45).</p> <p>«El tipo lingüístico de la primera parte de la <i>General estoria</i> desconoce los rasgos leonesizantes de la cuarta» (Fernández-Ordóñez 2004:408).</p>

Gran conquista de ultramar

Clave	GCU
Zona geográfica	Castilla central y oriental
Testimonio	Madrid: Biblioteca Nacional de España, MSS/1187 (BETA manid 1374)
Fuente	BiDTEA
Palabras	247830
Observaciones y apuntes	<p>Subcorpus <i>Textos cronísticos españoles</i>.</p> <p>Transcrito por: Louis Cooper.</p> <p>«códice de los últimos años del siglo XIII o los primeros del XIV», «parece provenir de la Cámara Real castellana» (Alvar y Lucía Megías 2002:605).</p>

Judizios de las estrellas

Clave	Judizios
Zona geográfica	Castilla central y oriental
Testimonio	Madrid: Biblioteca Nacional de España, MSS/3065 (BETA manid 1083)
Fuente	BiDTEA
Palabras	125.469
Observaciones y apuntes	Subcorpus <i>Obra en prosa de Alfonso X el sabio</i> . Transcrito por: Lloyd Kasten, John Nitti, and Wilhelmina Jonxis-Henkemans. Corregido por: Pedro Sánchez-Prieto Borja, Rocío Díaz Moreno, Elena Trujillo Belso y Francisco Gago Jover. Este testimonio «parece proceder del <i>scriptorium</i> alfonsí» (Alvar y Lucía Megías 2002:30). Según Fernández-Ordóñez (2004:405), «refleja una lengua castellana de tipo oriental», «aunque carece de los aragonesismos extremos de <i>Cruzes</i> ».

General Estoria IV

Clave	GE4
Zona geográfica	León y Castilla occidental
Testimonio	Roma: Biblioteca Apostólica Vaticana [BAV], Urb. lat. 539 (BETA manid 1077)
Fuente	BiDTEA
Palabras	456.336
Observaciones y apuntes	Subcorpus <i>Obra en prosa de Alfonso X el sabio; Textos cronísticos españoles</i> . Transcrito por: Lloyd Kasten, John Nitti, y Wilhelmina Jonxis-Henkemans. Corregido por: Pedro Sánchez-Prieto Borja, Rocío Díaz Moreno, Elena Trujillo Belso, y Francisco Gago Jover. Códice original del <i>scriptorium</i> alfonsí, 1280, copiado por Martín Pérez de Maqueda (Alvar y Lucía Megías 2002:49). Según Fernández-Ordóñez (2004:406), este testimonio «parece presentar un tipo occidental, que comparte algunos aspectos con el leonés oriental».

Libro de las cruces

Clave	Cruces
Zona geográfica	Navarra y Aragón
Testimonio	Madrid: Biblioteca Nacional de España, MSS/9294 (BETA manid 1088)
Fuente	BiDTEA
Palabras	130.671
Observaciones y apuntes	Subcorpus <i>Obra en prosa de Alfonso X el sabio</i> . Transcrito por: Lloyd Kasten, John Nitti, and Wilhelmina Jonxis-Henkemans.

Corregido por: Pedro Sánchez-Prieto Borja, Rocío Díaz Moreno, Elena Trujillo Belso, and Francisco Gago Jover.

Aunque es un códice original del *scriptorium* alfonsí, según Fernández-Ordóñez (2004:404), tiene «numerosas soluciones aragonesas o de estirpe occitánica».

Vidal Mayor

Clave	Vidal
Zona geográfica	Navarra y Aragón
Testimonio	Los Angeles: J. Paul Getty Museum, XIV 6 (BETA manid 2859)
Fuente	BiDTEA
Palabras	180.278
Observaciones y apuntes	Subcorpus <i>Textos Navarro-Aragoneses</i> . Transcrito por: Abraham Quintanar.
	Traducción aragonesa de mediados del XIII de la obra <i>In excelsis Dei thesauris</i> , de Vidal de Canellas, obispo de Huesca, por encargo de Jaime I (PhiloBiblon).

Siglo XIV

Leomarte: Sumas de la historia troyana

Clave	Sumas
Zona geográfica	Castilla central y oriental
Testimonio	Madrid: Biblioteca Nacional de España, MSS/9256 (BETA manid 1864)
Fuente	BiDTEA
Palabras	101.830
Observaciones y apuntes	Subcorpus <i>Textos cronísticos españoles</i> . Transcrito por: Robert G. Black.
	Códice de la segunda mitad del siglo XIV (Alvar y Lucía Megías 2002:749). Sobre el autor del texto, Rey (1932:13) afirma que «debió de ser castellano, puesto que no se hallan en el texto características lingüísticas de otras regiones. Los escasos galicismos que se encuentran se deben sin duda a que el texto procede en parte de una traducción del <i>Roman de Troie</i> aprovechada en la <i>General Estoria</i> ».

Mostrador de Justicia

Clave	Mostrador
Zona geográfica	Castilla central y oriental
Testimonio	Paris: Bibliothèque nationale de France, Espagnol 43 (BETA manid 2456)
Fuente	Edición de Mettmann (1996)
Palabras	389.990
Observaciones y apuntes	El autor, Abner de Burgos/Alfonso de Valladolid, nació alrededor de 1260 en Burgos. A los 60 años, en 1320, se

convierte al cristianismo. Su obra fue traducida del hebreo al castellano en el siglo XIV por el mismo autor (Szpiech 2016).

Crónica troyana promovida por Alfonso XI

Clave	CTroyana
Zona geográfica	León y Castilla occidental
Testimonio	San Lorenzo de El Escorial: Real Biblioteca del Monasterio de San Lorenzo de El Escorial (RBME), h.I.6 (BETA manid 1561)
Fuente	Edición de D'Ambruso (2012)
Palabras	242.072
Observaciones y apuntes	Traducción anónima del <i>Roman de Troie</i> de Benoît de Sainte-Maure encargada por Alfonso XI. Terminó de copiarse el 31 de diciembre de 1350 por un escribano identificado como Nicolás Gonçález. Solo el código Escorial h.I.6 conserva la obra completa (Alvar y Lucía Megías 2002:997; D'Ambruso 2012:LXXIX). Hay una edición de Parker (1977), aunque contiene numerosos errores de transcripción (D'Ambruso 2012:VII; Fernández-Ordóñez 2009:111; Rodríguez Molina 2010:846). La edición de D'Ambruso (2012) se basa en Escorial h.I.6, código para el que señala frecuentes leonesismos (2012:CXXXI).

Cuento muy fermoso de Otas de Roma

Clave	Otas
Zona geográfica	León y Castilla occidental
Testimonio	San Lorenzo de El Escorial: Real Biblioteca del Monasterio de San Lorenzo de El Escorial (RBME), h.I.13 (BETA manid 1303)
Fuente	<i>Textos y concordancias</i> de Madison
Palabras	43.469
Observaciones y apuntes	Transcrito por: Thomas D. Spaccarelli. Uno de los textos que conserva el código escurialense h.I.13 es <i>Otas</i> . Este testimonio «se ha datado o a mediados del siglo XIV o a comienzos del siglo XV» (Alvar y Lucía Megías 2002:856). Según <i>PhiloBiblon</i> , escrito en castellano occidental.

Libro de Alexandre

Clave	Alexandre
Zona geográfica	León y Castilla occidental
Testimonio	Madrid: Biblioteca Nacional de España, VITR/5/10 (BETA manid 1408)
Fuente	BiDTEA
Palabras	73.656
Observaciones y apuntes	Subcorpus <i>Textos poéticos españoles</i> . Transcrito por: Ivy A. Corfis. Código «del siglo XIV o muy a finales del siglo XIII copiado en León por Fray Lorenzo de Astorga» (Alvar y Lucía Megías 2002:756). Según Rodríguez Molina (2010:742), la lengua de

este testimonio «está fuertemente impregnada de rasgos occidentales, leoneses por más señas».

Poema de Alfonso Onceno

Clave	Onceno
Zona geográfica	León y Castilla occidental
Testimonio	San Lorenzo de El Escorial: Real Biblioteca del Monasterio de San Lorenzo de El Escorial (RBME), Y.III.9 (BETA manid 2228) +++
Fuente	CORDE
Palabras	42.856
Observaciones y apuntes	Edición: Juan Victorio, Madrid: Cátedra, 1991. Cordemáforo: ámbar (EC)

«El Poema de Alfonso XI podría recomendarse como un texto fiable para el estudio de la lengua de la primera mitad del siglo XIV a tenor de que el manuscrito más antiguo que conserva este texto compuesto hacia 1348, el ms. Esc. III-Y-9, fue copiado como muy tarde en el último tercio del trescientos, si no fuera porque el CORDE utiliza la edición de Victorio (1991), quien mezcla fragmentos del códice escurialense con el ms. 12 de la RAE, copiado a finales del siglo XV» (Rodríguez Molina y Octavio de Toledo y Huerta 2017b:12-13).

«En 1991, Juan Victorio publica una nueva edición; en ella tiene en cuenta los dos manuscritos, pero intenta realizar una reconstrucción, en todos los niveles, del texto original: "he procedido a hacer numerosos retoques en aras de ofrecer una lengua uniforme (...) ante las dos opciones que, para cualquier fenómeno, aparecen en el texto, retendré la más antigua", p. 36» (Alvar y Lucía Megías 2002:920).

«Tanto Menéndez Pidal (1957:298-299, 2005:558-559) como Diego Catalán (1953) señalaron años ha la existencia de numerosas soluciones lingüísticas occidentales en los manuscritos del Poema de Alfonso XI, lo que hace pensar en el origen leonés de su autor, quien escribía en un leonés occidental, muy mezclado con el gallego o el portugués» (Rodríguez Molina 2010:851).

Crónica de los conquistadores I

Clave	CQ1
Zona geográfica	Navarra y Aragón
Testimonio	Madrid: Biblioteca Nacional de España, MSS/2211 (BETA manid 1776)
Fuente	BiDTEA
Palabras	164.974
Observaciones y apuntes	Subcorpus <i>Textos navarro-aragoneses; Textos cronísticos españoles</i> . Transcrito por: John Nitti y Lloyd A. Kasten.

Obra del *scriptorium* de Juan Fernández de Heredia. Códice incompleto copiado en el siglo XIV (Alvar y Lucía Megías 2002:702). Sobre la heterogeneidad del aragonés de las obras de Heredia, cf. Fernández-Ordóñez (2006).

Crónica de los conquistadores II

Clave	CQ2
Zona geográfica	Navarra y Aragón
Testimonio	Madrid: Biblioteca Nacional de España (IGM), MSS/10134 (2) (BETA manid 3352)
Fuente	BiDTEA
Palabras	300.533
Observaciones y apuntes	Subcorpus <i>Textos navarro-aragoneses; Textos cronísticos españoles</i> . Transcrito por: John Nitti y Lloyd A. Kasten. Obra del <i>scriptorium</i> de Juan Fernández de Heredia. Esta segunda parte se conserva en un solo manuscrito (Alvar y Lucía Megías 2002:703). Sobre la heterogeneidad del aragonés de las obras de Heredia, cf. Fernández-Ordóñez (2006).

Grant crónica de Espanya I

Clave	GC1
Zona geográfica	Navarra y Aragón
Testimonio	Madrid: Biblioteca Nacional de España (IGM), MSS/10133 (BETA manid 1780)
Fuente	BiDTEA
Palabras	463.103
Observaciones y apuntes	Subcorpus <i>Textos navarro-aragoneses; Textos cronísticos españoles</i> . Transcrito por: John Nitti y Lloyd A. Kasten. Obra del <i>scriptorium</i> de Juan Fernández de Heredia. Códice «terminado de copiar por Alvar Pérez de Sevilla, canónigo de Jaén, el 13 de enero de 1385» (Alvar y Lucía Megías 2002:698). Según Enguita Utrilla y Lagüéns Gracia (2015:1793), este texto tiene «clara fisonomía aragonesa, en la que interfiere con mayor o menor fuerza el modelo castellano». Sobre la heterogeneidad del aragonés de las obras de Heredia, cf. Fernández-Ordóñez (2006).

Historia troyana

Clave	HTroyana
Zona geográfica	Navarra y Aragón
Testimonio	Madrid: Biblioteca Nacional de España (IGM), MSS/10801 (BETA manid 1785)
Fuente	BiDTEA
Palabras	58.889

Observaciones y apuntes	Subcorpus <i>Textos navarro-aragoneses; Textos cronísticos españoles</i> . Transcrito por: John Nitti y Lloyd A. Kasten. Obra del <i>scriptorium</i> de Juan Fernández de Heredia. Basada en la <i>Historia destructionis Troiae</i> , de Guido delle Colonne del siglo XIII. Copia de finales del XIV (Alvar y Lucía Megías 2002:710). Sobre la heterogeneidad del aragonés de las obras de Heredia, cf. Fernández-Ordóñez (2006).
--------------------------------	---

Siglo XV

Arcipreste de Talavera o Corbacho

Clave	Corbacho
Zona geográfica	Castilla central y oriental
Testimonio	San Lorenzo de El Escorial: Real Biblioteca del Monasterio de San Lorenzo de El Escorial (RBME), h.III.10 (BETA manid 1977)
Fuente	CORDE
Palabras	75.886
Observaciones y apuntes	Edición: Marcella Ciceri, Madrid: Espasa-Calpe, 1990. Cordemáforo: verde (EC). «Códice fechado el 10 de julio de 1466, fue copiado por Alfonso de Contreras» (Alvar y Lucía Megías 2002:169). «Es conocido que Martínez de Toledo se formó durante varios años de su vida en la Corona de Aragón, estancia que pudo influirle en su propio idiolecto que, por contacto lingüístico, adquiriría algunos rasgos propios del español de la zona, incluso del catalán-aragonés» (Garrido Martín 2016:668).

Atalaya de las corónicas

Clave	Atalaya
Zona geográfica	Castilla central y oriental
Testimonio	Palma de Mallorca: Fundación Bartolomé March, 23/10/1 (BETA manid 3482)
Fuente	Edición de Garrido Martín (2016)
Palabras	152.217
Observaciones y apuntes	«Alfonso Martínez de Toledo empezó a escribir su obra más extensa, la <i>Atalaya de las corónicas</i> , por encargo del rey Juan II de Castilla, en 1443» (Alvar y Lucía Megías 2002:171). «Es conocido que Martínez de Toledo se formó durante varios años de su vida en la Corona de Aragón, estancia que pudo influirle en su propio idiolecto que, por contacto lingüístico, adquiriría algunos rasgos propios del español de la zona, incluso del catalán-aragonés» (Garrido Martín 2016:668). El códice que se ha tomado como base para la edición ha sido caracterizado por Garrido Martín como un texto castellano con impronta oriental (2016:669-70, 811, 946).

Biblia de Arragel

Clave	Arragel
Zona geográfica	Castilla central y oriental
Testimonio	Madrid: Carlos Fitz-James Stuart y Martínez de Irujo, Duque de Alba, Vitrina de Alba (BETA manid 3324)
Fuente	<i>Biblia medieval</i>
Palabras	588.000
Observaciones y apuntes	Traducción llevada a cabo entre 1422 y 1430 por Mosé Arragel, natural de Guadalajara, por encargo de Luis de Guzmán. Varios estudios de Matute han propuesto que se trata de un texto con características del castellano oriental (2013a, 2013b).

Claros varones de Castilla

Clave	Varones
Zona geográfica	Castilla central y oriental
Testimonio	Madrid: Biblioteca Nacional de España, INC/566 (BETA manid 2083)
Fuente	<i>Textos y concordancias</i> de Madison
Palabras	49.437
Observaciones y apuntes	Transcrito por: Michael L. Dangerfield. Obra de Fernando del Pulgar, redactada en el último cuarto del siglo XV. Impresa en Sevilla en 1500. «Galería de retratos de nobles y preladados que participaron en la actividad política y cultural en tiempos de Juan II, de Enrique IV y de los Reyes Católicos» (Alvar y Lucía Megías 2002:522). «Aunque no se sabe con certeza la patria de Pulgar, existen indicios fundados para suponer que era oriundo de Toledo [...]. Si esta suposición fuera cierta, <i>Claros varones de Castilla</i> sería un representante lingüístico del habla toledana de finales del siglo XV con un elevado grado de fiabilidad» (Rodríguez Molina 2010:906).

Crónica de España de Diego de Valera

Clave	Valeriana
Zona geográfica	Castilla central y oriental
Testimonio	Madrid: Biblioteca Nacional de España, INC/1732 (BETA manid 2202)
Fuente	<i>Textos y concordancias</i> de Madison
Palabras	134.478
Observaciones y apuntes	Transcrito por: María Jesús García Toledano. Corregido por: Vicens Colomer. «Nació Diego de Valera en Cuenca en 1412, y murió, seguramente en 1488, en el Puerto de Santa María» (Alvar y Lucía Megías 2002:403). Compuesta por Valera directamente para la imprenta (Sevilla, 1482) (2002:422).

Istoria de las biendanzas e fortunas

Clave	Bienandanzas
--------------	--------------

Zona geográfica	Castilla central y oriental
Testimonio	Madrid: Real Academia de la Historia, 9/2100 (BETA manid 1810)
Fuente	CORDE
Palabras	587.240
Observaciones y apuntes	<p>Edición: Ana María Marín Sánchez, Madrid: Corde, 2000. Cordemáforo: verde (EC).</p> <p>Según Rodríguez Molina y Octavio de Toledo (2017b:11) es una de las «contadas ocasiones» en las que «la Academia ha encargado incluso una transcripción <i>ad hoc</i> para incorporarla al <i>CORDE</i>, suponemos que debido a la falta de edición moderna del texto o a causa de las insatisfacciones que despiertan algunas ediciones».</p> <p>Lope García de Salazar, «nominado primer cronista de Vizcaya» (Alvar y Lucía Megías 2002:824), empezó a escribir <i>Las Bienandanzas</i> en 1471 y debió de terminar poco antes de su muerte en 1476. Este es el testimonio conservado más antiguo, terminado de copiar el 16 de abril de 1492, encargado por Ochoa de Salazar, nieto de Lope (Alvar y Lucía Megías 2002:828-29).</p> <p>«Desde el punto de vista lingüístico, la figura de Lope García de Salazar resulta muy interesante, pues sirve para caracterizar el castellano hablado en la Vizcaya bajomedieval. García de Salazar nació, creció y vivió toda su vida en los valles vizcaínos, por lo que la adscripción geográfica del autor resulta segura» (Rodríguez Molina 2010:897).</p>

Libro de las Propiedades de las cosas

Clave	Propiedades
Zona geográfica	Castilla central y oriental
Testimonio	Madrid: Biblioteca Nacional de España, INC/1884 (BETA manid 1648)
Fuente	<i>Textos y concordancias</i> de Madison
Palabras	460.172
Observaciones y apuntes	<p>Preparado por: María Nieves Sánchez, María Teresa Herrera.</p> <p>Trasladado del francés al romance por Fray Vicente de Burgos. Impreso en 1494 (PhiloBiblon). «Habitualmente se dice que está hecha a partir de la [versión] francesa de Jean Corbechon, lo que es en parte cierto, pero no del todo, porque el traductor utilizó también una fuente latina, al menos en buena parte de su trabajo» (Sánchez González de Herrero 2008:362)</p> <p>«la variedad empleada pertenece sin duda al oriente, posiblemente a la zona aragonesa» [...] «el léxico muestra mezcla de elementos aragoneses, catalanes, quizá occitanos en algún caso, más los galicismos atribuibles al proceso de traducción» (2008:353)</p>

Traducción y glosas de la Eneida I-III

Clave	Eneida
Zona geográfica	Castilla central y oriental
Testimonio	Madrid: Biblioteca Nacional de España, MSS/17975 (BETA manid 1629) +++
Fuente	CORDE
Palabras	298.478
Observaciones y apuntes	Edición: Pedro M. Cátedra, Madrid: Turner Libros, 1994. Cordemáforo: ámbar (EC)
	Enrique de Villena (1384-1434), originario de Cuenca, poeta en las cortes de Castilla y Aragón, escribe obras en catalán y en castellano (Alvar y Lucía Megías 2002:454-55).

Virtuosas e claras mugeres

Clave	Virtuosas
Zona geográfica	Castilla central y oriental
Testimonio	Salamanca: Biblioteca General Histórica de la Universidad de Salamanca, 207 (BETA manid 1938)
Fuente	Edición de Pons Rodríguez (2008)
Palabras	110.716
Observaciones y apuntes	Obra redactada en 1446 por Álvaro de Luna, condestable de Castilla. El código que se utiliza como base en esta edición es del mismo siglo XV.

Visión deleytable

Clave	Visión
Zona geográfica	Castilla central y oriental
Testimonio	Lisboa: Biblioteca Nacional de Portugal, COD. 3337 (BETA manid 3947)
Fuente	CORDE
Palabras	75.909
Observaciones y apuntes	Edición: Jorge García López, Salamanca: Universidad de Salamanca, 1991. Cordemáforo: verde (EC)
	Los testimonios que transmiten la obra de Alfonso de la Torre se agrupan en dos familias, una oriental, con variantes lingüísticas aragonesas, navarras o incluso catalanas; y otra occidental, a la que pertenece el código de esta edición. «La familia occidental nos presenta casi con toda seguridad la prosa más próxima al autor, un castellano culto, lleno de expresiones latinizantes y con una capa ligeramente oriental» (Alvar y Lucía Megías 2002:131).

Crónica de 1344

Clave	Cronica1344
Zona geográfica	León y Castilla occidental
Testimonio	Madrid: Biblioteca Francisco de Zabálburu y Basabe, 11-109 (BETA manid 1506)

Fuente	BiDTEA
Palabras	209.005
Observaciones y apuntes	Subcorpus <i>Textos cronísticos españoles</i> . Transcrito por: José P. da Cruz.
	Solo incluye la primera parte de la crónica, los ff. 1r-106v (PhiloBiblon). Este códice de la primera mitad del siglo XV conserva la segunda redacción de la crónica, traducción al castellano desde el portugués (Alvar y Lucía Megías 2002:320-22).

El Victorial

Clave	Victorial
Zona geográfica	León y Castilla occidental
Testimonio	Madrid: Real Academia de la Historia, 9/5112 (BETA manid 4325)
Fuente	CORDE
Palabras	116.933
Observaciones y apuntes	Edición: Rafael Beltrán Llavador, Madrid: Taurus, 1994. Cordemáforo: ámbar (EC)
	Según Philobiblon, «Beltrán 1997 basa su ed. en RAH 9/5112 (B), controlando el texto con BNE MSS/17648 (A) y Santander Bibl. Menéndez y Pelayo M-103 (=328) (C).»
	Los tres testimonios son copias tardías. Se considera que la obra se redactó entre 1431 y 1435, mientras que el códice base, B, es de finales del siglo XV o de la primera mitad del XVI. A es de la segunda mitad del XV o comienzos del XVI y C de finales del XVI o del XVII (Alvar y Lucía Megías 2002:610). En definitiva, es una obra de transmisión textual compleja y en la edición se han mezclado distintos testimonios.
	Miranda (1993:125), en la introducción a su edición, afirma que «la lengua de <i>El Victorial</i> es el castellano de principios del siglo XV de la provincia de Valladolid, tierra de Díez de Games y de Pero Niño a lo largo de casi toda su vida».

Libro de buen amor

Clave	LBA
Zona geográfica	León y Castilla occidental
Testimonio	Salamanca: Biblioteca General Histórica de la Universidad de Salamanca, Ms. 2663 (BETA manid 2127)
Fuente	BiDTEA
Palabras	53.640
Observaciones y apuntes	Subcorpus <i>Textos poéticos españoles</i> . Transcrito por: Eric Naylor y Steve Kirby.
	Códice copiado por Alfonso de Paradinas, de Salamanca, hacia 1415. Aunque este testimonio es de un siglo después de la

elaboración de la obra, se han destacado su «fidelidad al modelo» (Alvar y Lucía Megías 2002:742) y su «fuerte impronta leonesa» (Rodríguez Molina 2010:838).

Exemplario por ABC

Clave	ABC
Zona geográfica	León y Castilla occidental
Testimonio	Madrid: Biblioteca Nacional de España, MSS/1182 (BETA manid 1526)
Fuente	<i>Textos y concordancias</i> de Madison
Palabras	112.982
Observaciones y apuntes	Transcrito por: J. Coca. <i>Exempla</i> compilados por Clemente Sánchez, bachiller en leyes por Salamanca, canónigo de León y arcediano de Valderas en la primera mitad del siglo XV. Este testimonio es del mismo periodo. «Todo el texto está en castellano, pero con rasgos leoneses debidos posiblemente al copista» (Alvar y Lucía Megías 2002:815).

Un sermonario castellano medieval

Clave	Sermonario
Zona geográfica	León y Castilla occidental
Testimonio	Salamanca: Biblioteca General Histórica de la Universidad de Salamanca, 1854 (BETA manid 2525)
Fuente	CORDE
Palabras	211.668
Observaciones y apuntes	Edición: Manuel Ambrosio Sánchez, Salamanca: Universidad de Salamanca, 1999. Cordemáforo: verde (EC) Sermonario salmantino.

Crónica de Aragón de Vagad

Clave	CroAra
Zona geográfica	Navarra y Aragón
Testimonio	Madrid: Biblioteca Nacional de España, INC/2258 (BETA manid 1816)
Fuente	DiCCA-XV
Palabras	229.678
Observaciones y apuntes	Subcorpus <i>Textos històrics i descriptius</i> . Obra de Gauberto Fabrizio de Vagad, cronista mayor del reino de Aragón. Traducción del latín de Gonzalo García de Santa María. Editado por Pablo Hurus en Zaragoza el 1499. «El aragonésismo de Vagad mana a borbotones por cada página de la obra, y se derrama a través de un retoricismo desbordante y un estilo ampuloso, artificioso, excesivo y muy latinizante» (Rodríguez Molina 2010:923).

Historia de los reyes godos

Clave	Godos
Zona geográfica	Navarra y Aragón
Testimonio	Madrid: Biblioteca Nacional de España, MSS/9055 (BETA manid 1382)
Fuente	DiCCA-XV
Palabras	1.488
Observaciones y apuntes	Subcorpus <i>Textos històrics i descriptius</i> . «Annex al final del llibre factici manuscrit <i>Tratados históricos y genealógicos</i> » (ff 74v-78v) (DiCCA-XV).

Tractado de Roma

Clave	Roma
Zona geográfica	Navarra y Aragón
Testimonio	Madrid: Biblioteca Nacional de España, INC/727 (BETA manid 1967)
Fuente	DiCCA-XV
Palabras	36.445
Observaciones y apuntes	Subcorpus <i>Textos històrics i descriptius</i> . Incunable editado por Pablo Hurus en Zaragoza el 1498 junto con el <i>Viaje de la Tierra Santa</i> .

Viaje de la Tierra Santa

Clave	Viaje
Zona geográfica	Navarra y Aragón
Testimonio	Madrid: Biblioteca Nacional de España, INC/727 (BETA manid 1967)
Fuente	DiCCA-XV
Palabras	138.552
Observaciones y apuntes	Subcorpus <i>Textos històrics i descriptius</i> . Incunable editado por Pablo Hurus en Zaragoza el 1498 junto con el <i>Tractado de Roma</i> .

Vida de los santos religiosos de Egipto

Clave	Vida
Zona geográfica	Navarra y Aragón
Testimonio	New York: Hispanic Society of America (HSA), (BETA manid 1821)
Fuente	DiCCA-XV
Palabras	33.498
Observaciones y apuntes	Subcorpus <i>Textos històrics i descriptius</i> . Incunable editado por Juan Hurus en Zaragoza hacia 1488.

Siglo XVI¹⁵***Belianís de Grecia***

Clave	Belianís
Zona geográfica	Castilla central y oriental
Testimonio	Madrid: Biblioteca Nacional de España, RI/113 (BETA manid 5825)
Fuente	CORDE
Palabras	321.828
Observaciones y apuntes	Edición: Lilia E. F. de Orduna, Kassel: Reichenberger, 1997. Autor: Jerónimo Fernández. Impreso en 1547 en Burgos por Martín Muñoz.

Crónica de la Nueva España

Clave	CNE
Zona geográfica	Castilla central y oriental
Fuente	CORDE
Palabras	379.099
Observaciones y apuntes	Edición: Manuel Magallón, Madrid: Atlas, 1971. Autor: Francisco Cervantes de Salazar.

Espejo de príncipes y caballeros o El caballero de Febo

Clave	Espejo
Zona geográfica	Castilla central y oriental
Fuente	CORDE
Palabras	463.134
Observaciones y apuntes	Edición: Daniel Eisenberg, Madrid: Espasa-Calpe, 1975. Autor: Diego Ortúñez de Calahorra, originario de Nájera (Logroño). «El lenguaje de Ortúñez es el castellano escrito de Castilla la Vieja a mediados del siglo XVI, con <i>h</i> muda y <i>s</i> y <i>z</i> sordas. Se hallan a la vez grafías cultas y populares. Como era de la Rioja, zona lingüística fronteriza, no sorprendería hallar formas dialectales, pero las únicas desviaciones del castellano normal -principalmente falta de diptongación- pueden ser obra del impresor zaragozano» (Eisenberg 2004).

Traducción de la Historia etiópica de los amores de Teágenes y Cariclea de Heliodoro

Clave	Etiópicas
Zona geográfica	Castilla central y oriental
Fuente	CORDE
Palabras	119.914

¹⁵ Todos los textos del siglo XVI son del CORDE. Debido a que este periodo no está completamente cubierto por los recursos utilizados para los siglos anteriores —PhiloBiblon, el Cordemáforo o el diccionario de Alvar y Lucía Megías (2002)—, la información aportada en las fichas es algo más limitada. La zona geográfica de cada uno de estos textos se ha asignado a partir de la localización de provincia que propone el CORDE en su nómina de autores y obras. Se ha prescindido de la información del testimonio en las fichas que no ha sido posible completar con PhiloBiblon.

Observaciones y apuntes	Edición: Francisco López Estrada, Madrid: Real Academia Española, 1954. Traducción de Fernando de Mena, natural de Ciudad Real, en 1587.
--------------------------------	---

Coloquios satíricos

Clave	CSatíricos
Zona geográfica	León y Castilla occidental
Fuente	CORDE
Palabras	89.769
Observaciones y apuntes	Edición: Lina Rodríguez Cacho, Madrid: Turner, 1994. Autor: Antonio de Torquemada, natural de Astorga, León.

Jardín de flores curiosas

Clave	Jardín
Zona geográfica	León y Castilla occidental
Fuente	CORDE
Palabras	127.414
Observaciones y apuntes	Edición: Lina Rodríguez Cacho, Madrid: Turner, 1994. Autor: Antonio de Torquemada, natural de Astorga, León.

Lisuarte de Grecia

Clave	Lisuarte
Zona geográfica	León y Castilla occidental
Fuente	CORDE
Palabras	150.717
Observaciones y apuntes	Edición: Emilio J. Sales Dasí, Alcalá de Henares: Centro de Estudios Cervantinos, 2002

Seis tratados muy devotos y útiles para cualquier fiel cristiano

Clave	Tratados
Zona geográfica	Navarra y Aragón
Fuente	CORDE
Palabras	30.423
Observaciones y apuntes	

Traducción de «Orlando furioso» de Ludovico Ariosto

Clave	Orlando
Zona geográfica	Navarra y Aragón
Fuente	CORDE
Palabras	258.256
Observaciones y apuntes	Edición: Francisco José Alcántara, Barcelona: Planeta, 1988

Para finalizar, en la Tabla 1 podemos encontrar un resumen con la lista de textos del corpus literario y la información básica asociada a cada uno: clave, siglo, zona geográfica, fuente y volumen de palabras.

	Clave	Siglo	GEO	Fuente	Palabras
<i>Biblia Escorial Li.6</i>	E6	XIII	CCO	Biblia Medieval	428697
<i>General Estoria I</i>	GE1	XIII	CCO	BiDTEA	535516
<i>Gran conquista de Ultramar</i>	GCU	XIII	CCO	BiDTEA	247830
<i>Judizios de las estrellas</i>	Judizios	XIII	CCO	BiDTEA	125469
<i>General Estoria IV</i>	GE4	XIII	LCO	BiDTEA	456336
<i>Libro de las cruces</i>	Cruces	XIII	NyA	BiDTEA	130671
<i>Vidal Mayor</i>	Vidal	XIII	NyA	BiDTEA	180278
<i>Leomarte: Sumas de la historia troyana</i>	Sumas	XIV	CCO	BiDTEA	101830
<i>Mostrador de Justicia</i>	Mostrador	XIV	CCO	Mettmann (1996)	389990
<i>Crónica troyana promovida por Alfonso XI</i>	CTroyana	XIV	LCO	D'Ambruoso (2012)	242072
<i>Cuento muy fermoso de Otas de Roma</i>	Otas	XIV	LCO	Madison	43469
<i>Libro de Alexandre</i>	Alexandre	XIV	LCO	BiDTEA	73656
<i>Poema de Alfonso Onceno</i>	Onceno	XIV	LCO	CORDE	42856
<i>Crónica de los conquistadores I</i>	CQ1	XIV	NyA	BiDTEA	164974
<i>Crónica de los conquistadores II</i>	CQ2	XIV	NyA	BiDTEA	300533
<i>Grant crónica de Espanya I</i>	GC1	XIV	NyA	BiDTEA	463103
<i>Historia troyana</i>	HTroyana	XIV	NyA	BiDTEA	58889
<i>Arcipreste de Talavera o Corbacho</i>	Corbacho	XV	CCO	CORDE	75886
<i>Atalaya de las corónicas</i>	Atalaya	XV	CCO	Garrido Martín (2016)	152217
<i>Biblia de Arragel</i>	Arragel	XV	CCO	Biblia Medieval	588000
<i>Claros varones de Castilla</i>	Varones	XV	CCO	Madison	49437
<i>Crónica de España de Diego de Valera</i>	Valeriana	XV	CCO	Madison	134478
<i>Istoria de las bienandanzas e fortunas</i>	Bienandanzas	XV	CCO	CORDE	587240
<i>Libro de las Propiedades de las Cosas</i>	Propiedades	XV	CCO	Madison	460172
<i>Traducción y glosas de la Eneida, I-III</i>	Eneida	XV	CCO	CORDE	298478
<i>Virtuosas e claras mugeres</i>	Virtuosas	XV	CCO	Pons Rodríguez (2008)	110716
<i>Visión deleytable</i>	Visión	XV	CCO	CORDE	75909
<i>Crónica de 1344</i>	Cronica1344	XV	LCO	BiDTEA	209005
<i>El Victorial</i>	Victorial	XV	LCO	CORDE	116933
<i>Libro de buen amor</i>	LBA	XV	LCO	BiDTEA	53640
<i>Exemplario por ABC</i>	ABC	XV	LCO	Madison	112982
<i>Un sermulario castellano medieval</i>	Sermulario	XV	LCO	CORDE	211668
<i>Crónica de Aragón de Vagad</i>	CroAra	XV	NyA	DiCCA-XV	229678
<i>Historia de los reyes godos</i>	Godos	XV	NyA	DiCCA-XV	1488
<i>Tractado de Roma</i>	Roma	XV	NyA	DiCCA-XV	36445
<i>Viaje de la Tierra Santa</i>	Viaje	XV	NyA	DiCCA-XV	138552
<i>Vida de los sanctos religiosos de Egipto</i>	Vida	XV	NyA	DiCCA-XV	33498
<i>Belianís de Grecia</i>	Belianís	XVI	CCO	CORDE	321828
<i>Crónica de la Nueva España</i>	CNE	XVI	CCO	CORDE	379099
<i>Espejo de príncipes y caballeros o El caballero del Febo</i>	Espejo	XVI	CCO	CORDE	463134
<i>Historia etiópica de los amores de Teágenes y Cariclea</i>	HEtiópica	XVI	CCO	CORDE	119914
<i>Coloquios satíricos</i>	CSatíricos	XVI	LCO	CORDE	89769
<i>Jardín de flores curiosas</i>	Jardín	XVI	LCO	CORDE	127414
<i>Lisuarte de Grecia</i>	Lisuarte	XVI	LCO	CORDE	150717
<i>Seis tratados muy devotos y útiles para cualquier fiel cristiano</i>	Tratados	XVI	NyA	CORDE	30423
<i>Traducción de "Orlando furioso" de Ludovico Ariosto</i>	Orlando	XVI	NyA	CORDE	258256

Tabla 1. Resumen de obras y fuentes del corpus de textos literarios.

5.3. Corpus documental auxiliar

El corpus base de esta tesis es el que se ha descrito en las secciones anteriores: los textos documentales del CODEA+ y la selección de textos literarios de diversas fuentes. Sin embargo, en algunas ocasiones se ha ampliado la muestra de textos.

Para completar el *continuum* del norte peninsular e incorporar el resto de lenguas iberorromances, ha sido necesario añadir dos corpus específicos: para el catalán, el *Corpus Informatizat del Català Antic* (CICA); y para el gallego, el *Tesouro Medieval Informatizado da Lingua Galega* (TMILG). En ambos casos he seleccionado las tipologías de tipo documental para poder localizar geográficamente los documentos con más precisión.

Además, en el estudio del reemplazo de *connusco*, *convusco* por *con nós*, *con vós* y, posteriormente, por *con nosotros*, *con vosotros*, ha sido necesario ampliar el corpus

documental para obtener un número suficiente de ejemplos. Para ello, he recurrido al *Corpus Hispánico y Americano en la Red: Textos Antiguos* (CHARTA), a documentos notariales del *Corpus diacrónico del español* (CORDE)¹⁶ y a colecciones documentales digitales o en papel¹⁷.

a) *Corpus Informatizat del Català Antic*

El CICA es un corpus en línea y de acceso abierto dirigido por Joan Torruella junto con Manuel Pérez Saldanya y Josep Martines. Este corpus fue creado con el objetivo de servir de apoyo para la redacción de la gramática del catalán antiguo. Así, se reunieron textos que permitieran apoyar las descripciones con «datos reales y cuantificables» (Torruella Casañas 2009).

Actualmente¹⁸, incluye 414 obras principalmente en catalán de entre los siglos XI y XVIII con un total de 8.656.847 palabras. En todo caso, como puede verse en la Figura 3, el grueso de palabras del corpus se concentra entre la segunda mitad del siglo XIII y el XVI, lo que coincide con el periodo analizado en esta tesis.

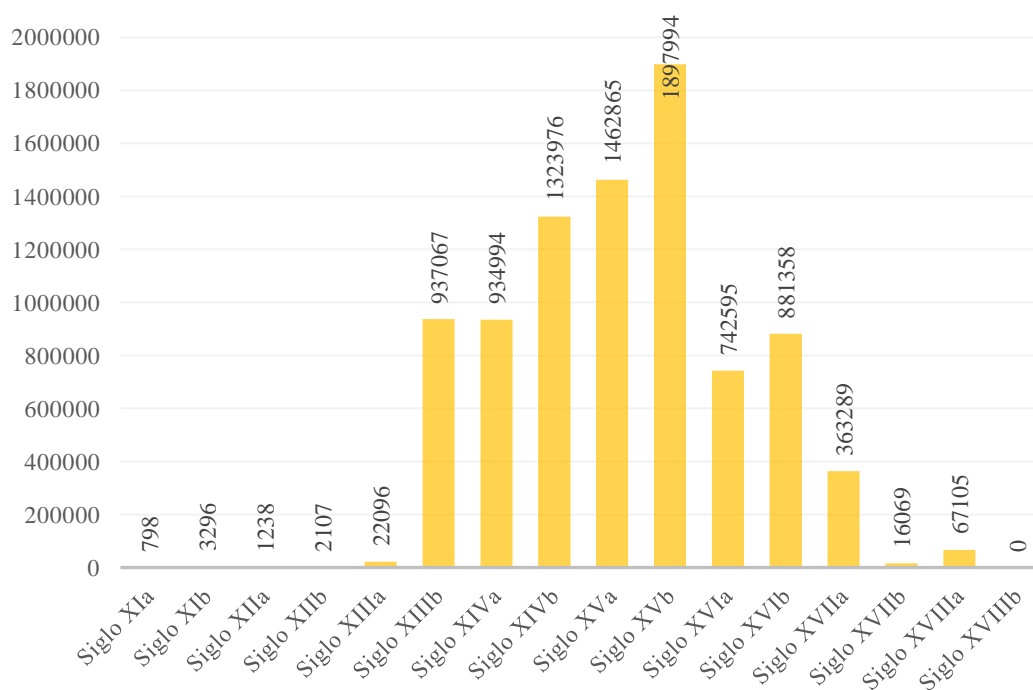


Figura 3. Volumen de palabras por cada medio siglo en el CICA.

Asimismo, el CICA clasifica sus textos en las siguientes tipologías:

- A. Prosa de ficción
- B. Crónicas y obras historiográficas
- C. Obras religiosas y morales
- D. Prosa de la cancillería
- E. Textos administrativos
- F. Textos jurídicos

¹⁶ Del CORDE se han utilizado principalmente los textos de la tipología identificada en el corpus como *10. Documentos notariales*. También se han añadido algunas cartas de compraventa de la tipología *16. Pintura*. En ambos casos, únicamente se han conservado los ejemplos que tuvieran algún tipo de localización geográfica en el índice de obras del CORDE.

¹⁷ Para más información sobre los datos y la metodología específica de cada fenómeno, véase la sección correspondiente en los capítulos 3, 4 y 5, § 3.

¹⁸ Última consulta: 30 de septiembre de 2020.

- G. Libros de la corte
- H. Textos científicos y técnicos
- I. Epistolarios y dietarios
- J. Poesía
- L. Obras gramaticales y lexicográficas

Los tipos de textos con una menor representación en el corpus son la prosa cancelleresca, la poesía y las obras gramaticales y lexicográficas. El resto de tipologías constan de al menos medio millón de palabras (Figura 4).

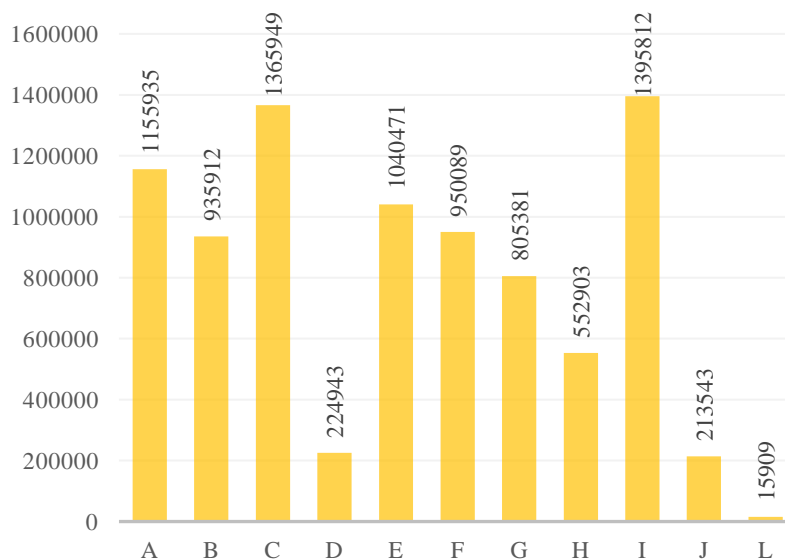


Figura 4. Volumen de palabras por tipología textual en el CICA.

El objetivo de utilizar el CICA es completar el mapa de la difusión de las variantes en el espacio geográfico. Por este motivo, he utilizado únicamente las tipologías textuales que he considerado similares a las del CODEA: *D – Prosa cancelleresca*, *E – Textos administratius*, *F – Textos jurídics*, *G – Llibres de cort*.

En cuanto a la información sobre el origen geográfico, el CICA no incluye la localización concreta del documento, pero sí los clasifica dependiendo del dialecto: catalán oriental (balear, central, septentrional y alguerés) y occidental (valenciano y noroccidental).

b) Tesouro Medieval Informatizado da Lingua Galega

Este corpus de textos medievales en gallego está dirigido por Xavier Varela Barreiro y se integra en el proyecto *Tesouro Medieval Galego-Portugués*, que tiene como objetivo reunir los textos producidos durante la Edad Media en Galicia y Portugal.

Actualmente, el TMILG incluye 82 obras desde finales del siglo VIII al XVII. El grueso del corpus, cerca del 80 % según Pichel y Varela Barreiro (2009:199), está constituido por la prosa notarial. Otro 15 % corresponde a lo que en la página del corpus se denomina como «prosa non notarial», que incluye prosa histórica, literaria, jurídica, religiosa y técnica. Finalmente, el 5 % restante lo componen los textos en verso. Así, del mismo modo que con el CICA, he filtrado los resultados por la tipología de *prosa notarial* que proporciona el mismo corpus con el objetivo de poder cartografiar el origen y la difusión de las variantes en el espacio geográfico.

5.4. Representatividad y equilibrio del corpus

En esta sección he presentado el catálogo de fuentes y testimonios del corpus principal de la tesis, formado por textos documentales y literarios de los siglos XIII al XVI, que he clasificado en tres zonas geográficas: Navarra y Aragón, Castilla central y oriental, y León y Castilla Occidental. Para estos segmentos cronológicos y geográficos dispongo de: 1273 textos documentales con un total de casi 800.000 palabras (CODEA+) y 46 textos literarios que suman más de 9.500.000 palabras. Como puede verse en las cifras, el volumen del corpus literario es mucho mayor que el del documental. Esto se debe a la necesidad de tener una muestra lo suficientemente variada en cada uno de los segmentos cronológicos y geográficos. Así, se evita extrapolar las características de un testimonio a las de un periodo.



Figura 5. Volumen de palabras en cada uno de los siglos y zonas geográficas en textos documentales (CODEA+) y literarios.

En la Figura 5 puede verse un resumen de la proporción del corpus base dependiendo del siglo, la zona geográfica y el tipo de texto. En los textos documentales vemos lo que ya habíamos avanzado en la Figura 2. Aunque es inevitable que la representación de cada zona geográfica oscile siglo a siglo, en general todos los segmentos tienen una representación mínima de un 20 % respecto del resto de zonas, con la excepción del área navarroaragonesa en siglo XVI, que disminuye hasta un 7 %. Por otro lado, los textos literarios muestran una representación más constante siglo a siglo y destaca Castilla central y oriental como el área con un mayor número de palabras. Esto es de esperar, ya que es la zona más amplia y de la que se han conservado más textos. El único momento

en el que esto cambia es en el siglo XIV, cuando la producción del *scriptorium* de Fernández de Heredia inclina la balanza hacia los textos más orientales.

Mediante los criterios establecidos en este capítulo, mi intención ha sido reunir una muestra representativa del habla de los distintos periodos, zonas y tipos de textos seleccionados. No obstante, la elaboración de cualquier corpus diacrónico debe enfrentarse a una paradoja:

Una paradoja de la composición de los corpus diacrónicos es que, por un lado, deben ser heterogéneos (tienen que incluir textos de diferentes autores, épocas, géneros, registros, dialectos) y a la vez deben ser homogéneos (es decir, los diferentes cortes sincrónicos representados en el corpus tienen que ser comparables entre sí) (Enrique-Arias 2012:96).

De un modo u otro, por el simple hecho de tener que poner límites al estudio o de depender de los testimonios conservados, es inevitable incumplir estos principios. Así, respecto a la heterogeneidad, es necesario recordar que, como he mencionado en los criterios y como puede verse al repasar la nómina de textos, he priorizado los textos cronísticos, mientras que solo he incluido unas pocas obras en verso y he dejado de lado los textos científico-técnicos. Así, la mayor parte de los textos literarios del corpus son obras en las que predomina la narración y el diálogo, contextos en los que la aparición de pronombres de primera y segunda persona es más frecuente.

Asimismo, en cuanto a la homogeneidad, uno de los posibles inconvenientes del corpus es el hecho de comparar crónicas medievales con novelas del siglo XVI: «A pesar de que en ambos casos se trata de discurso narrativo, nos encontramos ante obras con convenciones textuales muy diferentes, y en las que la distribución de narración, descripción y diálogo puede diferir significativamente» (Enrique-Arias 2012:96).

En definitiva, a pesar de los obstáculos e inconvenientes, espero haber conseguido un corpus filológicamente fiable y a la vez dialectalmente representativo que sea suficientemente robusto como para sostener conclusiones firmes.

6. Extracción de datos

Como se deduce a partir la descripción de las fuentes y los textos de esta tesis, la innovación metodológica en lingüística de corpus y, concretamente, el desarrollo de los corpus electrónicos son unos de los pilares esenciales de la metodología de este trabajo. La publicación de corpus en línea no solo pone los documentos y testimonios a disposición de los investigadores, sino que en muchos casos van acompañados de una interfaz de búsqueda. Esta es una de las ventajas más evidentes de los corpus digitales, ya que permite rastrear rápidamente las variantes lingüísticas objeto de estudio en vastas colecciones de textos. Sin embargo, esto también tiene un inconveniente: debemos saber exactamente lo que queremos buscar. Por tanto, solo encontraremos aquellas formas que ya conozcamos de antemano y existe la posibilidad que pasemos por alto variantes menos habituales o que han recibido menos atención.

En los corpus convencionales los usuarios introducen en la casilla de consulta las formas que son relevantes para rastrear el fenómeno que desean investigar. Esto significa que es necesario conocer tales formas de antemano por medio de gramáticas, diccionarios históricos o estudios monográficos. La principal desventaja de este proceder es que, por muy bien que hagamos nuestro trabajo previo de investigación de materiales de referencia, siempre existe el riesgo de pasar por alto alguna forma relevante por no haber sido estudiada con anterioridad: es decir, el valor heurístico del corpus queda severamente limitado pues la vía de acceso a los datos solamente nos permite rastrear lo que ya nos es conocido (Enrique-Arias 2012:88).

Esto queda parcialmente solucionado con los corpus lematizados. Los que incorporan búsquedas por lema, permiten extraer todas las formas de un paradigma a través de una sola consulta. No obstante, la lematización automática no siempre puede resolver la ambigüedad de palabras homógrafas o difíciles de clasificar:

los errores se producirán generalmente en los casos de usos más sorprendentes, que pueden también ser muy interesantes —incluso los más interesantes— para el investigador, por lo que en este caso la tecnología puede volverse en nuestra contra, impidiéndonos encontrar resultados dignos de ser considerados (generando falsos negativos, es decir, resultados deseados no encontrados) (De Benito Moreno 2019:5).

En lo que respecta a los fenómenos que tenemos entre manos, el paradigma de los pronombres personales tiene la ventaja de ser una clase cerrada o de cambio lento, por lo que las variantes son limitadas. Solo el caso de los pronombres con modificador como *nós todos* o *vós mismos* podría dar algún problema debido a que no hay una lista cerrada de refuerzos pronominales, pero tiene fácil solución: una vez extraídas todas las variantes gráficas de los pronombres simples *nós* y *vós*, pueden localizarse rápidamente todos los modificadores mediante la lectura del contexto inmediatamente posterior.

La extracción de los pronombres simples *nós* y *vós* ha sido diferente dependiendo del corpus y de su presentación de los textos. En los corpus con presentación crítica, como el CODEA+ y *Biblias Hispánicas*, para extraer los ejemplos de estas variantes de los pronombres tónicos ha sido suficiente con buscar las formas habituales con tilde. El CODEA+ ofrece algunos resultados para otras variantes gráficas como *bós* que también se han añadido al corpus de esta tesis, pero es un número mínimo de ejemplos. Asimismo, en algunas ocasiones, durante la lectura del contexto próximo se han localizado ejemplos de pronombres tónicos que no aparecían tildados en el texto crítico y que se han añadido manualmente durante la codificación.

Estos pequeños detalles del CODEA+ avanzan algunos obstáculos en la extracción de los ejemplos en otros corpus. La BiDTEA, por ejemplo, recoge los textos en una versión semi-paleográfica que no distingue gráficamente entre pronombres tónicos y átonos, y que, por supuesto, presenta la variación gráfica habitual de <v, u, b>. Para la extracción de ocurrencias del CORDE también ha sido imprescindible tener en cuenta todas las posibles variantes debido a la mezcla de transcripciones y ediciones sin un criterio único. Por tanto, en estos casos ha sido necesario realizar consultas que incluyan todas las formas gráficas <nós, nos, vós, vos, uós, uos, bós, bos>, así como realizar una primera codificación rápida para descartar los ejemplos de pronombres átonos mediante el contexto.

En cuanto a la extracción de las formas reforzadas *nosotros* y *vosotros*, además de la variación consonántica mencionada para las simples, se han tenido en cuenta las ocurrencias con <ss> y la flexión de género *-os/as*. No se han hecho búsquedas específicas para los casos en los que el refuerzo *otros* aparece separado gráficamente del pronombre, sino que se han localizado junto con el resto de modificadores durante la codificación de las formas simples.

Los datos extraídos mediante las búsquedas comentadas hasta ahora han sido la base de los tres capítulos de análisis de esta tesis. Además, para el Capítulo 5 —referente al reemplazo de *connusco*, *convusco* por *con nós*, *con vós* y posteriormente por *con nosotros*, *con vosotros*—, se han realizado búsquedas de todas las palabras terminadas en <usco> y <osco> para localizar las formas sintéticas de los pronombres. Aunque mediante este método se obtienen algunos falsos positivos como *busco* o *cognosco*, estos pueden descartarse muy fácilmente durante el análisis de los ejemplos. A cambio, con una búsqueda sencilla se cubren todas las opciones gráficas de los pronombres *connusco* y

convusco, incluso aquellas en las que la preposición aparece separada del resto de la forma.

En general, en los corpus que tienen su propia interfaz de búsqueda, he utilizado esta misma para extraer los datos. Este es el caso del CODEA+, el CICA, el TMILG y el CORDE. Sin embargo, aunque la BiDTEA y *Biblias Hispánicas* tienen su propia herramienta para realizar las búsquedas, también permiten la descarga del texto completo. En estos casos he optado por vaciar los ejemplos directamente de los archivos de texto mediante la herramienta *AntConc* (Anthony 2019), que permite personalizar la extensión de los contextos anterior y posterior, así como filtrar y ordenar los datos de forma avanzada, lo que facilita la posterior codificación de los datos. Este también ha sido el método utilizado para los *Textos y concordancias* de Madison y otras ediciones en PDF que han podido convertirse en archivos de texto.

En último lugar, es necesario mencionar que algunas de las colecciones documentales utilizadas como corpus auxiliar del Capítulo 5 (§ 3) no estaban disponibles digitalmente. Estas ediciones en papel se han escaneado utilizando un escáner de libros profesional¹⁹ y se han procesado con un software de reconocimiento de caracteres OCR²⁰. Soy consciente de que la extracción de las ocurrencias puede no haber sido perfecta debido a cierto margen de error inevitable del reconocimiento de caracteres. No obstante, este margen de error también existe en la transcripción y edición de los corpus digitales, en las interfaces de búsqueda o incluso en la lectura atenta de los testimonios. Confío en que el uso de tecnología adecuada y la selección exhaustiva y cuidadosa de los ejemplos hayan reducido la posibilidad de errores al mínimo.

7. Herramientas de análisis y visualización de datos

Antes de pasar a los capítulos de análisis es necesario mencionar algunas cuestiones técnicas sobre las herramientas de análisis y visualización de datos.

Una vez extraídos los ejemplos a partir de lo descrito en la sección anterior, se han codificado individualmente teniendo en cuenta las particularidades de cada fenómeno, tal y como se indica en la exposición de los datos y la metodología específica de cada capítulo. A partir de estos datos he realizado un análisis exhaustivo y cuantitativo. Para ello, he utilizado el lenguaje de programación estadística R (2021) a través del programa RStudio (2021) y el paquete ‘tidyverse’ (Wickham et al. 2019).

El tratamiento estadístico de los datos de esta tesis es principalmente descriptivo y generalmente me limito a mostrar frecuencias absolutas y relativas. Ocasionalmente, estos datos se presentan en tablas de contingencia, utilizadas para analizar la relación entre dos variables de tipo cualitativo; y se completan con las pruebas estadísticas básicas más habituales en los trabajos de lingüística con variables categóricas²¹. En primer lugar, utilizo la prueba χ^2 de Pearson, a partir de la que obtengo el valor de p , que utilizo para evaluar la significatividad estadística de los datos²². A partir de χ^2 se calculan el coeficiente de *phi* (ϕ), el coeficiente de contingencia y V de Cramer. Mientras que la prueba χ^2 de Pearson y el valor de p nos permiten detectar si dos variables categóricas son independientes o si están asociadas, los coeficientes nos indican la fuerza de esa

¹⁹ Modelo eScan Open System (i2s).

²⁰ ABBYY Finereader 12.

²¹ Para evaluar la significación estadística mediante las pruebas citadas de forma simultánea he utilizado la función `assocstats()` del paquete ‘vcd’ (Meyer, Zeileis, y Hornik 2020). La información obtenida será la que se presente junto con las tablas de contingencia.

²² Para descartar la hipótesis nula, considero el umbral de significatividad moderada ($p < 0.05$) y el de significatividad elevada ($p < 0.01$) (Levshina 2015:12).

asociación con un rango de 0 (sin asociación) a 1 (asociación perfecta) (Levshina 2015:200, 209). Para interpretar el tamaño del efecto, utilizaré los siguientes rangos:

- $0.1 \leq \phi < 0.3$: asociación débil
- $0.3 \leq \phi < 0.5$: asociación moderada
- $\phi \geq 0.5$: asociación fuerte

En cualquier caso, el uso de estas pruebas estadísticas está subordinado a la visualización de los datos. En todas las secciones de análisis me baso principalmente en la presentación visual, ya sea mediante gráficos de líneas o de barras, o mediante mapas. Por tanto, solo recorro a las pruebas estadísticas en aquellos casos en los que la representación gráfica no es suficiente para detectar tendencias claras y evidentes en los datos.

Para el cartografiado de datos en mapas lingüísticos he utilizado el programa QGIS (2021), programa de código abierto y gratuito. Previamente, ha sido necesario obtener las coordenadas geográficas de todos los textos documentales. Estas se han extraído automáticamente mediante el paquete ‘ggmap’ (Kahle y Wickham 2013), que utiliza la API de geolocalización de Google para recuperar las coordenadas geográficas de una lista de localizaciones. He presentado los datos en tres tipos de mapas dependiendo de la fuente de datos y del número de variantes lingüísticas en competición: a partir de los textos documentales, mapas de puntos para registrar el avance de una variante y mapas con gráficos de sectores para representar más de una variante por punto; y, a partir de los textos literarios, mapas con zonas geográficas amplias²³.

²³ Para más detalles sobre los criterios de categorización del *continuum* septentrional y sus criterios véase § 4.3 en este mismo capítulo.

Capítulo 3: La aparición de los pronombres reforzados *nosotros* y *vosotros*

1. Presentación¹

El primer fenómeno que voy a examinar es el del origen y el avance de las variantes pronominales reforzadas *nosotros* y *vosotros*, que reemplazan a las formas simples *nós* y *vós* durante la baja Edad Media². Este reemplazo será central para esta tesis, ya que a su alrededor gravitan los fenómenos de los capítulos siguientes: la presencia de otras variantes de los pronombres con refuerzos como *todos*, *mismos* y *ambos* (Capítulo 4); y el retroceso de las formas pronominales *connusco*, *convusco*, que fueron reemplazados por *con nós*, *con vós* y, posteriormente, por *con nosotros*, *con vosotros* (Capítulo 5).

El reemplazo de los pronombres simples *nós*, *vós* por los reforzados *nosotros*, *vosotros* ha sido ampliamente estudiado: el origen de las formas innovadoras (Alvar y Pottier 1983:122; Lausberg 1973:162), las causas de su aparición, el valor enfático o contrastivo que pudieron tener las variantes reforzadas en su origen (García et al. 1990:72; Gili Gaya 1946:111-12) y las ventajas que pudieron aportar las nuevas formas al paradigma pronominal (Alvar y Pottier 1983:123; García et al. 1990:75-77; Rini 1999:209). No obstante, la mayoría de los trabajos que han abordado esta cuestión se han centrado únicamente en textos literarios.

Además, uno de los aspectos que no ha recibido tanta atención ha sido el origen y la difusión de estas variantes en el espacio geográfico. En su momento, tanto Gili Gaya (1946:109) como Lenz (1925:245) plantearon la similitud del significado de las formas compuestas castellanas con las del francés, que sufren un proceso paralelo que no llega a gramaticalizarse, pero en las décadas posteriores no se desarrolló esta vertiente diatópica. Solo Fernández-Ordóñez (2011:78) ha profundizado más sobre la dimensión geográfica y ha situado el foco del cambio en la antigua Corona de Aragón. Esta autora destaca, además, la influencia directa del catalán y del occitano.

Mi objetivo principal es analizar el origen y la difusión en el tiempo y en el espacio de los pronombres reforzados *nosotros* y *vosotros*. Reexaminaré algunos de los aspectos estudiados anteriormente como la cronología, el orden de aparición y el contexto sintáctico a la luz de nuevos datos. El uso de un corpus que incluye tanto textos literarios como documentales anclados en el espacio geográfico permitirá conocer en detalle el recorrido de las variantes compuestas de los pronombres y establecer los factores que influyeron en su difusión.

La estructura de este capítulo será la siguiente: repasaré los trabajos previos (§ 2), presentaré la metodología y los datos específicos de esta sección (§ 3) y realizaré un análisis completo del reemplazo de las formas simples por las reforzadas (§ 4). En el análisis de este fenómeno revisaré algunas de las cuestiones más comentadas en la bibliografía previa, como las primeras ocurrencias (§ 4.1), la cronología (§ 4.2), el orden de aparición (§ 4.3) y el contexto sintáctico (§ 4.4). Asimismo, examinaré el origen y la difusión de las variantes por el espacio geográfico (§ 4.5) y exploraré en detalle cómo se propagan a través de los distintos ámbitos de emisión de los textos documentales (§ 4.6).

¹ En Gomila Albal (2016) puede encontrarse una versión parcial y preliminar de este capítulo, centrada únicamente en los datos de textos documentales a partir de la versión de CODEA 2011. Los datos actualizados a partir de CODEA+2015 aparecen en Enrique-Arias, Ribas Marí y Gomila Albal (En evaluación). En el presente capítulo presento por primera vez los datos extraídos de textos literarios, así como un análisis comparativo y exhaustivo.

² Es necesario señalar que esta evolución afecta solo a *nós* y a *vós* con referente plural. Estas variantes pueden utilizarse también como formas de tratamiento de cortesía con referente singular, pero estos casos quedan excluidos de esta investigación.

2. Estado de la cuestión

Los pronombres del castellano actual *nosotros* y *vosotros* tienen su origen en los pronombres latinos *NOS* y *VOS*, formas de nominativo y acusativo (Alvar y Pottier 1983:122), que con el tiempo fueron reforzadas en algunas lenguas romances con los continuadores de *ALTEROS* para el masculino y de *ALTERAS* para el femenino (Lausberg 1973:162).

2.1. Causas de la aparición de las formas reforzadas

Uno de los motivos a los que se suele recurrir para explicar la aparición de las formas reforzadas es la necesidad de aportar un valor enfático frente a las variantes simples. Para Nebrija (1492:192), *nós* y *vós* son los pronombres habituales no marcados para la primera y la segunda persona del plural, mientras que *nos otros* y *vos otros* son formas enfáticas. En estas variantes, el elemento *otros*, que aparece separado gráficamente de *nos* y *vos*, es simplemente una partícula enfática similar a *mesmo*:

Esta partezilla mesmo compónese con todos los otros pronombres, como io mesmo; él mesmo, sí mesmo, este mesmo, esse mesmo, él mesmo; mesmo no añade sino una expresión τ hemencia que los griegos τ gramáticos latinos llaman emphasi; τ por esta figura dezimos nos otros, vos otros (Nebrija 1492:192).

Menéndez Pidal (1982:351) se basa en este valor enfático para explicar que, en el momento de su aparición, *nosotros* y *vosotros* se utilizaban para «poner la primera o segunda persona en contraste con otra» hasta que, al generalizarse, fueron perdiendo este valor y se convirtieron en formas únicas, reemplazando las formas simples, que quedaron relegadas a un estilo elevado y cancelleresco.

Según Alvar y Pottier (1983:122), el valor enfático o de contraste de las formas reforzadas se ve favorecido por el propio significado del plural de los pronombres. Los pronombres de primera y segunda persona del plural no son la suma de sus correspondientes singulares. Es decir, *nós* y *vós* no significan ‘varios yo’ y ‘varios tú’, sino ‘yo y otra(s) persona(s)’ y ‘tú y otra(s) persona(s)’³. Por este motivo, el refuerzo – *otros* remarcaría este valor, que ya se encontraría en origen en las formas simples. El valor que aporta este modificador sería, por tanto, redundante (Nowikow 1994:284).

Según Gili Gaya (1946:108-9), la definición de Menéndez Pidal fue interpretada por autores posteriores de forma más restrictiva. Así, Lenz (1925:227 y ss.) plantea la posibilidad de que las formas reforzadas surgieran para distinguir los valores de inclusivo y exclusivo. Esta distinción la encontramos en lenguas que tienen pronombres de primera persona plural específicos para marcar cuándo se incluye o se excluye a la persona a la que se dirige el hablante. De este modo, según Lenz (1925:228), la forma compuesta *nosotros* aparece como variante exclusiva, mientras que *nós* pasa a tener un valor inclusivo. Es decir, *nós*, como variante inclusiva, englobaría tanto al hablante como a su interlocutor, mientras que *nosotros* excluiría al interlocutor, a la segunda persona. Como ya he comentado en el Capítulo 1 (§ 6.3.c), no creo que la distinción inclusivo/exclusivo sea adecuada para explicar la aparición de las formas compuestas, ya que no es aplicable a la segunda persona *vosotros*.

Gili Gaya (1946:110) precisa que es cierto que *nosotros* tiene valor de exclusión, pero no solo puede excluir la segunda persona, sino también una tercera o una cuarta. Por lo tanto, «la persona que hablase, en unión del grupo de que formaba parte, quería

³ Este es el motivo por el que, en castellano, del mismo modo que muchas otras lenguas, los plurales no tienen el mismo tema morfológico que los singulares (cf. Capítulo 1, § 6.3).

diferenciarse de otra persona cualquiera, individual o plural», y esta misma explicación puede aplicarse, ahora sí sin problemas, también a la segunda persona plural.

Las formas compuestas tienen además ciertas ventajas paradigmáticas respecto de las formas simples (Alvar y Pottier 1983:123; García et al. 1990:75-77; Rini 1999:209). La adición de *-otros* permite, en primer lugar, diferenciar estos pronombres tónicos respecto de los átonos *nos* y *vos*, que en su origen eran idénticos a pesar de que más tarde el átono *vos* daría lugar a *os*. En segundo lugar, *nosotros* y *vosotros* especifican el número del referente «tanto lexicalmente (el significado mismo de *otro*) como por su terminación nominal de plural (o sea, *-s*)» (García et al. 1990:75). Finalmente, también tienen la ventaja de la expresión del género (*-otros* y *-otras*).

En definitiva, las causas propuestas para la aparición de las variantes reforzadas *nosotros* y *vosotros* se resumen en la aportación de un valor enfático y en la distinción de rasgos gramaticales de número y género.

2.2. Cronología

Uno de los temas más destacados por todos los estudios previos es el de la cronología. Según Eberenz (2000:58-59) los pronombres *nós* y *vós* junto con el elemento *otros* empiezan a aparecer en los siglos XIII y XIV. Como es habitual, estas primeras apariciones eran esporádicas y tenían el valor enfático o contrastivo del que hemos hablado más arriba. En el siglo XV, en cambio, las formas *nosotros* y *vosotros* ya son formas gramaticalizadas y se van generalizando (2000:59).

Según Gili Gaya, que utiliza para su estudio textos literarios, en el siglo XV «ambas formas habían rebasado ya su significación originaria y tendían a convertirse en las únicas usuales para expresar el plural» (1946:112). El autor destaca, además, que el cambio es bastante brusco, de modo que en los textos del siglo XIV encontramos pronombres *nosotros* y *vosotros* con el valor enfático original, mientras que en los textos del XV ya aparecen directamente las formas gramaticalizadas: «los textos castellanos saltan casi inmediatamente —en poco más de un siglo— desde los plurales exclusivos originarios a la suplantación total de las formas simples por las compuestas» (Gili Gaya 1946:113). Por tanto, la fase de difusión tuvo lugar principalmente durante el siglo XV, de modo que a final de siglo las variantes simples *nós* y *vós* se consideran formas recesivas y marcadas (Eberenz 2000:83).

Finalmente, el cambio se considera consumado en las primeras décadas del XVI. En este punto, las formas reforzadas están generalizadas y han perdido restricciones de uso. Además, *nosotros* y *vosotros* empiezan a aparecer junto con otros modificadores que anteriormente acompañaban a las formas simples, como *ambos*, *mismos*, *todos* y *solos* (2000:62).

En resumen, en el siglo XIII se producen las primeras apariciones de las formas compuestas, durante el siglo XV se difunden y generalizan, y en el XVI ya son las formas habituales. Autores como Fernández-Ordóñez (2011:76), De Jonge y Nieuwenhuijsen (2009), Líbano Zumalacárregui (1991:112), García et al. (1990:65) y Gili Gaya (1946:112) coinciden a grandes rasgos con esta cronología.

2.3. Orden de aparición

Aunque hay bastante acuerdo en la cronología general del fenómeno, no lo ha habido tanto en el orden de aparición de las variantes. Según Lenz (1925:229), *nosotros* apareció en primer lugar con valor exclusivo, mientras que *vosotros* surgió posteriormente como calco analógico: «*Vosotros* no está en igual oposición a *vos* [...]. Su formación será esencialmente debida al modelo de la primera». Gili Gaya niega la hipótesis de Lenz y afirma que el empleo de las dos formas «venía a ser próximamente igual» (1946:112-13).

Para Gili Gaya, lo más probable es que la influencia entre las dos formas fuera recíproca, pero recalca la necesidad de profundizar en el estudio del siglo XIV para «determinar cuál de los dos pudo por su predominio numérico ayudar preferentemente a la propagación del otro» (1946:112).

En cambio, García *et al.* (1990:81-83) observan en sus datos que, por lo general, el refuerzo del pronombre está más extendido en la segunda persona. Según estos autores, esto se debe a que el pronombre que tuvo más motivos para añadir el elemento *-otros* es el de segunda persona. Uno de los motivos es semántico: el significado de la segunda persona *vós* (y, actualmente, *vosotros*) necesariamente excluye al emisor, mientras que la primera persona puede excluir o incluir al interlocutor dependiendo del contexto. Es decir, «la naturaleza de la 2ª p. es intrínsecamente contrastiva, ya que está definida por contraposición con la primera» (1990:80), mientras que *nosotros* es una forma neutra (De Jonge y Nieuwenhuijsen 2009). Esto hace que *vós* sea especialmente compatible con el elemento contrastivo *-otros*. Otro de los motivos es la necesidad de desambiguar el referente. Tanto *nós* como *vós* podían usarse con referentes singulares, ya fuera el *nós* mayestático o el *vós* de cortesía. Sin embargo, el uso de *vós* con referente singular era mucho más frecuente que el de *nós*, por lo que la necesidad de desambiguar era mayor en la segunda persona. Por lo tanto, García *et al.* (1990) concluyen que la primera forma que apareció fue *vosotros*, mientras que *nosotros* se creó por analogía tras la generalización de la segunda persona. El orden de aparición *vosotros* > *nosotros* es el más aceptado por estudios más recientes como el de Eberenz (2000:59) o el de De Jonge y Nieuwenhuijsen (2009).

2.4. Contexto sintáctico

Respecto a la posición sintáctica, no parece haber evidencias claras de una distribución de las variantes reforzadas.

Según Gili Gaya (1946:117), prácticamente la totalidad de los ejemplos tempranos de las formas compuestas aparecen como término de preposición, mientras que solo aparece algún ejemplo particular en función de vocativo y muy raros ejemplos en función de sujeto. A pesar de esto, él mismo justifica sus datos explicando que es normal que se elida el sujeto en castellano.

Eberenz (2000:71) afirma que en el siglo XV no hay diferencias entre la aparición de las formas reforzadas en posición de sujeto y en término de preposición. No obstante, si revisamos los datos de *nós* y *nosotros* para este siglo de su tabla 7, podemos ver que el 58 % de los ejemplos en posición de sujeto se corresponden con las formas reforzadas. Esta situación se invierte en término de preposición, donde predominan las formas simples con un 62 %. Por tanto, según los datos de Eberenz, la posición de sujeto parece favorecer ligeramente la aparición de la variante compuesta, aunque el autor no considera que sea una diferencia relevante.

Al margen de la posición sintáctica, Nowikow (1994:293) señala algunos contextos recurrentes de las variantes simples y reforzadas. Frecuentemente, *vós* aparece en estructuras especificadoras, «con un elemento singularizador», como *cualquiera de vós*, *cada uno de vós* o *ninguno de vós*. En cambio, *vosotros* tiende a aparecer en sintagmas generalizadores con un carácter menos específico, como *a vosotros*, *con vosotros* o *todos vosotros*.

2.5. Presencia de las variantes reforzadas en otras variedades

Además del castellano, otras lenguas y dialectos romances refuerzan sus pronombres de primera y segunda persona del plural⁴. De las variedades que incorporan este refuerzo, algunas lo terminan gramaticalizando, y otras simplemente lo mantienen con valor enfático.

En la actualidad, podemos encontrar las formas gramaticalizadas *nosautres* y *vosautres* en occitano, y *nosaltres* y *vosaltres* en catalán, aunque estas variantes no tienen distinción de género. Según Gili Gaya (1946:113), estas formas aparecen antes en catalán que en castellano, pero su extensión es más lenta y progresiva, por lo que el cambio se consume en el siglo XV de forma paralela en las dos lenguas. Sin embargo, según Fernández-Ordóñez (2011:77), el cambio empezó y culminó antes en catalán y en occitano. Así, según sus datos, el occitano usa *nosautres* y *vosautres* como formas no marcadas desde la Edad Media y el catalán registra *nosaltres* y *vosaltres* en la obra de Ramon Llull (1232-1315). *Vosaltres* es ya abundante a mediados del siglo XIII, momento en el que *vosotros* todavía aparece de forma esporádica. A esto hay que sumar la cantidad de ejemplos de formas compuestas que encuentra esta autora en el aragonés escrito a finales del siglo XIV, que es mucho mayor que los que presenta el castellano en la misma época (Fernández-Ordóñez 2011:nota 98).

Tras la gramaticalización, en ambas lenguas se documenta variación diatópica que atestigua el desgaste fonológico de la estructura. En catalán encontramos variantes como *nosatres*, *natres*, *nantres*, *nosaltros*, *noltros* o *naltros*; y en occitano, *nosaus*, *nautres*, *naltres*, *natres* o *nantres* (Nowikow 1994:288). Esta variación muestra que los dos elementos que componen las formas compuestas se han ido fundiendo hasta el punto de que, en algunos casos, el modificador *-altres* prácticamente no se puede identificar.

Las opciones reforzadas existen además en otras lenguas, pero estas las mantienen como opción enfática. Es el caso del francés (*nous autres*, *vous autres*), el italiano (*noialtri*, *voialtri*), el sardo (*noisáteros*, *boisáteros*) y el retorromance (*noaltris*, *voaltris*) (Fernández-Ordóñez 2011:76; Lausberg 1973:162; Meyer-Lübke 1974:103; Posner 1998:113).

Respecto al francés, el uso actual de *nous autres* y *vous autres*, que ha permanecido sin gramaticalizarse, es muy similar al que encontramos en los ejemplos contrastivos en el castellano del siglo XIV y de principios del XV (De Jonge y Nieuwenhuijsen 2009). En todo caso, es necesario preguntarse por qué en francés no llegaron a gramaticalizarse. Hay que tener en cuenta que, debido a la evolución fonética de las desinencias verbales del francés, el pronombre personal en función de sujeto empieza a suplir la función de las desinencias: «Tales pronombres quedan entonces átonos en proclisis, y en esta posición la resistencia a la añadidura de *autres* (necesariamente acentuado) tenía que ser muy grande» (Gili Gaya 1946:117).

En cuanto al italiano, según Nowikow (1994:288-289), las formas *noialtri* y *voialtri* tienen variación en distintas épocas y dialectos: *nojaltri* y *vojaltri* en Ancona; *ñaut* y *vjaut* en Piemonte; *nuàtri*, *nuàutri* y *vuàtri*, *vuàutri* en Sicilia; y *nuàttri* y *vuàttri* en Calabria. Además, *altro* o *altri* pueden aparecer también junto a pronombres de segunda y tercera persona: *tuatru*, *tuatru*, *tuàutro* en Calabria; *lùjatri*, *lujatri* en Liguria; y *loráut* y *lurác* en Piemonte. Es necesario destacar que algunos dialectos solo presentan como forma enfática la segunda persona *voialtri*. Esto sería un indicio de que el cambio tuvo su origen en la segunda persona.

⁴ Las lenguas romances no son las únicas que refuerzan los pronombres personales: «el fenómeno del reforzamiento pronominal rebasa los límites de lo románico y parece tener (aunque se trate de los empleos ocasionales y menos sistemáticos) un carácter más universal» (Nowikow 1994:283).

En contraste con estas lenguas y dialectos orientales, encontramos otros en la mitad occidental de la península ibérica que todavía conservan los pronombres *nós* y *vós*, como el bable occidental y las hablas de algunos puntos de Galicia y del norte de Portugal (Alvar y Pottier 1983:123; Fernández-Ordóñez 2011:78; Gili Gaya 1946:116). En estas zonas el cambio todavía se encuentra en proceso de difusión, por lo que conviven variantes simples y reforzadas (Alvar 1996:146; Álvarez Blanco 1994:25-26; Álvarez Fernández-Cañedo 1963; Lapesa 1981:487-88; Rodríguez Castellano 1952:121; Zamora Vicente 1970:169).

Todos estos hechos apuntan al origen oriental de los pronombres *nosotros* y *vosotros*, y cabe suponer que estas variantes llegaron al castellano desde el oriente de la península ibérica, a través del catalán y el occitano.

2.6. Recapitulación: preguntas de investigación

De las cuestiones que hemos repasado en este estado de la cuestión, podemos extraer las preguntas de investigación que trataré de responder durante mi análisis. Sobre el valor enfático que aportan las variantes innovadoras *nosotros* y *vosotros*, ¿en qué casos podemos detectarlo? ¿Qué tipos de contrastes se establecen en nuestros textos? Sobre la distinción de número que aporta la adición de *otros*, ¿hasta qué punto puede haber sido determinante en la aparición de las nuevas formas reforzadas? En cuanto a la cronología, toda la bibliografía coincide en que el reemplazo de *nós*, *vós* por *nosotros*, *vosotros* empieza y termina entre los siglos XIII y XVI. ¿Es así en todos los tipos de texto? Respecto al orden de aparición, los estudios más recientes concluyen que fue la segunda persona la que inició el cambio. ¿Hasta qué punto *vosotros* se extendió antes que *nosotros*? ¿Por qué la segunda persona adopta antes el refuerzo *otros*? En cuanto al contexto sintáctico, ¿hay realmente una posición sintáctica —sujeto o término de preposición— que favorezca el uso de las variantes de los pronombres reforzados? Desde un punto de vista más amplio, ¿hay algún tipo de estructura recurrente que favorezca la aparición de una u otra variante? Y, finalmente, respecto a la difusión de *nosotros*, *vosotros* desde el oriente peninsular, ¿cuánto tardan las variantes innovadoras en difundirse por todo el *continuum* septentrional? ¿En qué punto de la difusión se encuentran en catalán? ¿Hasta dónde llegan en gallego?

3. Datos y metodología específica

Tal y como hemos visto en el capítulo anterior (Capítulo 2, § 4), la mayoría de las investigaciones sobre el reemplazo de *nós*, *vós* por *nosotros*, *vosotros* citadas en el estado de la cuestión se basan únicamente en fuentes literarias y no se tiene en cuenta la procedencia dialectal de estas. Para este análisis, he hecho una selección rigurosa de los testimonios que forman parte del corpus de estudio. He priorizado testimonios originales o próximos al original y he tenido en cuenta su procedencia dialectal. Además, he incorporado textos documentales, lo que permitirá en cartografiar en detalle el avance de las variantes en el espacio geográfico. La variedad del corpus en combinación con la perspectiva diatópica aportará nuevos datos imprescindibles para completar el cuadro sobre el origen y la difusión de las formas *nosotros* y *vosotros* en castellano.

El corpus principal de textos documentales de este capítulo es el *Corpus de documentos anteriores a 1800* (CODEA+) (para una descripción del corpus cf. Capítulo 2, § 5.1). En primer lugar, he extraído todas las variantes gráficas de las formas simples *nós* y *vós*. He descartado todos aquellos casos que, aunque aparecieran tildados en el texto crítico, fueran en realidad átonos. Al mismo tiempo, durante la codificación he ido añadiendo manualmente algunos ejemplos que aparecían en el contexto próximo y que estaban sin tildar pese a ser tónicos.

Las formas compuestas presentan una mayor variación gráfica y, además, tienen distinción de género, por lo que realicé una búsqueda con las siguientes variantes: *nosotros*, *nosotras*, *vosotros*, *vosotras*, *bosotros*, *bosotras*, *nossotros*, *nossotras*, *vosotros*, *vossotras*, *bossotros*, *bossotras*. Asimismo, la búsqueda de las formas simples *nós* y *vós* también ha arrojado resultados de formas compuestas en las que el pronombre y el modificador *otros* aparecían separados en el texto crítico: *nós otros*, *vós otros*, *bós otros*.

Tras las búsquedas se obtuvieron un total de 9.203 ejemplos de todas las variantes entre los siglos XIII y XVII. Antes de empezar propiamente con la codificación, ha sido necesario desambiguar el referente de todas las formas simples consultando el contexto próximo para distinguir los que tuvieran referente plural. Así, para este análisis solo he utilizado los datos de *nós* y *vós* que tienen como referente a más de una persona y he dejado de lado el uso de estos pronombres como formas de tratamiento con referente singular.

Además de los ejemplos claramente plurales, se han hallado ejemplos con referentes colectivos, como «vós, el concejo» o «nós, el cabildo», que se han marcado como tales durante la codificación (261 casos). En el análisis, estos ejemplos se han englobado dentro de la etiqueta general de plurales por dos motivos: en primer lugar, la frecuencia de este tipo de referentes no es tan alta como para hacer un análisis separado, al menos por ahora; y, en segundo lugar, se ha observado que en muchas ocasiones otros elementos de la oración concordaban semánticamente en plural con los pronombres:

- a) que nós el cabildo seamos tenudos de dexar esta donación (CODEA-0073)
- b) E nós conuiento otorgamos que somos pagados con aquel heredamiento (CODEA-0709)

Una vez eliminados los casos con referente singular o que en realidad fueran átonos, quedaron 5.466 ejemplos.

Para el cartografiado de datos se han descartado todos los ejemplos que no tuvieran ningún tipo de localización. Si no había información concreta sobre la población del documento, se ha situado en la capital de la provincia. Los documentos que tuvieran una localización conjetural (marcada con un interrogante en el índice del corpus) se han mantenido en dicha localización.

El corpus de textos literarios utilizado para el análisis de este fenómeno es el descrito en el Capítulo 2 (§ 5.2). Debido a que muchos de los textos son versiones semipaleográficas, además de las búsquedas del CODEA, se han buscado las correspondientes variantes gráficas sin tildar o con <u> consonántica, con las que se han obtenido un total de 30.289 ejemplos. Después de descartar los casos de pronombres átonos o falsos positivos, han quedado 16.988 ejemplos.

Este alto número de casos ha imposibilitado codificar exhaustivamente el referente singular o plural de todas las ocurrencias, ya que esto suele requerir la lectura de un contexto amplio. Ante esto, he decidido extraer una selección de 2.000 ejemplos sobre la que hacer el análisis. Para ello, he buscado que cada texto tuviera una muestra proporcional a su peso en número de palabras dentro del corpus. De este modo, un texto como *E6*, de 428.697 palabras, supone un 4.46 % del corpus. Por tanto, extraería aleatoriamente el equivalente al 4.46 % de los 2.000 ejemplos que forman la muestra: 89 ocurrencias. En la Tabla 2 puede verse el cálculo de la muestra para cada una de las obras que forman el corpus de textos literarios.

La extracción se ha hecho obra a obra. Primero se han ordenado los ejemplos aleatoriamente y luego se han ido identificando los pronombres con referente plural o colectivo hasta que se ha alcanzado el número de la muestra. En algunas obras, el número

de ejemplos plurales ha sido inferior al establecido por los criterios. En ese caso, se ha indicado entre paréntesis en la Tabla 2 el número real de ocurrencias obtenidas. En definitiva, la muestra aleatoria de los textos literarios que será la base para el análisis de este capítulo está formada por 1.933 ejemplos y está distribuida de forma proporcional al peso de cada obra en el corpus.

<i>Siglo</i>	<i>Texto (clave)</i>	<i>Palabras</i>	<i>Peso en el corpus (%)</i>	<i>Ejemplos de la muestra</i>
<i>XIII</i>	E6	428.697	4.46	89
	GCU	247.830	2.58	52
	GE1	535.516	5.58	112
	Judizios	125.469	1.31	26 (10)
	GE4	456.336	4.75	95
	Cruzes	130.671	1.36	27 (5)
	Vidal	180.278	1.88	38 (35)
<i>XIV</i>	Mostrador	389.990	4.06	81
	Sumas	101.830	1.06	21
	Alexandre	73.656	0.77	15
	CTroyana	242.072	2.52	50
	Onceno	42.856	0.45	9
	Otas	43.469	0.45	9
	CQ1	164.974	1.72	34
	CQ2	300.533	3.13	63
	GC1	463.103	4.82	96
	HTroyana	58.889	0.61	12
<i>XV</i>	Arragel	588.000	6.12	122
	Atalaya	152.217	1.59	32 (16)
	Bienandanzas	587.240	6.12	122
	Corbacho	75.886	0.79	16
	Eneida	298.478	3.11	62
	Propiedades	460.172	4.79	96
	Valera	134.478	1.40	28
	Varones	49.437	0.51	10
	Virtuosas	110.716	1.15	23
	Visión	75.909	0.79	16
	ABC	112.982	1.18	24
	Cronica1344	209.005	2.18	44
	LBA	53.640	0.56	11
	Sermonario	211.668	2.20	44
	Victorial	116.933	1.22	24
	CroAra	229.678	2.39	48
	Godos⁵	1.488	0.02	0
	Roma	36.445	0.38	8 (5)
	Viaje	138.552	1.44	29
	Vida	33.498	0.35	7
<i>XVI</i>	Belianís	321.828	3.35	67
	CNE	379.099	3.95	79
	Espejo	463.134	4.82	96
	HEtiópica	119.914	1.25	25
	CSatíricos	89.769	0.93	19
	Jardín	127.414	1.33	27
	Lisuarte	150.717	1.57	31
	Orlando	258.256	2.69	54 (49)
	Tratados	30.423	0.32	6
		<i>9.603.145</i>	<i>100</i>	<i>2.000 (1.933)</i>

Tabla 2. Número de palabras, peso en el corpus y muestra aleatoria para cada texto literario.

⁵ A causa de su bajo peso en el corpus y de acuerdo con los criterios de representación de la muestra, no se ha extraído ningún ejemplo de *Godos* para este capítulo, pero se ha mantenido para los siguientes.

Además de los textos documentales y literarios en castellano, he utilizado el corpus auxiliar de textos en catalán y en gallego para completar el análisis del origen y la difusión de las variantes en el espacio geográfico: el *Corpus Informatitzat del Català Antic* (CICA) y el *Tesouro Medieval Informatizado da Lingua Galega* (TMILG).

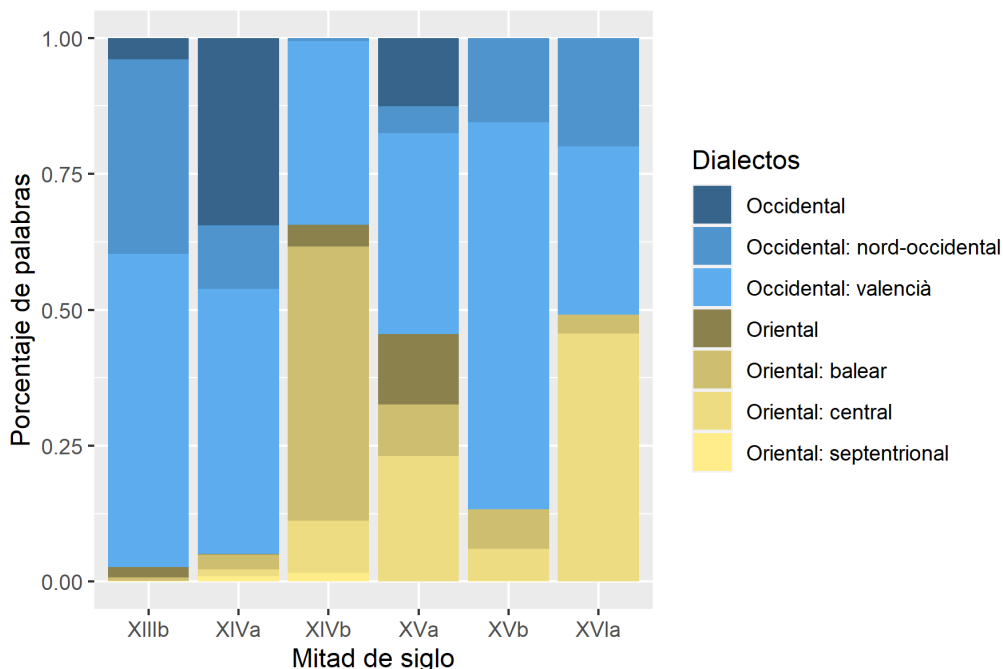


Figura 6. Volumen de palabras por dialecto en los textos documentales del CICA.

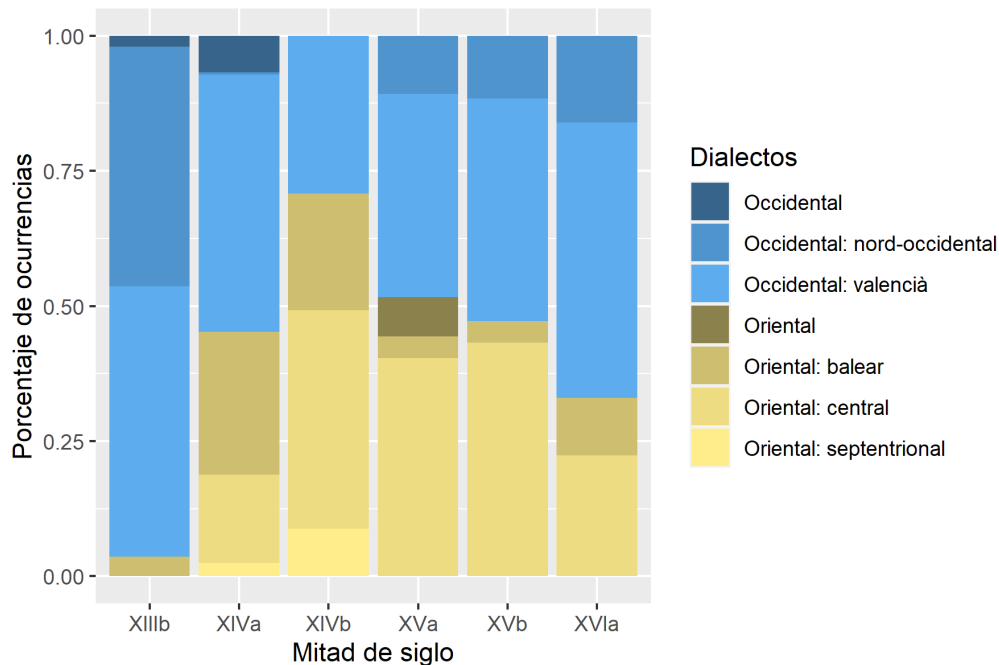


Figura 7. Volumen de ocurrencias por dialecto en la muestra aleatoria de los textos documentales del CICA.

En cuanto al CICA (*cf.* Capítulo 2, § 5.3.a), para este capítulo he utilizado únicamente las tipologías textuales que he considerado similares a las del CODEA: D – Prosa cancelleresca, E – Textos administratius, F – Textos jurídics, G – Llibres de cort. Tras buscar todas las variantes gráficas de los pronombres objeto de estudio se obtuvieron

9.133 ejemplos. Para poder codificar los ejemplos de forma más exhaustiva y desambiguar entre referentes singulares y plurales, decidí hacer una selección aleatoria de hasta 150 ejemplos por cada medio siglo. En un primer análisis observé que la muestra estaba desequilibrada y que los textos de los dialectos occidentales estaban sobrerrepresentados en contraste con los orientales. Esta asimetría está causada por la propia composición del corpus. Tal y como puede verse en la Figura 6, en la mayoría de los periodos cronológicos la representación de los textos occidentales, especialmente los valencianos, es considerablemente superior. Para solucionar esto y equilibrar un poco la muestra (Figura 7), además de los 150 ejemplos cada medio siglo ya mencionados, saqué otros 100 exclusivamente de textos orientales.

El resultado es una muestra de 1.422 ejemplos⁶ de entre los siglos XIII y XVI, de los cuales 605 tienen referentes plurales o colectivos (Tabla 3).

	XIIIa	XIIIb	XIVa	XIVb	XVa	XVb	XVIa	XVIb	Total
Muestra extraída	26	250	250	250	250	250	112	34	1.422
Variantes con referente plural	10	79	66	101	116	134	77	22	605

Tabla 3. Ejemplos extraídos aleatoriamente de los textos documentales del CICA y número de ocurrencias de pronombres con referente plural (*nós, vós, nosaltres, vosaltres*).

En cuanto al TMILG (*cf.* Capítulo 2, § 5.3.b), las búsquedas han sido similares a las de los corpus anteriores y se han filtrado por la tipología de «prosa notarial» que proporciona el mismo corpus. El mayor obstáculo que se ha encontrado durante el trabajo con el TMILG es que se ha detectado una elevada cantidad de ejemplos que no aparecían en los resultados de las búsquedas pero sí en el contexto anterior o posterior de las ocurrencias listadas y que han tenido que añadirse manualmente⁷.

El resultado han sido 1.378 ejemplos desde la segunda mitad del siglo XIII hasta finales del XV, por lo que se han analizado en su totalidad, sin necesidad de extraer una muestra. Después de descartar los casos con referente singular, ambiguos⁸ o falsos positivos, han permanecido 851 ejemplos.

4. Análisis

4.1. Primeras ocurrencias

Los primeros casos de formas reforzadas aparecen en el siglo XIII tanto en los textos documentales del CODEA+ como en el corpus de textos literarios, lo que coincide con lo establecido en el estado de la cuestión. En este siglo, únicamente encontramos un texto de cada corpus que documente estas variantes.

En el corpus de textos literarios, encontramos los primeros ejemplos en la *Gran Conquista de Ultramar*, que podemos situar en los últimos años del XIII (Alvar y Lucía Megías 2002:605):

(1)

- a) & si **nos otros** fuéremos muertos o presos. los moros tomaran las duennas (GCU, fol. 219v, *BiDTEA*)

⁶ Los periodos en los que no se llega a los 250 ejemplos (la primera mitad del siglo XIII y las dos mitades del XVI) no tienen más ocurrencias de las variantes pronominales.

⁷ Además, para el Capítulo 4 realicé búsquedas expresamente de los pronombres acompañados por modificadores distintos a *otros/outros*, como *todos, mismos* y *ambos*. El corpus devolvió ejemplos que no habían aparecido en las consultas originales de las formas simples hechas para este capítulo. Estos casos fueron incorporados a los datos que aquí se presentan, pero esto plantea la posibilidad de que haya ejemplos de las variantes simples *nós* y *vós* que no hayan aparecido en ninguna de las búsquedas.

⁸ Este ha sido el único corpus en el que no ha sido posible desambiguar el referente de todas las ocurrencias, ya que la interfaz limita en gran medida el acceso al contexto.

- b) & sera grant desondra del Rey de francia & de nos. [...] & **uos otros** que dezides a esto (GCU, fol. 249r, *BiDTEA*)

Algo más tempranas son las ocurrencias de formas reforzadas del CODEA+ que encontramos en un documento de 1265. Se trata de una carta emitida por el obispo de Pamplona⁹:

- (2)
- a) Entendimos que algunos de **vosotros** capellanos e clérigos non queredes recibir a los demandadores de Sant Pedro de Ribas (CODEA-0979)
 - b) aquella cosa amamos especialment sobre todas las otras e **vós otros** la devedes amar (CODEA-0979)
 - c) e coando las avrán feitas mandamos a **vosotros** capellanos que prengades uno o dos de vuestros parroquianos (CODEA-0979)
 - d) e si por aventura fuesse ninguno de **vosotros** capellanos o clérigos rebelles
 - e) assí que por el bien ni la mercé que **vosotros** a ellos faredes podades ganar la gloria de paraíso (CODEA-0979)

En ninguno de estos ejemplos podemos argumentar que haya valor exclusivo. La oposición inclusivo/exclusivo se usa para diferenciar aquellos contextos en los que se incluye al interlocutor en la primera persona del plural. Es decir, *nosotros* puede hacer referencia a todos los participantes del acto comunicativo o puede excluir a la segunda persona. El único ejemplo de *nosotros* que encontramos en estas primeras ocurrencias es el de (1a). Sin embargo, el contraste que establece este pronombre no es respecto a una segunda persona, sino a una tercera, «los moros». Por tanto, podríamos hablar de valor enfático o contrastivo en un sentido amplio, no de valor exclusivo.

El valor enfático no siempre está claro y no ha habido hasta la fecha estudios que hayan sido capaces de captar inequívocamente el valor contrastivo de los ejemplos más tempranos. Es difícil argumentar que el contraste se deba específicamente a la adición del modificador o simplemente a la aparición del sujeto explícito, como es el caso de a) y b) en (1) o b) y e) en (2).

Además, el contraste puede establecerse de diferentes maneras o en diferentes direcciones. Por un lado, en ejemplos como en (2b) vemos una clara oposición entre el referente de *vosotros*, los «capellanos e clérigos», y la primera persona, el obispo que escribe el documento. El obispo se dirige específicamente a aquellos que se han negado a recibir a los monjes de San Pedro de Ribas que piden limosna, a diferencia de como hubiera hecho él mismo. Por otro lado, si observamos el uso de la forma simple *vós* en el mismo texto, vemos que también puede definirse un contraste entre el referente de *vosotros*, un grupo reducido al que se dirige el obispo; y el de *vós*, un grupo más amplio no solo de clérigos, sino también de abades y rectores:

- (2)
- f) vos mandamos a **vós** los abades, e a los rectores, e a los capellanos e a todos los clérigos establecidos por todo nuestro bispado en virtud de obediencia e deius pena d'escomulgamiento que coando acayecieren a los devantditos cuestores que **vós** los recebades benignament en vuestras eglesias e en vuestros logares (CODEA-0979)

En resumen, las primeras documentaciones tienen lugar en la segunda mitad del XIII tanto en el corpus literario como en el documental. En estas primeras apariciones no siempre

⁹ Doc. 979, “Carta por la que el obispo de Pamplona da licencia para que los monjes del monasterio de San Pedro de Ribas pidan limosna por todo el obispado” (17/08/1265). Archivo Histórico Nacional.

se refleja con claridad el valor contrastivo, aunque, dado el caso, el contraste puede establecerse tanto hacia alguno de los interlocutores como hacia una tercera persona.

4.2. Cronología

Los dos ejemplos de formas reforzadas citados en (1) se corresponden con el 0.5 % de pronombres de primera y segunda persona de plural del siglo XIII de la muestra del corpus literario. A continuación, en la Tabla 4 podemos ver la evolución completa del reemplazo de las variantes simples *nós* y *vós* por las reforzadas *nosotros* y *vosotros* en este tipo de textos.

	XIII	XIV	XV	XVI
<i>Nós, vós</i> (plural)	99.5 % (395/397)	82.1 % (320/390)	57.3 % (428/747)	12.3 % (49/399)
<i>Nosotros, vosotros</i>	0.5 % (2/397)	17.9 % (70/390)	42.7 % (319/747)	87.7 % (350/399)

Tabla 4. Frecuencias de *nós*, *vós*, *nosotros* y *vosotros* desde el siglo XIII al XVI en la muestra de textos literarios (N = 1933).

A grandes rasgos, la evolución de las variantes en la muestra refleja una evolución similar a la ya descrita por los estudios anteriores. Tras las primeras ocurrencias en el siglo XIII, las formas reforzadas avanzan progresivamente y rebasan la fase inicial del cambio en el siglo XIV con un 18 % de aparición. La fase intermedia la podemos situar en el siglo XV, momento en el que se aproxima al 50 %. Finalmente, en el XVI el cambio se puede considerar completo con un 88 % de formas reforzadas. De acuerdo con Gili Gaya (1946:113), el cambio es rápido y relativamente brusco, especialmente en los siglos XV y XVI.

	1251-1300	1301-1350	1351-1400	1401-1450	1451-1500	1501-1550
<i>Nós, vós</i> (plural)	99.4 % (994/1000)	99.4 % (809/814)	97.3 % (974/1001)	95.6 % (922/964)	90.2 % (572/634)	82.4 % (488/592)
<i>Nosotros, vosotros</i>	0.6 % (6/1000)	0.6 % (5/814)	2.7 % (27/1001)	4.4 % (42/964)	9.8 % (62/634)	17.6 % (104/592)

Tabla 5. Frecuencias de *nós*, *vós*, *nosotros* y *vosotros* desde 1251 a 1550 en CODEA+ (N = 5005)¹⁰.

En cambio, el avance de los pronombres *nosotros* y *vosotros* en los textos documentales sigue un ritmo diferente. En el CODEA+ se han obtenido 5.005 formas entre 1251 y 1550. De estas, un 95 % se corresponden con las formas simples, mientras que las formas compuestas apenas llegan a un 5 % en total. Tal y como puede verse en la Tabla 5, el avance de las nuevas variantes es radicalmente más lento en este tipo de textos, de modo que la fase inicial del cambio se prolonga hasta el siglo XVI. En ese momento, en los textos literarios el cambio se ha completado, mientras que en los textos documentales las variantes reforzadas no llegan al 18 % de aparición.

Por tanto, parece evidente que el cambio no se difunde a la vez en todos los tipos de texto (Figura 8). Estas diferencias pueden estar causadas por el tipo de documentos incluidos en el CODEA+: textos legislativos, cartas de compraventa, actas y declaraciones, etc.; que favorecen la aparición de ciertas fórmulas fijas, especialmente en el encabezado o el cierre, que en muchos casos incluyen pronombres. De este modo, las

¹⁰ Estos datos difieren ligeramente de los presentados en Enrique-Arias, Ribas Marí y Gomila Albal (En evaluación) para el estudio de *nós*, *vós*, *nosotros*, *vosotros*, ya que en esta ocasión he decidido mantener las ocurrencias sin data tópica en las secciones en las que este dato no es imprescindible.

variantes innovadoras tardarían más en difundirse en contextos como los que vemos en (3), extraídos de documentos del siglo XVI:

- (3)
- a **vós** el concejo, justicia, regidores, caballeros, escuderos, oficiales e omnes buenos de la villa de Vilvao, salud e gracia (CODEA-1425, Bilbao, 1506).
 - a vuestro alcalde en el dicho oficio e a cadaúno de **vós** a quien esta mi carta fuere mostrada, salud e gracia (CODEA-0296, Burgos, 1515).
 - a los que agora son como a los que serán de aquí adelante e a cadaúno de **vós** a quien esta nuestra carta fuere mostrada (CODEA-0277, Valladolid, 1523).
 - e a cada uno e cualquier de **vós** en vuestros lugares e jurisdicciones, e a otras cualesquier personas de cualquier estado e condición (CODEA-0306, Madrid, 1552).

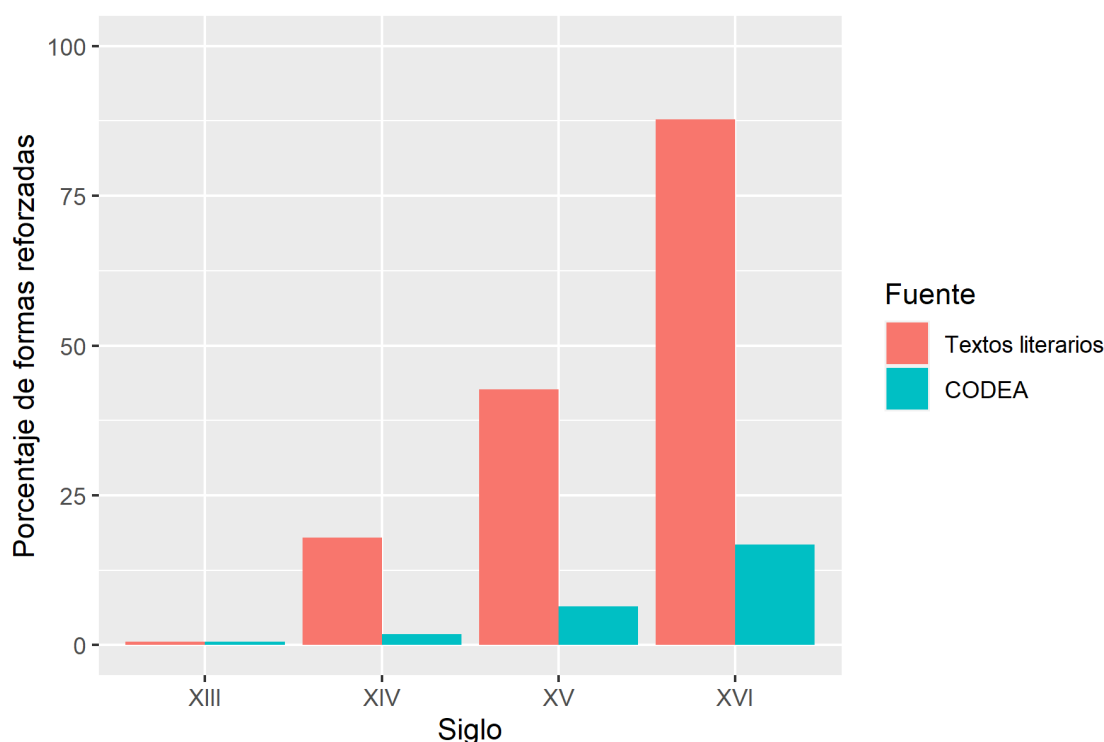


Figura 8. Evolución de las variantes *nosotros* y *vosotros* en la muestra de los textos literarios y en el CODEA+.

Uno de los muchos ejemplos de este fenómeno lo podemos encontrar en un documento del siglo XVI (CODEA-1425), donde prácticamente la totalidad de pronombres de segunda persona del plural están expresados mediante formas compuestas:

- (4)
- a) **bosotros** fazíades cierto nombramiento de escrivanos
 - b) **bosotros** nin los dichos escrivanos vayades nin pasedes
 - c) seyendo para ello **vosotros** e los dichos escrivanos
 - d) los quales sean esaminados por **vosotros**
 - e) an de ser elegidos por **vosotros**
 - f) por el mismo fecho **vosotros** perdáis
 - g) los escrivanos que por **vosotros** fueren elegidos

En cambio, *vós* se mantiene en fórmulas del texto, típicas de documentos cancillerescos como este:

(5)

- a) a **vós** el concejo, justicia, regidores, caballeros, escuderos, oficiales e omnes buenos de la villa de Vilvao, salud e gracia
- b) fue acordado que de aquí adelante **vós** el dicho concejo juntamente con el nuestro corregidor

En resumen, en los textos literarios el reemplazo de *nós*, *vós* por *nosotros*, *vosotros* se inicia y culmina entre el siglo XIII y XVI, de acuerdo con los análisis realizados en estudios previos, que también utilizaban esta tipología textual como fuente de datos. En cambio, el avance de las variantes innovadoras en los textos documentales es considerablemente más lento. En términos porcentuales, en el siglo XVI este tipo de textos se encontraría todavía en una fase inicial del cambio. La resistencia a la adopción de los pronombres reforzados puede explicarse por el carácter formulístico de la documentación notarial y por otros factores que presentaremos en las secciones siguientes.

4.3. Orden de aparición

Si volvemos a las primeras ocurrencias citadas en (1) y (2) podemos ver que la gran mayoría de los primeros ejemplos de formas reforzadas son de segunda persona. En los textos literarios encontramos dos ejemplos: uno de *nosotros* y otro de *vosotros*. No obstante, en los textos documentales del CODEA+, en el periodo que va desde las primeras ocurrencias en el siglo XIII hasta el último cuarto del siglo XIV solo se documenta la segunda persona *vosotros*. Es decir, la primera ocurrencia de *nosotros* en los textos documentales no aparece hasta más de 100 años después de la primera aparición de *vosotros*.

Al margen de los primeros casos del siglo XIII, es imprescindible comparar el avance de las variantes de primera y segunda persona a través de los siglos XIV, XV y XVI en cada uno de los tipos de texto. Tal y como veremos a continuación, en ambos casos la segunda persona es la que se difunde más rápidamente.

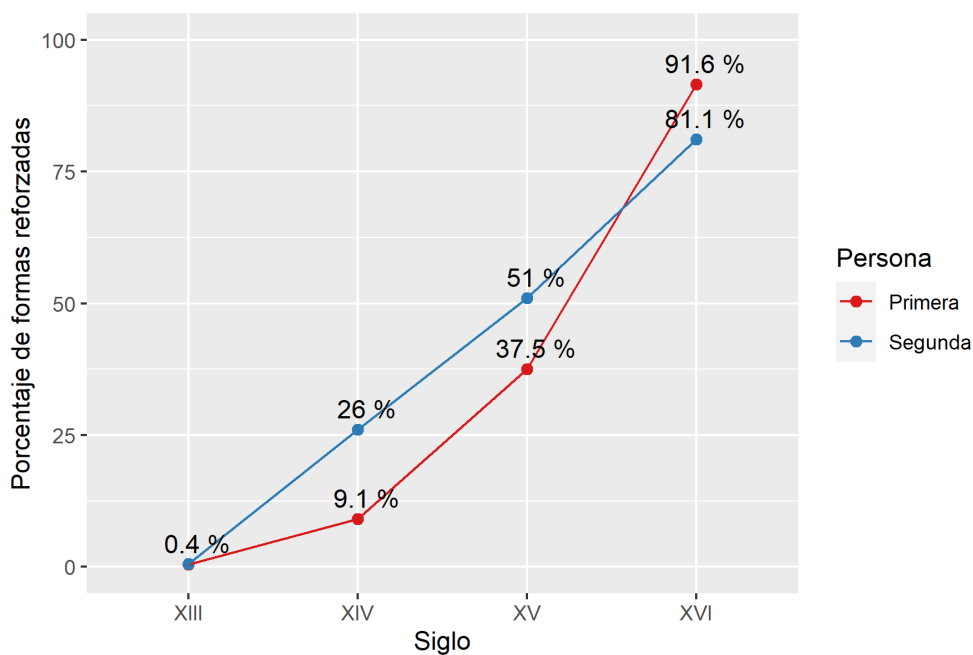


Figura 9. Porcentaje de aparición de las formas reforzadas de primera y segunda persona del plural *nosotros* y *vosotros* en la muestra de textos literarios.

En los textos literarios (Figura 9) llama especialmente la atención el pronunciado ascenso en el uso de la segunda persona desde el inicio del proceso de cambio. En lugar de formar una curva en S, prácticamente forma una diagonal. Esto sugiere que *vosotros* podría haber estado presente desde mucho antes en la lengua oral, y que penetró rápidamente en los textos de tipo literario. El avance de la primera persona, en cambio, es algo más lento y el porcentaje de aparición se mantiene por debajo de la segunda persona tanto en el siglo XIV como en el XV. En el siglo XVI, cuando el cambio ya se ha completado en los textos literarios, el uso de *nosotros* asciende rápidamente y alcanza —e incluso supera— a *vosotros*.

En los textos documentales (Figura 10), tal y como hemos visto más arriba, el avance de ambas variantes es mucho más lento. Sin embargo, no solo *nosotros* se documenta un siglo después que *vosotros*, sino que la difusión de las dos formas reforzadas se produce en paralelo con un siglo de diferencia durante todo el período analizado.

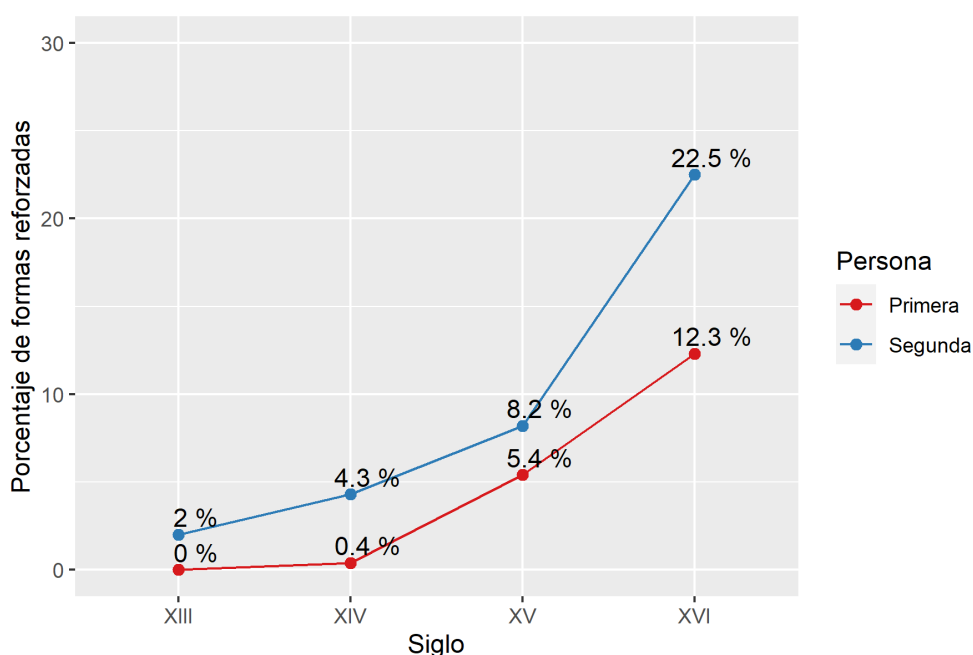


Figura 10. Porcentaje de aparición de las formas reforzadas de primera y segunda persona del plural *nosotros* y *vosotros* en el CODEA+.

Por tanto, a partir de estos datos, parece evidente que la segunda persona *vosotros* es la que encabeza el cambio, mientras que la primera persona *nosotros* se forma siguiendo el modelo de esta. El hecho de que sea la segunda persona la que inicie el proceso nos puede dar una idea de qué lo causa. El motivo del refuerzo de los pronombres debe de estar de alguna manera en la propia naturaleza del pronombre de segunda persona. En este sentido, mi hipótesis se basa en la necesidad de especificar el número del referente.

La principal aportación del refuerzo desde el punto de vista gramatical, dentro del paradigma de pronombres personales tónicos, es la distinción explícita del número. Como hemos comentado en páginas anteriores, tanto *nós* como *vós* pueden utilizarse con referente tanto plural como singular. Sin embargo, el uso de la segunda persona *vós* como forma de tratamiento de cortesía es mucho más frecuente que el de la primera *nós*. Si examinamos el uso de los pronombres *nós*, *vós* según su referente en el CODEA+ (Figura 11), veremos que el empleo de *vós* como forma de cortesía con referente singular es siempre más habitual —alrededor de un 60 %— que el del mismo pronombre con valor plural. En cambio, el uso de *nós* con una sola persona como referente es mucho menos

habitual que su correspondiente plural. El porcentaje más elevado es un 39 % en el siglo XIII, pero en los siglos siguientes es incluso inferior. Por tanto, la necesidad de marcar el plural en el pronombre de segunda persona *vós* era mayor que en la primera persona *nós*.

Además, esto encaja con la jerarquía implicacional de persona establecida para el parámetro de número (Croft 2003:161) (Capítulo 2, § 6.2). En esta escala, la primera persona es la que tiene menos tendencia a presentar flexión de número. De este modo, si debe marcarse estructuralmente el número alguno de los pronombres, primero se marcará la 3ª persona, luego la 2ª y en último lugar la 1ª (1 < 2 < 3). Por tanto, el avance de la marca de número en los pronombres de plural del castellano medieval habría seguido esta escala.

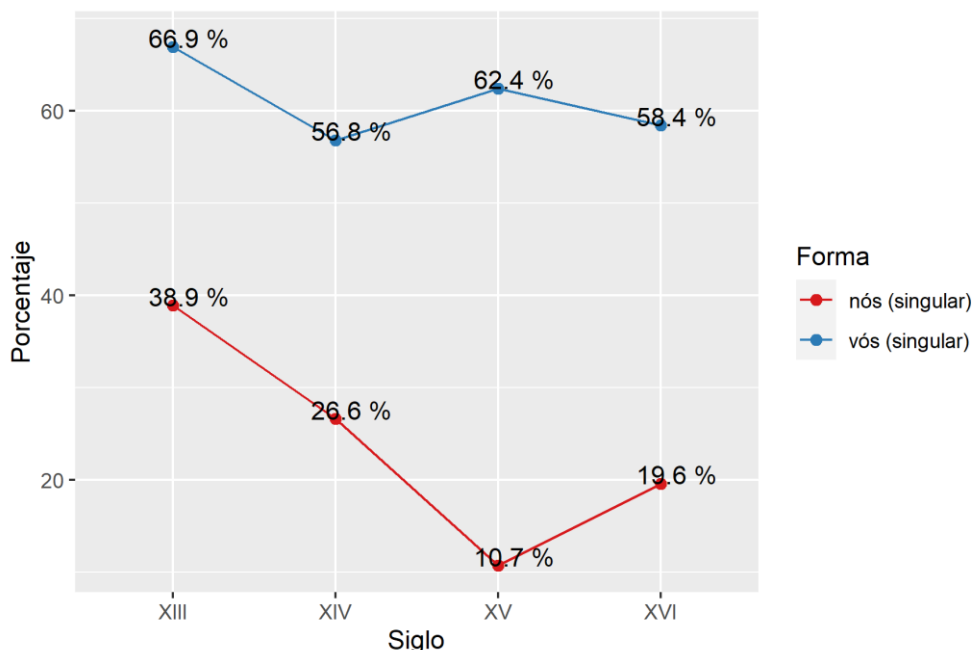


Figura 11. Porcentaje de uso de las formas simples *nós* y *vós* con referente singular frente a las formas correspondientes con referente plural en el CODEA+.

4.4. Contexto sintáctico

En cuanto al contexto sintáctico de las variantes, la mayoría de los estudios que han tenido en cuenta esta variable han distinguido entre la posición de sujeto y la de término de preposición (Eberenz 2000:71; Gili Gaya 1946:117), pero ninguno parece haber encontrado diferencias concluyentes entre ambos contextos.

A continuación, en las Figuras 12 y 13 podemos ver la distribución de las variantes en las dos principales posiciones sintácticas¹¹. A primera vista, no parece haber diferencias entre la posición de sujeto y la de término de preposición. A pesar de que tanto en el corpus literario como en el documental el contexto de término de preposición parece favorecer muy ligeramente las formas reforzadas durante la fase de difusión, es difícil saber hasta qué punto la diferencia es relevante.

¹¹ En la tercera categoría «otros» se agrupan los ejemplos que no se ajustan a las dos posiciones principales, como vocativos o términos de comparación. Estos contextos no se han tenido en cuenta para los cálculos estadísticos.

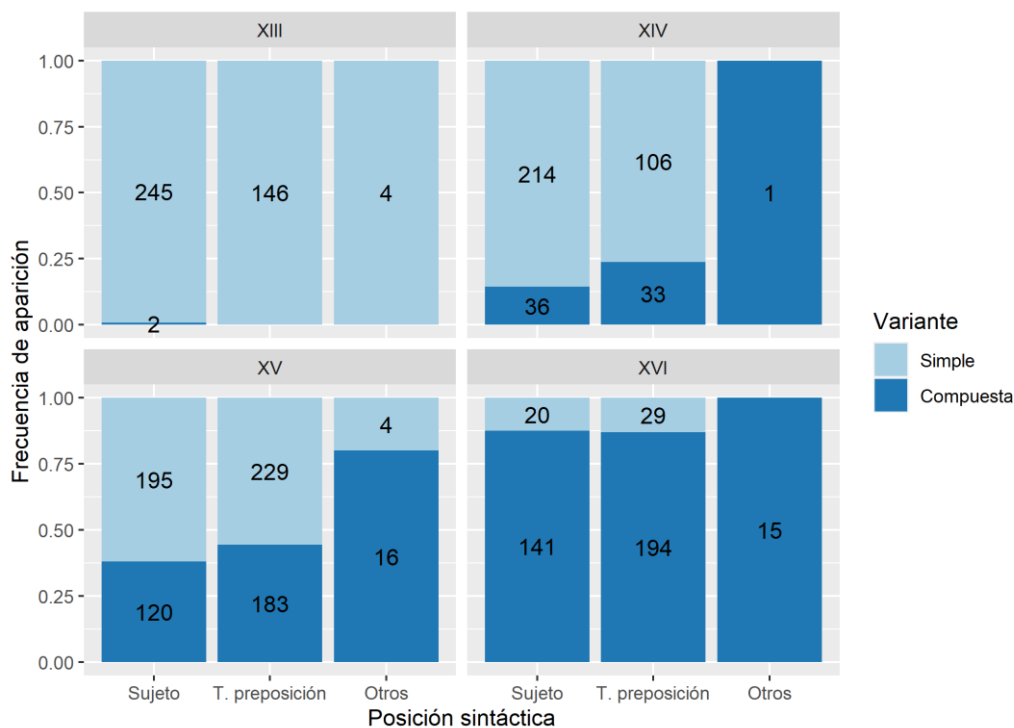


Figura 12. Frecuencia de aparición y número de ocurrencias de las variantes simples y compuestas en posición de sujeto, en término de preposición o en otros contextos (textos literarios).

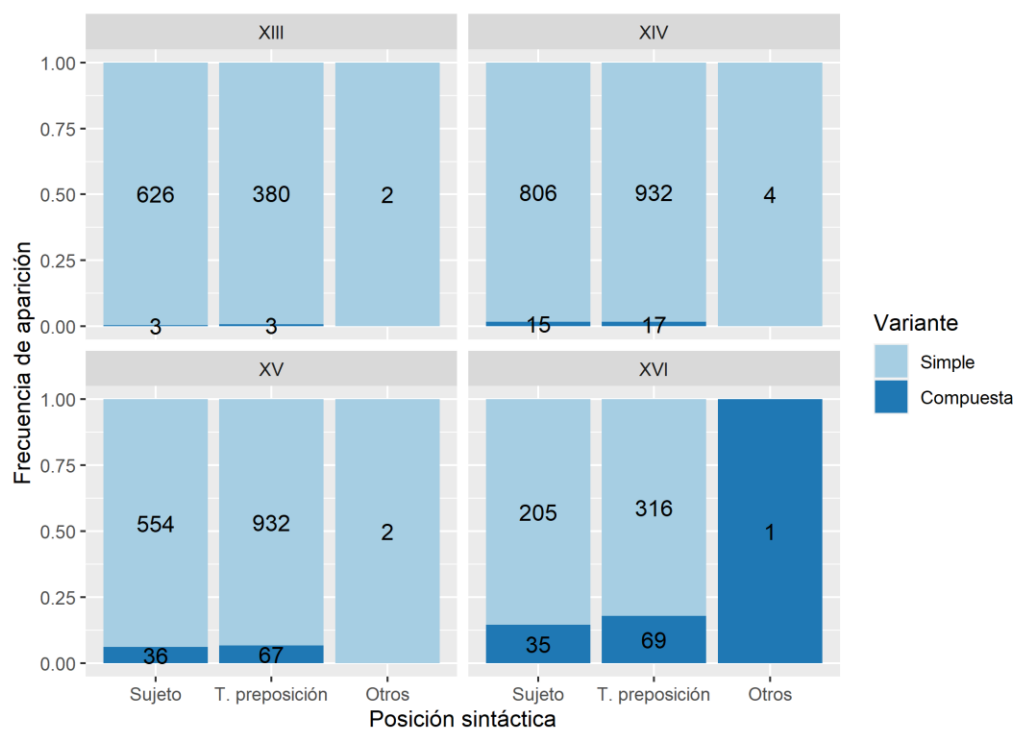


Figura 13. Frecuencia de aparición y número de ocurrencias de las variantes simples y compuestas en posición de sujeto, en término de preposición o en otros contextos (textos documentales).

Para detectar si estas diferencias son significativas desde el punto de vista estadístico nos centraremos en los textos literarios, ya que son los que presentan el cambio completo. En la Tabla 6 podemos ver la frecuencia de aparición de las formas simples y reforzadas

según su posición sintáctica en la muestra de los textos sin distinción de siglo. En ambos contextos encontramos que *nós* y *vós* son las variantes mayoritarias. No obstante, en la tabla de contingencia se aprecia, igual que en las figuras, una pequeña diferencia: mientras que casi un 70 % de los pronombres en posición de sujeto se corresponde con las formas simples, en término de preposición la presencia de *nós* y *vós* desciende a un 55 %. Esta diferencia es estadísticamente significativa ($p < 0.01$), aunque la relación entre ambas variables es muy débil (V de Crámer = 0.143)¹².

	Sujeto	Término de preposición
Simple	674 (69.3 %)	510 (55.4 %)
Compuesta	299 (30.7 %)	410 (44.6 %)
Total	973 (100 %)	920 (100 %)

Pearson's x-squared = 38.641, df = 1, p = 5.0948e-10

Phi-coefficient: 0.143

Contingency coeff.: 0.141

Cramer's V: 0.143

Tabla 6. Frecuencia de las variantes simples y compuestas en posición de sujeto y en término de preposición en textos literarios.

Por tanto, aunque la posición sintáctica no parece jugar un papel determinante en la adopción de las variantes reforzadas *nosotros*, *vosotros*, sí nos puede servir para entender mejor las vías de difusión de estas formas. A partir de los datos, deducimos que los nuevos pronombres se difunden por escala de relaciones gramaticales desde las posiciones más bajas. En esta jerarquía, la posición de sujeto es la menos marcada y la de oblicuo la más marcada. Los elementos en posición de sujeto son algo más resistentes al cambio, por lo que es esperable que los pronombres reforzados, como formas más marcadas, se difundan algo antes tras preposición.

Otra de las explicaciones que se han propuesto para la distribución de formas simples y reforzadas según el contexto sintáctico es la de Nowikow (1994). A partir de los ejemplos expuestos por Gili Gaya (1946), Nowikow propone que las variantes *nós*, *vós* tienden a aparecer en estructuras especificadoras, con elementos singularizadores, como *cualquiera de vós*, *cada uno de vós* o *ninguno de vós*; mientras que *nosotros*, *vosotros* se emplean en estructuras más generalizadoras (1994:293). Si nos fijamos, lo que tienen en común los contextos en los que el pronombre va precedido por las estructuras especificadoras que señala Nowikow es que su referente siempre va a ser claramente plural. Ninguna de estas estructuras es ambigua en cuanto al número del pronombre. Si entendemos que el refuerzo *otros* aparece junto al pronombre como estrategia para desambiguar el referente, aparecerá antes en contextos donde esta desambiguación sea necesaria. Por ese motivo, es posible que los pronombres reforzados tarden más en llegar a estas estructuras.

Otros elementos que realizan la función de desambiguar el referente son las aposiciones. En la documentación notarial es especialmente frecuente la aparición de los pronombres simples *nós* y *vós* seguidos de una aposición con el nombre completo de las personas a las que se hace referencia; o con algún sintagma como *los dichos compradores*, *las dichas partes* o *los sobredichos*.

Precisamente, la combinación de ambos recursos, es decir, el uso tanto de estructuras especificadoras —tipo *cualquier de vós*— como de aposiciones, tiene un papel

¹² Si examinamos cada una de las preposiciones individualmente, tampoco se observan diferencias señalables. La única preposición que parece favorecer de forma más o menos sistemática la aparición de *nosotros* y *vosotros* es *con*, tanto en textos literarios como en textos documentales. Preposiciones como *a* o *de*, que quizá son las marcas de caso más genéricas, se mantienen en la media. El resto muestran una evolución irregular o tienen un número insuficiente de ejemplos.

inegable en la retención de los pronombres simples en la documentación notarial. En (6) y (7) pueden verse ejemplos de dos documentos del siglo XVI. En (6) aparecen tanto aposiciones (a, e) como estructuras especificadoras (b, c, d, f, g). Este caso es especialmente paradigmático, ya que usa varias veces *vós* y *vosotros* en la misma línea (6d, f, g). En los ejemplos en los que aparecen las dos variantes, el pronombre simple va acompañado de *cualquier de*, *algunos de* o *cadaúno de*; pero el pronombre compuesto no lleva ningún otro especificador.

En (7) vemos la repetición sistemática de *dichos compradores* o de los nombres propios cada vez que aparece el pronombre simple *vós*, mientras que con *vosotros* esto no es necesario (b, k, o, p, r). No es imposible encontrar ejemplos del pronombre reforzado con aposición (q), pero este uso está lejos de ser sistemático.

(6) CODEA-0384

- a) a **vós** los mis arrendadores e recabdadores mayores
- b) e a cadaúno e cualquier de **vós** ante quien esta mi carta fuere mostrada
- c) vos mando a todos e a cadaúno de **vós** que dexéis e consintáis al dicho prior
- d) e que **vós otros** o cualquier de **vós** o quien vuestro poder hobiere
- e) e que **vós** las dichas guardas e alvaleros de las dichas salinas d'Espartinas toméis
- f) E si contra lo en mi carta contenido o contra alguna cosa o parte d'ello **vós otros** o algunos de **vós** fuéredes o pasáredes
- g) que fagan e manden fazer en **vós otros** e en cadaúno de **vós** que contra ello

(7) CODEA-0835

- a) libramos y desemparamos a **vós** los honorables Pedro Ximeno e Isabel Crespo, cónjuges
- b) y ad aquellos quien **vosotros** de aquí adelante querredes
- c) a **vós** dichos Pedro Ximeno e Isabel Crespo
- d) los cuales de **vós** dichos compradores atorgamos
- e) recibido de **vós** dichos compradores
- f) querientes que **vós** dichos compradores
- g) aquéllos y aquéllas en **vós**, dichos compradores
- h) y reconocemos por **vós**, dichos compradores
- i) intemptadas a **vós** dichos compradores
- j) finque en obción de **vós** dichos compradores
- k) levar y prosseguir por **vosotros** mismos los dichos pleitos
- l) renunciando a **vós** dichos compradores
- m) lo que Dios no mande, **vós** dichos compradores o los vuestros
- n) prometemos y nos obligamos a **vós** dichos compradores
- o) queremos que **vosotros** y los vuestros siades creídos
- p) obligamos a **vosotros** y a los vuestros todos nuestros bienes
- q) si por mala voz imposada o imponedera a **vosotros**, dichos compradores
- r) la sobredicha pieça que a **vosotros** vendemos

Podríamos pensar que el mantenimiento de las variantes simples en estos contextos se debe al carácter formulístico de los documentos notariales. Esto es en parte cierto. Ejemplos como (6a, b) y (7a) forman parte de la cabecera del documento, caracterizada por usar fórmulas fijas. No obstante, cabe destacar que estas estructuras no solo aparecen en la cabecera o el cierre, sino que se repiten una y otra vez en el cuerpo del texto. La razón de ser de esta especificación constante es la necesidad de aclarar quiénes son los intervinientes en un proceso de compraventa, los beneficiarios de un testamento, etc. Este tipo de texto puede favorecer, por su registro, el uso de *nós* y *vós* como formas de tratamiento con referente singular, y al mismo tiempo necesita que se elimine totalmente la ambigüedad del referente de los pronombres. Esta ambigüedad puede resolverse mediante el uso de estructuras especificadoras, de aposiciones o mediante la adición del refuerzo pronominal *otros*.

Como hemos ido viendo en el análisis de este capítulo, los textos documentales castellanos del siglo XVI usan todavía de forma minoritaria los pronombres reforzados *nosotros*, *vosotros*, mientras que utilizan otras estrategias desambiguadoras con frecuencia. Sin embargo, tal y como veremos en la sección siguiente, los textos documentales del *Corpus Informatizat del Català Antic* (CICA) se comportan de forma totalmente diferente. Esperaríamos que, por ser textos de una misma tipología, del mismo registro, con la misma tendencia a repetir fórmulas fijas, tuvieran un comportamiento similar. Sin embargo, el uso de aposiciones es mucho menos frecuente, mientras que los pronombres compuestos *nosaltres*, *vosaltres* se generalizan mucho antes. Esto apoya la idea de que estas distintas estrategias tienen el mismo fin, el de identificar el número del referente. Es decir, el catalán opta por el uso del modificador de forma mucho más temprana, por lo que puede prescindir del resto de estrategias.

En definitiva, tanto el uso de estructuras especificadoras y aposiciones como la adición del modificador son mecanismos que permiten marcar explícitamente el número del pronombre. En mi opinión, más que el valor contrastivo o enfático que pudieran aportar los modificadores, la clave de su uso como refuerzos pronominales estaría en la desambiguación del referente. Por tanto, los pronombres reforzados aparecerán más frecuentemente en aquellos contextos en los que se necesite marcar explícitamente el número.

4.5. Origen y difusión de las variantes en el espacio geográfico

La presencia de equivalentes a *nosotros* y *vosotros* gramaticalizados en catalán y en occitano, frente a la conservación de *nós* y *vós* en el bable occidental y en gallego-portugués sugiere que los pronombres reforzados llegaron al castellano desde el oriente peninsular. En esta sección pretendo trazar el origen y la difusión de estas variantes en los textos literarios y documentales en castellano. Asimismo, con el objetivo de completar el *continuum* del norte peninsular, incorporaré datos de textos documentales catalanes y gallegos.

Como hemos visto en las secciones anteriores, la muestra de textos literarios refleja el cambio completo de las formas simples *nós* y *vós* por las reforzadas *nosotros* y *vosotros*. Es decir, entre los siglos XIII y XVI se produce el origen de las formas reforzadas, su difusión y el reemplazo de las formas simples. En cualquier caso, el avance de las formas innovadoras no se produce a la vez en todo el territorio peninsular, sino que podemos encontrar diferencias notables entre las distintas zonas geográficas.

Los primeros ejemplos en textos literarios, citados en (1), son del mss. 1187 de la *Gran conquista de ultramar*, testimonio de los últimos años del XIII, de la Cámara Real castellana de Sancho IV (Alvar y Lucía Megías 2002:603-5). Hasta el siglo XIV no documentamos más ejemplos de los pronombres compuestos. A partir de este momento, puede verse una clara difusión de las variantes reforzadas desde el oriente peninsular hacia el occidente (Figura 14). Durante el siglo XIV, en Navarra y Aragón se alcanza rápidamente un 28 % de variantes reforzadas, seguidas de Castilla central y oriental, con un 13 %. En cambio, León y Castilla occidental no las documentan hasta el XV. De este modo, en la fase inicial del cambio encontraríamos una difusión escalonada de las variantes reforzadas desde el este al oeste del norte peninsular.

En el siglo XV, toda la mitad central y oriental se sitúa en la fase intermedia del cambio (NyA 56 %; CCO 47 %). Mientras, el extremo más occidental inicia el proceso con un siglo de diferencia (20 %). Sin embargo, en el siglo XVI las variantes reforzadas avanzan rápidamente en León y Castilla occidental hasta ponerse prácticamente a la altura del resto del territorio peninsular: 82 % frente al 95 % de Castilla central y oriental.

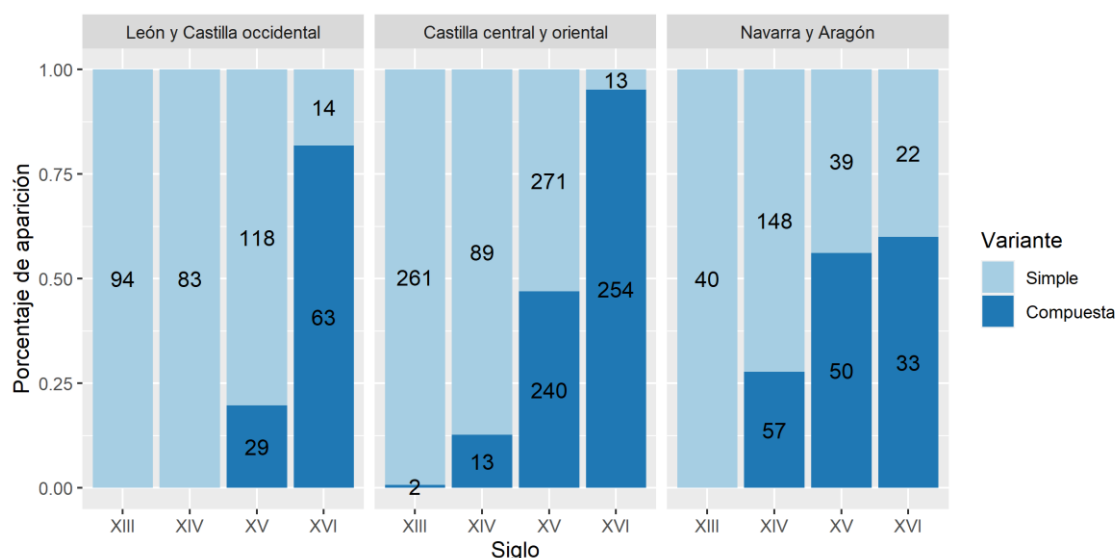


Figura 14. Presencia de formas simples y compuestas según zona geográfica (textos literarios).

En este último siglo destaca el bajo número de ocurrencias en Navarra y Aragón (55 en total) y, especialmente, el bajo porcentaje de formas reforzadas (60 %). Los 22 ejemplos de formas simples que se registran en este segmento pertenecen a una misma obra, el *Orlando*. La persistencia de los pronombres *nós* y *vós* en este texto puede deberse a que se trata de un poema épico caballeresco. Por tanto, la métrica o la voluntad arcaizante que puede observarse en (8) pueden haber influido en la elección de estas variantes.

(8)

- a) Si cristianísimos queréis llamaros, / y **vós** también católicos nombrados, / ¿por qué matáis los de Cristo tan caros? (*Orlando*, pág. 256, CORDE)
- b) No seáis mujeres vanas ni orgullosas / por decir que de **vós** todos nacemos: / que de espinas nacer se ven las rosas (*Orlando*, pág. 476, CORDE)
- c) «Mas bien te pido que antes que aquí sea / lid entre **nós**, que clara se me entienda / mi justísima excusa [...]» (*Orlando*, pág. 543, CORDE)
- d) Decirles suele: «Bien que hayáis **vós** sido / cada cual aquí en armas un lucero [...]» (*Orlando*, pág. 555, CORDE)

Una de las diferencias entre textos literarios y documentales es el lugar de las primeras documentaciones. Los primeros ejemplos de los pronombres reforzados, citados en (2), aparecen en un documento de mediados del siglo XIII en Pamplona (CODEA-0979). Así, estas primeras ocurrencias, las únicas de este siglo, son más tempranas que las de la *Gran conquista de ultramar* y nos sitúan todavía más al este.

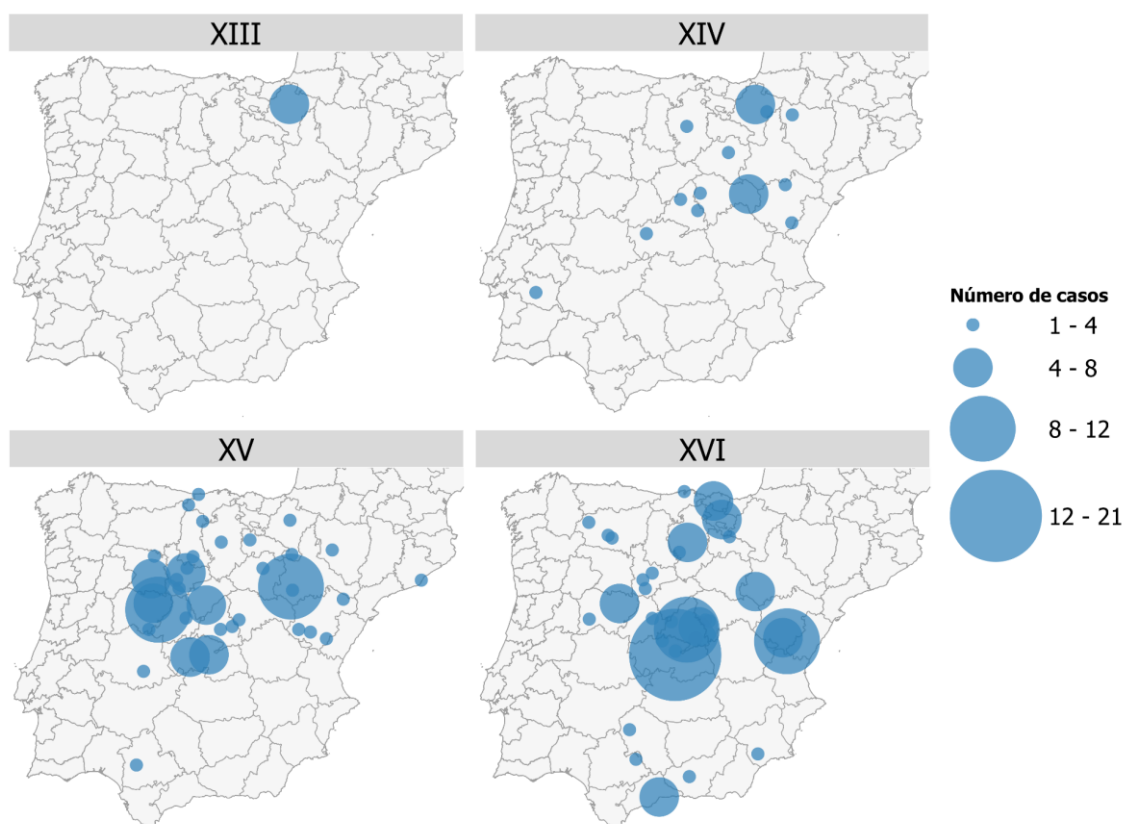
En contraste con los textos literarios, los textos documentales presentan un avance de las formas reforzadas considerablemente más lento. Por tanto, las formas simples *nós* y *vós* son las mayoritarias en todos los periodos cronológicos y en las tres zonas. Sin embargo, los textos documentales nos permiten situar las ocurrencias de las formas reforzadas de forma más precisa y cartografiar el avance del cambio.

Tal y como podemos ver en el Mapa 2, en el siglo XIV la actual provincia de Navarra sigue siendo uno de los puntos en los que encontramos más apariciones de las formas compuestas, concretamente 10 ocurrencias en 3 documentos¹³. El otro punto en el que encontramos un número destacable de casos de formas compuestas es en Molina de Aragón, localidad situada en el extremo oriental de la actual provincia de Guadalajara.

¹³ CODEA-0938, CODEA-0937 y CODEA-0940.

Los 8 casos de este punto se corresponden con 4 documentos¹⁴ de entre los años 1372 y 1382. Además de las ocurrencias mencionadas, en el siglo XIV se documentan las formas compuestas en toda la mitad oriental del norte peninsular¹⁵. En este momento, igual que ocurría en los textos literarios, todavía no aparece ninguna de las variantes reforzadas en las provincias que se corresponderían con León y Castilla occidental.

En el siglo XV, a pesar de que ya hemos visto que el porcentaje de las formas *nosotros* y *vosotros* todavía no es demasiado alto en este tipo de documentos, en el Mapa 2 puede verse claramente que se han extendido prácticamente por toda la mitad norte de la península, exceptuando Asturias y León, donde sí que encontramos otros documentos que únicamente utilizan las formas simples *nós* y *vós*. En el siglo XVI, aunque hay muchas provincias sin documentos que registren ninguna de las variantes, la extensión de los pronombres compuestos es prácticamente la misma e incluso encontramos algunos casos de *nosotros* y *vosotros* en León.



Mapa 2. Localización de las formas compuestas *nosotros* y *vosotros* en los siglos XIII-XVI en los textos documentales del CODEA+ .

Aunque el cambio está lejos de completarse en los textos de tipo documental, el cartografiado de los datos evidencia un claro paralelismo con el avance de las variantes en los textos literarios: en ambos tipos de documentos los pronombres reforzados se originan en la mitad oriental de la península (siglo XIII) y se difunden desde ahí hacia el

¹⁴ CODEA-2064, CODEA-2324, CODEA-2321 y CODEA-2318.

¹⁵ Es necesario destacar la falta de documentos del CODEA+ para las localidades y provincias más orientales. Entre los siglos XIII-XVI, para las provincias de Navarra, Huesca, Zaragoza y Teruel, el CODEA+ tiene un total de 238 documentos (187.684 palabras), mientras que para las provincias que hemos incluido las otras dos zonas geográficas (Castilla central y oriental; León y Castilla occidental) tiene 1033 documentos (609.500 palabras). Es decir, el número de palabras de los documentos de Castilla y León triplica el de las de Navarra y Aragón.

centro (siglo XIV). Hasta el siglo XV no empiezan a documentarse en León y Castilla occidental. Por tanto, los pronombres *nosotros* y *vosotros* se difunden en el espacio geográfico al mismo ritmo en ambos tipos de texto, al margen de cuál sea su presencia en términos porcentuales. Más adelante, en § 4.6, me detengo en la cuestión del registro y la variación textual.

Si hacemos una búsqueda en el CICA sin filtro tipológico, los primeros ejemplos de formas reforzadas aparecen en la primera mitad del siglo XIII¹⁶. Serían, por tanto, las documentaciones más tempranas entre todos los corpus utilizados. Estas primeras ocurrencias, que vemos en (9), tienen lugar en la *Carta de franquesa*, un texto jurídico balear. En este texto aparece cuatro veces la forma reforzada de segunda persona *vosaltres*, mientras que solo aparece dos veces la forma simple *vós*. *Nós* es mucho más frecuente —no aparece *nosaltres*—, aunque la mayoría de las veces tiene un valor mayestático, ya que hace referencia al rey Jaume I.

(9)

juren totes aquestes coses axí con demunt són escrites a **vosaltres** tenguen e observent [sic] [37] Prometem encara a **vosaltres** que que no darem ne excambiarem vos ha alcun[a] perssona [...] mas tots temps tendrem **vosaltres** a la Corona d' Aragó e amarem e defensarem **vosaltres** en tots lochs axí con a feels nostres (*Carta de franquesa*, pág. 120, CICA)

La primera persona *nosaltres* se documenta por primera vez en la segunda mitad del XIII y de forma mucho más esporádica que *vosaltres*, que es ya bastante abundante en los textos del área del catalán oriental. La variante reforzada de primera persona aparece en obras en las que la segunda persona parece más consolidada, como la *Crònica* de Bernat Desclot o las *Vides de Sants Rosselloneses*.

En el TMILG, los primeros ejemplos del pronombre compuesto (10) no se registran hasta un siglo y medio después, en la segunda mitad del siglo XIV. Todos los ejemplos de este siglo son de prosa notarial y están dispersos en tres documentos. Una vez más, la primera persona *nosoutros* se documenta varias décadas más tarde, ya en el siglo XV.

(10)

- a) a hun fillo ou filla, qual nomear o pustrimeiro de **bosoutros** a seus pasamentos (Lugo, 1368, TMILG)
- b) et senllas voces que **vosoutros** nomeardes daran quarto (1382, Ourense, TMILG)
- c) apelo da dicta sentença et mandamento por **uosoutros** dada (1395, A Coruña, TMILG)
- d) Et outrosy demandar et auer de **uosoutros** et de cada hun de uos todos los dapnos (1395, A Coruña, TMILG)
- e) toda cosa que por **vosoutros** seia feyta en perjuyzo nen en fauor de mjn (1395, A Coruña, TMILG)

A continuación, comparamos la evolución de los pronombres compuestos en los documentos notariales del CICA y el TMILG frente a los datos ya vistos más arriba del CODEA+. En la Figura 15 podemos ver que, pesar de contrastar textos de tipologías similares, el avance de las formas innovadoras en los textos documentales del CICA es claramente mucho más rápido que en los demás corpus. En el siglo XIII la aparición de *nosaltres* y *vosaltres* llega a un 18 %, mientras que en el CODEA+ no se supera el 0.5 %. Un siglo después, el cambio se encuentra en una fase avanzada en catalán con casi un 40

¹⁶ Es necesario decir que, aunque el corpus tiene documentos de periodos anteriores, el volumen de textos y palabras en los siglos XI y XII es mucho menor. Por ejemplo, la primera mitad del XIII tiene algo más de 20.000 palabras, mientras que el XII al completo no llega a las 3.500.

%, y en el siglo XV ya son formas claramente mayoritarias (88 %). En el XVI, las formas simples son residuales. De este modo, el cambio concluye no solo antes que en los textos documentales castellanos, sino también antes que en los literarios, que en el siglo XV todavía se encuentran en una fase intermedia.

En definitiva, a diferencia de Gili Gaya, que afirma que el cambio culmina a la vez en catalán y en castellano (1946:113), nuestros datos corroboran los de Fernández-Ordóñez (2011:77). Es decir, en castellano, las formas reforzadas aparecen, se generalizan y se convierten en formas no marcadas más tarde que en catalán.

En cuanto al TMILG, aunque *nosoutros* y *vosoutros* se documentan con un siglo de diferencia respecto a los textos castellanos, no observamos una diferencia notable en su evolución. Además, la falta de datos del siglo XVI dificulta hacer afirmaciones sobre la tendencia del cambio.

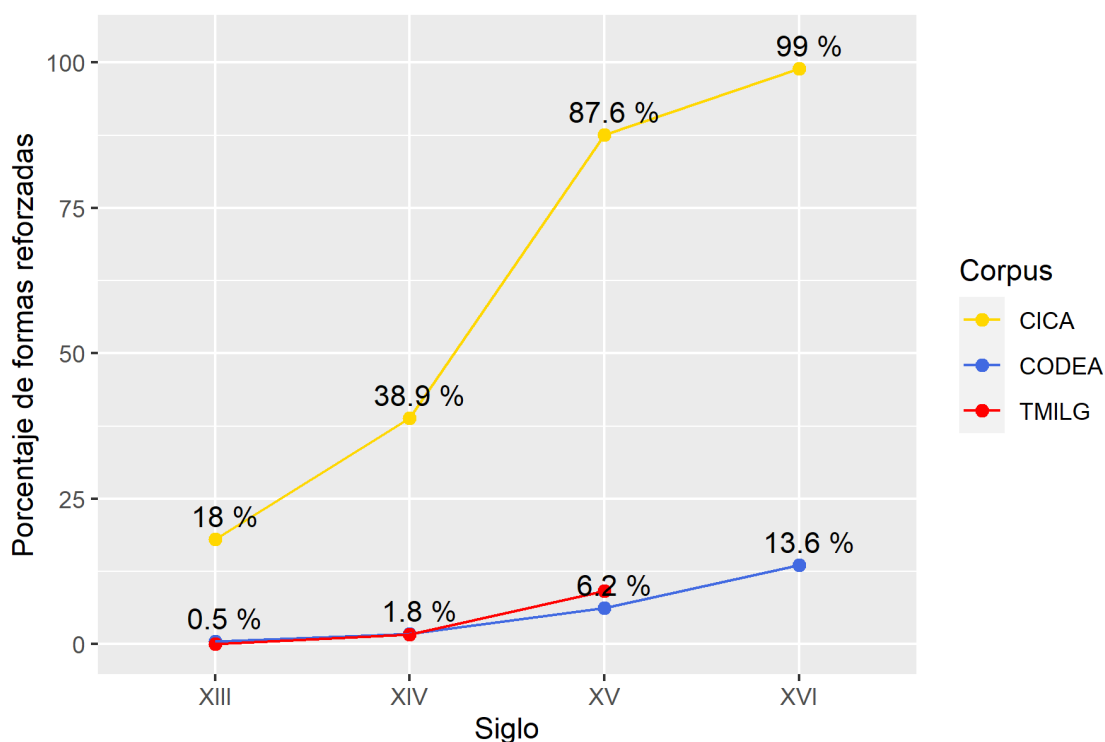


Figura 15. Difusión de los pronombres *nosotros* y *vosotros* desde el siglo XIII al XVI en los textos documentales del CODEA+ 2015, el CICA y el TMILG.

En el Mapa 3 puede verse la evolución completa de los pronombres reforzados por zonas, de este a oeste: catalán oriental, catalán occidental¹⁷, Navarra y Aragón, Castilla central y oriental, León y Castilla occidental¹⁸, Galicia oriental y Galicia occidental¹⁹.

Desde el siglo XIII hasta mediados del XIV, los pronombres reforzados se difunden claramente desde el oriente peninsular. El cambio es notablemente más rápido en catalán, especialmente en las variedades orientales, que superan el 65 % (49 casos) de *nosaltres* y *vosaltres* en el siglo XIV. El ascenso es más lento en el catalán occidental (17 %; 16 ocurrencias), pero sigue habiendo una diferencia considerable respecto a los textos

¹⁷ La ilustración de estas dos zonas en el Mapa 3 únicamente es una ayuda visual y no pretende ser totalmente fiel a las fronteras de los dialectos del catalán, que, por supuesto, van más allá de las fronteras políticas de Cataluña y se prolongan hacia el oeste en Aragón, hacia el norte en Francia y hacia el sur en Valencia.

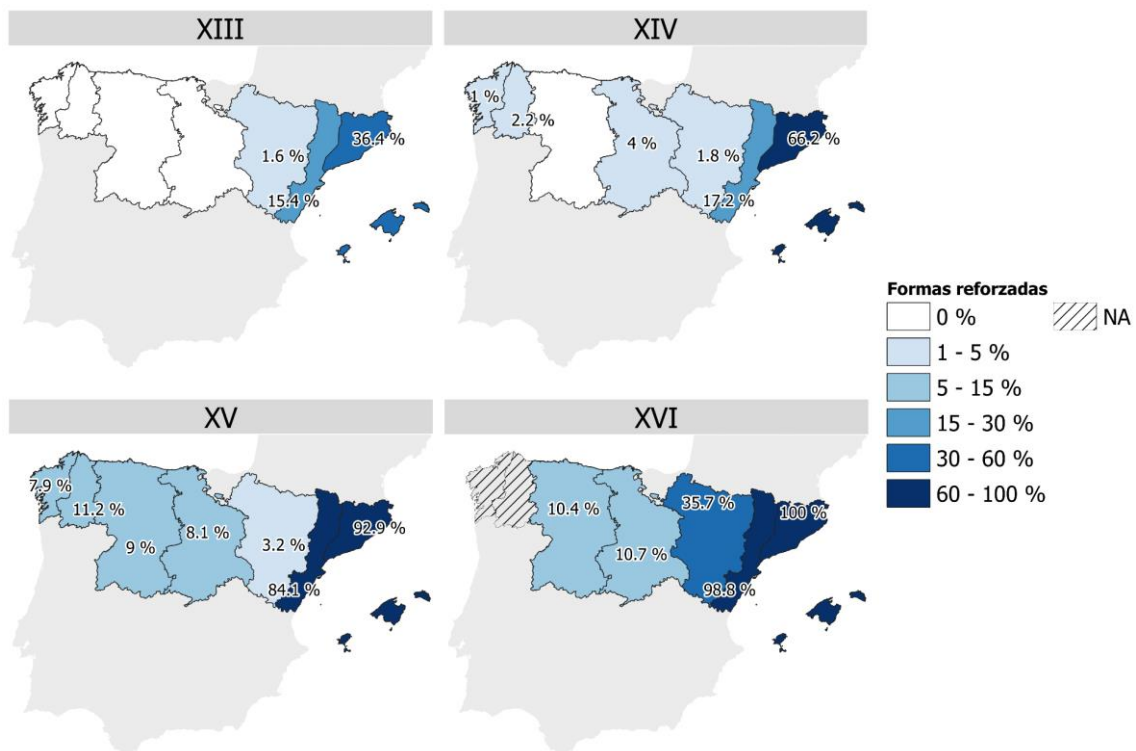
¹⁸ Para la correspondencia entre provincias y zonas, cf. Capítulo 2, § 4.3.

¹⁹ Galicia oriental: provincias de Lugo y Ourense; Galicia occidental: provincias de A Coruña y Pontevedra.

castellanos, donde su presencia es totalmente esporádica y no supera el 5 %. A través de Navarra y Aragón, estas variantes entran en el área de Castilla central y oriental, tal y como hemos descrito previamente en el Mapa 2.

Uno de los puntos que hemos destacado en el análisis de los datos del CODEA+ es que en el siglo XIV todavía no se documenta ningún ejemplo de *nosotros/vosotros* en la zona de León y Castilla occidental hasta el siglo XV. Los documentos de esta zona registran un total de 492 pronombres de primera y segunda persona plural, pero todos ellos se presentan con la forma simple. Esto parecía una muestra más de la difusión de las formas reforzadas desde el oriente. Sin embargo, a finales del XIV encontramos los cinco ejemplos ya citados en los textos gallegos, que se pueden deber a la influencia de los modelos de prestigio castellanos desde el siglo XIII (Mariño Paz 2017; Tabernero, Pichel, y Enrique-Arias 2020:179)²⁰.

En el siglo XV las variantes se han extendido por todo el norte peninsular. A excepción de en catalán, donde el reemplazo de las formas simples por las reforzadas está muy avanzado, en el resto del territorio se mantiene en la fase inicial, que se extiende al XVI.



Mapa 3. Presencia de formas compuestas en los documentos notariales del CODEA+, el CICA y el TMILG desde el siglo XIII al XVI.

En definitiva, como ya hemos visto en las anteriores secciones, la vitalidad de los pronombres reforzados depende en gran medida del tipo de texto. Sin embargo, al margen del porcentaje de aparición, el avance de estas variantes en el espacio geográfico se

²⁰ Sobre el cambio inducido por contacto castellanizante en el gallego medieval: «parece prudente non considerar os castellanismos visibles nos nosos textos medievais como formas necesariamente usadas daquela con regularidade por grandes continxentes de individuos galegofalantes. [...] Os castellanismos léxicos, fonéticos, morfolóxicos e sintácticos dos documentos notariais [...], deberanse imputar exclusivamente á responsabilidade dos seus autores, que durante a segunda metade do século XV e os principios do XVI estaban plenamente inmersos xa nunha dinámica de substitución do galego polo castelán no exercicio da súa profesión» (Mariño Paz 2017).

produce al mismo ritmo en textos documentales y literarios. Se registran por primera vez en el corpus documental castellano en un texto navarro de 1265. En los últimos años del XIII aparecen por primera vez en un texto literario, en la *Gran conquista de ultramar*. En ambas tipologías textuales, en el siglo XIV se documentan los pronombres *nosotros*, *vosotros* en todo el centro y el este peninsular, y no llegan a León y Castilla occidental hasta el siglo XV. En el XVI ya podemos encontrar los pronombres compuestos en cualquier punto del *continuum* norte.

4.6. Variación textual y de registro

En cada una de las secciones de este capítulo hemos ido analizando cómo afectan distintos factores internos y externos a la aparición, difusión y adopción de los pronombres compuestos *nosotros* y *vosotros*. En cada caso, hemos ido comentando por separado la evolución en textos documentales frente a textos literarios. Esto ha probado que las variables que resultan esenciales para explicar el refuerzo de los pronombres como la persona —difusión de 2ª a 1ª— o el espacio geográfico —difusión de este a oeste— afectan de igual manera a los dos tipos de texto.

Sin embargo, hay una diferencia fundamental en la frecuencia de aparición y en el grado de difusión de los pronombres *nosotros*, *vosotros* según la tipología textual. En los textos literarios el reemplazo de *nós*, *vós* por *nosotros*, *vosotros* se inicia y culmina entre los siglos XIII y XVI, mientras que el avance de las variantes innovadoras en los textos documentales es considerablemente más lento. La diferencia en la difusión de estas formas es tal que en el siglo XVI este tipo de textos se encontraría todavía en una fase inicial del cambio.

Estas diferencias se explican por el carácter conservador de los textos de tipo notarial, por la presencia de fragmentos formulísticos en la cabecera y el cierre de los documentos y por el uso de estrategias diferentes al refuerzo de los pronombres para desambiguar el número del referente, como el empleo de estructuras especificadoras y de aposiciones junto al pronombre. Asimismo, la rápida adopción de los pronombres compuestos *nosotros*, *vosotros* por parte de los textos literarios puede deberse a la influencia de los modelos literarios orientales, tanto el provenzal como el catalán (Deyermond y Rico 1980:23, 49, 126, 331; Menéndez Pidal 1962:103-4).

En cualquier caso, el uso de diferentes tipologías textuales nos permite acceder de forma limitada al factor del registro. La sociolingüística suele prestar atención a factores externos como el género, la edad o el nivel educativo de los hablantes con el objetivo de establecer la dirección del cambio, de distinguir si se trata de un cambio desde arriba o un cambio desde abajo (Labov 1994:78; Silva-Corvalán y Enrique-Arias 2017:270-71). En el estudio del castellano medieval, es difícil acceder a este tipo de variables, y únicamente podemos acceder al registro a través de textos de diferentes niveles de formalidad.

El CODEA+ distingue entre cinco ámbitos de emisión que se encuentran en distintos niveles de formalidad, de más a menos formal: cancilleresco, judicial, municipal, eclesiástico y particular (Sánchez-Prieto Borja 2012:11). A continuación, pretendo analizar la evolución de los pronombres reforzados *nosotros*, *vosotros* en cada uno de estos ámbitos de emisión con el objetivo de trazar la dirección del cambio.

De los 5.005 ejemplos extraídos del CODEA+ entre 1251 y 1550 que forman el corpus documental de este capítulo (*cf.* Tabla 5), 2.672 pertenecen al ámbito eclesiástico, 1.387 al particular, 612 al cancilleresco, 197 al municipal y 137 al judicial. Para el análisis de esta variable he seleccionado los tres ámbitos con un mayor número de ejemplos: para el extremo más formal, el cancilleresco; y para el menos formal, el particular y el eclesiástico (Figura 16). El eclesiástico es el primer ámbito de emisión que documenta

los pronombres reforzados y mantiene una evolución progresiva y constante hasta el siglo XVI. El cancilleresco no registra las formas reforzadas hasta la primera mitad del siglo XIV, pero, de los tres ámbitos de emisión, es el que presenta un mayor uso de los pronombres hasta el siglo XVI. Finalmente, los textos particulares no solo documentan las variantes *nosotros*, *vosotros* más tarde, en la segunda mitad del XIV, sino que también son los que tienen una menor presencia de estas variantes.

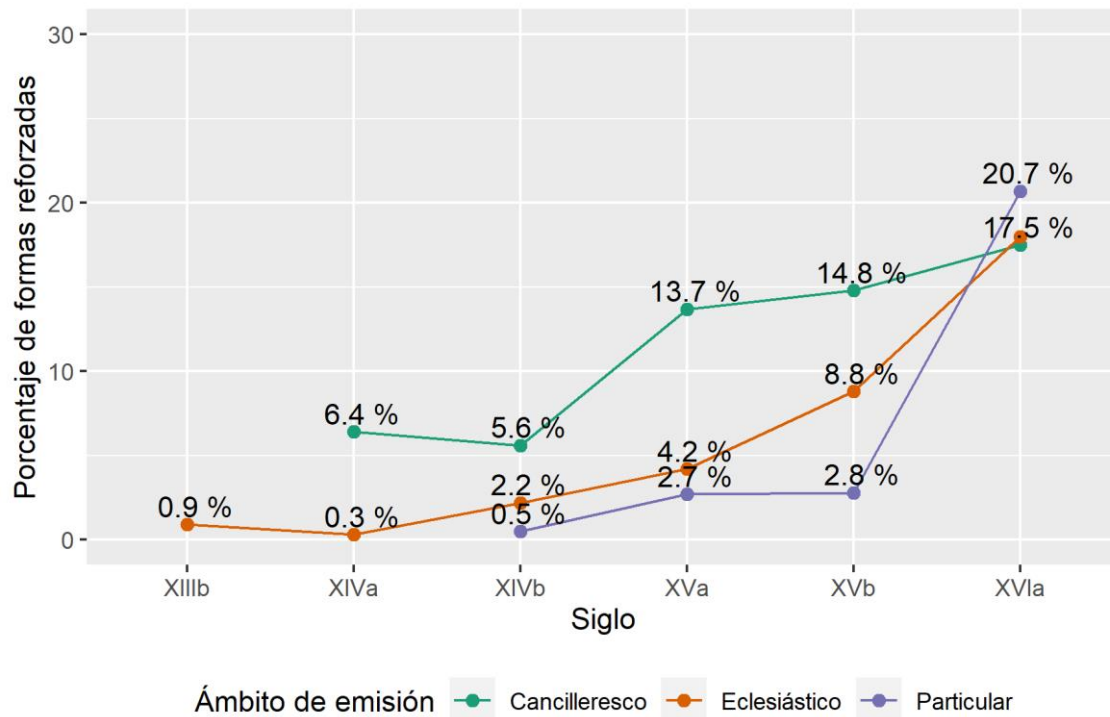


Figura 16. Porcentaje de aparición de las formas reforzadas *nosotros* y *vosotros* en los textos cancillerescos, eclesiásticos y particulares del CODEA+ por mitad de siglo.

Estos datos apuntarían a una difusión de las variantes *nosotros*, *vosotros* desde el ámbito de emisión más formal —el cancilleresco— hasta los menos formales —eclesiástico y particular. Esto, junto a la rápida difusión de los pronombres compuestos en los textos literarios, sugiere que el reemplazo de las formas simples *nós* y *vós* fue un cambio producido desde arriba hacia abajo, de modo que las variantes reforzadas habrían sido introducidas conscientemente por la clase dominante.

No obstante, otras características del reemplazo de *nós*, *vós* por *nosotros*, *vosotros* entran en contradicción con la difusión de las variantes innovadoras desde arriba. Ya hemos visto que la difusión de los pronombres reforzados *nosotros*, *vosotros* se produce desde el oriente peninsular hacia el occidente. Sin embargo, los cambios desde arriba en principio no muestran una distribución geográfica definida (Sánchez-Prieto Borja 2012:30-31). Además, también hemos visto que un factor interno como la persona es relevante para la descripción del proceso de reemplazo, algo que es habitual para los cambios desde abajo (Labov 1994:78; Silva-Corvalán y Enrique-Arias 2017:270-71). Todo esto hace pensar que en la difusión de los pronombres reforzados *nosotros*, *vosotros* interactúan una mezcla de factores desde arriba y desde abajo.

Si incorporamos la perspectiva geográfica al análisis de la difusión de las formas *nosotros*, *vosotros* en los distintos ámbitos de emisión, observamos que el papel del ámbito cancilleresco no es igual de relevante en las tres zonas geográficas distinguidas (Figura 17). Como ya sabemos, los pronombres *nosotros*, *vosotros* aparecen primero en

un documento navarro de tipo eclesiástico, ámbito de emisión que se encuentra en el extremo menos formal. En los textos navarroaragoneses, los textos cancillerescos no presentan ni un solo ejemplo de los pronombres reforzados. Es cierto que en Navarra y Aragón no encontramos tantos documentos de este tipo, pero en todos ellos se utilizan exclusivamente los pronombres simples *nós* y *vós*. Asimismo, de entre todos los periodos y zonas, los textos eclesiásticos y particulares navarroaragoneses del XVI son los que alcanzan con diferencia un mayor porcentaje de formas reforzadas, 23 % y 37 % respectivamente. Por tanto, a partir de los datos de los textos documentales extraemos que en Navarra y Aragón el cambio se difunde desde abajo, desde los ámbitos de emisión más informales. En cambio, en Castilla y en León los pronombres reforzados *nosotros*, *vosotros* acceden antes a los textos de cancillería, mientras que tienen una presencia menor en el resto de ámbitos.

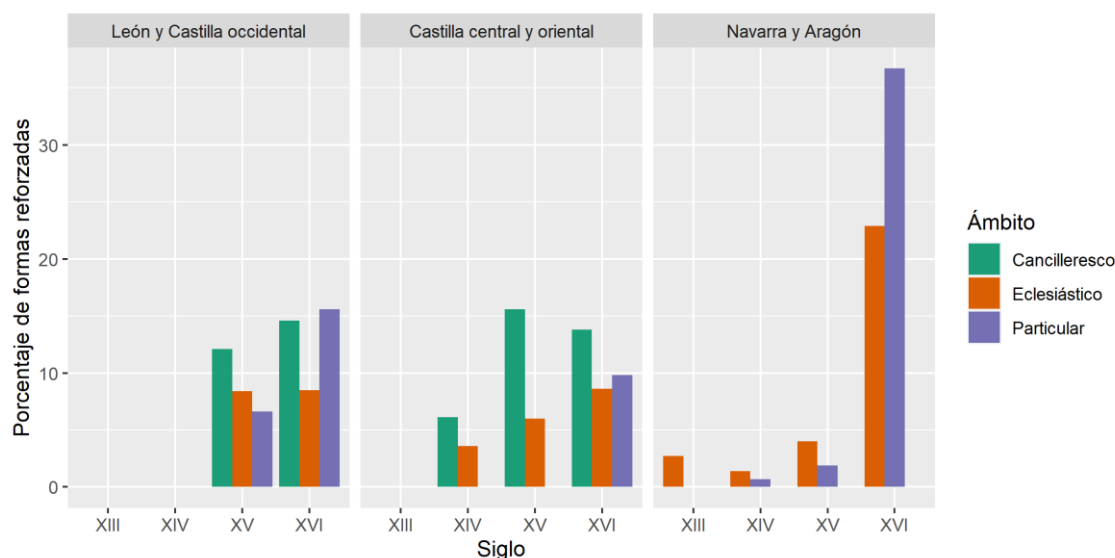


Figura 17. Porcentaje de aparición de las formas reforzadas *nosotros* y *vosotros* en los textos cancillerescos, eclesiásticos y particulares del CODEA+ por siglo y por zona geográfica.

Esto sugiere que, en las zonas más alejadas del foco del cambio, las variantes innovadoras llegan como formas de prestigio que se difunden primero a través de textos literarios y que, dentro de los documentales, acceden primero a los textos de la cancillería, luego a los eclesiásticos y en último lugar a los particulares.

Por tanto, podemos distinguir dos vías de entrada de los pronombres reforzados *nosotros* y *vosotros* en castellano. En el área oriental, donde empieza el cambio, las nuevas formas se difunden desde abajo, desde los textos más informales. Esto podría suceder de este modo debido al contacto de lenguas prolongado fruto del bilingüismo de zonas de frontera con el catalán y el occitano. En cambio, en las áreas central y occidental, más alejadas del foco geográfico del cambio, los pronombres *nosotros* y *vosotros* se difunden desde arriba, desde los ámbitos cultos, como formas prestigiosas que entran en los textos de la cancillería y los literarios. Siguiendo la metáfora de Chambers y Trudgill (2004:166), las formas innovadoras se difunden como una piedra que es lanzada y salta por la superficie del estanque del *continuum* septentrional de este a oeste²¹.

²¹ El hecho de que los pronombres reforzados accedieran ya desde el siglo XIV al modelo de prestigio castellano pudo haber sido la clave para la difusión completa de los pronombres. Esto es diferente de lo que ocurre con los demostrativos tipo *aqueste*, que se difunden en las variedades aragonesas hasta ser variantes categóricas a causa de la influencia del modelo lingüístico catalán y, posteriormente, con la fluctuación del

5. Resumen

En este capítulo he analizado el origen y el avance de las variantes pronominales reforzadas *nosotros* y *vosotros*, que reemplazan a las formas simples *nós* y *vós* durante la baja Edad Media. Este reemplazo es el fenómeno central de esta tesis, ya que a su alrededor gravitan los fenómenos de los capítulos siguientes.

A partir de las preguntas de investigación presentadas al final del estado de la cuestión (§2) hemos obtenido las siguientes respuestas:

Las primeras documentaciones de los pronombres *nosotros*, *vosotros* tienen lugar en la segunda mitad del XIII tanto en el corpus literario como en el documental. En estas primeras apariciones no siempre se refleja con claridad el valor contrastivo de los pronombres, aunque, dado el caso, el contraste puede establecerse tanto hacia alguno de los interlocutores como hacia una tercera persona.

En cuanto a la cronología, en los textos literarios el reemplazo de *nós*, *vós* por *nosotros*, *vosotros* se inicia y culmina entre el siglo XIII y XVI, de acuerdo con los análisis realizados en estudios previos, que también utilizaban esta tipología textual como fuente de datos. En cambio, el avance de las variantes innovadoras en los textos documentales es considerablemente más lento. En términos porcentuales, en el siglo XVI este tipo de textos se encontraría todavía en una fase inicial del cambio.

Respecto al orden de aparición, parece evidente que la segunda persona *vosotros* es la que encabeza el cambio, mientras que la primera persona *nosotros* se forma siguiendo el modelo de esta. Esto se podría explicar por la mayor necesidad de desambiguar el número del referente en la segunda persona, que también se utilizaba frecuentemente como forma de tratamiento de cortesía con referente singular.

En cuanto a la posición sintáctica, el contexto tras preposición parece favorecer ligeramente la aparición de los pronombres reforzados *nosotros*, *vosotros* frente a la posición de sujeto. Por tanto, las formas innovadoras se habrían difundido desde las posiciones más bajas de la escala de relaciones gramaticales. En cualquier caso, este no parece haber sido un factor especialmente determinante para la adopción de las formas. Un factor más relevante puede haber sido la presencia de estructuras especificadoras y aposiciones como mecanismos que permiten marcar explícitamente el número del pronombre y que, en consecuencia, habrían favorecido la retención de los pronombres simples *nós* y *vós*.

La difusión de *nosotros*, *vosotros* en el espacio geográfico se produce al mismo ritmo en textos documentales y literarios. Se registran por primera vez en el oriente peninsular en la segunda mitad del siglo XIII. En ambas tipologías textuales, al margen del porcentaje de aparición, en el siglo XIV se documentan los pronombres *nosotros*, *vosotros* en todo el centro y el este peninsular, y no llegan a Castilla occidental y León hasta el siglo XV. En el XVI ya podemos encontrar los pronombres compuestos en cualquier punto del *continuum* norte.

Si atendemos a la difusión de los pronombres reforzados a través de los ámbitos de emisión de los textos documentales, observamos dos vías de entrada de los pronombres *nosotros*, *vosotros*: una “desde abajo” a través de los registros menos formales en las zonas más orientales, fruto de un contacto de lenguas a largo plazo; y otra “desde arriba” a través de los textos más formales en las zonas más lejanas, donde los nuevos pronombres llegan como formas de prestigio.

modelo de prestigio al castellano, retroceden hasta desaparecer (Enrique-Arias et al. En evaluación). En el caso de los pronombres reforzados *nosotros* y *vosotros*, cuando el modelo de prestigio fluctúa al castellano, las formas innovadoras ya han accedido a los contextos formales, por lo que no se produce un retroceso.

Capítulo 4: Los refuerzos pronominales enfáticos

1. Presentación¹

Durante la Edad Media, los pronombres personales *nós* y *vós* podían aparecer reforzados con modificadores distintos a *otros*, como *todos*, *mismos* o *ambos*. Habitualmente, cuando se habla de los pronombres compuestos o reforzados de primera y segunda persona del plural surgidos durante la Edad Media se hace referencia a *nosotros* y *vosotros*. La aparición de estas variantes ha sido ampliamente estudiada: su origen (Alvar y Pottier 1983; Lausberg 1973), la cronología (Eberenz 2000) y las causas del reemplazo (García et al. 1990; Gili Gaya 1946; Menéndez Pidal 1982) han sido aspectos tratados en profundidad. Más recientemente, el fenómeno también ha sido abordado desde el punto de vista diatópico y se ha situado el foco del cambio en el oriente de la península ibérica (Enrique-Arias et al. En evaluación; Fernández-Ordóñez 2011; Gomila Albal 2016).

Sin embargo, se ha prestado mucha menos atención a las combinaciones de los pronombres *nós* y *vós* con modificadores distintos de *otros*. Para algunos autores (Eberenz 2000:59; Rini 1999), antes de que *nosotros* y *vosotros* se generalizaran, estos compuestos convivían con otras variantes enfáticas que no llegaron a gramaticalizarse en castellano, como *vós todos* o *vós mismos*. Incluso, Rini (1999:211) afirma que los textos más tempranos presentan una clara superioridad de *vós todos* frente a *vós otros*.

En este capítulo me propongo examinar en detalle todos los posibles factores que pudieron influir en la difusión de unas variantes reforzadas y en el retroceso de otras. Tomaré los principales aspectos que se han tenido en cuenta para la descripción del fenómeno de reemplazo de *nós*, *vós* por *nosotros*, *vosotros* —cronología, orden de aparición, distribución geográfica— y los trasladaré al resto de modificadores registrados en mi corpus para presentar un cuadro completo de la variación de las formas pronominales reforzadas.

De este modo, mi objetivo principal es conocer la relevancia que pudieron tener las variantes de los pronombres personales *nós* y *vós* reforzados con elementos como *todos*, *mismos* y *ambos* antes de que *nosotros* y *vosotros* se convirtieran en formas no marcadas. Pretendo comprobar la vitalidad de estas construcciones en los textos más tempranos y averiguar qué contextos podían favorecer la aparición de una u otra variante. Especialmente, incorporaré la perspectiva diatópica para ver si, del mismo modo que *nosotros* y *vosotros*, alguna de las otras variantes tiene una distribución geográfica determinada.

Mi hipótesis de partida es que la frecuencia de aparición de las variantes reforzadas con *todos*, *mismos* y *ambos* disminuirá a medida que *nosotros* y *vosotros* se generalicen. Del mismo modo, dado que *nosotros* y *vosotros* llegan al castellano desde el catalán y el occitano a través de las variedades navarroaragonesas, y se extienden por la península ibérica desde ahí hacia el oeste y el sur (cf. Capítulo 3, § 4.5), es probable que las variantes con *ambos*, *mismos* y *todos* se conserven más tiempo donde *nosotros* y *vosotros* tarden más en llegar, es decir, en la mitad occidental del norte peninsular.

Así, pretendo contribuir al estudio de los pronombres reforzados incluyendo todas las variantes que pudieron participar en el proceso de reemplazo de los pronombres simples *nós* y *vós*. Analizaré su evolución cronológica en textos literarios y documentales, además de su distribución geográfica. Para ello, antes de nada, repasaré algunos puntos clave sobre la aparición de *nosotros*, *vosotros*, así como la escasa información que se tiene sobre el resto de los modificadores (§ 2); y presentaré los datos y la metodología de

¹ Este capítulo es una versión ampliada de Gomila Albal (2021). Respecto de esa versión, se han añadido todos los datos correspondientes a los textos literarios.

esta sección (§ 3). El análisis (§ 4) se dividirá en cuatro partes: la localización de las primeras apariciones de las variantes (§ 4.1), el establecimiento de la cronología (§ 4.2), el orden de aparición (§ 4.3) y la distribución geográfica (§ 4.4).

2. Estado de la cuestión

2.1. Puntos clave del refuerzo con *otros*²

Entre los procesos de cambio lingüístico analizados en esta tesis, el que ha recibido más atención en los manuales de historia de la lengua es el reemplazo de *nós, vós* por *nosotros, vosotros*. Además, también es el que se ha analizado más exhaustivamente (cf. Capítulo 3). En esta sección pretendo repasar muy someramente algunos de los puntos clave extraídos de ese análisis que pueden ser especialmente relevantes para el examen del resto de modificadores: las primeras documentaciones, la cronología, el orden de aparición y la distribución geográfica.

Las primeras documentaciones de los pronombres reforzados con *otros* que encontramos en el corpus, tanto en los textos documentales y como en los literarios, se sitúan en la segunda mitad del siglo XIII. No obstante, los dos corpus difieren en cuanto a la velocidad de la difusión durante los siglos posteriores. Los textos literarios presentan el mayor aumento de las variantes *nosotros* y *vosotros* en el siglo XV. En el siglo XVI estos pronombres se asientan como formas habituales no marcadas. Este proceso de difusión y adopción de las nuevas variantes coincide a grandes rasgos con el descrito por los estudios anteriores. En cambio, en los documentales el proceso de difusión es más lento y se prolonga durante más tiempo, de modo que en el XVI todavía son formas minoritarias.

En cuanto al orden de aparición, actualmente todo apunta a que fue la segunda persona *vosotros* la que apareció primero, la que lideró el cambio, mientras que el modificador se extendió a la primera *nosotros* por analogía. El hecho de que la segunda persona sea la que encabece el cambio se explica por la mayor necesidad de marcar explícitamente el número del pronombre frente al frecuente uso de *vós* como forma de tratamiento de cortesía. Además, el avance de 2ª a 1ª sigue la jerarquía implicacional de persona establecida para el parámetro de número (Croft 2003:161) (Capítulo 2, § 6.2), de modo que la marca avanzaría desde las posiciones más bajas.

Respecto a la distribución geográfica, todas las evidencias presentadas en el capítulo anterior apoyan que los pronombres reforzados *nosotros, vosotros* llegaron al castellano desde el oriente peninsular, desde el catalán y el occitano, y a través de las variedades navarroaragonesas. En el siglo XIV, las variantes ya estaban presentes en la mitad oriental de la península ibérica, y en el XV ocupaban todo el norte. Esta difusión geográfica se presenta del mismo modo en textos documentales y literarios. Es decir, el avance en el espacio geográfico se produce al mismo ritmo en ambos tipos de texto al margen de la presencia global de estas variantes en cada uno.

2.2. El uso de pronombres reforzados en otras lenguas

Además del castellano, el catalán y el occitano también reemplazaron los pronombres personales *nós* y *vós* por las formas reforzadas con continuadores de *alteros*, que hoy en día son formas no marcadas en estas lenguas. Otras variedades conservaron las formas simples y mantuvieron las compuestas como opciones enfáticas, como es el caso del francés (*nous autres, vous autres*), algunos dialectos del norte de Italia, el sardo y el retorromance (Fernández-Ordóñez 2011:76; Lausberg 1973:162; Meyer-Lübke 1974:103; Posner 1998:113).

² Para un estado de la cuestión más detallado sobre el reemplazo de *nós, vós* por *nosotros, vosotros* cf. Capítulo 3, § 2.

Spitzer (1947) conecta la existencia de los pronombres reforzados con *alteros* de las lenguas romances con las formas reforzadas del pronombre *you* del inglés estadounidense: *you two you three, you people...* y destaca la que considera más frecuente, *you all* o *y'all*, característica del sur de Estados Unidos³. Ante esto, se plantea la siguiente pregunta:

On pourrait se demander pourquoi, dans la même situation linguistique de base, la nécessité de distinguer un ou plusieurs allocutés, l'espagnol (et le cat., le port., le prov., le rhétoroman etc.) emploie *otro*, l'anglais *all 'tout'* comme moyen de distinction de pluriels (Spitzer 1947:171)

A partir del artículo de Spitzer (1947) y, concretamente, a partir de esta pregunta, Rini (1999) se plantea si, del mismo modo que en la actualidad los hablantes de inglés americano experimentan con distintas formas de marcar *you* como plural, quizá los hablantes del castellano de la Edad Media pudieron experimentar con maneras diferentes de marcar *vós* como plural antes de decantarse por *vosotros*.

2.3. *Nós* y *vós* con modificadores distintos a *otros* en castellano

Hasta donde sabemos, el de Rini (1999) es el único estudio que rastrea exhaustiva y cuantitativamente la aparición de elementos diferentes de *otros* junto a *nós* y *vós* para marcarlos como plural. Para ello, utiliza obras de entre los siglos XIII y XV, la mayoría textos cronísticos y muchos de ellos del *scriptorium* alfonsí (para la lista completa, cf. Rini 1999: n. 2). En su análisis encuentra apariciones muy esporádicas de *vós dos*, *vós amos* y *vós tres*, pero destaca la frecuencia de *vós todos*, «parallel to the American English *y'all*» (Rini 1999:210), que analiza en detalle en el resto del artículo. Así, llega a la conclusión de que, antes de que *vosotros* se generalizara, la variante reforzada con *todos* fue temporalmente la más frecuente.

No obstante, el estudio presenta varios problemas que obligan a tratar sus conclusiones con cautela. En primer lugar, a pesar de su exhaustividad, el número de ocurrencias de las variantes analizadas es muy reducido en todas las obras. Este es un inconveniente que señala el mismo autor (Rini 1999:211), pero esto no le impide concluir categóricamente unas páginas después que «*vos todos* had become established as an analytic pronominal syntagm and had begun to replace tonic *vos*» (1999:216).

En segundo lugar, cabe destacar que este autor únicamente compara la aparición de *todos* frente a *otros*. Sin embargo, para medir la relevancia real de *nós todos* y *vós todos*, sería necesario incluir en sus cálculos las ocurrencias de las formas simples *nós* y *vós*. En el siglo XIII, que es cuando, en principio, *todos* superaría en frecuencia a *otros* como modificador de los pronombres, la variante mayoritaria todavía es la simple.

Para terminar, cabe señalar que encontramos otras menciones superficiales a los modificadores. En García *et al.* (1990:70) se nombran las variantes *vós mismos* y *vós todos* como paralelos de *vosotros*. Eberenz (2000:59, 62-63) añade *ambos* y *solos* a la lista de modificadores que podían acompañar a los pronombres *nós* y *vós*. Además, señala que más tarde empezaron a aparecer junto a *nosotros* y *vosotros*, lo que sería un indicador de que estas variantes ya estarían gramaticalizadas.

2.4. Recapitulación: preguntas de investigación

De las cuestiones que hemos repasado en este breve estado de la cuestión, podemos extraer las preguntas de investigación que trataré de responder durante mi análisis. ¿Cuál

³ Rini (1999:210), haciendo referencia a Spitzer, añade además la forma *you guys*, que señala como característica de la zona centro-occidental de Estados Unidos, «used even if the group of addressees is composed entirely of females».

fue realmente la vitalidad de modificadores como *todos*, *mismos* y *ambos* antes de que *nosotros* y *vosotros* se generalizaran? ¿*Vós todos* llegó a reemplazar a *vós* en algún contexto tal y como propone Rini (1999)? De igual modo que *otros* se extiende antes en la segunda persona, ¿alguno de los modificadores tiene preferencia por *nós* o *vós*? ¿Cuáles de los refuerzos pronominales enfáticos presentan una distribución geográfica?

3. Datos y metodología específica

El corpus documental de este capítulo está formado por los textos del *Corpus de documentos españoles anteriores a 1800* (CODEA+). Para el análisis de los modificadores he utilizado como base la codificación realizada para el análisis del reemplazo de *nós*, *vós* por *nosotros*, *vosotros* el Capítulo 3⁴. Por tanto, he partido de una base de datos conformada por 5.466 ejemplos de todas las variantes de los pronombres de primera y segunda persona con referente plural o colectivo. La única diferencia respecto del capítulo anterior es que no solo he distinguido entre variantes con y sin el refuerzo *otros*, sino que he codificado cualquier otro elemento que apareciera modificando el pronombre. Los refuerzos que se han documentado en este corpus han sido *otros* (303 casos), *todos* (23), *mismos* (64) y *ambos* (20). A pesar de que Eberenz (2000:59) menciona el uso del modificador *solos*, este no se ha registrado en el CODEA+.

En cuanto a los textos literarios, se ha utilizado el corpus completo. En el Capítulo 3 se señalaba que uno de los mayores obstáculos para la codificación de los pronombres *nós* y *vós* era la necesidad de identificar y separar los casos en los que el pronombre se usaba con un referente singular, es decir, los ejemplos de *nós* mayestático o *vós* como forma de tratamiento. Debido al exceso de ocurrencias de formas simples y reforzadas extraídas del corpus literario (16.988 en total), se decidió hacer una selección aleatoria de hasta 2.000 ejemplos. Para el análisis de los pronombres con modificador, sin embargo, el tratamiento de los datos ha sido diferente. En este caso se han codificado todas las variantes reforzadas del corpus al completo. El uso de modificadores como *todos*, *mismos* o *ambos* permite desambiguar el referente sin necesidad de acudir a un contexto más amplio. Por tanto, de la muestra total de variantes pronominales, se han localizado 3.421 ejemplos de los refuerzos *otros*, *todos*, *mismos*, *ambos*, *solos*, *ensemble* o numerales.

No obstante, para conocer el alcance real de la difusión de las formas reforzadas es indispensable tener en cuenta las variantes simples *nós* y *vós*. Con la intención de salvar el impedimento de codificar el referente de los 13.566 ejemplos restantes, he decidido calcular la estimación de formas sin modificador a partir de la muestra del corpus utilizada para el Capítulo 3. Esta selección está formada por 1.933 ocurrencias de los pronombres con referente plural. La extracción de estos ejemplos se ha hecho de forma aleatoria y cada uno de los textos literarios está representado por un número de ejemplos proporcional a su peso dentro del corpus (cf. Tabla 2). De este modo, a partir de esta muestra, descrita y analizada en el capítulo anterior, podemos conocer la proporción de formas sin modificador y extrapolarla al total del corpus de textos literarios.

Variantes	XIII	XIV	XV	XVI	Total
<i>Nós, vós</i> (plural)	388	315	417	48	1168
<i>Nosotros, vosotros</i>	2	70	319	350	741
<i>Nós, vós</i> + otros modificadores	7	5	11	1	24
Total	397	390	747	399	1933

Tabla 7. Recuentos de las variantes simples, de las reforzadas con *otros* y de las reforzadas con otros modificadores en la muestra del corpus literario.

⁴ Para el proceso de extracción y limpieza de los datos, cf. Capítulo 3, § 3. Datos y metodología específica.

Variantes	XIII	XIV	XV	XVI
<i>Nós, vós</i> (plural)	¿?	¿?	¿?	¿?
<i>Nosotros, vosotros</i>	2	444	1344	1435
<i>Nós, vós + otros modificadores</i>	95	55	44	2

Tabla 8. Recuento de las variantes reforzadas en el corpus completo de textos literarios.

Por tanto, tenemos por un lado una muestra de casi 2.000 ocurrencias en la que se han seleccionado únicamente ejemplos con referente plural. En esta muestra conocemos las frecuencias absolutas de todas las variantes (Tabla 7). Por otro lado, tenemos el corpus completo de 16.988 ocurrencias, en el que solo conocemos el número de variantes reforzadas, no el de formas simples con referente plural (Tabla 8). A partir de estos datos, podemos estimar el número de formas simples de cada siglo mediante una regla de tres:

Muestra	<i>Nós, vós</i> (plural)	<i>Nosotros, vosotros</i>
	N	N
Corpus completo	X	N

Tabla 9. Ilustración de la regla de tres utilizada para estimar el número de formas simples del corpus literario completo a partir de la muestra.

Variantes	XIII	XIV	XV	XVI	Total
<i>Nós, vós</i> (plural)	[388]	[1998]	[1757]	[197]	[4340]
<i>Nosotros, vosotros</i>	2	444	1344	1435	3225
<i>Nós, vós + otros modificadores</i>	62	24	36	1	123
Total	[452]	[2466]	[3137]	[1633]	[7688]

Tabla 10. Estimación de las variantes *nós* y *vós* con referente plural del corpus de textos literarios a partir de la muestra.

En la Tabla 10 aparecen las ocurrencias estimadas para las variantes *nós* y *vós* con referente plural a partir de las cifras obtenidas en la muestra aleatoria. Esta estimación únicamente se utilizará para examinar la cronología de las variantes de forma general. El análisis en detalle de la difusión de los modificadores se centrará en las cifras reales obtenidas y en cálculos de frecuencia por 100.000 palabras.

Con el objetivo de completar el mapa peninsular, he acudido a corpus de otras variedades iberromances: el *Corpus Informatizat del Català Antic* (CICA) y el *Tesouro Medieval Informatizado da Lingua Galega* (TMILG). En ambos casos he seleccionado las tipologías de tipo documental para poder localizar geográficamente los documentos con más precisión. En cuanto al CICA (cf. Capítulo 2, § 5.3.a), he seleccionado las tipologías de *D – Prosa cancelleresca*, *E – Textos administratius*, *F – Textos jurídics* y *G – Llibres de cort*. En cuanto al TMILG (cf. Capítulo 2, § 5.3.b), se han filtrado los resultados por la tipología de *Prosa notarial* que proporciona el mismo corpus. El mayor obstáculo que se ha encontrado durante el trabajo con el TMILG es que se ha detectado una elevada cantidad de ejemplos que no aparecían en los resultados de las consultas pero sí en el contexto anterior o posterior de las ocurrencias listadas y que han tenido que añadirse manualmente⁵.

Debido a que el corpus auxiliar tiene como principal finalidad rastrear geográficamente las variantes con modificador, únicamente se han codificado y analizado los ejemplos de *nós* y *vós* que fueran acompañados de un refuerzo. En el CICA se han localizado 1761 ejemplos de pronombres reforzados con *altres*, 2 con *ambdós*, 4 con

⁵ Además, para el presente capítulo realicé búsquedas expresamente de los pronombres acompañados por modificadores distintos a *otros/outros*, como *todos*, *mesmos* y *ambos*. El corpus devolvió ejemplos que no habían aparecido en las búsquedas originales de las formas simples hechas para el Capítulo 3. Estos casos fueron incorporados a los datos presentados en dicho capítulo, pero esto plantea la posibilidad de que haya ejemplos de las variantes simples *nós* y *vós* que no hayan aparecido en ninguna de las búsquedas.

mateixos y 1 con *tots*; en el TMILG, 193 con *ambos*, 60 con *todos*, 51 con *outros* y 23 con *mesmos*.

4. Análisis

4.1. Primeras ocurrencias

En los textos documentales castellanos del siglo XIII encontramos las primeras documentaciones de todos los modificadores que se registran en el CODEA+: *otros*, *todos*, *misimos* y *ambos*. En (1) podemos ver el primer registro de cada variante:

(1)

- a) **nós todos** tres, con auctoritat e otorgamiento de los tredze e del sobredicho cabildo, **nós todos** sobredichos, el chanceler e el maestro con estos comendadores (CODEA-1200, Burgos, 1239)
- b) nuestros propincos o de los estraños o **nós misimos** contra esta carta d'esta vendeción (CODEA-0335, León, 1244)
- c) E **nós ambos** cabezales del arcipreste vos somos fiadores de fazérvoslo sano (CODEA-0156, Burgos, 1247)
- d) Entendimos que algunos de **vosotros** capellanos e clérigos non queredes recibir a los demandadores [...]. La coal cosa pesa mucho a nós porque aquella cosa amamos especialment sobre todas las otras e **vós otros** la devedes amar (CODEA-0979, Navarra, 1265)

La primera en aparecer es *nós todos*, y lo hace en un documento de Burgos de 1239 con dos ocurrencias de esta variante. Una de estas ocurrencias además va acompañada por el numeral *tres*. La aparición de un numeral junto a *todos* es bastante frecuente en el corpus. Concretamente, de las 23 ocurrencias de pronombre + *todos*, 12 van seguidas de *tres* o *cuatro*. Pocos años después, en un documento de León de 1244 se registra por primera vez *nós misimos*; y en otro de Burgos de 1247 aparece el primer caso de *nós ambos*. La última variante en documentarse es *vós otros* o *vosotros*, en un documento de Pamplona (Navarra) de 1265, que ya ha sido citado en el capítulo correspondiente al análisis de estas variantes. Aunque en el ejemplo aparecen dos ocurrencias, en este primer documento aparecen hasta 5 casos del pronombre, todos ellos en segunda persona.

De este modo, todas las primeras documentaciones se encuentran en un margen de menos de 30 años. Entre estos modificadores, *todos* no solo es el primero en aparecer, sino que en los documentos más tempranos también es el que tiene un número de ocurrencias ligeramente mayor.

En los textos literarios no es posible precisar del mismo modo la fecha de las primeras documentaciones. En cualquier caso, para este tipo de texto en el siglo XIII encontramos un total de 97 pronombres con modificador: 62 ejemplos de *misimos*, 32 de *todos*, 2 de *otros* y 1 de *solos*. De estas cifras extraemos que los modificadores más frecuentes en los textos más tempranos son, en primer lugar, *misimos* y, en segundo lugar, *todos*. No obstante, cabe destacar que ninguno de estos ejemplos proviene de textos navarroaragoneses. En estas primeras documentaciones tampoco parece haber una preferencia por la primera o la segunda persona en ninguno de los refuerzos pronominales.

4.2. Cronología

A pesar de las primeras documentaciones comentadas en § 4.1, la presencia de los modificadores es bastante limitada en ambos corpus.

En el CODEA+ ninguna de las formas reforzadas supera el 1 % en el siglo XIII, donde las formas simples *nós* y *vós* son la variante mayoritaria con diferencia. Además, tal y como vemos en la Tabla 11, las variantes sin modificador seguirán siendo las más utilizadas en los siglos siguientes. Como ya se ha mencionado en el capítulo

correspondiente, el reemplazo de las formas *nós* y *vós* por las formas reforzadas *nosotros* y *vosotros* en los textos documentales es más lento que en los literarios.

Variante	XIII		XIV		XV		XVI		XVII	
	N	%	N	%	N	%	N	%	N	%
<i>Otros</i>	6	0.51	32	1.8	103	6.47	139	17.03	23	21.5
<i>Todos</i>	10	0.85	9	0.51	3	0.19	1	0.12	0	0
<i>Mismos</i>	8	0.68	34	1.92	12	0.75	10	1.23	0	0
<i>Ambos</i>	8	0.68	6	0.34	5	0.31	1	0.12	0	0
Sin mod.	1146	97.28	1693	95.43	1468	92.27	665	81.5	84	78.5

Tabla 11. Ocurrencias y porcentajes de las variantes con y sin modificador en el CODEA+.

En la Tabla 11 podemos ver que las formas reforzadas con *todos*, *mismos* y *ambos* tienen una presencia bastante reducida en el corpus documental, de modo que ninguno de ellos supera el 2% en ninguno de los siglos⁶. Si solo miramos el número absoluto de ocurrencias, de todos los modificadores que se registran en el CODEA+ junto con los pronombres *nós* y *vós*, *todos* es el más frecuente del siglo XIII. No obstante, las 23 ocurrencias que se registran —20 de *nós todos* y 3 de *vós todos*— en todo el corpus están muy lejos de permitirnos confirmar que en algún momento pudieran empezar a reemplazar a *nós* y *vós* tal y como afirmaba Rini (1999:216).

En los textos literarios la vitalidad de los modificadores es todavía más limitada. En esta tipología textual, como bien sabemos, las variantes *nosotros* y *vosotros* se difunden más rápidamente y el cambio se completa en el siglo XVI. En el siglo XIV, a diferencia de lo que vemos en los textos documentales, los pronombres reforzados con *otros* ya han empezado a difundirse, por lo que la presencia de otros modificadores prácticamente queda acotada al XIII.

Variante	XIII		XIV		XV		XVI	
	N	%	N	%	N	%	N	%
<i>Otros</i>	2	0.41	444	17.78	1344	42.73	1435	87.82
<i>Todos</i>	32	6.6	24	0.96	5	0.16	1	0.06
<i>Mismos</i>	62	12.78	15	0.6	31	0.99	1	0.06
<i>Ambos</i>	0	0	9	0.36	5	0.16	0	0
<i>Solos</i>	1	0.21	1	0.04	1	0.03	0	0
Numeral	0	0	6	0.24	2	0.06	0	0
Sin mod. [estimación]	[388]	[80]	[1998]	[80.02]	[1757]	[55.87]	[197]	[12.06]

Tabla 12. Ocurrencias y porcentajes de las variantes con y sin modificador en el corpus de textos literarios⁷.

La principal diferencia respecto de los textos documentales es que, tal y como puede verse en la Tabla 12, en los textos más tempranos la presencia de los modificadores es mayor en los literarios. En los textos de este periodo del CODEA+, ningún modificador superaba

⁶ Tras una primera búsqueda en el corpus de los pronombres *nós* y *vós* junto con las variantes de *todos*, *mismos* y *ambos* puede parecer que la frecuencia de aparición es mayor que lo que se refleja en la tabla. No obstante, durante la codificación se han tenido que descartar muchos casos, dado que se ha considerado que, en realidad, el modificador no acompañaba al pronombre, sino a otro elemento posterior. Uno de los ejemplos más frecuentes es el de «*nós*, ambas las dichas partes», que aparece con algunas diferencias gráficas más de medio centenar de veces en documentos desde el siglo XIV al XVI. Lo mismo ocurre con frases similares a «*obligo a vós todos los dichos bienes*».

⁷ Debido a la imposibilidad de desambiguar el referente de todos los ejemplos de *nós* y *vós* de los textos literarios, se ha calculado una estimación a partir de los resultados de la muestra aleatoria realizada para el Capítulo 3. Los datos presentados para esta variante en la Tabla 14 únicamente tienen el objetivo de servir de comparación para el resto de variantes. Para más detalles sobre esta estimación cf. § 3. Datos y metodología específica.

el 1 % de aparición. En cambio, en el corpus literario, *mismos* alcanza casi un 13 % de aparición y *todos* supera el 6 %. En cualquier caso, en los siglos siguientes su presencia es marginal.

Para analizar en detalle la evolución de cada uno de los modificadores, he excluido las formas simples y he calculado la frecuencia de aparición del resto de variantes por cada 100.000 palabras en ambos corpus.

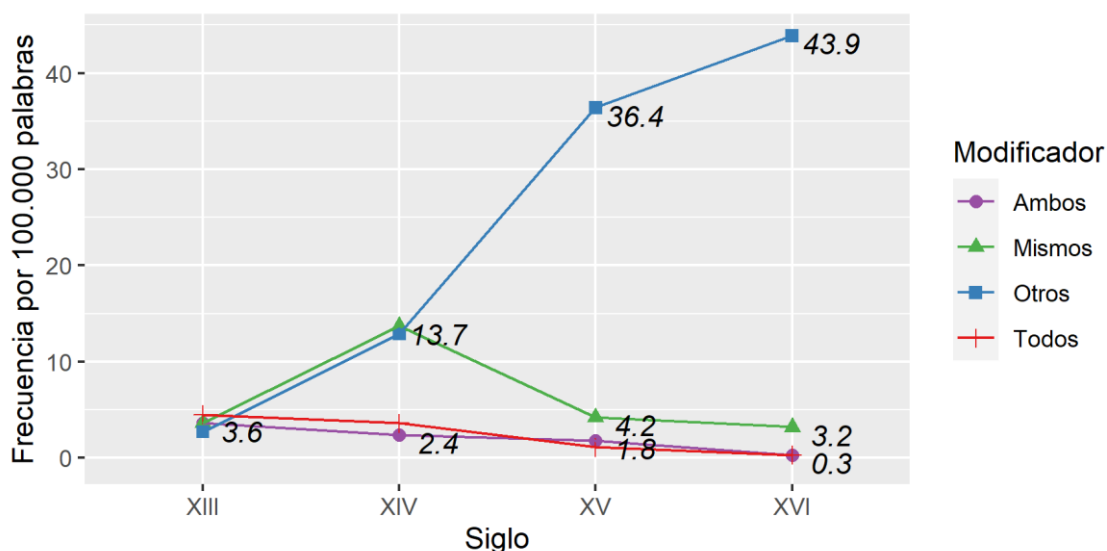


Figura 18. Frecuencia por cada 100.000 palabras de cada uno de los modificadores en el CODEA+ (siglos XIII-XVI).

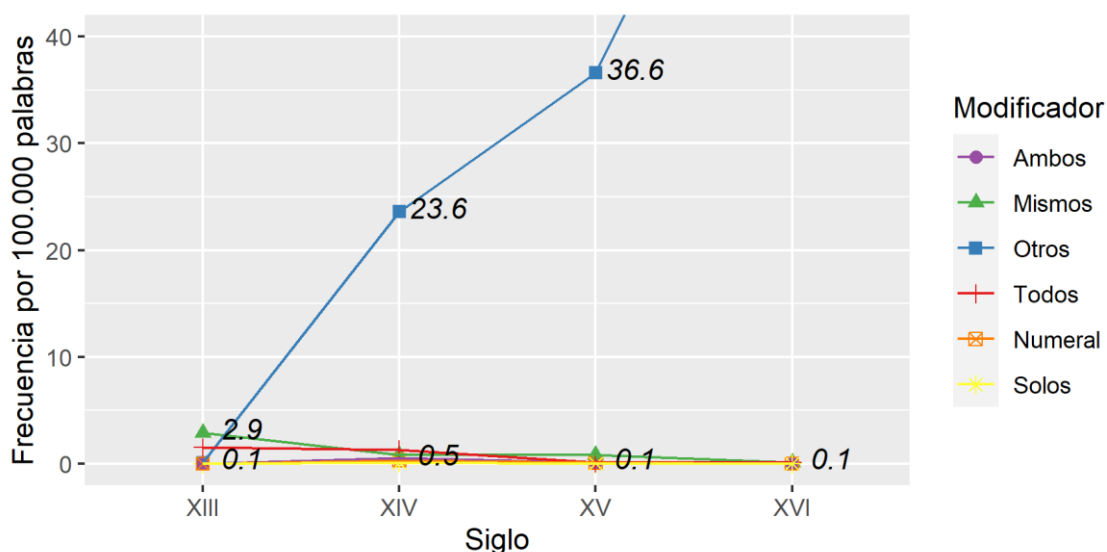


Figura 19. Frecuencia por cada 100.000 palabras de cada uno de los modificadores en el corpus de textos literarios (siglos XIII-XVI).

En los textos documentales del CODEA+ (Figura 18) podemos ver que los cuatro modificadores parten prácticamente del mismo punto en el siglo XIII y, en el siglo XIV, *otros* y *mismos* incrementan su frecuencia en la misma medida. Sin embargo, en los siglos siguientes *otros* sigue creciendo de forma lenta y constante, mientras que *mismos* declina y en los siglos XV y XVI se mantiene en un uso esporádico. Por otro lado, *todos* y *ambos* en ningún momento aumentan su frecuencia de uso, sino que van decreciendo lentamente hasta que en el siglo XVI solo se registra una ocurrencia de cada uno de estos dos

modificadores. Finalmente, en el siglo XVII ya no encontramos ningún caso de *nós* y *vós* con *todos*, *mismos* o *ambos*.

En los textos literarios, en cambio, hay mucha menos variación (Figura 19). Excepto *otros*, la presencia del resto de modificadores es mucho más marginal: el uso de *ambos*, *solos* o de algún numeral es totalmente esporádico y casi anecdótico, y estos refuerzos se documentan por última vez en los textos del XV; la presencia de *todos* y *mismos*, que destacan porcentualmente en el siglo XIII (Tabla 12), se prolonga hasta el XVI, pero en este siglo encontramos solo una ocurrencia de cada uno.

Además, es necesario destacar que ya no vemos el ligero aumento de *mismos* en el siglo XIV que observábamos en el corpus documental. En su lugar, el uso de *nosotros* y *vosotros* aumenta rápidamente hasta una frecuencia de 23.6 cada 100.000 palabras, casi el doble que en el CODEA+. En este mismo periodo, en el siglo XIV, empezamos a registrar casos en los que *mismos* y *todos* modifican ya no a *nós* o *vós*, sino a las formas reforzadas *nosotros* y *vosotros*. En el CODEA+ esto no ocurre hasta el siglo XV:

(2)

- a) dios ha ordenado vos otros seyer preueres et auos dado poder de iutgar **vosotros mismos** (CQ2)
- b) nosotros reconocemos en **nos otros mismos** que hauemos fallido (GCI)
- c) por donde a **nos otros todos** seya apparellada Razon (HTroyana)
- d) E obligamos a **nosotros mesmos** e a todos nuestros bienes de nós e de cada uno de nós (1424, CODEA-0551)⁸

Como ya hemos mencionado antes, según Eberenz (2000:59, 62-63) la presencia de los modificadores ante las formas ya reforzadas con *otros* sería un indicio de su gramaticalización. Por tanto, a partir de la fecha de los ejemplos de (2) y del descenso de *mismos* en las Figuras 18 y 19 podemos determinar que los pronombres compuestos *nosotros* y *vosotros* empezaron a gramaticalizarse en el siglo XIV. Esto se refleja en primer lugar en los textos de tipo literario (2a-c). Los textos de tipo documental, más conservadores, empiezan a evidenciar la gramaticalización más tarde, en el siglo XV (2d).

Es llamativo que encontremos indicios de la gramaticalización de *nosotros*, *vosotros* en los textos del CODEA+ del siglo XV. En este periodo y en este corpus, los pronombres compuestos no superan el 10 % de aparición (Tabla 11). Es decir, se encontrarían todavía en una fase inicial del cambio. Por tanto, aunque el cambio avanza mucho más lento en los textos documentales respecto de los literarios y los pronombres reforzados están lejos de ser formas mayoritarias, la aparición de variantes como *nosotros mesmos* en una etapa tan temprana de la difusión sugeriría que estos pronombres han entrado en este tipo de textos ya gramaticalizados.

La idea de que modificadores como *todos* o *mismos* pueden acompañar a no solo las formas simples *nós* y *vós*, sino también a los pronombres compuestos *nosotros* y *vosotros* puede plantear dudas sobre si realmente estos refuerzos estaban en competición con *otros*. Si así fuera, si *todos* y *mismos* fueran simplemente elementos enfáticos que pudieran aparecer junto con cualquier pronombre de primera y segunda persona del plural simple o compuesto, su uso no tendría por qué descender tan abruptamente como lo hace en cuanto *nosotros* y *vosotros* empiezan a extenderse, sino que podrían haber seguido funcionando como refuerzos enfáticos de *nós* y *vós* hasta su reemplazo total. Incluso, esperaríamos que su frecuencia se mantuviera más o menos estable, al margen de cuál de las variantes pronominales —simple o compuesta— fuera la mayoritaria.

⁸ Otros ejemplos los podemos encontrar en los documentos CODEA-0414, CODEA-1364, CODEA-0449, CODEA-0835, CODEA-0645, CODEA-1342.

En la Figura 20 podemos ver la frecuencia por cada 100.000 de cada uno de los modificadores, pero no solo como refuerzos de *nós* y *vós*, sino también de *nosotros* y *vosotros*. En líneas generales, encontramos un descenso en el uso de refuerzos pronominales entre el siglo XIII y XVI. El hecho de que su uso no sea constante demuestra que no podían utilizarse indistintamente tras cualquiera de las variantes pronominales. Es decir, como muestran los ejemplos de (2), existía la posibilidad de que estos refuerzos acompañaran a los pronombres compuestos, pero no eran tan habituales como con los simples. La frecuencia de aparición de alguno de estos modificadores junto a un pronombre de primera o segunda persona de plural es mayor en el XIII, con las variantes simples. Por tanto, estos modificadores pudieron usarse como estrategias para desambiguar el número de *nós* y *vós*. Tras la difusión de *otros*, los demás refuerzos se descartan y quedan relegados a funciones enfáticas ocasionales. *Ambos*, *solos* y los numerales tienen una frecuencia absolutamente mínima que se mantiene por debajo del 0.5 por cada 100.000 palabras durante todo el periodo analizado. *Mismos*, después de un considerable descenso en el XIV y de ser desechado como modificador de las formas simples, es el único que se mantiene con una frecuencia destacable reforzando los pronombres compuestos *nosotros*, *vosotros* en el siglo XVI. Sin embargo, el empleo de *mismos* en este contexto no es tan común como tras las formas simples. A partir de este momento, se extiende a los singulares (p. e. «vós mesmo no lo podéis creer», *Espejo*) o a otros pronombres del paradigma (p. e. «ella mesma le cortava las conservas» o «contra sí mesmos», *Espejo*). La evolución más interesante es la del modificador *todos*, que presenta un descenso en forma de curva en S. De entre todos los modificadores analizados, *todos* es el que muestra un reemplazo más claro por *otros*. Una vez que *nosotros* y *vosotros* se generalizan, *todos* deja de aparecer como modificador pospuesto, incluso tras las formas reforzadas. Posteriormente, lo encontramos en posición antepuesta: «que tu señor y todos vosotros os salvéis» (*CNE*).

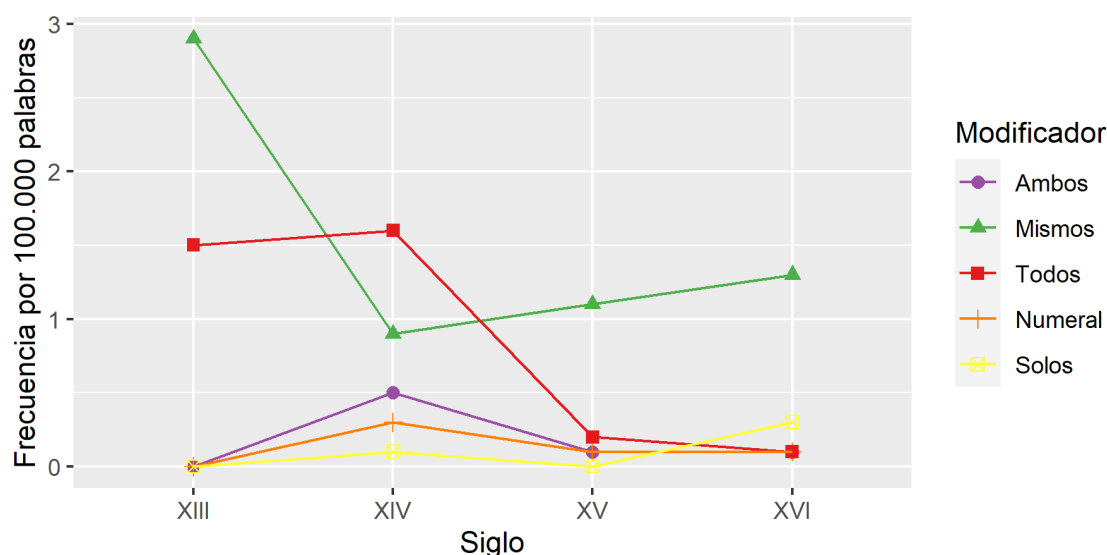


Figura 20. Frecuencia por cada 100.000 palabras de cada uno de los modificadores junto con *nós*, *vós*, *nosotros*, *vosotros* en el corpus de textos literarios (siglos XIII-XVI)

En definitiva, durante el siglo XIII, los hablantes experimentan con distintos refuerzos pronominales enfáticos que desambiguan el número de las variantes simples *nós* y *vós*. Una vez que *nosotros* y *vosotros* empiezan a difundirse y se gramaticalizan, antes incluso de que se conviertan en formas mayoritarias, el resto de modificadores se descartan y quedan relegados a otras funciones enfáticas ocasionales. *Mismos* se usa con valor

reflexivo y se extiende a otros pronombres del paradigma. Por ese motivo, se mantiene también tras *nosotros* y *vosotros*. *Todos* desaparece como refuerzo pronominal pospuesto y en la actualidad lo encontramos antepuesto. Los únicos que mantienen una frecuencia de uso similar en todo el periodo son aquellos refuerzos con un empleo más marginal que impide ver patrones de difusión o retroceso concretos: *ambos*, *solos* y los numerales.

Cabe destacar que, aunque la superioridad de *todos* respecto de *otros* es muy breve y está limitada al siglo XIII, es el refuerzo que por su comportamiento (Figura 20) parece competir más claramente con *otros*. Es difícil hablar de cambio fallido, ya que en su momento de mayor difusión no sobrepasaría lo que entenderíamos como una fase muy incipiente del cambio. No obstante, tal y como se verá en las secciones siguientes, es una variante que vale la pena incluir en la descripción del reemplazo de *nós* y *vós* por *nosotros* y *vosotros*.

4.3. Primera y segunda persona

Si volvemos a las primeras documentaciones de las variantes reforzadas citadas en (1), es necesario mencionar que *todos*, *mismos* y *ambos* aparecen en su primera documentación con la primera persona *nós*, mientras que *otros* aparece antes con la segunda persona *vós*. Esto no parece casual, ya que el corpus documental no registra la variante de primera persona *nosotros* hasta un siglo después, en 1378 (CODEA-2321); mientras que antes de esta fecha sí encontramos abundantes casos de *nós* con *todos*, *mismos* y *ambos*. Esto podría sugerir que la aparición de uno u otro modificador podría estar determinada por la variable persona.

Respecto a los textos documentales (Figura 21), en la segunda persona, la variante más frecuente durante todo el periodo analizado es *vosotros*, que apenas deja espacio para el resto de modificadores. En cambio, en la primera persona no solo no encontramos *otros* hasta finales del siglo XIV, sino que *ambos*, *mismos* y *todos* tienen una vitalidad mayor en los textos más tempranos y su presencia se prolonga hasta la primera mitad del XVI. Así, el 89.7 % de los pronombres reforzados con *ambos*, *mismos* y *todos* en el CODEA+ son de primera persona.

La situación es totalmente diferente en los textos literarios (Figura 22), en los que no se aprecian diferencias entre la primera y la segunda persona. En el siglo XIII, *todos* y *mismos* son los modificadores más frecuentes en ambas personas. En el siglo XIV, *otros* se extiende al mismo ritmo tanto tras *nós* como tras *vós*, y el resto de refuerzos quedan relegados a un uso esporádico, que se prolonga en los siglos siguientes.

La diferencia entre los textos documentales, que muestran una distribución clara de los modificadores según persona, y los literarios, que presentan un desarrollo parejo en ambas personas, puede deberse a la rapidez con la que se produjo el cambio. En los textos literarios, el reemplazo de *nós*, *vós* por *nosotros*, *vosotros* fue más rápido que en los documentales. Por tanto, el periodo de variación y experimentación con otros modificadores fue menor. En los textos documentales el refuerzo *otros* tardó un siglo en extenderse de la segunda persona a la primera. Esto pudo abrir la puerta a una mayor presencia de los otros modificadores. En los textos literarios, el avance de *nosotros* va algo por detrás de *vosotros*, pero la diferencia no es tan grande y de hecho documentamos ambas variantes desde el siglo XIII (Capítulo 3, § 4.1 y 4.3).

De estos datos deducimos que la variable persona es relevante en las etapas más tempranas del refuerzo de *nós* y *vós*. En los textos literarios, donde el modificador *otros* margina rápidamente a todos los demás en el siglo XIV, esta diferencia no se aprecia. En cambio, los textos documentales, que presentan un cambio más lento, nos permiten ampliar y ver con un mayor detalle las fases iniciales, lo que posibilita distinguir aquellos contextos a los que acceden primero cada uno de los refuerzos. En este tipo de textos,

vemos que *vós* favorece claramente el modificador *otros*, mientras que *nós* presenta más variación.

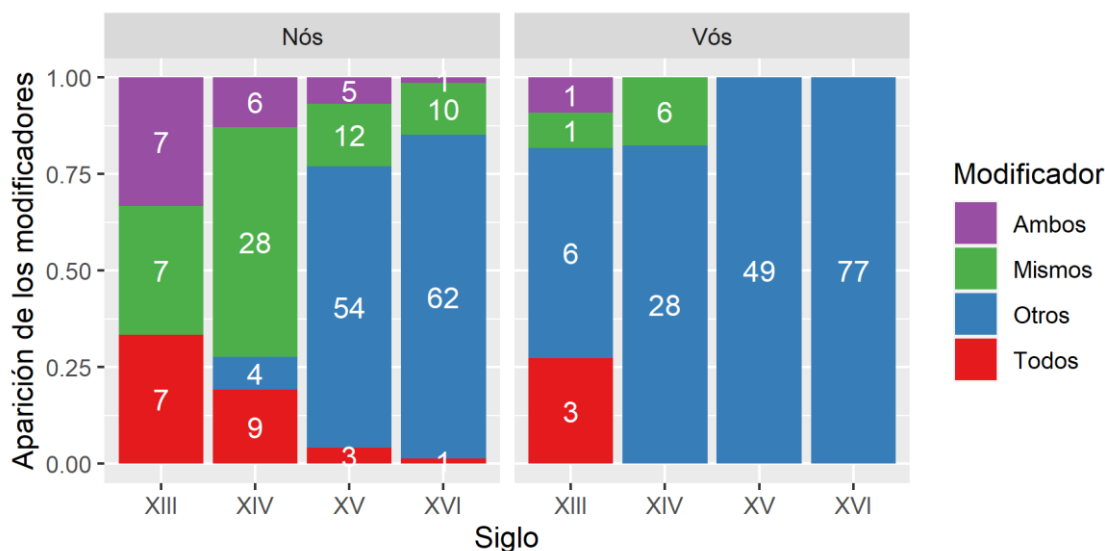


Figura 21. Aparición de los pronombres con modificador en primera (*nós*) y segunda (*vós*) persona en el CODEA+.

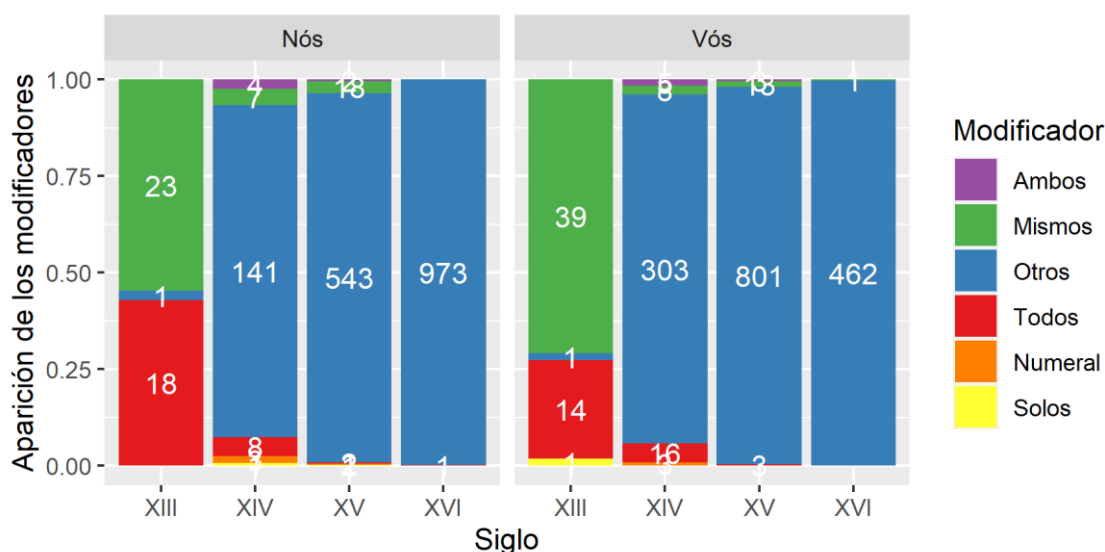
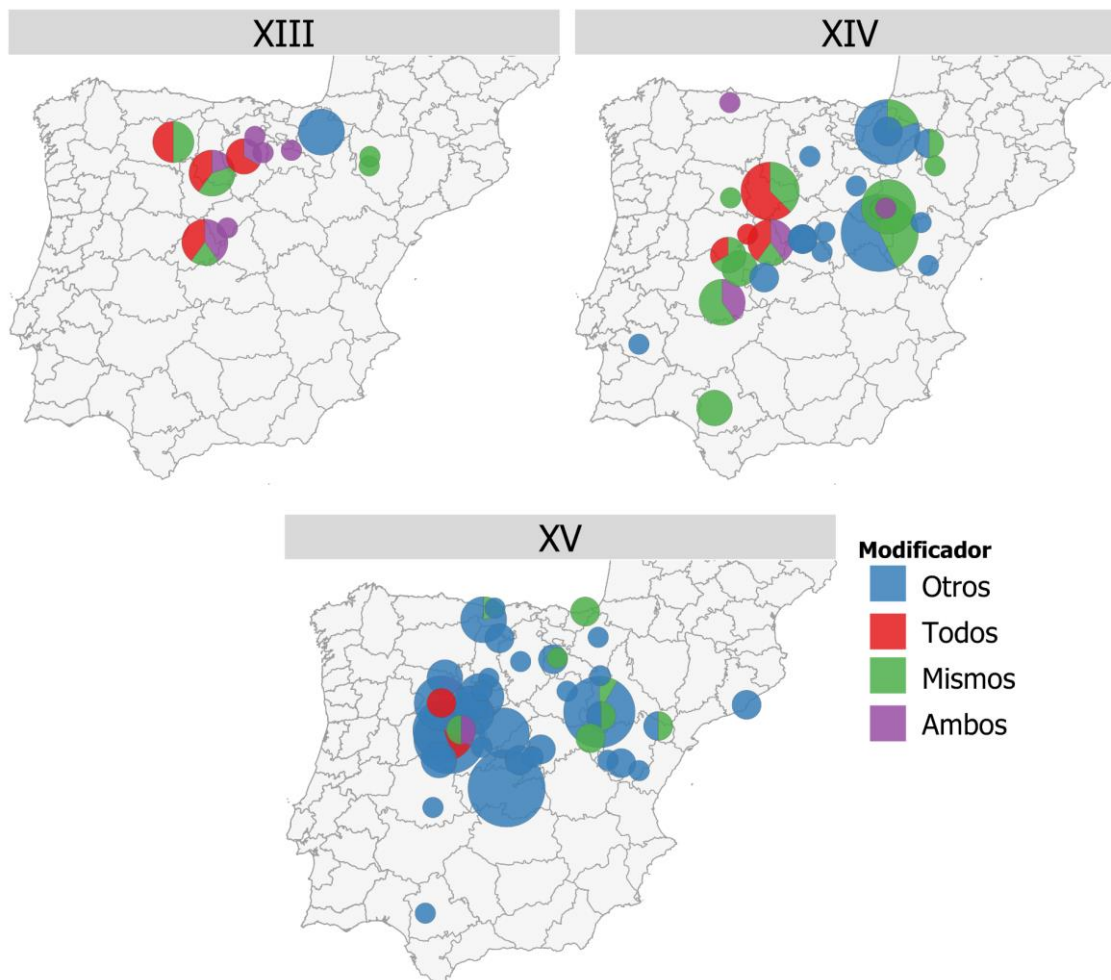


Figura 22. Aparición de los pronombres con modificador en primera (*nós*) y segunda (*vós*) persona en el corpus de textos literarios.

4.4. Distribución geográfica

En el Capítulo 3 (§ 4.5) hemos probado el origen oriental de las variantes pronominales reforzadas *nosotros* y *vosotros*, que llegaron al castellano desde el catalán y el occitano a través de las variedades navarroaragonesas, y que se extendieron por la península ibérica desde ahí hacia el oeste y el sur. En consecuencia, es posible que las variantes con otros modificadores se conserven más tiempo en el occidente peninsular.

En esta sección pretendo trazar el origen y la difusión de cada una de las variantes reforzadas en los textos literarios y documentales en castellano. Asimismo, con el objetivo de completar el *continuum* del norte peninsular, incorporaré datos de textos documentales catalanes y gallegos.



Mapa 4. Localización de cada una de las variantes pronominales reforzadas en el CODEA+ (siglos XIII-XV).

El cartografiado de las variantes de los textos documentales permite distinguir algunos patrones de difusión en el espacio peninsular. En cuanto a *otros*, como bien sabemos, en el siglo XIII todas sus ocurrencias pertenecen al mismo documento y se concentran en la actual provincia de Navarra. En el siglo XIV se extienden rápidamente hacia el sur por Aragón hasta Guadalajara, de modo que la mayoría de las ocurrencias se encuentran en la mitad oriental del norte peninsular. Al mismo tiempo, empiezan a aparecer casos de forma más esporádica en el norte y el centro de la península, en Burgos, Madrid y Toledo. En el siglo XV, a pesar de que ya hemos visto que el porcentaje de las formas *nosotros* y *vosotros* todavía no es demasiado alto, en el Mapa 4 puede verse claramente que se han extendido prácticamente por toda la mitad septentrional de la península, exceptuando Asturias y León, donde sí que encontramos otros documentos que únicamente utilizan las formas simples *nós* y *vós*.

En cambio, *mismos* y *ambos* no parecen tener una distribución geográfica definida. Podemos encontrar ocurrencias de *mismos* desde el extremo más oriental al más occidental en todos los siglos; y los casos de *ambos* son demasiado esporádicos y dispersos para poder distinguir un patrón.

Finalmente, a pesar, de nuevo, del bajo número de ocurrencias con el que estamos trabajando, *nós todos* y *vós todos* sí parecen ubicarse en una zona geográfica concreta. Desde el siglo XIII al XV, todas las apariciones de este modificador se concentran en la

mitad occidental del norte peninsular. En el siglo XIII esta variante se registra en Burgos, Palencia, León y Ávila. En el siglo XIV, *todos* se mantiene en el occidente, en Valladolid, Salamanca y Ávila. En último lugar, en el siglo XV, aunque ya solo aparecen 3 casos de esta variante, estos se encuentran en el extremo más occidental, arrinconados en Zamora y Salamanca por *nosotros* y *vosotros*, que ya se han extendido por toda la mitad septentrional de la península. En definitiva, las formas que parecen sujetas a variación geográfica son *nosotros*, *vosotros*, pero también *nós todos* y *vós todos*.

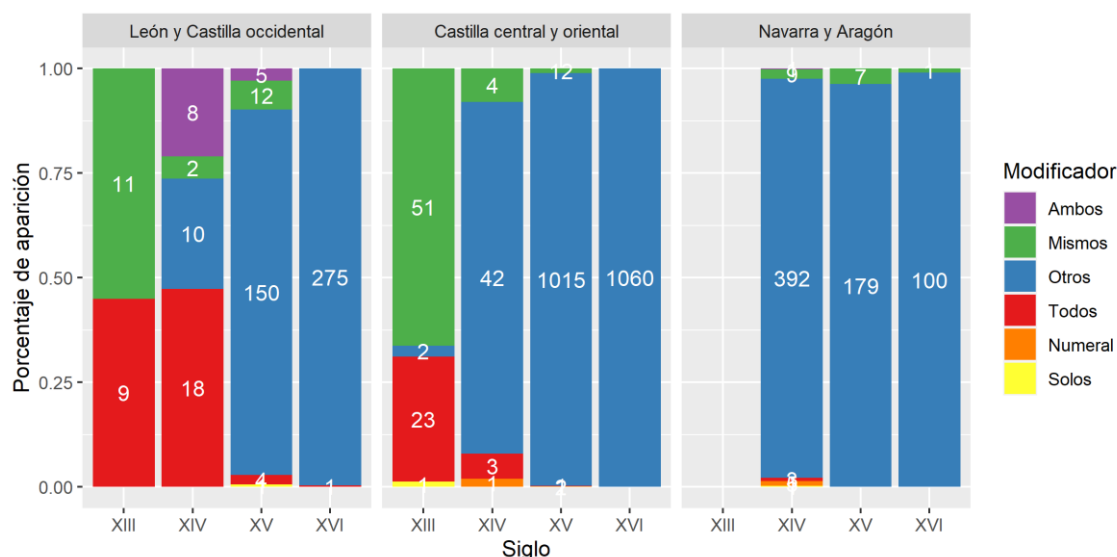


Figura 23. Distribución de las variantes pronominales reforzadas por zona en los textos literarios (siglos XIII-XVI).

La vinculación de las variantes pronominales reforzadas con el modificador *todos* y el occidente peninsular se mantiene en los textos literarios. Así, en la Figura 23 puede observarse que la vitalidad de este modificador disminuye a medida que nos acercamos al oriente. En León y Castilla occidental, *todos* es una de las principales variantes reforzadas durante los siglos XIII y XIV. Además, es la zona en la que sobreviven más tiempo, ya que encontramos el último ejemplo en el siglo XVI. En Castilla central y oriental se documentan por última vez en el XV; y en Navarra y Aragón en el XIV.

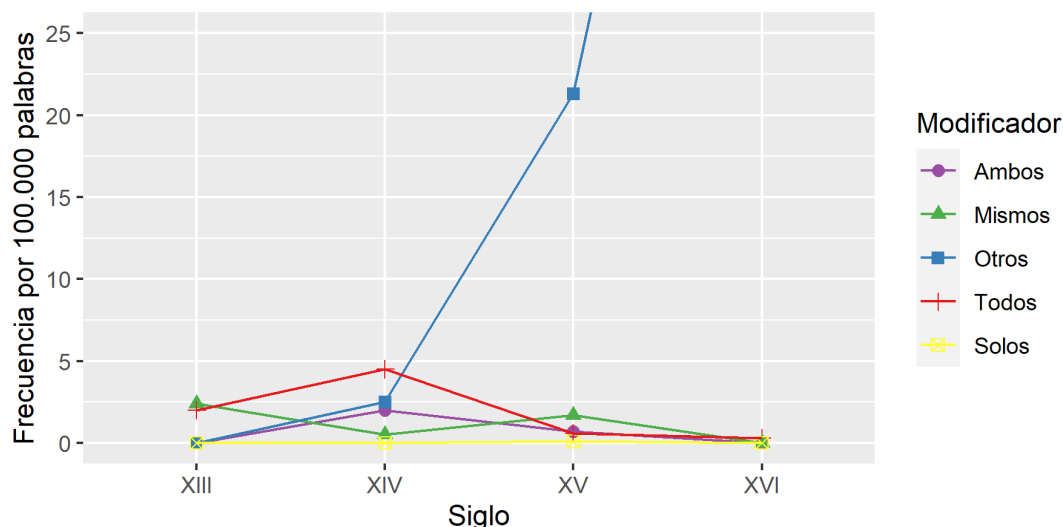


Figura 24. Frecuencia por cada 100.000 palabras de cada uno de los modificadores en los textos literarios de León y Castilla occidental (siglos XIII-XVI).

Si consideramos únicamente los datos de los textos literarios occidentales, podemos observar que en esta zona sí hubo una predominancia clara de los pronombres reforzados con *todos* durante la mitad del periodo analizado, en los siglos XIII y XIV. A partir del XV, en cambio, la frecuencia de uso de *nosotros* y *vosotros* se dispara, mientras que la de *nós todos* y *vós todos* se desploma. En la Figura 24 podemos llegar a distinguir cierta curva en campana propia de los cambios incompletos o fallidos.

Otra cuestión relevante respecto de la distribución de los modificadores en los textos literarios es el uso de *ambos* y el numeral *dos*. Frente a los textos documentales, donde los ejemplos de *ambos* estaban dispersos en el espacio geográfico, en los literarios este modificador está localizado en los textos de León y Castilla occidental (3). En cambio, en el resto del norte peninsular, para expresar el valor dual se utiliza el numeral *dos* (4).

(3)

- a) Mas lydiat lo **uos ambos**; pensatt delo liurar (*Alexandre*)
- b) Pero, porque a **vos amos** amo mucho, non me vos puedo en todo escusar (*CTroyana*)
- c) Sy quier aun non es partida la batalla de **nos anbos**; ante vos conviene a fazer más que cuydades (*Otas*)
- d) vna buena medjanera que sepa sabiamente andar esta carrera que entienda de **vos anbos** byen la vuestra manera (*LBA*)
- e) Rey don garcia sallid aparte & apartar se ha esta ljd por **nos amos** (*Cronica1344*)

(4)

- a) de lo que el ffablo vna palabra entendemos **nos dos** (*Mostrador*)
- b) e non con nosotras otro ageno en la casa estava, salvo **nós dos** solas en casa (*Arragel*)
- c) Et sepas que ninguno non puede saber lo que **nos dos** fazemos solos (*HTroyana*)

Con el objetivo de completar el mapa peninsular, he acudido a corpus de otras variedades iberorromances: el *Corpus Informatizat del Català Antic* (CICA) y el *Tesouro Medieval Informatizado da Lingua Galega* (TMILG). Como ya se ha mencionado en la metodología de este capítulo, en ambos casos he seleccionado las tipologías de tipo documental para poder localizar geográficamente los documentos con más precisión.

Como podemos ver en la Figura 25, la situación del catalán y el gallego es muy diferente. En el CICA prácticamente las únicas variantes reforzadas que se utilizan son *nosaltres* y *vosaltres*, probablemente porque en catalán estas se extendieron y gramaticalizaron muy tempranamente. Así, las ocurrencias del resto de modificadores son anecdóticas. Además, otra muestra de que las variantes reforzadas con *otros* están gramaticalizadas en catalán es que desde la primera mitad del siglo XIV podemos encontrar algunos ejemplos de *nosaltres* y *vosaltres* reforzados con los principales modificadores, como los que vemos en (5):

(5)

- a) volem que **vosaltres abdosos**, el dit mísser Ramon e En Guillem Çabadia donets recapte per part de la nostra cort (*Documents de la Cancelleria d'Alfons III*, s. XIV)
- b) com encara en profits de **vosaltres mateixs** e de la ciutat, axí que quant a aquesta part contra **vosaltres mateixs** nos par que hajats scrit (*Lletres reials a la ciutat de Girona*, s. XV)
- c) Pregam e exortam-vos per ço, a **vosaltres tots** e a cascun de vós (*Epistolari de la València Medieval*, s. XV)

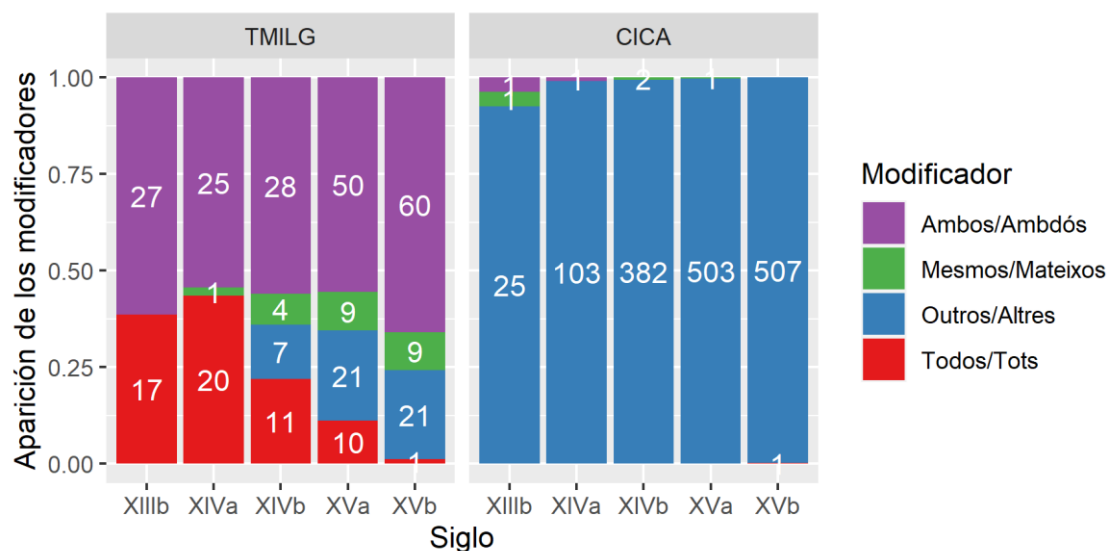


Figura 25. Aparición de las formas pronominales reforzadas en el TMILG y el CICA por mitad de siglo.

En cambio, en el TMILG la evolución de las variantes es totalmente diferente. Lo más llamativo a primera vista es el uso constante y estable de *nós ambos* y *vós ambos* durante todo el periodo analizado. La mayoría de los casos de estas variantes reforzadas con *ambos* tienen lugar en cartas de compraventa y donaciones, y hacen referencia al matrimonio que participa en la transacción.

En cuanto al resto de modificadores que aparecen en el corpus gallego, *mesmos* ocupa un lugar marginal, mientras que *todos* destaca con una frecuencia bastante mayor que la que encontramos en el CODEA+ hasta mediados del siglo XIV. En la segunda mitad de este siglo empiezan a documentarse casos de *outros*, que en los siglos siguientes irá reemplazando a *todos*.

En definitiva, los datos de los tres corpus analizados —CODEA+, textos literarios, TMILG y CICA— parecen indicar que *nós todos* y *vós todos* tuvieron una mayor vitalidad en la mitad occidental de la península ibérica. Para concluir, en la Tabla 13 podemos ver el porcentaje de las variantes pronominales reforzadas con *todos* frente al resto de formas con modificador en todos los corpus documentales (en el cálculo de estos porcentajes no se incluyen las formas simples). He separado los resultados en cinco categorías que pretenden reflejar el *continuum* del norte de la península, de oeste a este: los resultados del TMILG; los del CODEA+2015 separados en tres zonas, 1) León y Castilla occidental, 2) Castilla central y oriental y 3) Navarra y Aragón⁹; y los del CICA.

Siglo	TMILG	León y Castilla occidental	Castilla central y oriental	Navarra y aragón	CICA
XIII	38.6 %	42.9 %	28.6 %	0.0 %	0.0 %
XIV	32.3 %	47.4 %	0.0 %	0.0 %	0.0 %
XV	6.1 %	5.5 %	0.0 %	0.0 %	0.1 %

Tabla 13. Porcentaje de aparición de *nós todos* y *vós todos* en comparación con el resto de formas reforzadas en el TMILG, el CODEA+ y el CICA.

Así, las variedades navarroaragonesas y el catalán no registran *nós todos* y *vós todos*, que en el siglo XIII se documentan tanto en castellano como en gallego. En el siglo XIV

⁹ En León y Castilla occidental he incluido los datos del CODEA+ adscritos a las actuales provincias de Asturias, León, Zamora, Salamanca, Valladolid, Palencia y Ávila; y en Castilla central y oriental, los de Cantabria, Segovia, Madrid, Burgos, Guadalajara, Soria y La Rioja.

empiezan a retroceder y desaparecen de Castilla central y oriental. En el XV solo queda algún rastro esporádico en Galicia, en León y en Castilla occidental. Por tanto, además de *nosotros* y *vosotros*, de origen oriental, las únicas variantes que parecen claramente vinculadas a una zona geográfica son *nós todos* y *vós todos*, con una frecuencia mayor y más prolongada en las localidades más occidentales. Los datos del CICA y el TMILG confirman esta tendencia: mientras que en catalán *todos* tiene una presencia prácticamente nula, en gallego *todos* es más frecuente que *outros* durante los siglos XIII y XIV.

En cuanto a *mismos* y *solos*, ninguno de los dos parece tener una distribución geográfica concreta y queda todavía por identificar hasta qué punto estos modificadores compiten por los mismos contextos que *otros* y *todos*. *Nós solos* y *vós solos* tienen una frecuencia muy baja en todos los documentos castellanos y es probable que el significado tan específico de estos sintagmas haya impedido que se generalizaran. *Ambos* está en una situación similar, con la diferencia de que tiene una vitalidad notoria en los textos gallegos. Por otro lado, *nós mismos* y *vós mismos* podrían haber sido mejores candidatos, ya que tienen una frecuencia comparable a *nosotros* y *vosotros* en los textos más tempranos y su significado es más amplio que el de *ambos* y *solos*. Sin embargo, es una variante que no se ve afectada por la variable diatópica y que se extiende a otros pronombres dentro del paradigma. Además, en la actualidad participa en un cambio en marcha dentro del paradigma de los pronombres personales con el aparente reemplazo de *consigo* por *con sí mismo*¹⁰.

En definitiva, para el refuerzo de los pronombres *nós* y *vós* surgen dos candidatos en zonas distintas de la península ibérica: *todos* en occidente y *otros* en oriente. La variación se resuelve en las fases iniciales del cambio. De este modo, los resultados aquí presentados matizan los de Rini (1999). En términos de frecuencias y porcentajes, *vós todos* estuvo lejos de reemplazar a la forma simple *vós*, pero es necesario incluir las variantes pronominales reforzadas con *todos* para comprender de forma íntegra el reemplazo de las simples en todas las variedades iberorromances. Solo teniendo en cuenta todas las variantes que intervienen en el proceso de cambio y trazando el origen y la difusión de cada una en el espacio geográfico podemos completar el puzzle del reemplazo de los pronombres simples por los reforzados.

5. Resumen

En este capítulo he analizado el origen y la distribución de refuerzos pronominales enfáticos como *todos*, *mismos* y *ambos*, que podían modificar a los pronombres *nós* y *vós* antes de que *nosotros* y *vosotros* se generalizaran. La incorporación de estas variantes al estudio sobre el reemplazo de los pronombres simples *nós* y *vós* completa el análisis realizado en el Capítulo 3.

A partir de las preguntas de investigación presentadas al final del estado de la cuestión (§2) hemos obtenido las siguientes respuestas:

La vitalidad de modificadores como *todos*, *mismos* y *ambos* está limitada, en general, a los textos más tempranos. En ningún caso el uso de alguno de estos refuerzos pronominales supera al de las formas simples *nós* y *vós*.

¹⁰ García (1989) explica que, del mismo modo que durante la Edad Media se perdieron las variantes pronominales sintéticas *connusco*, *convusco* que analizaremos en el capítulo siguiente, en la actualidad *connmigo*, *contigo* y *consigo* se encuentran en un proceso similar. El CUM enclítico de estos pronombres se estaría perdiendo a favor de variantes analíticas con la preposición y el correspondiente pronombre tónico singular. En el caso de la tercera persona, además de los pronombres *él* o *sí*, se añadiría el refuerzo *mismo*, dando lugar a variantes como *con él mismo* y *con sí mismo*.

Durante los siglos XIII y XIV, los hablantes experimentan con distintos elementos enfáticos que modifican a los pronombres *nós* y *vós*, y cuya función común es marcarlos como plurales. Mi hipótesis, por tanto, es que estos modificadores pudieron usarse como estrategias para desambiguar el número de *nós* y *vós*.

El uso de estos refuerzos pronominales se mantiene durante más tiempo en los contextos en los que el reemplazo de *nós*, *vós* por *nosotros*, *vosotros* es más lento, como en los textos documentales —con especial resistencia en la primera persona— y en las zonas geográficas más occidentales.

Una vez que *nosotros* y *vosotros* empiezan a difundirse y se gramaticalizan, antes incluso de que se conviertan en formas mayoritarias, el resto de modificadores se descartan y quedan relegados a otras funciones enfáticas ocasionales. *Mismos* se usa únicamente con valor reflexivo y acaba extendiéndose a todos los pronombres del paradigma. Por ese motivo, se mantiene también tras *nosotros* y *vosotros*. *Todos* desaparece como refuerzo pronominal pospuesto y en la actualidad lo encontramos antepuesto. *Ambos*, *solos* y los numerales se mantienen con un empleo marginal en los textos castellanos.

Las variantes lingüísticas que más claramente compiten con *nosotros* y *vosotros* son *nós todos* y *vós todos*. *Todos* es el único refuerzo además de *otros* que tiene una distribución geográfica definida: en los textos documentales y literarios castellanos de León y Castilla occidental, así como en los textos documentales en gallego, *nós / vós todos* es una de las principales variantes reforzadas durante los siglos XIII y XIV. Así, para el refuerzo de los pronombres *nós* y *vós* surgen dos candidatos en zonas distintas de la península ibérica: *todos* en occidente y *otros* en oriente. La variación se resuelve en las fases iniciales del cambio a favor de *nosotros* y *vosotros*, mientras que la frecuencia de *nós todos* y *vós todos* desciende con un patrón con forma de curva en S.

Capítulo 5: El retroceso de *connusco* y *convusco*

1. Presentación¹

Otro de los cambios sufridos por el paradigma pronominal del español a lo largo de la historia es la pérdida de los pronombres personales de primera y segunda persona *connusco* y *convusco*, formas propias del español medieval. Este es un fenómeno estrechamente relacionado con la aparición de *nosotros* y *vosotros*, analizada en el capítulo correspondiente. En este caso, no solo trataremos con una posición sintáctica mucho más concreta, sino con un proceso más complejo que involucra un mayor número de variantes.

Igual que los pronombres *nós* y *vós*, los pronombres sintéticos *connusco* y *convusco* solían tener un referente plural, aunque ambos podían utilizarse también con referentes singulares como formas de tratamiento. Posteriormente, las formas sintéticas fueron reemplazadas por las analíticas *con nós* y *con vós*, y, finalmente, por las formas analíticas reforzadas actuales *con nosotros* y *con vosotros*, que solo pueden tener un referente plural (García 1997:74).

Las variantes *connusco* y *convusco* aparecen mencionadas frecuentemente en los manuales cuando se trata el paradigma pronominal del castellano medieval (Cano 2005:615; Eberenz 2000:72; Lloyd 1993:444; Penny 2006:161; Torrens Álvarez 2007:90). Sin embargo, no han recibido tanta atención como el cambio general de *nós*, *vós* a *nosotros*, *vosotros*. Uno de los motivos es que, simplemente, no son variantes centrales en el paradigma de los pronombres personales tónicos. Además, hay que tener en cuenta que la naturaleza del fenómeno tiene algunas implicaciones metodológicas que dificultan su estudio. En primer lugar, la aparición de estas variantes está limitada al uso del comitativo, por lo que el número de ejemplos extraíbles de cualquier texto se reduce considerablemente. Esto tiene como principal consecuencia la necesidad de trabajar con un corpus más amplio. Asimismo, el número de variantes es mayor —*connusco*, *convusco*, *con nós*, *con vós*, *con nosotros*, *con vosotros*— y las cuatro primeras pueden tener referentes tanto singulares como plurales.

Generalmente, la bibliografía que menciona la pérdida de *connusco* y *convusco* se limita a describir las formas, a discutir su origen y a dar la fecha de su pérdida. No es hasta finales de los años 80 cuando Líbano Zumalacárregui (1988) traza el avance cronológico de las formas innovadoras. Solo García (1997:75) trata el orden de desaparición de los pronombres, una parte esencial del establecimiento de la cronología. Otros trabajos que han abordado esta cuestión se han centrado en las causas de su sustitución (García 1991; Rini 1990) y la cronología del cambio (Eberenz 2000:72-73), mientras que se ha dejado de lado el estudio de los patrones de difusión geográfica de las formas.

En este capítulo pretendo hacer una contribución al estudio de *connusco* y *convusco* analizando su evolución cronológica en textos literarios y documentales; así como su distribución geográfica. Para ello, antes de nada, repasaré los estudios previos (§ 2) y presentaré los datos y la metodología de esta sección (§ 3). El análisis (§ 4) se dividirá en cinco partes: la localización de las primeras apariciones de las variantes (§ 4.1), el

¹ Este capítulo es una versión corregida y ampliada de Gomila Albal (2018). He perfilado el corpus de estudio y he codificado los datos más exhaustivamente. Destaca la incorporación de más colecciones documentales (§ 3. Datos y metodología específica). A diferencia de lo que se presenta en Gomila Albal (2018), en este caso he descartado del análisis general los pronombres con referente singular. Asimismo, el cartografiado de los datos ya no es por provincia, sino por población; y se incorpora el análisis de las variables de persona y de verbo.

establecimiento de la cronología (§ 4.2), el orden de aparición (§ 4.3), el contexto sintáctico (§ 4.4) y la distribución geográfica (§ 4.5).

Mi hipótesis de partida es que en muchas de estas cuestiones va a haber una correspondencia clara con lo que ya se ha visto en el Capítulo 3 sobre el reemplazo de *nós* y *vós* por *nosotros* y *vosotros*. El periodo cronológico (siglos XIII-XVI) y las tipologías textuales son los mismos, y, respecto a la variación geográfica, volvemos a tener indicios de una distribución oriente-occidente: ya hemos establecido el origen oriental de (*con*) *nosotros*, (*con*) *vosotros*; y *connosco*, *convosco* se conservan todavía hoy en gallego y en portugués (Álvarez Blanco 2004). Sin embargo, está por ver hasta qué punto la pérdida de *connusco*, *convusco* a favor de *con nós* y *con vós*, y la posterior aparición de *con nosotros*, *con vosotros* forman parte de un mismo cambio lingüístico en varias fases o si son en realidad dos cambios diferentes que afectan en poco tiempo a la misma parcela del paradigma de los pronombres personales.

2. Estado de la cuestión

2.1. Origen de las formas

Los pronombres *connusco* y *convusco* tienen su origen en las formas latinas NOBISCUM y VOBISCUM, formadas a partir del ablativo junto a la adposición CUM enclítica. En latín tardío los pronombres NOBIS y VOBIS son reemplazados por los acusativos NOS y VOS, lo que da lugar a las variantes NOSCUM y VOSCUM, que ya aparecen censuradas en el *Appendix Probi* (Alvar y Pottier 1983:124; De Jonge y Nieuwenhuijsen 2009:1615; Lausberg 1973:164; Líbano Zumalacárregui 1988:501; Menéndez Pidal 1982:251; Rini 1990:59).

El cambio de ablativo a acusativo suele explicarse por la pérdida de los casos y la generalización del acusativo como caso oblicuo (Lloyd 1987:278). No obstante, Rini (1990:59) plantea otras explicaciones, ya que precisamente hoy en día encontramos otros pronombres que sí han conservado el caso, como ocurre en los dativos *mí*, *ti*, *sí*. Este autor propone que los pronombres NOBISCUM y VOBISCUM fueron reemplazados por *noscum* y *voscum* por analogía con las formas en singular MECUM, TECUM y SECUM; y que este reemplazo tuvo lugar antes de la pérdida de los casos. Los pronombres singulares ME, TE y SE eran iguales para el acusativo y para el ablativo. Por tanto, los hablantes pudieron interpretarlos como acusativos y, por extensión, utilizaron los acusativos para el plural (De Jonge y Nieuwenhuijsen 2009:1615; Rini 1990:59, 1992:140). Otro factor que pudo influir en este cambio es el número de sílabas. No son extraños los cambios hechos en latín para igualar el número de sílabas dentro de un paradigma y, en este caso, tanto las formas de singular (*mecum*, *tecum*, *secum*) como las de plural (*noscum*, *voscum*) pasarían a tener dos sílabas (Rini 1990:59-60).

Las formas *nosco* y *vosco* se difunden tempranamente en romance y se documentan en textos medievales leoneses, portugueses, italianos y sardos. Según Rini (1990:60), las primeras apariciones de estos pronombres en documentos castellanos son con la forma de las variantes *connusco* y *convusco*. Se deduce que en un estadio intermedio de estas variantes se utilizaron **connosco* y **convosco*. De este modo, en primer lugar se incorporó la preposición *con*, que se añadió de nuevo a pesar de conservarse como posposición; y, en segundo lugar, se produjo una disimilación vocálica en *connosco* y *convosco* por la influencia del cierre de las vocales en las formas singulares *-migo* y *-tigo* (Lloyd 1993:444; Menéndez Pidal 1982:251).

Rini (1990:60-61, 1992:36-38) repasa distintas hipótesis de la bibliografía previa que buscan justificar la adición de *con-* tanto en las formas singulares como en las plurales. Una de las primeras propuestas es la búsqueda de claridad, aunque el mismo autor considera que es una justificación demasiado amplia que no acaba de dar respuesta

al fenómeno. Otros autores explican que pudo existir un deseo de eliminar las estructuras originales con la posposición –CUM. Sin embargo, en ese caso, lo esperable hubiese sido mover el elemento de sitio, lo que en singular hubiese dado lugar a **con mí*, **con ti*, **con sí*. Otra de las propuestas es que, tras los cambios fonéticos sufridos por los pronombres, ya no se distinguía el elemento *con*, sino que se veía como una forma indivisible. Sin embargo, Rini critica este argumento, ya que, según él, esto explicaría por qué se conservó el –CUM, pero no por qué se tuvo que añadir de nuevo.

Estas propuestas se unen en una sola, que actualmente es la más apoyada: la analogía. Para entender este cambio hay que tener en cuenta el proceso sufrido por las formas singulares MECUM, TECUM, SECUM. Estos pronombres sonorizaron la /k/ intervocálica y pasaron de una terminación en –UM a una en –o. Así, la posposición –CUM pasó a ser –go, elemento irreconocible por los hablantes, por lo que fue necesario añadir la preposición CUM. Las formas del plural siguieron un proceso muy parecido, aunque sin sonorizar la /k/, ya que esta no se encontraba entre vocales en *nusco* y *vusco*. En todo caso, también se añadió la preposición *con*– al plural por analogía con las formas en singular (De Jonge y Nieuwenhuijsen 2009:1615; Lloyd 1987:278; Rini 1990:61-62).

En resumen, NOBISCUM y VOBISCUM pasan a ser *noscum* y *voscum* (ablativo > acusativo), *nosco* y *vosco* (-um > -o), *connosco* y *convosco* (adición de *con*-) y, finalmente, *connusco* y *convusco* (disimilación sugerida por la vocal tónica cerrada de *connigo* y *contigo*).

El origen de las formas analíticas simples *con nós* y *con vós* y las formas analíticas compuestas *con nosotros* y *con vosotros* tuvo algo de controversia en los primeros estudios. Al principio, García (1989) defiende que las formas *con nós* y *con vós* eran formas apocopadas de *connusco* y *convusco* que aparecieron por la necesidad de añadir el modificador *otros*. Incluso, afirma que *con nosotros* y *con vosotros* surgieron con anterioridad a *con nós* y *con vós*:

Estos contrincantes no surgen, empero, al mismo tiempo: *con nos otros* / *vos otros* aparecen en textos tempranos donde aún no se observa *con nos* / *vos*, lo que sugiere que el primer estímulo para la omisión de –co probablemente haya sido la adición de modificadores léxicos al pronombre (García 1991:11).

Es decir, según esta hipótesis, la apócope y la adición del modificador fueron simultáneas. De este modo, en un primer momento, la variación se produjo entre las formas sintéticas *connusco*, *convusco* y las formas analíticas *con nós*, *con vós* únicamente con modificadores como *otros*, *mismos* o *todos*. La autora justifica esto con el portugués, que elimina la posposición –co cuando el pronombre va acompañado por un modificador, ya sea antes o después: *com nós todos*, *com ambos nós*. Así, «lo que determinaba el uso de una alternativa u otra era la necesidad (o no) de recurrir al modificador léxico, cuyo valor semántico independiente es indiscutible» (García 1991:13).

No obstante, Rini (1992:57 y ss.) niega esta hipótesis y afirma que *connusco* y *convusco* fueron sustituidos primero por *con nós* y *con vós* y, posteriormente, se añadió el modificador para dar lugar a *con nosotros* y *con vosotros*. Por tanto, la pérdida de *connusco* y *convusco* no pudo estar causada directamente por la adición de –*otros*. Tiempo después, García (1997) acabó aceptando esta última hipótesis.

2.2. Cronología

Establecer la cronología del avance y la desaparición de *connusco*, *convusco*, *con nós*, *con vós*, *con nosotros* y *con vosotros* no es tarea fácil, ya que no solo debemos tener en cuenta que se entrelazan tres pares de variantes: sintéticas, analíticas simples y analíticas compuestas de primera y segunda persona del plural; sino que las formas en –*usco* y las

formas analíticas simples también pueden tener un referente singular. Así, *connusco* y *con nós* pueden tener valor de plural mayestático, mientras que *convusco* y *con vós* pueden utilizarse como formas de tratamiento de cortesía. Las únicas que no muestran ambigüedad en este sentido son las formas analíticas compuestas *con nosotros* y *con vosotros* (De Jonge y Nieuwenhuijsen 2009:1616; Líbano Zumalacárregui 1988:508).

Tras los cambios que hemos ido señalando en la sección anterior, *connusco* y *convusco* aparecen a finales del siglo XII o principios del XIII (De Jonge y Nieuwenhuijsen 2009:1616; Líbano Zumalacárregui 1988:502; Rini 1990:60). Durante el siglo XIII predominan las formas sintéticas, aunque es en este siglo en el que podemos encontrar las primeras apariciones de las formas analíticas simples *con nós* y *con vós*. A lo largo del siglo XIV *con nós* y *con vós* empiezan a extenderse ocupando progresivamente los contextos de uso de *connusco* y *convusco*. De este modo, las formas recesivas y las innovadoras compiten de forma más o menos igualada durante el último cuarto de este siglo y hasta principios del XV (García 1997:77; Rini 1990:63).

Ya en el siglo XV el uso de las formas en *-usco* desciende radicalmente y solo aparecen de forma esporádica. Aun así, el uso de las formas *connusco* y *convusco* con referente singular se prolonga durante más tiempo (García 1997:74). Según Eberenz (2000:72), «su extinción en la lengua general debía estar ya sellada hacia 1400 y probablemente mucho antes en la comunicación oral». Así, generalmente, las apariciones de *connusco* y *convusco* a finales del siglo XV pueden considerarse arcaizantes (García 1997:76). A su vez, se considera que *con vosotros* ya se utiliza de forma regular a finales del siglo XV (Líbano Zumalacárregui 1991:506).

En cuanto al orden de desaparición de la primera y la segunda persona, según García (1997:75), la primera en desaparecer fue *connusco*, mientras que *convusco* se mantuvo durante más tiempo. Este hecho es llamativo teniendo en cuenta que el pronombre tónico *vosotros* apareció antes que *nosotros* (cf. Capítulo 3 § 4.3). Esto nos podría hacer pensar que quizá ambos fenómenos no están tan relacionados como podríamos suponer en un primer momento. De todos modos, García (1997) justifica el orden de desaparición por los matices semánticos de cada una de las formas: mientras que *convusco*, la forma que se utilizó durante más tiempo, tenía un valor de cortesía claro respecto al interlocutor, *connusco* podía resultar arrogante:

While *convusco* may flatter the hearer, with *connusco* the speaker runs the risk of appearing arrogant. Real politeness would be lavish with *convusco* (particularly in singular reference!), but very sparing with *connusco* (García 1997:75).

No parece haber acuerdo con la fecha en la que empiezan a aparecer las formas analíticas compuestas *con nosotros* y *con vosotros*. Según Eberenz (2000:73), podemos encontrar los primeros casos en el siglo XV, especialmente en el marco de la narrativa historiográfica y ficcional. Cabe destacar que esta fecha es considerablemente posterior a la que se propone para la aparición de *nosotros* y *vosotros* en otras posiciones sintácticas. Según Líbano Zumalacárregui (1988:504), el primer ejemplo de *con nosotros* aparece en la última década del siglo XV. En cambio, según los datos de De Jonge y Nieuwenhuijsen (2009:1616) las variantes *con nosotros* y *con vosotros* ya estarían completamente establecidas en este siglo.

En cualquier caso, en el siglo XVI *connusco* y *convusco* ya se consideran arcaísmos y sobreviven como tales hasta el siglo XVII (García 1991:11). Así, *con nosotros* y *con vosotros* ya son las formas más frecuentes, aunque todavía se conservan las formas analíticas simples con referente singular (Líbano Zumalacárregui 1988:506).

En resumen, las formas analíticas simples *con nós* y *con vós* empiezan a aparecer en el siglo XIII y en el XV reemplazan a las sintéticas *connusco* y *convusco*, que

sobreviven en contextos arcaizantes. Al mismo tiempo, durante el siglo XV se empiezan a difundir las variantes analíticas compuestas *con nosotros* y *con vosotros* que hemos descrito en el capítulo correspondiente. Por otro lado, el orden de aparición y difusión de las variantes innovadoras en el reemplazo de *connusco*, *convusco* no parece coincidir con el que hemos presentado hasta ahora en los capítulos anteriores. Todos los datos extraídos de nuestro corpus apuntan que la segunda persona *vosotros* fue la que lideró el cambio. Sin embargo, según García (1997:75), en este caso el cambio se inicia en la primera persona.

2.3. Causas del cambio

Generalmente, para explicar la pérdida de las variantes *connusco* y *convusco* se recurre a su mala integración en el paradigma pronominal. Tanto Rini (1990:63) como García (1991:10-11, 1997:59 y ss.) argumentan que las variantes *connusco* y *convusco*, tras los cambios fonéticos y analógicos sufridos desde las formas latinas NOBISCUM y VOBISCUM, no se integran bien en el paradigma pronominal y resultan extrañas para los hablantes. En palabras de Rini (1990:63), «paradoxically, excessive analogical change may have eventually served as an impetus for the abandonment of *connusco*, *convusco* for *con nos*, *con vos*». Así, *-CUM* ha pasado a ser *-go* en singular y *-co* en plural. Además, los morfemas propiamente pronominales *-nos-* y *-vos-* se han quedado en el interior de la palabra, entre una preposición y una posposición, y han pasado a ser *-nus-* y *-vus-*, formas que no se encuentran en el resto del paradigma pronominal. Es cierto que *-me-*, *-te-* y *-se-* también cerraron sus vocales, pero *-mi-*, *-ti-* y *-si-* se pueden identificar con otros pronombres del paradigma, por lo que siguen siendo reconocibles (1990:63). El resultado es una forma morfológicamente poco transparente, que formalmente no parece un plural y que se diferencia del resto de pronombres del paradigma.

García (1991), además, justifica semánticamente la aparición de las variantes analíticas. Según esta autora, las primeras apariciones de *con vós* tendrían lugar en contextos con un matiz de contraste en los que interesaba diferenciar entre la segunda persona y el resto, mientras que *convusco* aparecería en contextos de unión solidaria. De este modo, encontraríamos ejemplos como *combatir convusco*, en el que emisor y receptor son aliados; frente a *combatir con vós*, en el sentido de ‘contra vosotros’ (García 1991:16). En el momento en el que coexisten las formas sintéticas con las formas analíticas simples y complejas, es difícil diferenciar los valores de cada una de las variantes. En este sentido, García (1991:14-15) defiende la existencia de un *continuum* que va desde *convusco* hasta *con vosotros*, pasando por *con vós*. De este modo, en un extremo encontraríamos la negación del contraste (*connusco*, *convusco*), en el centro la neutralidad (*con nós* y *con vós*) y, en el otro extremo, el máximo de contraste (*con nosotros* y *con vosotros*). Esta complejidad podría haber favorecido la desaparición de *connusco* y *convusco*.

Eberenz (2005:615) cuestiona las diferencias semánticas entre *convusco* y *con vós* y afirma que, en cualquier caso, este contraste ya se había perdido en el siglo XV:

Para explicar la diferencia entre *convusco* y el nuevo sintagma *con vós* se ha atribuido a *convusco* un valor de solidaridad entre hablante(s) e interlocutor(es), mientras que *con vós* indicaría, por el contrario, un contraste entre los dos grupos. Sin embargo, en el siglo XV tal distinción, si existió alguna vez, debía estar ya bastante difuminada (Eberenz 2005:615).

La última explicación que se ha propuesto para la pérdida de las variantes en *-usco* es la de Sánchez Vicente (2020), que destaca que por encima de la falta de transparencia de las formas está la unidad del paradigma. Compara diferentes lenguas y dialectos romances,

y llega a la conclusión de que la adición de la preposición tiene lugar en los casos en los que los cambios fonéticos rompen la unidad del paradigma. Para esto cita las formas del pamesano *mek*, *tek* y *nosk*, en las que la posposición ya no es transparente, pero no vuelve a añadirse porque mantiene la coherencia dentro del paradigma.

Lo que es importante en estos casos [...] es que el paradigma funcione de manera unitaria o no, puesto que el hecho de que no lo haga es lo que produce la desestabilización del sistema y la aparición de formas supletivas y diferentes al resto de las de su paradigma, como ocurre en el español estándar actual (Sánchez Vicente 2020:245).

Por tanto, las posibles explicaciones del reemplazo de *connusco* y *convusco* por las variantes analíticas se resumen en tres: la necesidad de expresar matices semánticos de contraste y solidaridad, la falta de transparencia de las formas en *-usco* debido a los reiterados cambios fonéticos y analógicos que han sufrido y, finalmente, la presión para mantener la unidad del paradigma.

2.4. Presencia de *connusco* y *convusco* en otras variedades romances

Actualmente, en gallego estándar coexisten las formas *connosco* y *convosco*, *con nós* y *con vós*, y *con nosoutros* y *con vosoutros*. Todas son correctas, aunque en su gramática se recomienda el uso de las primeras por ser las tradicionales (Álvarez Blanco, Regueira Fernández, y Monteagudo 1989). En todo caso, se señala que *connosco* y *convosco* son incorrectas en los casos en que el pronombre lleve algún elemento en aposición: «polo que hay que dicir: *non falou connosco* mais *non falou con nó-las dúas*, *viñeran convosco*, mais *viñeran con vós mesmos*» (1989:162). Sin embargo, en el uso real de la lengua, *connosco* y *convosco* han dejado de usarse en gran parte de Galicia (Álvarez Blanco 2004).

El portugués conservó también durante mucho tiempo las formas *nosco* y *vosco*, que fueron reemplazadas durante el siglo XIV por las actuales *connosco* y *convosco*. Asimismo, en la actualidad, tal y como ya hemos mencionado más arriba, se usan las variantes analíticas simples cuando el pronombre va acompañado por un modificador, ya sea antes o después: *com nos todos*, *com ambos nos*. Además, en el habla familiar están siendo reemplazadas por expresiones como *com a gente*, *con vocês*, *com os senhores*, *com as meninas*, etc. (Vázquez Cuesta y Mendes Da Luz 1987:41-42).

El hecho de que en portugués y en gallego se hayan conservado durante más tiempo las variantes *connosco* y *convosco*, en un estadio previo al cierre de la vocal, refuerzan la idea de que *connusco* y *convusco* pudieron ser reemplazadas en castellano por el hecho de que *-nus-* y *-vus-* no tuvieran un paralelo en el resto del paradigma pronominal. Así, la variante con *-nos-* y *-vos-* que se conservó en gallego sí era reconocible para los hablantes y quizá haya favorecido su supervivencia durante más tiempo.

2.5. Recapitulación: preguntas de investigación

De las cuestiones que hemos repasado en este estado de la cuestión, podemos extraer las preguntas de investigación que trataré de responder durante mi análisis. Sobre la cronología en distintos tipos de texto, de igual modo que en el reemplazo de *nós*, *vós* por *nosotros*, *vosotros* es más rápido en los textos literarios que en los documentales, ¿ocurre igual en el reemplazo de *connusco* y *convusco*? ¿Cuándo se completa la pérdida de estos pronombres en cada tipo de texto? Respecto al orden de aparición, ¿el cambio avanza antes en la segunda persona igual que con la aparición de *vosotros* o en la primera tal y como afirma García (1997)? En cuanto a las causas del cambio, ¿qué lleva a reemplazar las formas sintéticas en *-usco* por las analíticas? Sobre los valores semánticos que pudieron haber motivado la aparición de las formas innovadoras, ¿en qué casos podemos

detectar valor de solidaridad en *connusco*, *convusco* y valor de contraste en las variantes analíticas tipo *con nós*, *con nosotros*? ¿Podemos encontrar algún tipo de distribución de las variantes según el significado del verbo? ¿Qué papel tiene la presión del resto del paradigma en la pérdida de *connusco* y *convusco*? En cuanto a la variación diatópica, ¿cuáles de las variantes que intervienen en este proceso de cambio están adscritas a una zona geográfica determinada? Sabemos que *nosotros* y *vosotros* se difunden desde el oriente peninsular y, a partir de la conservación de las formas en *-usco* en las variedades romances más occidentales podemos deducir que estas variantes fueron retrocediendo hacia el oeste. Sin embargo, ¿qué ocurre con las formas *con nós* y *con vós*?

Todas estas preguntas serán imprescindibles para responder la más general: ¿Hasta qué punto la pérdida de *connusco*, *convusco* a favor de *con nós* y *con vós*, y la posterior aparición de *con nosotros*, *con vosotros* forman parte de un mismo cambio lingüístico en varias fases o son en realidad dos cambios diferentes que afectan en poco tiempo a la misma parcela del paradigma de los pronombres personales?

3. Datos y metodología específica

Debido a la baja frecuencia de aparición de las variantes que intervienen en este cambio lingüístico, ha sido necesario utilizar un corpus específico, más amplio. En el Capítulo 3, el exceso de resultados en el corpus de textos literarios y en el *Corpus Informatizat del Català Antic* (CICA) era un problema metodológico que se resolvía mediante selecciones aleatorias. En cambio, en esta ocasión, el problema es el opuesto: la falta de ejemplos en el corpus principal de textos documentales, el *Corpus de documentos anteriores a 1800* (CODEA+ 2015), ha obligado a recurrir a otros corpus y a colecciones documentales digitales o en papel (para una descripción del corpus cf. Capítulo 2, § 5.1 y 5.3).

Los datos del CODEA+ se han completado con el *Corpus Hispánico y Americano en la Red: Textos Antiguos* (CHARTA), con documentos notariales del *Corpus diacrónico del español* (CORDE)² y con las siguientes colecciones documentales:

- a) *Colección diplomática de la Almunia de doña Godina (1176-1395)* (Canellas López 1962)
- b) *Colección diplomática de Santa María la Mayor de Calatayud* (Lafoz Rabaza 2000)
- c) *Colección diplomática del concejo de Zaragoza (1276-1285)* (Canellas López 1975)
- d) *Colección documental del monasterio de San Esteban de Nogales (1149-1498)* (Cavero Domínguez 2001)
- e) *Colección documental del monasterio de Trianos (1111-1520)* (Fuente Crespo 2000)
- f) *Documentación del monasterio de la Trinidad de Burgos (1198-1400)* (García Aragón 1985)
- g) *Documentación medieval del archivo municipal de Pamplona I (1129-1356)* (Cierbide y Ramos 1998)
- h) *Documentación medieval del archivo municipal de Pamplona II (1357-1512)* (Cierbide y Ramos 2000)
- i) *Documentos lingüísticos del Alto Aragón* (Navarro Tomás 1957)
- j) *El primer cartulario de los reyes de Navarra* (Ramírez Vaquero, Herreros Lopetegui, y Beroiz Lazcano 2013)
- k) *Platería y joyería en Zaragoza (1413-1513)* (Gómez de Valenzuela 2017)
- l) *Selección de documentos de la villa aragonesa de Sos (1202-1533)* (Abella Samitier 2009)

² Del CORDE se han utilizado principalmente los textos de la tipología identificada en el corpus como *10. Documentos notariales*. También se han añadido algunas cartas de compraventa de la tipología *16. Pintura*. En ambos casos, únicamente se han conservado los ejemplos que tuvieran algún tipo de localización geográfica en el índice de obras del CORDE.

Para las formas analíticas simples y reforzadas *con nós*, *con vós*, *con nosotros* y *con vosotros* se ha realizado el mismo proceso de extracción que en el capítulo sobre las formas correspondientes (cf. Capítulo 3, § 3). Para las sintéticas, se han extraído todas las palabras terminadas en *-usco* y *-osco* para delimitar el número de variantes gráficas. Tras las búsquedas de todas las variantes se obtuvieron un total de 2.698 ejemplos.

Como bien se sabe, una muestra de textos del CODEA+ se incluye en CHARTA. Asimismo, se han detectado en el CORDE algunos documentos del Archivo Nacional que ya estaban incluidos en el CODEA+. En ambos casos se han descartado los duplicados y se han conservado los ejemplos del CODEA+. Esto es especialmente relevante para los duplicados del CORDE, corpus que ofrece menos información sobre el origen geográfico³ de los documentos y que tiene un acceso mucho más limitado al texto.

Después de descartar los ejemplos repetidos y los falsos positivos como *busco* y *cognosco*, el corpus ha quedado reducido a 1.033 ejemplos entre los siglos XIII y XVI. De estos, solo 531 tenían referente plural o colectivo. En la Tabla 14 pueden verse los datos extraídos para cada una de las variantes según fuente.

<i>Fuente</i>	<i>Connusco, convusco</i>	<i>Con nós, con vós</i>	<i>Con nosotros, con vosotros</i>	<i>Total</i>
CODEA+	45	34	10	89
CHARTA	94	2	2	98
CORDE	113	116	50	279
<i>Colecciones documentales</i>	25	31	9	65
Total	277	183	71	531

Tabla 14. Resumen de ejemplos con referente plural extraídos del corpus ampliado de textos documentales según fuente.

El corpus de textos literarios utilizado para el análisis de este fenómeno es el descrito en el Capítulo 2 (§ 5.2). En esta ocasión, no se ha hecho ninguna selección aleatoria, sino que se ha usado el corpus al completo. Mediante las búsquedas se han obtenido 1.254 ejemplos. Después de descartar los falsos positivos y los pronombres con referente singular, han quedado 852 ocurrencias. En la Tabla 15 pueden verse los datos recogidos de cada una de las obras del corpus de textos literarios.

<i>Siglo</i>	<i>Clave</i>	<i>Connusco, convusco</i>		<i>Con nós, con vós</i>		<i>Con nosotros, con vosotros</i>		<i>Total</i>
		%	N	%	N	%	N	
<i>XIII</i>	<i>E6</i>	91.3	105	8.7	10	–	–	115
	<i>GCU</i>	85.7	6	14.3	1	–	–	7
	<i>GE1</i>	100	70	–	–	–	–	70
	<i>GE4</i>	100	50	–	–	–	–	50
<i>XIV</i>	<i>Alexandre</i>	100	7	–	–	–	–	7
	<i>CQ1</i>	–	–	66.7	10	33.3	5	15
	<i>CQ2</i>	–	–	63.2	12	36.8	7	19
	<i>CTroyana</i>	100	49	–	–	–	–	49
	<i>GC1</i>	–	–	69.6	16	30.4	7	23
	<i>HTroyana</i>	42.9	3	42.9	3	14.3	1	7
	<i>Mostrador</i>	100	33	–	–	–	–	33

³ Al utilizar datos de colecciones documentales y del CORDE, el número de ejemplos con información geográfica solo de provincia es mayor. Igual que se ha hecho con el resto de fenómenos analizados, estos casos se han situado en la capital de provincia para el cartografiado.

	<i>Onceno</i>	100	7	–	–	–	–	7
	<i>Otas</i>	100	6	–	–	–	–	6
	<i>Sumas</i>	–	–	–	–	100	6	6
XV	<i>ABC</i>	60	3	40	2	–	–	5
	<i>Arragel</i>	3.3	6	37.4	68	59.3	108	182
	<i>Atalaya</i>	33.3	1	–	–	66.7	2	3
	<i>Bienandanzas</i>	54.2	13	16.7	4	29.2	7	24
	<i>Corbacho</i>	50	1	–	–	50	1	2
	<i>CroAra</i>	–	–	–	–	100	2	2
	<i>Cronica1344</i>	100	7	–	–	–	–	7
	<i>Eneida</i>	50	1	–	–	50	1	2
	<i>LBA</i>	100	1	–	–	–	–	1
	<i>Propiedades</i>	–	–	14.3	1	85.7	6	7
	<i>Sermonario</i>	100	18	–	–	–	–	18
	<i>Valera</i>	–	–	33.3	1	66.7	2	3
	<i>Varones</i>	–	–	–	–	100	1	1
	<i>Viaje</i>	–	–	20	1	80	4	5
	<i>Victorial</i>	–	–	28.6	2	71.4	5	7
	<i>Vida</i>	–	–	14.3	1	85.7	6	7
	<i>Virtuosas</i>	–	–	66.7	2	33.3	1	3
	<i>Visión</i>	–	–	–	–	100	1	1
XVI	<i>Belianís</i>	–	–	11.1	3	88.9	24	27
	<i>CNE</i>	–	–	25.6	11	74.4	32	43
	<i>CSatíricos</i>	–	–	–	–	100	5	5
	<i>Espejo</i>	–	–	11.8	2	88.2	15	17
	<i>HEtiópica</i>	–	–	–	–	100	24	24
	<i>Jardín</i>	–	–	–	–	100	5	5
	<i>Lisuarte</i>	–	–	33.3	2	66.7	4	6
	<i>Orlando</i>	–	–	25	1	75	3	4
	<i>Tratados</i>	–	–	–	–	100	27	27
	TOTAL	45.4	387	18	153	36.6	312	852

Tabla 15. Resumen de porcentajes y ocurrencias de cada una de las variantes en los testimonios del corpus de textos literarios.

4. Análisis

4.1. Primeras ocurrencias

Las primeras apariciones de las formas analíticas simples con referente plural en el corpus de textos documentales tienen lugar en la primera mitad del siglo XIII. Estos primeros ejemplos aparecen en una cesión de bienes de Silos de 1237 (*Cartulario de Silos*, CORDE). Ambos se corresponden con la primera persona *con nós*:

(1)

- a) Et a tal pleito ponedes vos don Rodrigo **con nos** [el abbat e el convento de Sancto Domingo], sub iura sobre los sanctos Evangelios, que vos non fagades (1237, Burgos, *Cartulario de Silos*, CORDE)
- b) Et otrosi ponedes vos **con nos** [el abbat e el convento de Sancto Domingo sobredichos] que, con la merced de Dios, que fagades vuestra vida ordenada e honesta (1237, Burgos, *Cartulario de Silos*, CORDE)

Apenas hay distancia entre la primera y la segunda persona, que se documenta en las siguientes décadas. La mayoría de estas primeras ocurrencias no tienen un referente claramente plural, sino colectivo:

- (2)
- a) por esta abenencia que **con vós** conuiento de Santa María de Nágera ovi (1249, sin localización, CODEA-0644)
 - b) fazemos tal pleito **con vós** el concejo de Salinas de Añana (1262, Burgos, CORHEN-0088)
 - c) non ayan razón de se escusar de me lo dar **con vós** [el concejo de la noble cibdat de Castiella] (1279, Sevilla, CODCAR-0440)

En cuando a los textos literarios, en el siglo XIII encontramos un ejemplo de *con nós* (3a) y 10 de *con vós* (3b-j), aunque todos los ejemplos de segunda persona son de la misma obra y se encuentran en la misma fórmula. El uso de la forma analítica simple en *E6* está claramente favorecido por la inserción del determinante antes del pronombre.

- (3)
- a) en saluo de la çipdad **con nos** nuestros cuerpos. & que nos fuessemos pora tierra de xristianos (*GCU*)
 - b) Ro:15:33 Dios de paz sea **con todos uos** amen. (*E6*)
 - c) Ro:16:24 La gracia del nuestro sennor ihesu christo sea **con todos uos** amen. (*E6*)
 - d) Cor1:16:24 e la mi amor **con todos uos** en ihesu christo. Amen (*E6*)
 - e) Cor2:13:13 La gracia de nuestro sennor ihesu christo y ell amor de dios e la comunidat del santo espiritu sea **con todos uos** amen. esto es. el santo spiritu que es ayuntamiento del padre e del fijo sea **con todos uos**. (*E6*)
 - f) Te2:3:16 Aquel sennor dios de paz de auos paz sin fin en todo logar. Nuestro sennor sea **con todos uos**. (*E6*)
 - g) Te2:3:18 La gracia de nuestro sennor ihesu christo sea **con todos uos** Amen. (*E6*)
 - h) Tit:3:15 Saludan te todos los que son comigo. Saluda alos que nos aman en la fe. La gracia de dios sea **con todos uos**. Amen. (*E6*)
 - i) Heb:13:25 Gracia sea **con todos uos** A. M. E. N. (*E6*)
 - j) Ap:22:21 La gracia del nuestro sennor ihesu christo sea **con todos uos**. A.M.E.N. (*E6*)

En cuanto a las formas *con nosotros* y *con vosotros*, estas se registran un siglo después, en el XIV, tanto en los textos documentales (4) como en los literarios (5):

- (4)
- a) de no auer manera nj partiçipation nj uso alguno **con nos otros** (1342, Jaca, *Documentación medieval del archivo municipal de Pamplona I*)
 - b) los vjnt jurados de la dicha villa del ayño passado que **con vos otros** ensemble se pleguen (1346, Pamplona, *Documentación medieval del archivo municipal de Pamplona I*)
 - c) que nos pleguemos **con uos otros** es en vuestro poder (1357, Pamplona, *Documentación medieval del archivo municipal de Pamplona II*)
 - d) es de nuestra voluntad de nos no **con uos otros** en el dicho logar plegar nin fazer vintena **con nos otros** njn dar [ayzma] nin logar de vos con nos njn nos **con uos otros** de peleyar nin mouer escandalo con uos (1357, Pamplona, *Documentación medieval del archivo municipal de Pamplona II*)
- (5)
- a) yo he pensado de partir **con uos otros** el trasoro que yo he aplegado (*CQ1*)
 - b) los de dentro han hablado **con nos otros** et dizen que si vos los queredes tomar a mercet (*CQ2*)
 - c) partit uos de aqui & leuat **con uos otros** todo lo que auedes (*GCI*)
 - d) ellos seyan **con nos otros** et nos ayuden potentment (*HTroyana*)

- e) E que el vos bien quisiese & guardase lo que **con vos otros** posiese (*Sumas*)

Los ejemplos citados dejan entrever algunas cuestiones que pueden ser relevantes para el análisis que realizaremos en las siguientes secciones. A diferencia de lo que hemos visto en el análisis de *nós*, *vós*, *nosotros* y *vosotros* (Capítulo 3, § 4.1), apenas hay distancia entre las documentaciones de la primera y la segunda persona. Es decir, la distinción de persona no parece ser tan determinante en los primeros ejemplos. Además, en cuanto a la distribución geográfica de estas primeras ocurrencias, las variantes analíticas simples (1-3) no parecen estar concentradas en una misma región, sino que están notablemente dispersas. En cambio, las analíticas compuestas, de acuerdo con los datos ya analizados (Capítulo 3, § 4.5), se concentran en la mitad oriental.

4.2. Cronología

En la Figura 26 podemos ver el avance y el retroceso de las formas pronominales en el corpus documental entre los siglos XIII y XVI. En el siglo XIII, *connusco* y *convusco* son claramente predominantes, mientras que el avance de *con nós* y *con vós* se encuentra en su fase inicial. En el siglo XIV, el uso de las formas sintéticas en *-usco* desciende rápidamente y baja del 60 %. En este momento, a partir de los datos, vemos que estas variantes están empezando a ser reemplazadas por las analíticas simples tipo *con vós*, que han ascendido a un 36 %. Sin embargo, simultáneamente, empiezan a aparecer las analíticas compuestas tipo *con vosotros*, lo que modificará la trayectoria de las simples en los siglos siguientes. A continuación, el uso de las variantes sintéticas se desploma.

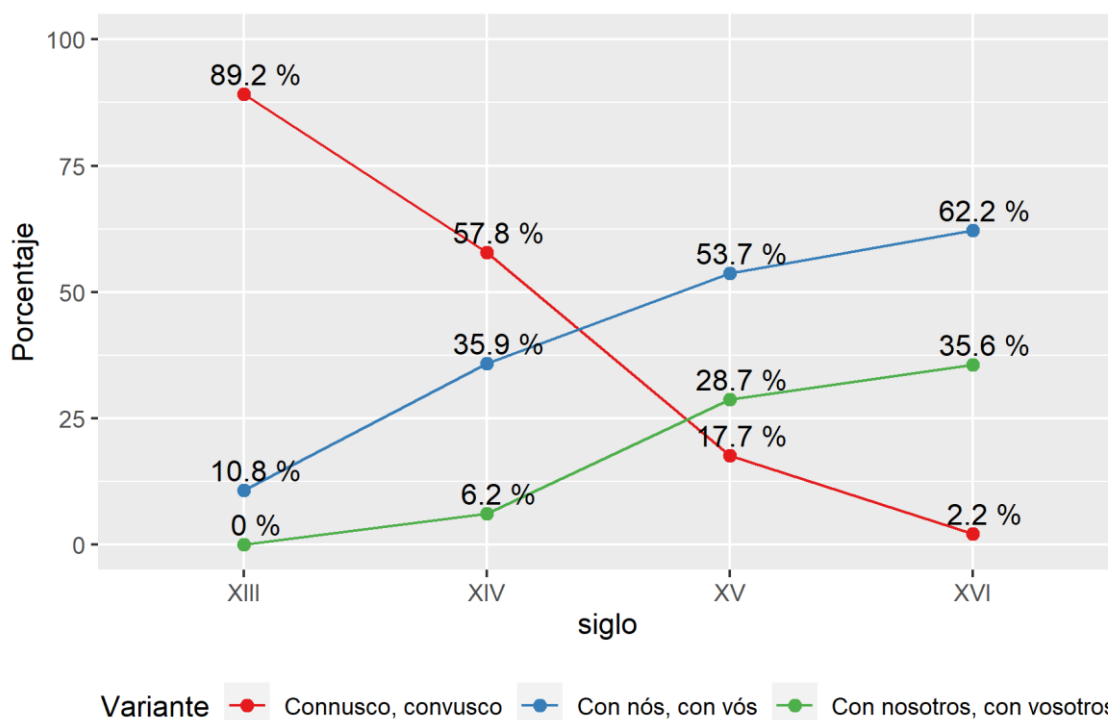


Figura 26. Evolución de las tres variantes con referente plural en textos documentales desde el siglo XIII al XVI.

A partir de mediados del XV, *connusco* y *convusco* prácticamente desaparecen de los textos documentales. El único ejemplo posterior que encontramos es de un documento cántabro fechado entre 1489 y 1522, que se corresponde con el 2 % de la Figura 25. Normalmente, suele indicarse que estas variantes pronominales siguen apareciendo

durante más tiempo con referente singular. No obstante, este no es el caso en los textos documentales, que muestran un descenso paralelo en la frecuencia de uso con referentes plurales y singulares.

Aunque en el siglo XVI *connusco* y *convusco* están claramente en desuso, en los textos documentales no se ha producido un reemplazo claro. *Con nós* y *con vós* son formas mayoritarias, pero su velocidad de difusión disminuye a medida que *con nosotros* y *con vosotros* van avanzando.

Habitualmente, en las innovaciones que presentan una difusión con forma de curva en S, encontramos que el avance de las variantes es más rápido en el punto intermedio, alrededor del 50 %. Sin embargo, el avance de *con nós* y *con vós* se ve frenado justo en esta fase, en los siglos XV y XVI. Esta retención anuncia el declive que sufrirán estas variantes en los siglos siguientes, cuando sean reemplazadas por *con nosotros* y *con vosotros*. Por tanto, lo que vemos en la evolución de las variantes analíticas simples tipo *con vós* es la mitad de una curva en campana. Al mismo tiempo, vemos el inicio de la difusión de las analíticas compuestas tipo *con vosotros*, que serán las que finalmente se extenderán.

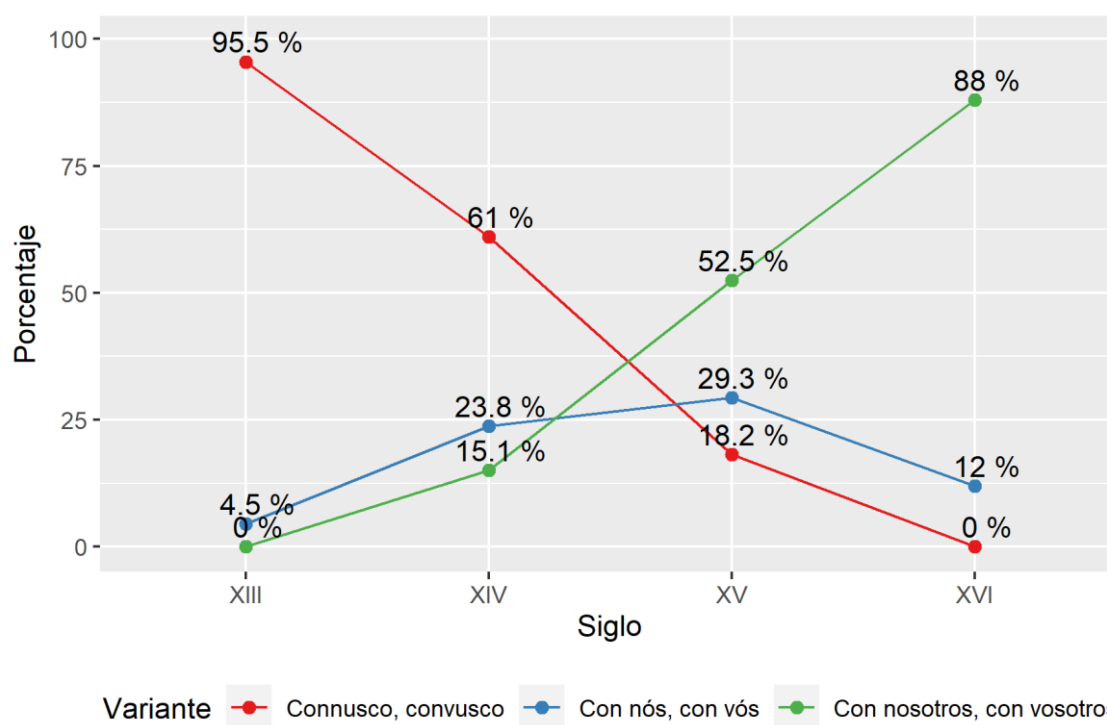


Figura 27. Evolución de las tres variantes con referente plural en textos literarios desde el siglo XIII al XVI.

En cuanto a los textos literarios, tal y como podemos ver en la Figura 27, la cronología de las variantes en esta tipología textual tiene algunas semejanzas con la de los textos documentales, pero también algunas diferencias sustanciales. El retroceso de *connusco* y *convusco* es prácticamente idéntico al que hemos visto en la Figura 26. En ambos casos el uso de estas variantes desciende de forma pronunciada y desaparecen del corpus en el

siglo XVI. Además, es necesario señalar que tampoco se registra ningún ejemplo de estos pronombres con referente singular en el siglo XVI⁴.

Otras similitudes las encontramos en el momento de aparición de las variantes analíticas, tal y como hemos detallado en la sección anterior (§ 4.1). *Con nós* y *con vós* empiezan a aparecer esporádicamente en el siglo XIII; y *con nosotros* y *con vosotros* en el siglo XIV. Cabe destacar que, según Líbano Zumalacárregui (1988), las formas analíticas compuestas aparecen en la última década del siglo XV. Sin embargo, tanto en el corpus documental como en el literario encontramos documentaciones más tempranas.

La diferencia más notable entre los textos literarios y los documentales la encontramos en la difusión de las variantes analíticas. Los datos presentados en la Figura 27 coinciden con lo que ya sabemos sobre el avance de *nosotros* y *vosotros* (Capítulo 3, § 4.2), de modo que la difusión de las formas analíticas reforzadas es considerablemente más rápida en textos literarios. Por tanto, los textos de esta tipología de entre los siglos XIII y XVI reflejan el reemplazo completo de *connusco/convusco* por *con nosotros/con vosotros*. En este caso, a diferencia de los textos de tipo documental, *con nós* y *con vós* no llegan a ser formas mayoritarias y no superan en ningún momento el 30 % de aparición. Así, las variantes analíticas simples muestran una difusión con forma de curva en campana, una forma que solo llegábamos a intuir en los textos documentales (Figura 26).

Además, es necesario precisar que, según García (1991:11), uno de los principales motivos por los que las formas en *-usco* desaparecen es la aparición de *nosotros*, *vosotros*. Sin embargo, el hecho de que el descenso de *-usco* sea el mismo al margen de una mayor o menor presencia de *-otros* descarta totalmente esta explicación. Por lo que se ve en los datos, una mayor o menor presencia de *nosotros*, *vosotros* no acelera ni frena el descenso de *connusco*, *convusco*. Este motivo nos hace pensar que probablemente las formas en *-usco* hubieran desaparecido igualmente a favor de las formas analíticas *con nós* y *con vós*, sin necesidad del modificador *otros*.

En definitiva, el descenso de *connusco* y *convusco* se produce al mismo ritmo en ambos tipos de texto. Se inicia en el siglo XIII y se completa en el XVI, al margen de cuál de las variantes analíticas sea la mayoritaria en ese momento, algo que sí depende de la tipología textual. El hecho de que esta variable afecte de forma desigual al retroceso de *connusco*, *convusco* y al avance de *con nosotros*, *con vosotros* sugiere que estamos ante dos cambios lingüísticos de naturaleza distinta.

4.3. Primera y segunda persona

Uno de los factores más relevantes en el reemplazo de *nós* y *vós* por *nosotros* y *vosotros* es el de persona. Las primeras documentaciones de las formas reforzadas son de segunda persona tanto en el corpus de textos literarios como en el de textos documentales. Además, en ambos tipos de texto podemos observar que *vosotros* es la forma que lidera el cambio, de modo que el modificador *otros* debió de extenderse a la primera persona por analogía (Capítulo 3, § 4.3).

Esta variable no ha recibido tanta atención en el análisis de *connusco* y *convusco*. Según García (1997:75), en este caso, la forma que lidera el cambio es la de primera persona *con nós*. Comprobar la relevancia de esta variante en el retroceso de las formas analíticas puede ser útil para delimitar la relación entre la pérdida de estas variantes y la aparición de *nosotros*, *vosotros*.

⁴ De cualquier modo, sabemos que estas formas sobrevivieron como arcaísmos hasta el siglo XVII (García 1991), tal y como se puede ver en los ejemplos de Alvar y Pottier (1987: 124 n. 30): «[las formas sintéticas] llegaron a Quevedo que utilizó *connusco* en un romance de carácter burlesco y *convusco* en un romance del Cid».

Tanto en los textos documentales (Figura 28) como en los literarios (Figura 29) las diferencias entre la primera y la segunda persona parecen mínimas. En el único punto en el que quizá podemos ver un mayor descenso de *connusco* frente a *convusco* es en los textos documentales del siglo XIV. Respecto al siglo anterior, el descenso en el uso de la primera persona (43.8 %) es considerablemente más abrupto que en la segunda persona (66.2 %). No obstante, esta desigualdad no se mantiene ni en los siglos siguientes ni en los textos literarios.

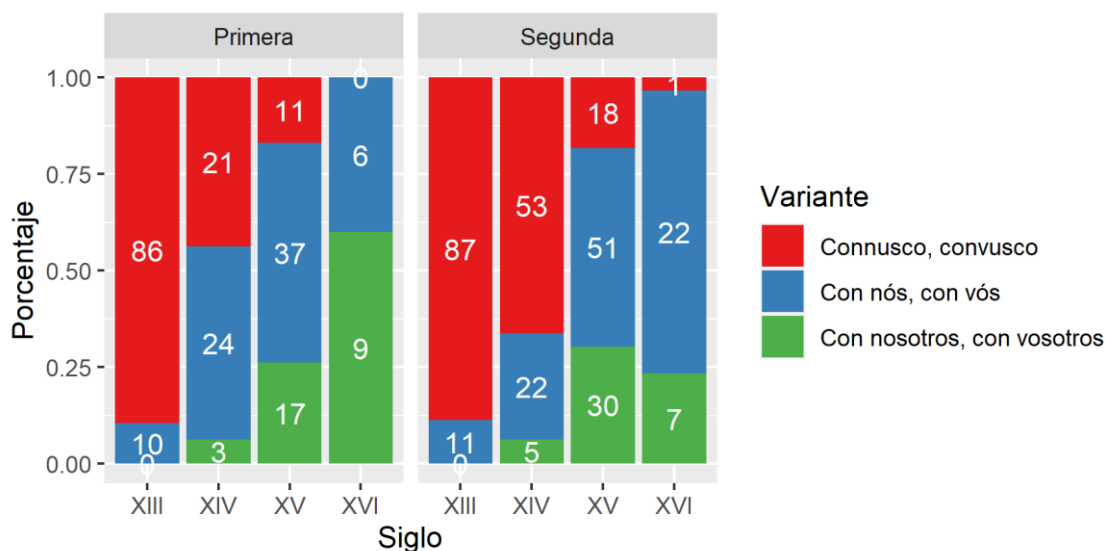


Figura 28. Porcentaje de aparición y ocurrencias de cada una de las variantes de primera y segunda persona en textos documentales de los siglos XIII al XVI.

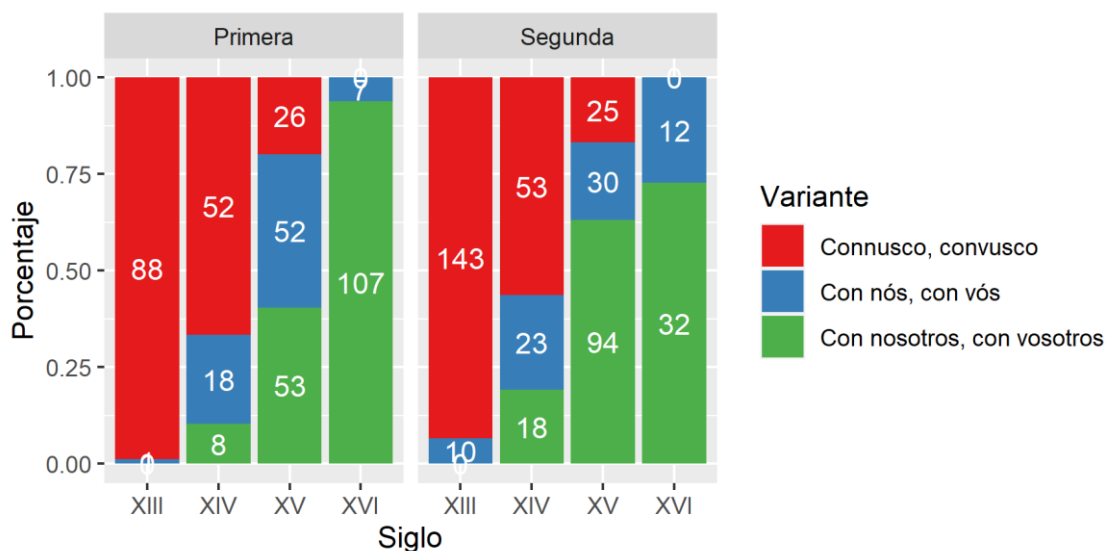


Figura 29. Porcentaje de aparición y ocurrencias de cada una de las variantes de primera y segunda persona en textos literarios de los siglos XIII al XVI.

Desde el punto de vista estadístico, las pequeñas diferencias que podemos observar entre los datos de primera y segunda persona no son significativas. Si vemos los datos de los siglos XIII al XVI de forma global, en los textos documentales (Tabla 16) encontramos prácticamente las mismas frecuencias relativas para ambas personas, de modo que no

parece haber ninguna relación entre esta variable y la aparición de cada una de las variantes.

En los textos literarios (Tabla 17), el porcentaje de las formas sintéticas es ligeramente más bajo en la primera persona, lo que podría sugerir que el cambio podría haber empezado ahí. No obstante, estadísticamente estos valores siguen sin ser significativos.

	Primera persona	Segunda persona
Sintética	118 (52.7 %)	159 (51.8 %)
Analítica simple	77 (34.4 %)	106 (34.5 %)
Analítica compuesta	29 (12.9 %)	42 (13.7 %)
Total	224	307

Pearson's x-squared = 0.072642, df = 2, p = 0.96433
Contingency coeff.: 0.012
Cramer's V: 0.012

Tabla 16. Frecuencias de cada una de las variantes en primera y segunda persona (textos documentales).

	Primera persona	Segunda persona
Sintética	166 (40.3 %)	221 (50.2 %)
Analítica simple	78 (18.9 %)	75 (17 %)
Analítica compuesta	168 (40.8 %)	144 (32.7 %)
Total	412	440

Pearson's x-squared = 8.8108, df = 2, p = 0.012211
Contingency coeff.: 0.101
Cramer's V: 0.102

Tabla 17. Frecuencias de cada una de las variantes en primera y segunda persona (textos literarios).

Por tanto, no parece que haya ninguna diferencia relevante entre la primera y la segunda persona. Una vez más, esto distingue este cambio lingüístico frente al más general de la aparición de *nosotros* y *vosotros*.

4.4. Valor contrastivo y contexto sintáctico

Una de las causas señaladas para el reemplazo de *connusco*, *convusco* por *con nós*, *con vós* y, más tarde, por *con nosotros*, *con vosotros* es la adición de un matiz de contraste. Según García (1991:14-15), en sus primeras ocurrencias, *con vós* aparecería en contextos contrastivos en los que interesaría diferenciar entre la segunda persona y el resto, mientras que *convusco* aparecería en contextos de unión solidaria. La autora destaca ejemplos como el del verbo *combatir*, que seleccionaría la variante *convusco* en contextos de alianza, frente a *con vós* en contextos de enemistad. Es decir, las apariciones de *con vós* tendrían lugar no con el valor de compañía habitual de la preposición *con*, sino con valor de 'contra'.

En la Figura 30 podemos ver la frecuencia relativa de cada una de las variantes junto con los verbos más frecuentes del corpus de textos literarios⁵. En el siglo XIII, el único verbo que no tiene las formas analíticas como únicas variantes pronominales es el verbo *ser*. Estos 10 ejemplos se corresponden con los ya citados en (3). Todos pertenecen a E6 y forman parte de la misma estructura: «sea con todos vós». Por tanto, en estos primeros ejemplos no podemos observar ningún valor contrastivo.

⁵ En la Figura 30 únicamente aparecen representados los verbos con más de 5 ejemplos en ese siglo. El tamaño del círculo es proporcional al número de ocurrencias.

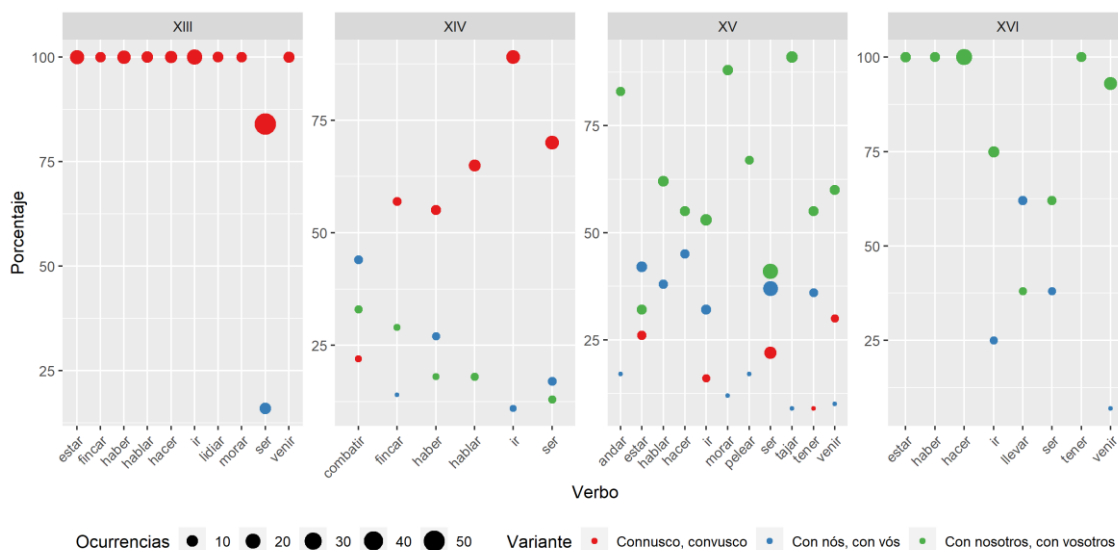


Figura 30. Ocurrencias y porcentajes de aparición de cada una de las variantes en los verbos más frecuentes del corpus de textos literarios (siglos XIII-XVI).

En el siglo XIV, en cambio, la situación es diferente. Las formas sintéticas siguen siendo, en casi todos los verbos, las variantes mayoritarias, pero conviven también con las analíticas simples y las reforzadas. Solo un verbo tiene *connusco*, *convusco* como variantes minoritarias: *combatir*. En los 9 ejemplos de este verbo que encontramos en (6) vemos, por lo general, un uso contrastivo de las variantes pronominales. Es decir, el verbo *combatir* suele ir acompañado de la persona *contra* la que se combate. De acuerdo con la hipótesis de García (1991), podríamos pensar que este es el motivo por el que la presencia de formas analíticas, tanto simples como reforzadas, es mayor en este verbo.

(6)

- a) aquestos que cada día salen a combaterse **coñusco** a grand porfía (*CTroyana*)
- b) e los de dentro non se trabajarán agora de salir acá nin de se combatir **coñusco** (*CTroyana*)
- c) son aquellos mismos que se combatieron **con uos otros** en las yslas gradaynes (*CQI*)
- d) et fabricio crido / O ercules fasta que pirus combata **con nos** aura oportunitat de pensar tales doctrinas (*CQI*)
- e) si el grant alexandre fuesse uenido en ytalia et huuyes combatido **con nos** et con nuestros parientes non seria famoso (*CQI*)
- f) Et sabet que ellos se quieren combatir **con uos** non pax por ardidez mas por necessitat (*CQI*)
- g) nuestro senyor natural se combate en semble **con vos otros** por la fe de ihesu xpisto (*CQ2*)
- h) son aquellos mismos que se combatieron **con nos otros** en las yslas gardaynes (*GCI*)
- i) Et sabet que ellos se querien combatir **con uos** non pas por ardidez mas por necessitat (*GCI*)

No obstante, si prestamos más atención a los ejemplos de (6), deberemos descartar esta explicación. Las dos ocurrencias de *connusco* (6a-b) aparecen en la *CTroyana*, testimonio que solo registra formas sintéticas, al margen del contexto. Los siete ejemplos restantes (6c-i) pertenecen a *CQI*, *CQ2* y *GCI*, que solo registran formas analíticas. Por tanto, la elección de cada una de las variantes no se debe tanto al valor contrastivo o solidario, sino a la presencia general de esas formas dentro de cada obra. Además, el único ejemplo en el que podemos identificar un valor de solidaridad (6g) lo hace usando *con vosotros*, no

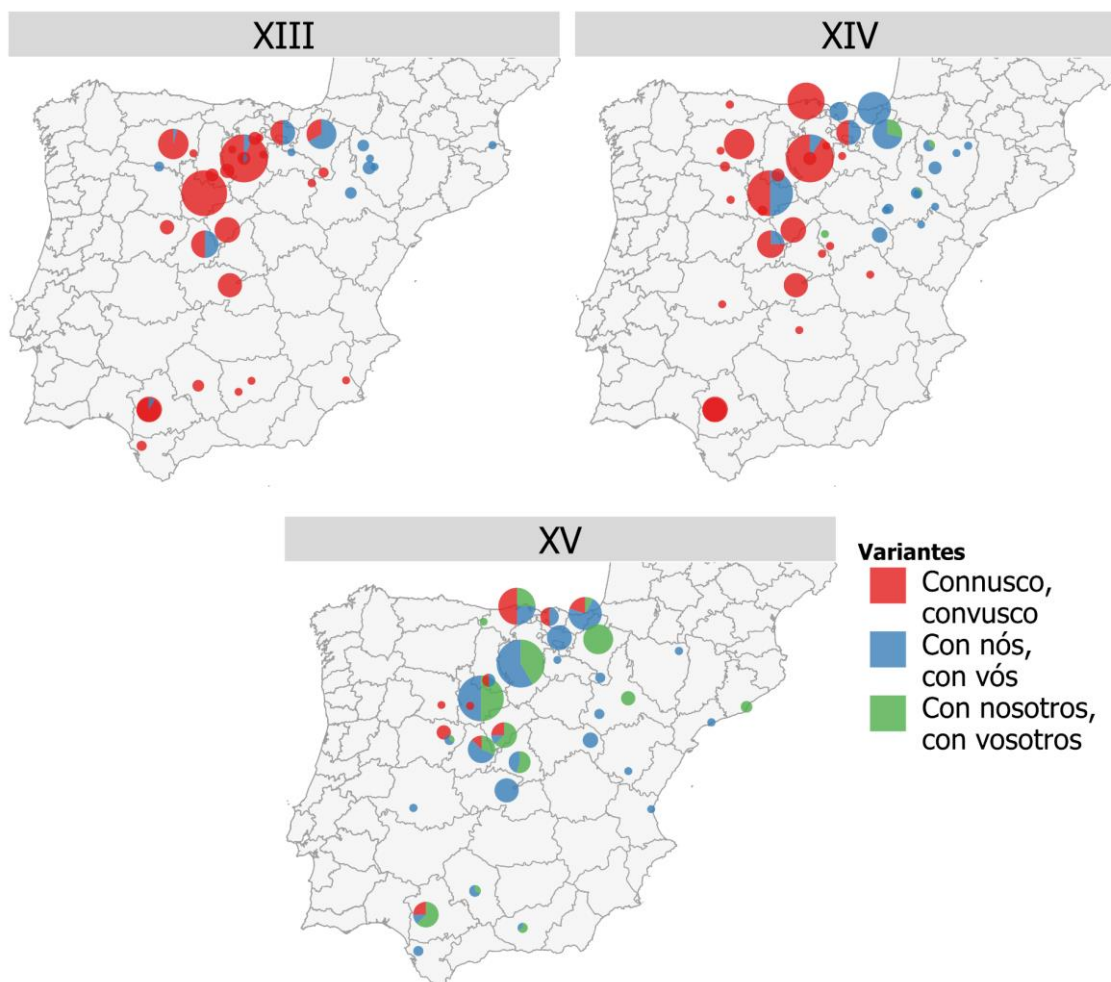
connusco. Para enfatizar el valor de unión no se utiliza la forma sintética de los pronombres, sino el adverbio *ensemble*.

En el siglo XV podemos agrupar los verbos entre los que mantienen las variantes *connusco*, *convusco*, aunque sea como formas minoritarias —*estar*, *ir*, *ser*, *tener*, *venir*—; y los que ya solo utilizan las analíticas —*andar*, *hablar*, *hacer*, *morar*, *pelear*, *tajar*. En ninguno de los dos casos parece haber una distribución basada en el significado de los verbos.

En definitiva, no hay ninguna evidencia que apoye una distribución de las variantes según el valor contrastivo o solidario del contexto próximo.

4.5. Distribución geográfica

La presencia de las variantes analíticas *connusco* y *convusco* en gallego y portugués actuales frente a su ausencia en el resto de variedades iberorromances sugiere una distribución occidente-oriente. Para comprobar la trayectoria de estas formas en el espacio geográfico podemos cartografiarlas a partir de los datos de los textos documentales. En el Mapa 5 puede verse la distribución de las tres variantes entre los siglos XIII y XV.



Mapa 5. Distribución de las variantes *connusco/convusco*, *con nós/con vós* y *con nosotros/con vosotros* en los siglos XIII-XV (textos documentales).

En el siglo XIII, de acuerdo con lo que hemos visto en la cronología, *connusco* y *convusco* son las variantes predominantes. Asimismo, las zonas con más ejemplos de estas formas

son Burgos, Valladolid y León. Las formas analíticas simples *con nós* y *con vós* también se documentan en todo el norte peninsular, aunque en menor medida. En cualquier caso, es importante destacar que incluso en los documentos más tempranos no encontramos ni un solo ejemplo de pronombres en *-usco* en el extremo más oriental, en Aragón.

En el siglo XIV, Navarra y Aragón presentan un uso exclusivo de variantes analíticas, que van empujando a las sintéticas hacia el occidente peninsular. En cuanto a *con nosotros* y *con vosotros*, empiezan a difundirse claramente desde el extremo oriental. *Con nós* y *con vós*, en cambio, aunque tengan una mayor vitalidad en esta misma área, no dejan de documentarse en otros puntos del *continuum* norte, como Burgos, Valladolid y Ávila. En el siglo XV, las formas analíticas simples y compuestas se han extendido rápidamente por el oeste peninsular y han aumentado su frecuencia de uso. Las formas en *-usco* se conservan principalmente en dos núcleos: uno en el norte, en Cantabria y el País Vasco; y otro en el oeste, en Zamora, Salamanca, Valladolid, Ávila y Segovia. En ambos casos, conviven con el resto de variantes.

En resumen, en el siglo XIII *connusco* y *convusco* empiezan siendo las formas mayoritarias en todo el norte peninsular excepto Navarra y Aragón. En el siglo XV, quedan arrinconadas en el norte y en el oeste. *Con nós* y *con vós* están presentes desde el inicio en todo el territorio peninsular, aunque sea de forma muy minoritaria. Debido a la ausencia de formas en *-usco* en aragonés, gozan de una mayor vitalidad en el extremo oriental, desde donde avanzan junto con las reforzadas hacia el resto de la península.

En cuanto a los textos literarios, debido a que no podemos localizar los datos en un punto concreto, compararemos la distribución de las variantes en las tres zonas geográficas principales que hemos presentado en la descripción del corpus (*cf.* Capítulo 2, § 4.3), de oeste a este: León y Castilla occidental, Castilla central y oriental, Navarra y Aragón. En la Figura 31 se muestra la distribución de cada una de las formas pronominales en los textos literarios de los siglos XIII-XVI.

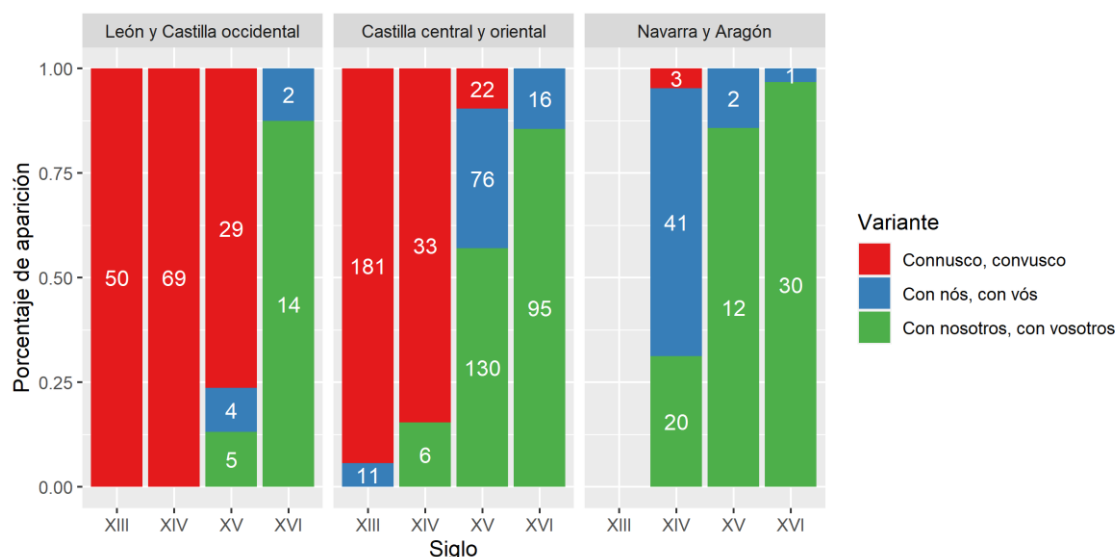


Figura 31. Ocurrencias y porcentaje de aparición de las variantes tras *con* en cada una de las zonas geográficas en textos literarios (siglos XIV-XVI).

Debido a que el corpus de textos literarios no tiene ningún texto de Navarra o Aragón para el siglo XIII, no podemos comparar este periodo directamente con el Mapa 5. En cualquier caso, los datos del siglo XIV son suficientemente elocuentes. En el trescientos, igual que ocurría en los textos documentales, *connusco* y *convusco* son las formas mayoritarias en todo el territorio peninsular excepto en el extremo más oriental, en los

textos navarroaragoneses. Navarra y Aragón destacan por tener únicamente tres ejemplos de los pronombres sintéticos. Además, estos tres casos pertenecen a un solo testimonio, la *HTroyana*. Las 61 ocurrencias restantes, que suponen el 95 % de los casos, se corresponden con variantes sintéticas y analíticas. En contraste, las otras dos zonas geográficas tienen como mayoritarias las variantes en *-usco*, con un 100 % en León y Castilla Occidental, y un 85 % en Castilla central y oriental.

En cambio, en el siglo XV *connusco* y *convusco* retroceden y solo se mantienen como predominantes en el extremo occidental. En Castilla central y oriental estas variantes han pasado a ser residuales de un siglo al otro, y no se documenta ningún ejemplo en los textos navarroaragoneses.

En el siglo XVI, como hemos ido comentando en las secciones anteriores, ya no hay rastro de *connusco* y *convusco*. La distribución del resto de variantes es similar en las tres zonas geográficas: *con nosotros* y *con vosotros* son formas mayoritarias, mientras que *con nós* y *con vós* aparecen de forma esporádica.

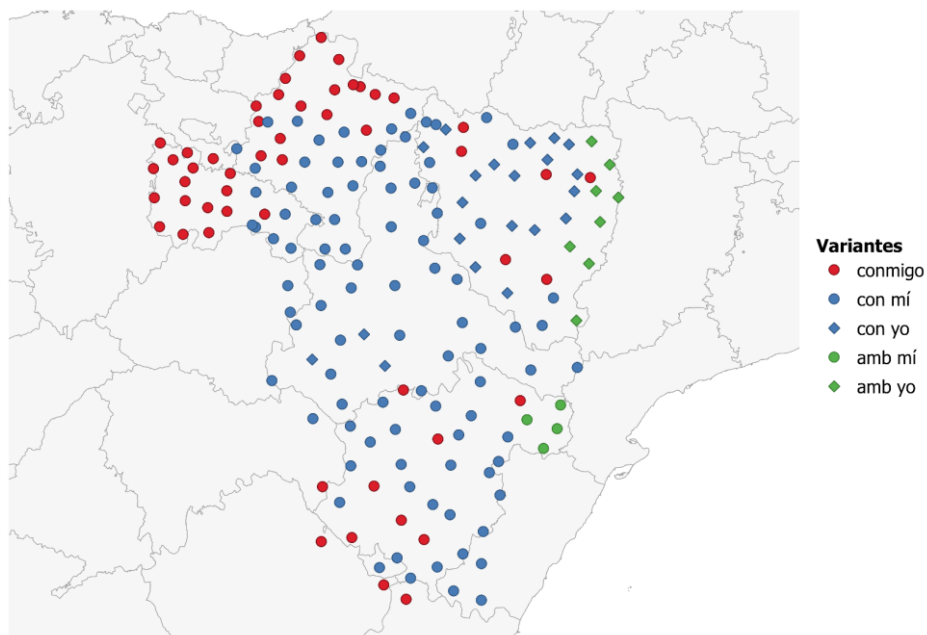
En Gomila Albal (2018) se concluía que *con nós* y *con vós* se extendían desde el oriente de la península ibérica. A la vista de los datos, esta afirmación es matizable. Tal y como se ve en el Mapa 5 y en la Figura 31, en los textos del XIII *con nós* y *con vós* no son exclusivos del oriente peninsular. En los textos documentales encontramos ejemplos de estas variantes en Navarra, Huesca y Zaragoza, pero también en Ávila, León, Burgos y Cantabria. En los textos literarios de este siglo también las registramos en Castilla central y oriental. Por tanto, no se puede afirmar que estas variantes se originaran en Navarra y Aragón. Además, hay que tener en cuenta que *nós* y *vós* existían y eran las formas habituales en el resto de contextos sintácticos de los pronombres personales tónicos. Es decir, *con nós* y *con vós*, aunque esporádicas, eran opciones posibles en todo el norte peninsular.

Lo verdaderamente relevante es que las formas sintéticas están totalmente ausentes de los textos navarroaragoneses documentales y solo aparecen en uno de los textos literarios del corpus. Una explicación para esta ausencia es que la preposición CUM no tuvo continuación en el latín de la Galia y Cataluña, sino que, en su lugar, estas variedades tomaron la preposición APUD (> *ab* > *amb*) (Batlle et al. 2016:289). De este modo, frente a la evolución de NOBISCUM y VOBISCUM que encontramos en castellano, que mantiene una posposición irreconocible para los hablantes y añade de nuevo la preposición, el catalán y el francés reemplazan directamente las formas sintéticas latinas por las analíticas *ab nós* y *ab vós* de forma muy temprana.

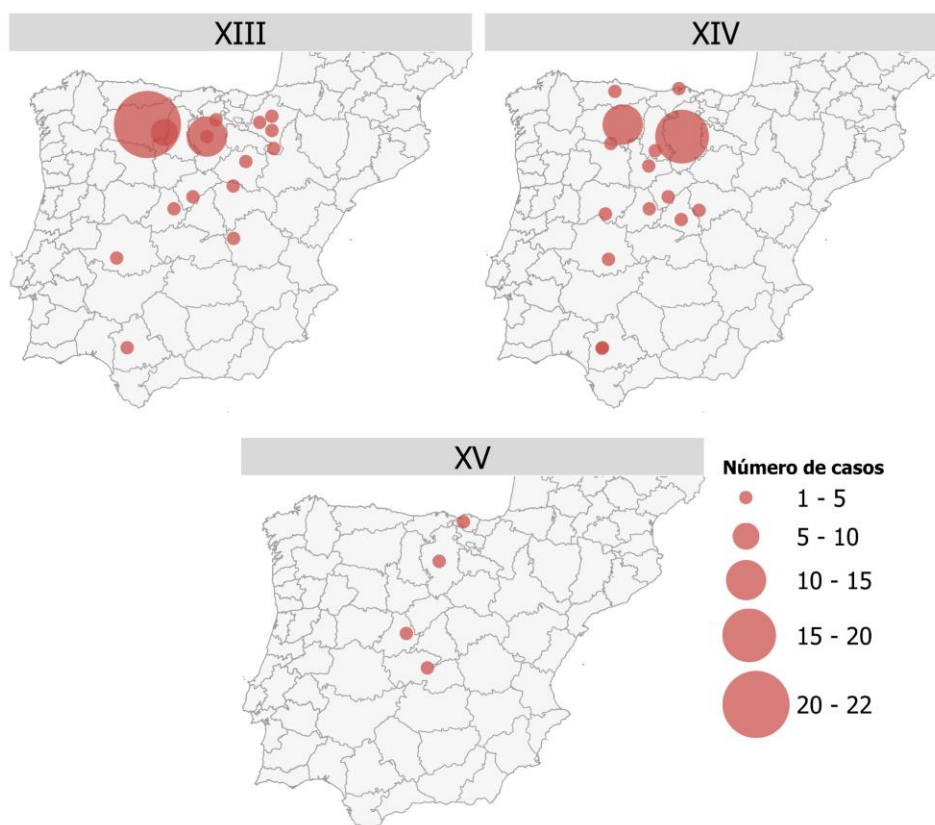
Las formas pronominales de estas variedades romances no pasaron por el proceso de desgaste y refuerzo de NOBISCUM y VOBISCUM, sino que, al reemplazar CUM por APUD, nivelan todo el paradigma pronominal y usan los pronombres NOS y VOS como ante cualquier otra preposición. Así, el uso de formas analíticas en el extremo más oriental de la península ibérica pudo extenderse tempranamente al aragonés. Más tarde, desde ahí, el uso de *con nós*, *con vós* pudo generalizarse al resto de la península al mismo tiempo que se difundían las reforzadas *con nosotros*, *con vosotros*.

La influencia del catalán sobre las variedades navarroaragonesas en el uso de variantes pronominales analíticas tipo *con* + pronombre tónico se extiende hasta la actualidad y hasta a las formas singulares del paradigma. En el Mapa 6 podemos ver las distintas realizaciones del comitativo de primera persona singular en el ALEANR. Las formas mayoritarias en todo el territorio son *con yo* y *con mí*. *Conmigo* aparece muy esporádicamente, pero en general queda reservado para el noroeste de Navarra y el oeste

de La Rioja. Además, en la franja oriental de Aragón, en la zona catalanoparlante, destaca el uso de distintas realizaciones fonéticas de la preposición *amb* (<APUD)⁶.



Mapa 6. Realizaciones de las variantes de *conmigo* en el ALEANR.



Mapa 7. Documentaciones de las variantes *conusco* y *convusco* con referente singular en el corpus de textos documentales.

⁶ Para simplificar el cartografiado he lematizado las distintas transcripciones fonéticas de *amb* usadas en el ALEANR, como [en], [ən], [əm], [am], etc. bajo la forma <amb>.

Otro de los factores que apoya la idea de que la pérdida de *connusco* y *convusco* no está tan directamente relacionada con la aparición de los pronombres reforzados de plural *nosotros* y *vosotros* como se había pensado hasta ahora es el recorrido geográfico de estas mismas variantes con referente singular (Mapa 7). Ya hemos visto más arriba que el descenso de las formas *connusco* y *convusco* ocurre al mismo ritmo en textos documentales y literarios. En los textos documentales, en los que la presencia de *nosotros*, *vosotros* es muy reducida, las formas en *-usco* retroceden a la misma velocidad que en los textos literarios, en los que las formas reforzadas son más abundantes. Además, si cartografiamos los ejemplos de *connusco* y *convusco* con referente singular, es decir, utilizados con valor mayestático o de cortesía, el retroceso hacia occidente es el mismo que hemos visto en el Mapa 5.

La presión del paradigma pronominal parece, por tanto, el factor determinante para el reemplazo de las variantes sintéticas por las analíticas. Esto distancia el fenómeno de pérdida de *connusco*, *convusco* de la aparición de los pronombres de primera y segunda persona plural *nosotros*, *vosotros*. A su vez, el retroceso de las formas en *-usco* se enmarca en uno mayor que afecta también a las formas en singular como *connmigo* (Mapa 6). De este modo, estaríamos ante un proceso de reajuste de las variantes pronominales que habría empezado en el siglo XIII y todavía no habría concluido.

5. Resumen

En este capítulo he analizado la pérdida de los pronombres personales de primera y segunda persona plural *connusco* y *convusco*, formas propias del español medieval, por las variantes analíticas *con nós* y *con vós*, y, finalmente, por las analíticas reforzadas actuales *con nosotros* y *con vosotros*. Una vez más, este es un fenómeno estrechamente relacionado con la aparición de *nosotros* y *vosotros*, analizada en el Capítulo 3.

A partir de las preguntas de investigación presentadas al final del estado de la cuestión (§2) hemos obtenido las siguientes respuestas:

Sobre la cronología en distintos tipos de texto, vemos que, mientras que en el reemplazo de *nós*, *vós* por *nosotros*, *vosotros* es más rápido en los textos literarios que en los documentales, el descenso de *connusco* y *convusco* se produce al mismo ritmo en ambos tipos de texto. La pérdida de las variantes en *-usco* se inicia en el siglo XIII y se completa en el XVI, al margen de cuál de las variantes analíticas —ya sean las formas tipo *con nós* o *con nosotros*— sea la mayoritaria en ese momento, algo que sí depende de la tipología textual. El hecho de que esta variable afecte de forma desigual al retroceso de *connusco*, *convusco* y al avance de *con nosotros*, *con vosotros* sugiere que estamos ante dos cambios lingüísticos de naturaleza distinta.

Respecto al orden de aparición, no parece que haya ninguna diferencia relevante entre la primera y la segunda persona. Una vez más, esto contrasta con la aparición de *nosotros*, *vosotros*, donde las formas de segunda persona aparecieron antes y las de primera persona después.

Sobre los valores semánticos que pudieron haber motivado la aparición de las formas innovadoras, no hay ninguna evidencia que muestre el uso de *connusco* y *convusco* en contextos de unión solidaria frente al uso de *con nós* o *con nosotros* en contextos de enemistad. No encontramos ningún tipo de distribución de las variantes según el significado del verbo y la elección de una forma u otra se debe más a los usos generales de cada obra.

En cuanto a la variación diatópica, en el siglo XIII *connusco* y *convusco* eran las formas mayoritarias en todo el norte peninsular excepto Navarra y Aragón. En el siglo XV retroceden y quedan arrinconadas en el norte y en el oeste. *Con nós* y *con vós* están presentes desde el inicio en todo el territorio peninsular, aunque en diferente grado:

mientras que en el centro y el oeste su presencia es limitada, tienen una mayor vitalidad en las variedades orientales. Esto es debido a la ausencia de formas en *-usco* en aragonés, gozan de una mayor vitalidad en el extremo oriental, desde donde avanzan junto con las reforzadas hacia el resto de la península.

La distribución geográfica de *connusco*, *convusco*, más abundantes en el oeste; y *con nós*, *con vós*, con una mayor presencia en el este, se basa en la ausencia de los continuadores de NOBISCUM y VOBISCUM en catalán y en francés, que reemplazaron directamente las formas sintéticas latinas por las analíticas *ab nós* y *ab vós* de forma muy temprana. El uso de formas analíticas en las variedades orientales pudo extenderse tempranamente al aragonés. Más tarde, desde ahí, el uso de *con nós*, *con vós* pudo generalizarse al resto de la península al mismo tiempo que se difundían las reforzadas *con nosotros*, *con vosotros*.

En Gomila Albal (2018) se concluía que *con nós* y *con vós* se originaban en el oriente de la península ibérica. A la vista de los nuevos datos, esta afirmación es matizable. En los textos del XIII *con nós* y *con vós* no son exclusivos del oriente peninsular. Por tanto, no se puede afirmar que estas variantes se “originaran” en Navarra y Aragón. *Con nós* y *con vós*, aunque esporádicas, eran opciones posibles en todo el norte peninsular.

Durante todo el análisis de este capítulo se han mencionado algunas características del reemplazo de *connusco*, *convusco* que separan este proceso de cambio del que afecta a las formas de primera y segunda persona plural *nós*, *vós*. El reemplazo de *nós*, *vós* por *nosotros*, *vosotros* se ve considerablemente afectado por dos variables: el tipo de texto y la persona gramatical. Sin embargo, estos dos parámetros no parecen intervenir en la pérdida de *connusco*, *convusco*, y todo sugiere que dichas variantes hubieran desaparecido incluso si *nosotros*, *vosotros* no hubieran surgido. Por tanto, el factor determinante para el reemplazo de las variantes sintéticas en *-usco* por las analíticas de *con* + pronombre tónico parece ser la tendencia a la nivelación del paradigma pronominal. A su vez, el retroceso de las formas en *-usco* se enmarca en un proceso de cambio mayor que afecta también a las formas en singular como *conmigo*. De este modo, estaríamos a un proceso de reajuste de las variantes pronominales que habría empezado en el siglo XIII y todavía no habría concluido.

Summary and conclusions

The present study aimed to re-examine the processes of change that affect the paradigm of stressed pronouns of the first- and second-person plural in Old Spanish (i.e. *nós*, *vós*, *nosotros*, *vosotros*, *connusco*, *convusco*) with particular focus on the diachronic and diatopic aspects of their historical evolution. Thus, my main purpose has been to trace the origin and diffusion in time and space of each of the relevant linguistic variants.

In collecting the corpus for this study, I have gathered texts of varied typology through the rigorous and careful selection of sources. In doing so I have paid special attention to the philological quality of the texts, avoiding those transmitted in late copies, or the ones for which there is no information on date of composition or geographical origin. The establishment of a solid empirical base, which enabled me to trace the chronological and geographical origin of the changes, allowed me to test new hypotheses and to examine in detail the diffusion patterns of the aforementioned morphosyntactic variants throughout the Iberian Peninsula. The use of a wide corpus of varied textual typologies has allowed me to conduct a thorough re-examination of aspects analyzed in previous studies.

I started by examining the replacement of the plural pronouns *nós* ‘we’, *vós* ‘you-PL’ by the compound forms *nosotros*, *vosotros* (Chapter 3). In this regard, I have analyzed the following aspects: the first attestations, the chronology, the order of appearance, the syntactic context, the diffusion of the new variants in the geographical space and their propagation through different levels of formality.

The pronouns *nosotros*, *vosotros* are first attested in the second half of the 13th century in both the literary and legal corpus. In these first appearances, the contrastive value that has been attributed in the literature to the newly emerging pronouns is not always clearly reflected in each individual example; in a few cases, however, a contrastive use can be perceived, in which the reinforced pronoun is used to establish a contrast with one of the interlocutors in an interaction or with a third person.

Regarding the chronology, my data from literary texts concurs with the findings from previous work; that is, the replacement of *nós*, *vós* by *nosotros*, *vosotros* takes place between the 13th and 16th centuries. In contrast, the progress of the innovative variants in legal texts—a textual type not explored in previous studies—is considerably slower. In percentage terms, in the 16th century this type of text was still in an initial phase of change.

Regarding the order of appearance, it seems clear that the second person *vosotros* is the one that leads the process of change, while the first person *nosotros* is formed by analogy. This could be explained by the greater need to disambiguate the number of the second person *vós* pronoun, which was also frequently used as a form of polite address with a singular referent.

As for the syntactic position, I have considered two contexts: subject and prepositional complement. The context after preposition seems to slightly favor the appearance of the reinforced pronouns *nosotros*, *vosotros* as opposed to the subject position. Therefore, the innovative forms would have spread from the lowest positions in the grammatical hierarchy (i.e. oblique) to the higher ones (i.e. subject). In any case, the syntactic position does not seem to have been a particularly determining factor in the adoption phase of the new forms. A more relevant factor may have been the presence of specifying structures and appositions; these alternative mechanisms mark explicitly the number of the pronoun and, consequently, would have favored the retention of the simple pronouns *nós* and *vós*.

The diffusion of *nosotros*, *vosotros* in the geographical space occurs at the same pace in legal and literary texts. The new forms are first documented in the eastern areas

in the second half of the 13th century. Albeit with rather low percentages of appearance in the legal corpus, by the 14th century the pronouns *nosotros*, *vosotros* already appear throughout the center and east of the peninsula, not reaching Leon and western Castile until the 15th century. In the 16th century we can finally find the compound pronouns in any point of the northern dialect *continuum*.

As for stylistic variation, if we consider the diffusion of the new pronouns through texts with different levels of formality, we can distinguish two contrasting paths for the introduction of the reinforced pronouns *nosotros*, *vosotros*. In the east, where the change first took place, the innovative forms spread “from below”, from the less formal texts. This can be interpreted as a change that takes place in areas where there is prolonged language contact via bilingualism with Catalan and Occitan. In the center and west, on the other hand, the new pronouns spread “from above”; that is, in the areas furthest from the focus of change, the innovative variants arrived as prestigious forms that were first disseminated through literary texts.

In Chapter 4 I have extended the analysis of the replacement of the plural pronouns *nós*, *vós* by *nosotros*, *vosotros* with the incorporation of other emphatic pronominal reinforcements, such as *todos* ‘all’, *mismos* ‘selves’, and *ambos* ‘both’. In this chapter the following aspects are considered: the first occurrences, the chronology, the influence of the grammatical person variable and the geographical distribution.

The data shows that the productivity of modifiers such as *todos*, *mismos* and *ambos* is limited, in general, to the earliest texts. At any rate, in no case does the use of any of said pronominal reinforcements surpass that of the simple forms *nós* and *vós*.

During the 13th and 14th centuries, speakers experimented with different emphatic elements that modified the pronouns *nós* and *vós*, and whose shared purpose appears to be to explicitly mark the pronouns as plural. Hence, my hypothesis is that these modifiers could be used as disambiguating strategies of *nós* and *vós*.

The use of these pronominal reinforcements is maintained for a longer period of time in those contexts in which the replacement of *nós*, *vós* by *nosotros*, *vosotros* is slower, such as legal texts—with special resistance in the first person—produced in the westernmost regions.

Once *nosotros* and *vosotros* begin to spread and become grammaticalized, the other modifiers are abandoned and relegated to occasional emphatic functions. *Vós* / *nós* *mismos* ends up being used with a purely reflexive value and *mismo* eventually spreads to all the pronouns in the paradigm. *Vós* / *nós* *todos* disappears as a postposed pronominal reinforcement and nowadays we find it preposed (i.e. *todos nosotros*). Likewise, *vós* / *nós* *ambos* or *solos* ‘alone’ as well as the numerals (i.e. *vós* / *nós* *dos*) are kept with a marginal use in the Castilian texts.

The linguistic variants that most clearly compete with *nosotros* and *vosotros* are *nós todos* ‘we all’ and *vós todos* ‘you all’. Moreover, it is the only reinforcement besides *otros* that has a defined geographical distribution: in Old Spanish legal and literary texts of León and western Castile, as well as in Galician legal texts, *vós todos* is one of the main reinforced variants during the 13th and 14th centuries. Thus, two candidates arise in different areas of the Iberian Peninsula for the reinforcement of the pronouns *nós* and *vós*: *todos* in the West and *otros* in the East. The variation is resolved in the initial phases of the change in favor of *nosotros* and *vosotros* while the frequency of *nós todos* and *vós todos* decreases with an S-curve pattern.

In Chapter 5 I have focused on the pronominal variants after the preposition *con* ‘with’ (*connusco*, *con nós*, *con nosotros* ‘with us’, and *convusco*, *con vós* *con vosotros* ‘with you-PL’). The Old Spanish pronouns *connusco* and *convusco* were replaced by *con nós*, *con vós* first, and later by *con nosotros*, *con vosotros*. In this chapter I examined the

following aspects: the first documentations, the chronology, the influence of the grammatical person variable, the syntactic context, and the geographical distribution.

Regarding the chronology of the pronouns in different text types, we see that, while the replacement of *nós*, *vós* by *nosotros*, *vosotros* is faster in literary texts than in legal texts, the decline of *connusco* and *convusco* occurs at the same rate in both types of texts. The loss of the *-usco* variants begins in the 13th century and ends in the 16th in all text types and regardless of which analytical variant (whether *con vós* or *con vosotros* type forms) is the most common at that time. The fact that the regression of *connusco*, *convusco* takes place evenly in all text types while the diffusion of *con nosotros*, *con vosotros* is strongly affected by textual parameters suggests that we are facing two linguistic changes of different nature.

Regarding the order of replacement and loss of *connusco* and *convusco*, there does not seem to be any relevant difference between the first and the second person forms. Again, this is in stark contrast with the emergence of *nosotros* and *vosotros*, in which second-person forms appeared earlier, and first-person ones last.

Concerning the semantic values that could have motivated the appearance of the innovative forms, I have reviewed the claim that *connusco* and *convusco* are used in contexts of solidarity ('with you') as opposed to the use of *con nós*, *con vós* in contexts of enmity ('against you'). An extensive review of the lexical verb types that appeared with *con* + pronoun in the data did not yield any kind of distribution of the variants according to the meaning of the verb. The choice of one form or another seems determined by the general uses of each text.

As for the diatopic variation, in the 13th century *connusco* and *convusco* were the most common forms throughout the north of the peninsula except for Navarre and Aragon. In the 15th century, they receded, cornered in the north and west. *Con nós* and *con vós* are also present from the beginning throughout the peninsular territory but with varying degrees of intensity: while in the center and west their numbers are very limited, they enjoy a greater vitality in eastern varieties. This is due to the absence of the *-usco* forms in Aragonese. From there, *con nós* and *con vós* advance together along with the reinforced forms (i.e. *vosotros* type forms) towards the east of the peninsula.

The different geographical distribution of *connusco* and *convusco*, more prevalent in the west, and *con nós* and *con vós*, in the east, results from the absence of the descendants of *NOBISCUM* and *VOBISCUM* in Catalan and French, which directly replaced the Latin synthetic forms by the analytic ones *ab nós* and *ab vós* very early on. The use of analytical forms in Catalan varieties may have extended early to Aragonese, and from there, the use of *con nós*, *con vós* could be generalized to the rest of the peninsula at the same time that *con nosotros*, *con vosotros* spread.

In Gomila Albal (2018) it was concluded that *con nós* and *con vós* originated in the east of the Iberian Peninsula and from there spread towards the west. In view of the new expanded data, this statement should be nuanced: since *con nós* and *con vós* are not exclusive to the east of the Iberian Peninsula in 13th century texts, it cannot be said that these variants "originated" in Navarre and Aragon. Although rare, *con nós* and *con vós*, as opposed to *connusco* and *convusco*, were possible options throughout the north of the peninsula, including the western areas.

Recall that we already mentioned some characteristics of the replacement of *connusco*, *convusco* that separate this process of change from that undergone by the first- and second-person plural pronouns *nós*, *vós*. The replacement of *nós*, *vós* by *nosotros*, *vosotros* is strongly affected by two variables: textual typology and grammatical person. However, these parameters do not seem to intervene in the loss of *connusco*, *convusco*, and everything suggests that the forms in *-usco* would have disappeared even if *nosotros*,

vosotros had not originated. Therefore, the determining factor for the replacement of the synthetic variants (i.e. *convusco*, etc.) by the analytic ones (i.e. *con* + pronoun) seems to be the tendency to level the pronominal paradigm. In turn, the regression of the forms in *-usco* appears to be part of a larger process of change that also affects the singular forms, such as *conmigo* ‘with me’. This way, we would be witnessing a process of readjustment of the pronominal variants that would have started in the 13th century and has not yet concluded.

In sum, the data presented in this dissertation demonstrates the complex interaction of a wide range of factors in the diffusion of linguistic variants. This can only be revealed through the use of a corpus with precise chronological and geographical coordinates and with enough register variety to access, even indirectly, the stylistic factor.

Geographical and chronological variation has been crucial in all three processes of change presented here. The expansion of *nosotros*, *vosotros* originates in the east of the peninsula, as a result of language contact with Catalan, and the new forms are later adopted by Castilian writers as prestigious variants. Likewise, the exhaustive mapping of the reinforced pronominal variants reveals the relevance of two forms that had been previously overlooked, *nós todos* and *vós todos*, associated to the western regions. These forms, due to their brief vitality, have generally gone unnoticed in previous studies. Lastly, although *con nós* and *con vós* do not originate in the eastern regions, the absence of *-usco* variants in the Aragonese documents leads us to assume that the use of *con* + pronoun type forms could have spread earlier throughout the eastern peninsula due to contact with the variants *ab nós* and *ab vós* of Catalan.

Ultimately, this research underscores the possibilities afforded by the careful examination of data from a variety of registers, and the need to include analysis of geographic variation, in order to better understand the diffusion of linguistic innovations in language-contact situations.

Referencias bibliográficas

- Abella Samitier, Juan. 2009. *Selección de documentos de la villa aragonesa de Sos (1202-1533)*. Zaragoza: Institución «Fernando el Católico».
- Almeida Cabrejas, Belén. 2015. «Evolución de las fórmulas en las cartas de venta del “CODEA”». Pp. 601-15 en *Actas del IX Congreso Internacional de Historia de la Lengua Española (Cádiz, 2012)*, editado por J. M. García Martín, T. Bastardín Candón, y M. Rivas Zancarrón. Iberoamericana / Vervuert.
- Almeida Cabrejas, Belén, Pedro Sánchez-Prieto Borja, y Delfina Vázquez Balonga. 2018. «Para una geografía diacrónica: CODEA+ 2015 como atlas lingüístico dinámico del español». Pp. 1537-51 en *Actas del X Congreso Internacional de Historia de la Lengua Española: Zaragoza, 7-11 de septiembre de 2015*, editado por M. L. Arnal Purroy, R. M. Castañer Martín, J. M. Enguita Utrilla, V. Lagüéns Gracia, y M. A. Martín.
- Alvar, Carlos, y José Manuel Lucía Megías. 2002. *Diccionario filológico de literatura medieval española: textos y transmisión*. Madrid: Castalia.
- Alvar, Manuel. 1996. *Manual de dialectología hispánica: el español de España*. Barcelona: Ariel.
- Alvar, Manuel, y Bernard Pottier. 1983. *Morfología histórica del español*. Madrid: Gredos.
- Álvarez Blanco, Rosario. 1994. «Gramática descriptiva e gramática prescriptiva». *Cadernos de Lingua* 10:19-35.
- Álvarez Blanco, Rosario. 2004. «A variación nosco, connosco, con nós en galego medieval». *Verba* 31:45-72.
- Álvarez Blanco, Rosario, Xosé Luís Regueira Fernández, y Henrique Monteagudo. 1989. *Gramática galega*. 2ª. Vigo: Galaxia.
- Álvarez Fernández-Cañedo, Jesús. 1963. *El habla y la cultura popular de Cabrales*. Madrid: Instituto Miguel de Cervantes.
- Anthony, Lawrence. 2019. *AntConc (Versión 3.5.8) [Software]*. Tokio: Waseda University.
- Del Barrio de la Rosa, Florencio. 2016. «La distribución de las variantes -mente, -miente y -miente en el CODEA (1221-1420): espacio variacional y cambio lingüístico». *Scriptum Digital* 5:85-102.
- Del Barrio de la Rosa, Florencio. 2018. *Espacio variacional y cambio lingüístico en español*. Madrid: Visor Libros.
- Batlle, Mar, Joan Martí i Castell, Josep Moran i Ocerinjauregui, y Joan Anton Rabella. 2016. *Gramàtica històrica de la llengua catalana*. Barcelona: Publicacions de l'Abadia de Montserrat.
- De Benito Moreno, Carlota. 2015. *Las construcciones con «se» desde una perspectiva variacionista y dialectal*. Tesis doctoral: Universidad Autónoma de Madrid.
- De Benito Moreno, Carlota. 2019. «Los corpus del español desde la perspectiva del usuario lingüista». *Scriptum Digital* 8:1-21.
- Blas Arroyo, José Luis. 2004. «El español actual en las comunidades del ámbito lingüístico catalán». Pp. 1065-86 en *Historia de la lengua española*, editado por R. Cano Aguilar. Barcelona: Ariel.
- Blas Arroyo, José Luis. 2007. «El contacto de lenguas como factor de retención en procesos de variación y cambio lingüístico: Datos sobre el español en una comunidad bilingüe peninsular». *Spanish in Context* 4(2):263-91. doi: 10.1075/sic.4.2.07bla.
- Blythe, Richard A., y William Croft. 2012. «S-curves and the mechanisms of propagation

- in language change». *Language* 88(2):269-304.
- Breban, Tine. 2014. «What is secondary grammaticalization? Trying to see the wood for the trees in a confusion of interpretations». *Folia Linguistica* 48(2):469-502. doi: 10.1515/flin.2014.016.
- Bybee, Joan L. 2003. «Mechanisms of change in grammaticalization: the role of frequency». Pp. 602-23 en *The Handbook of Historical Linguistics*, editado por B. D. Joseph y R. D. Janda. Malden: Blackwell.
- Bybee, Joan L., Revere Perkins, y William Pagliuca. 1994. *The evolution of grammar: Tense, aspect and modality in the languages of the world*. Chicago / London: University of Chicago Press.
- Calderón Campos, Miguel. 2010. «Variantes formales y valores semánticos de (v)os(otros) en la diacronía del español». Pp. 135-47 en *De moneda nunca usada: Estudios dedicados a José M^a Enguita Utrilla*, editado por R. M. Castañer Martín y V. Llagüéns Gracia. Zaragoza: Insituto Fernando El Católico.
- Campbell, Lyle. 1999. *Historical linguistics: An Introduction*. Cambridge: MIT Press.
- Campbell, Lyle. 2001. «What's wrong with grammaticalization». *Language Sciences* 23:113-61. doi: 10.1016/S0388-0001(00)00019-X.
- Canellas López, Ángel. 1962. *Colección diplomática de la Almunia de Doña Godina*. Zaragoza: Institución «Fernando el Católico».
- Canellas López, Ángel. 1975. *Colección diplomática del Concejo de Zaragoza (1276-1285)*. Zaragoza: Cátedra Zaragoza en la Universidad.
- Cano, Rafael. 2005. *Historia de la Lengua Española*. Barcelona: Ariel.
- Cavero Domínguez, Gregoria. 2001. *Colección documental del monasterio de San Esteban de Nogales (1149-1498)*. León: Centro de estudios e investigación «San Isidoro».
- Chambers, J. K. 2013. «Patterns of Variation including Change». Pp. 297-323 en *The Handbook of Language Variation and Change*, editado por J. K. Chambers y N. Schilling. Oxford: Wiley-Blackwell.
- Chambers, J. K., y Peter Trudgill. 2004. *Dialectology*. 2^a edición. Cambridge: Cambridge University Press.
- Cierbide, Ricardo, y Emiliana Ramos. 1998. *Documentación medieval del Archivo Municipal de Pamplona I (1129-1356)*. Donostia: Eusko Ikaskuntza.
- Cierbide, Ricardo, y Emiliana Ramos. 2000. *Documentación medieval del Archivo Municipal de Pamplona II (1357-1512)*. Donostia: Eusko Ikaskuntza.
- Company, Concepción. 2003. «La gramaticalización en la historia del español». *Medievalia* 35:3-62.
- Company, Concepción. 2008. «Gramaticalización, género discursivo y otras variables en la difusión del cambio sintáctico». Pp. 17-52 en *Sintaxis histórica del español y cambio lingüístico. Nuevas perspectivas desde las tradiciones discursivas*, editado por J. Kabatek. Frankfurt / Madrid: Iberoamericana / Vervuert.
- Comrie, Bernard. 1989. *Language universals and linguistic typology*. 2^a edición. Chicago: University of Chicago Press.
- Croft, William. 2003. *Typology and universals*. 2^a edición. Cambridge: Cambridge University Press.
- Cysouw, Michael. 2001. *The paradigmatic structure of person marking*. Tesis doctoral: Radboud University Nijmegen.
- Cysouw, Michael. 2013. «Inclusive/Exclusive Distinction in Independent Pronouns». en *The World Atlas of Language Structures Online*, editado por M. Dryer y M. Haspelmath. Leipzig: Max Planck Institute for Evolutionary Anthropology.
- D'Ambruso, Claudia. 2012. *Edición crítica y estudio de la Crónica troyana promovida*

- por Alfonso XI. Tesis doctoral: Universidade de Santiago de Compostela.
- Daniel, Michael. 2013. «Plurality in Independent Personal Pronouns». en *The World Atlas of Language Structures Online*, editado por M. Dryer y M. Haspelmath. Leipzig: Max Planck Institute for Evolutionary Anthropology.
- Denison, David. 2003. «Log(istic) and simplistic S-curves». Pp. 54-70 en *Motives for Language Change*, editado por R. Hickey. Cambridge: Cambridge University Press.
- Deyermond, Alan, y Francisco Rico. 1980. *Historia y crítica de la literatura española, 1. Edad Media*. Barcelona: Editorial Crítica.
- Dixon, Robert M. 1979. «Ergativity». *Language* 55:59-138.
- Dworkin, Steven N. 2011. «La variación y el cambio léxico: algunas consideraciones». Pp. 155-70 en *Así se van las lenguas variando: Nuevas tendencias en la investigación del cambio lingüístico en español*.
- Eberenz, Rolf. 2000. *El español en el otoño de la Edad Media*. Madrid: Gredos.
- Eberenz, Rolf. 2005. «Cambios morfosintácticos en la Baja Edad Media». Pp. 613-42 en *Historia de la Lengua Española*, editado por R. Cano. Barcelona: Ariel.
- Eisenberg, Daniel. 2004. *Introducción a «Espejo de Príncipes y Cavalleros [El Cavallero del Febo]» de Diego Ortúñez de Calahorra*. Alicante: Biblioteca Virtual Miguel de Cervantes.
- Elvira, Javier. 1998. *El cambio analógico*. Madrid: Gredos.
- Enguita Utrilla, José María, y Vicente Lagüéns Gracia. 2015. «El aragonés medieval en la “Grant Crónica de Espanya” a través de algunas muestras fonéticas y morfosintácticas». Pp. 1791-1812 en *Actas del IX Congreso Internacional de Historia de la Lengua Española (Cádiz, 2012)*, editado por J. M. García Martín, F. J. de Cos Ruiz, y M. Franco Figueroa. Iberoamericana: Vervuert.
- Enrique-Arias, Andrés. 2009. «Ventajas e inconvenientes del uso de Biblia medieval (un corpus paralelo y alineado de textos bíblicos) para la investigación en lingüística histórica del español». Pp. 269-83 en *Diacronía de las lenguas iberorrománicas: nuevas aportaciones desde la lingüística de corpus*, editado por A. Enrique-Arias. Frankfurt/Madrid: Vervuert/Iberoamericana.
- Enrique-Arias, Andrés. 2010a. «Introducción». Pp. 15-18 en *La Biblia Escorial I.I.6: transcripción y estudios*, editado por A. Enrique-Arias. Logroño: Cilengua. Centro Internacional de Investigación de la Lengua Española.
- Enrique-Arias, Andrés. 2010b. «La traducción del código Escorial I.I.6 en el contexto de los romanceamientos bíblicos medievales». Pp. 67-86 en *La Biblia Escorial I.I.6: transcripción y estudios*, editado por A. Enrique-Arias. Logroño: Cilengua. Centro Internacional de Investigación de la Lengua Española.
- Enrique-Arias, Andrés. 2010c. «On language contact as an inhibitor of language change: the Spanish of Catalan bilinguals in Majorca». Pp. 97-118 en *Continuity and change in grammar*, editado por A. Breitbarth, C. Lucas, S. Watts, y D. Willis. Filadelfia / Amsterdam: John Benjamins.
- Enrique-Arias, Andrés. 2012. «Dos problemas en el uso de corpus diacrónicos del español: perspectiva y comparabilidad». *Scriptum Digital* 1:85-106.
- Enrique-Arias, Andrés. 2014. «Efectos del contacto de lenguas en el castellano de Mallorca: una perspectiva histórica». Pp. 271-97 en *Perspectives in the study of Spanish language variation*, editado por A. Enrique-Arias, M. J. Gutiérrez, A. Landa, y F. Ocampo. Santiago de Compostela: Anexos de Verba, 72.
- Enrique-Arias, Andrés. 2018. «Factores diatópicos en la variación entre “este” y “aqueste” en la historia del español». Pp. 1553-69 en *Actas del X Congreso Internacional de Historia de la Lengua Española: Zaragoza, 7-11 de septiembre de 2015*, editado por M. L. Arnal Purroy, R. M. Castañer Martín, J. M. Enguita Utrilla,

- V. Lagüéns Gracia, y M. A. Martín Zorraquino.
- Enrique-Arias, Andrés. 2019a. «Los corpus informatizados aplicados al estudio del libro antiguo. Técnicas, recursos, problemas». Pp. 335-63 en *La producción del libro en la Edad Media: una visión interdisciplinar*, editado por G. Avenoz, L. Fernández Fernández, y M. L. Soriano Robles. Madrid: Sílex.
- Enrique-Arias, Andrés. 2019b. «Testing contact-induced change in the Spanish of Mallorca: Insights from a historical perspective». Pp. 168-83 en *Lusophone, Galician, and Hispanic Linguistics. Bridging Frames and Traditions*, editado por G. Rei-Doval y F. Tejedo-Herrero.
- Enrique-Arias, Andrés, Patricia Ribas Marí, y Marina Gomila Albal. En evaluación. «Algunos trasvases lingüísticos entre Castilla y el oriente peninsular en la baja Edad Media». en *La lengua española en las fuentes documentales del siglo XIII al XIX. Estudios sobre el corpus CODEA*, editado por B. Almeida Cabrejas y P. Sánchez-Prieto Borja.
- Faulhaber, Charles B. 1997. *PhiloBiblon*. Berkeley: Bancroft Library. University of California.
- Fernández-Ordóñez, Inés. 2001. «Hacia una dialectología histórica: reflexiones sobre la historia del leísmo, el laísmo y el loísmo». *Boletín de la Real Academia Española* 81(284):389-464.
- Fernández-Ordóñez, Inés. 2004. «Alfonso X el Sabio en la historia del español». Pp. 182-422 en *Historia de la lengua española*, editado por R. Cano. Barcelona: Ariel.
- Fernández-Ordóñez, Inés. 2006. «La historiografía medieval como fuente de datos lingüísticos: tradiciones consolidadas y rupturas necesarias». Pp. 1779-1808 en *Actas del VI Congreso Internacional de Historia de la Lengua española: Madrid, 29 de septiembre-3 octubre 2003*, editado por J. L. Girón Alconchel y J. J. de Bustos Tovar. Madrid: Arco Libros.
- Fernández-Ordóñez, Inés. 2009. «Manuscritos historiográficos “de autor”». Pp. 91-125 en *Los códices literarios de la Edad Media: interpretación, historia, técnicas y catalogación*, editado por P. M. Cátedra García, E. B. Carro Carbajal, y J. Durán Barceló.
- Fernández-Ordóñez, Inés. 2011. *La lengua de Castilla y la formación del español*. Madrid: Real Academia Española.
- Fernández-Ordóñez, Inés. 2012. *Transmisión y metamorfosis: hacia una tipología de mecanismos evolutivos en los textos medievales*. Salamanca: Seminario de Estudios Medievales y Renacentistas.
- Fuente Crespo, Josefa de la. 2000. *Colección documental del monasterio de Triano (1111-1520)*. León: Centro de estudios e investigación «San Isidoro».
- Gago Jover, Francisco. 2015. «La Biblioteca Digital de Textos del Español Antiguo (BiDTEA)». *Scriptum Digital* 4:5-36.
- Gago Jover, Francisco, y Francisco Javier Pueyo Mena. 2018a. «El Old Spanish Textual Archive, diseño y desarrollo de un corpus de textos medievales: el corpus textual». *Cuadernos del Instituto Historia de la Lengua* 11:165-209.
- Gago Jover, Francisco, y Francisco Javier Pueyo Mena. 2018b. «El Old Spanish Textual Archive, diseño y desarrollo de un corpus de textos medievales: lematización y etiquetado gramatical». *Scriptum Digital* 11:25-35.
- Garachana, Mar. 1997. *Los procesos de gramaticalización: una aplicación a los conectores contraargumentativos*. Tesis doctoral: Universitat de Barcelona.
- García Aragón, Lucía. 1985. *Documentación del monasterio de la Trinidad de Burgos 1198-1400*. Burgos: Garrido Garrido.
- García, Erica C. 1989. «La historia se repite con sigo». Pp. 113-24 en *Actas del IX*

- Congreso de la Asociación Internacional de Hispanistas 18-23 agosto 1986*, editado por S. Neumeister. Frankfurt am Main: Vervuert.
- García, Erica C. 1991. «Variación sincrónica y equivalencia diacrónica: el caso de -usco». *Cahiers de linguistique hispanique médiévale* 16:5-51.
- García, Erica C. 1997. «Convusco: a casualty of analogy, or of differential “(un)fitness”?» *Linguistics* 35:57-87.
- García, Erica C., Bob De Jonge, Dorien Nieuwenhuijsen, y C. Lechner. 1990. «(V)os- (otros): ¿dos y el mismo cambio?» *Nueva Revista de Filología Hispánica* 38(1):63-132. doi: 10.24201/NRFH.V38I1.774.
- Garrido Martín, Blanca. 2016. *La Atalaya de las corónicas (1443) de Alfonso Martínez de Toledo: edición crítica y estudio lingüístico*. Tesis doctoral: Universidad de Sevilla.
- van Gelderen, Elly. 2013. «The Linguistic Cycle and the Language Faculty». *Language and Linguistics Compass* 7(4):233-50. doi: 10.1111/lnc3.12017.
- Gili Gaya, Samuel. 1946. «Nos-otros, vos-otros». *Revista de Filología Española* 30:108-17.
- Gómez de Valenzuela, Manuel. 2017. *Platería y joyería en Zaragoza (1413-1513)*. Zaragoza: Institución «Fernando el Católico».
- Gomila Albal, Marina. 2016. «Sobre el origen y la difusión geográfica de las formas nosotros y vosotros en castellano». *Iberoromania* 2016(83):103-25.
- Gomila Albal, Marina. 2018. «Variación diacrónica y diatópica de con + pronombre personal de 1ª y 2ª persona del plural». *Bulletin of Hispanic Studies* 95(8):801-23.
- Gomila Albal, Marina. 2021. «Variación y cambio de los pronombres personales nós y vós junto con los modificadores otros, todos, mismos y ambos». Pp. 97-114 en *Variación diatópica y morfosintaxis en la historia del español (Anexo 82 de Verba)*, editado por B. Garrido Martín, M. del C. Moral del Hoyo, y M. Raab. Santiago de Compostela: Servizo de Publicacións e Intercambio Científico.
- Greenberg, Joseph H. 1966a. *Language universals: with special reference to feature hierarchies*. Berlin / New York: Mouton de Gruyter.
- Greenberg, Joseph H. 1966b. «Some universals of grammar with particular reference to the order of meaningful elements». Pp. 73-113 en *Universals of language*, editado por J. H. Greenberg. Cambridge: MIT Press.
- Heine, Bernd. 2003. «Grammaticalization». Pp. 575-601 en *The Handbook of Historical Linguistics*, editado por B. D. Joseph y R. D. Janda. Malden: Blackwell.
- Heine, Bernd, Ulrike Claudi, y Friederike Hünemeyer. 1991. *Grammaticalization: a conceptual framework*. Chicago: University of Chicago Press.
- Heine, Bernd, y Kyung-An Song. 2011. «On the grammaticalization of personal pronouns». *Journal of Linguistics* 47(3):587-630. doi: 10.1017/S0022226711000016.
- Hernández-Campoy, Juan Miguel, y Juan Camilo Conde-Silvestre. 2012. *The handbook of historical sociolinguistics*. Oxford: Wiley-Blackwell.
- Hock, Hans Henrick, y Brian D. Joseph. 2009. *Language History, Language Change, and Language Relationship*. 2ª edición. Berlin / New York: Mouton de Gruyter.
- Hopper, Paul J. 1991. «On some principles of grammaticization». Pp. 17-35 en *Approaches to grammaticalization. Volume I*, editado por E. C. Traugott y B. Heine. Amsterdam / Philadelphia: John Benjamins.
- Hopper, Paul J., y Elizabeth Closs Traugott. 2003. *Grammaticalization*. 2ª edición. Cambridge: Cambridge University Press.
- De Jonge, Bob, y Dorien Nieuwenhuijsen. 2009. «Formación del paradigma pronominal de las formas de tratamiento». Pp. 1593-1672 en *Sintaxis histórica de la lengua*

- española. Segunda parte, la frase nominal*, editado por C. Company Company. México: Fondo de Cultura Económica.
- Joseph, Brian D., y Richard D. Janda. 2003. *The Handbook of Historical Linguistics*. Malden: Blackwell.
- Kabatek, Johannes. 2013. «¿Es posible una lingüística histórica basada en un corpus representativo?» *Iberoromania* 77:8-28.
- Kabatek, Johannes. 2018. *Lingüística coseriana, lingüística histórica, tradiciones discursivas*. editado por C. Bleorțu y D. P. Gerards. Frankfurt / Madrid: Iberoamericana / Vervuert.
- Kahle, David, y Hadley Wickham. 2013. «ggmap: Spatial Visualization with ggplot2». *The R Journal* 5(1):144-61.
- Kawasaki, Yoshifumi. 2014. «Datación crono-geográfica de documentos medievales españoles». *Scriptum Digital* 3:29-63.
- Kuteva, Tania, Bernd Heine, Bo Hong, Haiping Long, Heiko Narrog, y Seongha Rhee. 2019. *World Lexicon of Grammaticalization*. 2ª edición. Cambridge: Cambridge University Press.
- Labov, William. 1972. «Some principles of linguistic methodology». *Language in Society* 1(1):97-120.
- Labov, William. 1994. *Principles of Linguistic Change. Volume 1: Internal Factors*. Oxford: Blackwell.
- Labov, William. 2001. *Principles of Linguistic Change. Volume 2: Social Factors*. Oxford: Blackwell.
- Lafoz Rabaza, Herminio. 2000. *Colección diplomática de Santa María la Mayor de Calatayud*. Zaragoza: Institución «Fernando el Católico».
- Lapesa, Rafael. 1981. *Historia de la lengua española*. Madrid: Gredos.
- Lausberg, Heinrich. 1973. *Lingüística románica*. Madrid: Gredos.
- Lehmann, Christian. 2015. *Thoughts on grammaticalization*. 3ª edición. Berlin: Language Science Press.
- Lenz, Rodolfo. 1925. *La oración y sus partes*. 2ª edición. Madrid: Centro de Estudios Históricos.
- Levshina, Natalia. 2015. *How to do Linguistics with R. Data exploration and statistical analysis*. Amsterdam / Philadelphia: John Benjamins.
- Libano Zumalacárregui, Ángeles. 1988. «De las formas medievales connusco, convusco a las modernas con nosotros con vosotros». Pp. 501-8 en *Actas del I Congreso Internacional de Historia de la Lengua Española*. Cáceres: Arco/Libros.
- Libano Zumalacárregui, Ángeles. 1991. «Morfología histórica del español: las formas de tratamiento». *Revista de Filología Española* 71:107-21.
- Lightfoot, David. 1991. *How to set parameters: arguments from language change*. Cambridge: The MIT Press.
- Lloyd, Paul M. 1987. *From Latin to Spanish*. Philadelphia: American Philosophical Society.
- Lloyd, Paul M. 1993. *Del latín al español. Vol. 1. Fonología y morfología históricas de la lengua española*. Madrid: Gredos.
- López Serena, Araceli. 2021. «La tradicionalidad discursiva como materia y las tradiciones discursivas como objeto de estudio». *Verba* 48. doi: 10.15304/verba.48.6864.
- Mariño Paz, Ramón. 2017. «Breve reflexión histórica sobre el cambio inducido por contacto castellanizante en galego». Pp. 47-62 en *Estudios sobre el cambio lingüístico no galego actual*, editado por X. L. Regueira y E. Fernández Rei. Santiago de Compostela: Consello da Cultura Galega.

- Matute, Cristina. 2013a. «Hacia una caracterización dialectal de la interpolación en el castellano de la Edad Media». Pp. 155-64 en *Actas del VIII Congreso Internacional de Lingüística y de Filología Románicas*, editado por E. Casanova y C. Calvo Rigual. Berlín: De Gruyter.
- Matute, Cristina. 2013b. «La variedad dialectal del castellano en la Edad Media a través del corpus digitalizado Biblia Medieval». Pp. 502-7 en *La Lingüística Aplicada en l'era de la globalització / La Lingüística aplicada en la era de la globalización / Applied Linguistics in the Age of Globalization*. Lleida: Edicions de la Universitat de Lleida.
- Matute, Cristina, y Enrique Pato. 2010. «Morfología y sintaxis en el código Escorial I.I.6». Pp. 45-66 en *La Biblia Escorial I.I.6: transcripción y estudios*, editado por A. Enrique-Arias. Logroño: Cilengua. Centro Internacional de Investigación de la Lengua Española.
- Meillet, Antoine. 1912. «L'évolution des formes grammaticales». Pp. 130-48 en *Linguistique historique et linguistique générale*, editado por A. Meillet. París: Champion.
- Menéndez Pidal, Ramón. 1919. *Documentos lingüísticos de España. Reino de Castilla*. Madrid: Centro de Estudios Históricos.
- Menéndez Pidal, Ramón. 1962. *Poesía juglaresca y juglares: aspectos de la historia literaria y cultural de España*. 7ª ed. Madrid: Espasa-Calpe.
- Menéndez Pidal, Ramón. 1982. *Manual de Gramática Histórica Española*. 17.ª ed. Madrid: Espasa-Calpe.
- Mettmann, Walter. 1996. *Alfonso de Valladolid, Mostrador de justicia*. Opladen: Westdeutscher Verlag.
- Meyer-Lübke, Wilhelm. 1974. *Grammaire des langues romanes*. Marseille: Laffitte Reprints.
- Meyer, David, Achim Zeileis, y Kurt Hornik. 2020. *vcd: Visualizing categorical Data*. Software: R package version 1.4-8.
- Miranda, Alberto. 1993. *Gutierre Díez de Games, El Victorial*. Madrid: Cátedra.
- Moral del Hoyo, María del Carmen. 2015. «Hacia una dialectología gramatical del castellano medieval: cuestiones morfológicas del imperfecto y futuro de subjuntivo». *Scriptum Digital* 4:143-64.
- Navarro Tomás, Tomás. 1957. *Documentos lingüísticos del alto Aragón*. New York: Syracuse University Press.
- Nebrija, Antonio de. 1492. *Gramática de la lengua castellana*. 3ª ed. editado por Estudio y edición de Antonio Quilis. Madrid: Centro de Estudios Ramón Areces.
- Nevalainen, Terttu. 2015. «Descriptive adequacy of the S-curve model in diachronic studies of language change». *Varieng. Studies in Variation, Contacts and Change in English* 16.
- Nevalainen, Terttu, y Helena Raumolin-Brunberg. 2016. *Historical Sociolinguistics: Language Change in Tudor and Stuart England*. New York: Routledge.
- Nowikow, Wiaczeslaw. 1994. *Sobre la pluralización de personas gramaticales en las lenguas románicas: «nos», «vos» - «nos alteros», «vos alteros»*. Universidad de Valladolid.
- Octavio de Toledo y Huerta, Álvaro S. 2016. *Los relacionantes locativos en la historia del español*. Berlín: De Gruyter.
- Octavio de Toledo y Huerta, Álvaro S. 2018a. «¿Tradiciones discursivas o tradicionalidad? ¿Gramaticalización o sintactización? Difusión y declive de las construcciones modales con infinitivo antepuesto». Pp. 79-134 en *Procesos de textualización y gramaticalización en la historia del español*, editado por J. L. Girón

- Alconchel, F. J. Herrero Ruiz de Loizaga, y D. M. Sáez Rivera. Madrid / Frankfurt am Main: Iberoamericana / Vervuert.
- Octavio de Toledo y Huerta, Álvaro S. 2018b. «De un occidentalismo cortesano y una transfusión fallida: historia de es(t)otro». *Estudios de Lingüística del Español* 39:305-61.
- Oesterreicher, Wulf. 2011. «Conquistas metodológicas en la lingüística diacrónica actual. La historicidad del lenguaje». Pp. 305-34 en *Así se van las lenguas variando: Nuevas tendencias en la investigación del cambio lingüístico en español*, editado por M. Castillo Lluch y L. Pons Rodríguez. Peter Lang.
- Parker, Kelvin M. 1977. *La versión de Alfonso XI del Roman de Troie. Ms. H-j-6 del Escorial. Critical Edition*. Illinois: Applied Literature Press.
- Penny, Ralph. 2004. *Variación y cambio en español*. Madrid: Gredos.
- Penny, Ralph. 2006. *Gramática histórica del español*. 2ª edición. Barcelona: Ariel.
- Pichel, Ricardo, y Xavier Varela Barreiro. 2009. «O corpus do galego medieval Tesouro Medieval Informatizado da Lingua Galega (TMILG)». Pp. 195-216 en *Diacronía de las lenguas iberorrománicas: nuevas aportaciones desde la lingüística de corpus*, editado por A. Enrique-Arias. Madrid / Frankfurt am Main: Iberoamericana / Vervuert.
- Pons Rodríguez, Lola. 2006. «Canon, edición de textos e historia de la lengua cuatrocentista». Pp. 69-125 en *Historia de la lengua y crítica textual*, editado por L. Pons Rodríguez. Madrid / Frankfurt am Main: Iberoamericana / Vervuert Verlag.
- Pons Rodríguez, Lola. 2008. *Álvaro de Luna, Virtuosas e claras mugeres (1446)*. Burgos: Instituto Castellano y Leonés de la Lengua.
- Posner, Rebecca. 1998. *Las lenguas romances*. New York: Syracuse University Press.
- Postma, Gertjan. 2010. «The impact of failed changes». Pp. 269-402 en *Continuity and change in grammar*, editado por A. Breitbarth, C. Lucas, S. Watts, y D. Willis. Amsterdam / Filadelfia: John Benjamins.
- QGIS, Development Team. 2021. *QGIS Geographic Information System*. QGIS Association: 3.16.8-Hannover.
- R, Core Team. 2021. *R: A Language and Environment for Statistical Computing*. Vienna: R Foundation for Statistical Computing.
- Ramírez Vaquero, Eloísa, Susana Herreros Lopetegui, y Marcelino Beroiz Lazcano. 2013. *El primer cartulario de los reyes de Navarra: el valor de lo escrito*. Pamplona: Gobierno de Navarra.
- Rey, Agapito. 1932. *Leomarte, Sumas de Historia Troyana*. Madrid: S. Aguirre.
- Ribas Marí, Patricia. 2022. *Las traducciones bíblicas en la Castilla medieval: caracterización lingüística desde una perspectiva dialectal*. Tesis doctoral: Universitat de les Illes Balears.
- Rini, Joel. 1990. «Excessive analogical change as an impetus for lexical loss: old spanish conusco, convusco». *Romanische Forschungen* 102(1):58-64.
- Rini, Joel. 1992. *Motives for linguistic change in the formation of the Spanish object pronouns*. Newark: Juan de la Cuesta.
- Rini, Joel. 1999. «The Rise and Fall of Old Spanish Y' all: vos todos vs. vos otros». Pp. 209-21 en *Essays in Hispanic Linguistics: Dedicated to Paul M. Lloyd*, editado por R. J. Blake, D. L. Ranson, y R. Wright. Newark: Juan de la Cuesta.
- Rodríguez Castellano, Lorenzo. 1952. «El pronombre personal en asturiano». *Boletín del Instituto de Estudios Asturianos* 6(15):119-30.
- Rodríguez Molina, Javier. 2010. *La gramaticalización de los tiempos compuestos en español antiguo: Cinco cambios diacrónicos*. Tesis doctoral: Universidad Autónoma de Madrid.

- Rodríguez Molina, Javier. 2018. «La estratigrafía de los manuscritos medievales castellanos: logros y perspectivas». *Medioevo romanzo* 42(1):93-127.
- Rodríguez Molina, Javier, y Álvaro S. Octavio de Toledo y Huerta. 2017a. «Acceso a la base de datos CORDEMÁFORO». *Scriptum Digital* 6.
- Rodríguez Molina, Javier, y Álvaro S. Octavio de Toledo y Huerta. 2017b. «La imprescindible distinción entre texto y testimonio: el CORDE y los criterios de fiabilidad lingüística». *Scriptum Digital* 6:5-68.
- Rojo, Guillermo. 2016. «Corpus textuales del español». Pp. 285-96 en *Enciclopedia Lingüística Hispánica*, editado por J. Gutiérrez Rexach. London: Routledge.
- RStudio, Team. 2021. *RStudio: Integrated Development Environment for R*. Boston: RStudio, PBC.
- Ruiz Antón, J. Carlos. 1998. *Curso de tipología lingüística: enfoque funcional*. Valencia: Universitat Jaume I.
- Sánchez-Prieto Borja, Pedro. 2011. *La edición de textos españoles medievales y clásicos. Criterios de presentación gráfica*. San Millán de la Cogolla: Cilengua.
- Sánchez-Prieto Borja, Pedro. 2012. «Desarrollo y explotación del corpus de documentos españoles anteriores a 1700 (CODEA)». *Scriptum Digital* 1:5-35.
- Sánchez-Prieto Borja, Pedro. 2016. «Tiempo y memoria en las fuentes documentales de la Edad Media». *Cuadernos del CEMYR* 24:11-30.
- Sánchez-Prieto Borja, Pedro, y María Jesús Torrens Álvarez. 2010. «Escorial I.I.6». Pp. 33-44 en *La Biblia Escorial I.I.6: transcripción y estudios*, editado por A. Enrique-Arias. Logroño: Cilengua. Centro Internacional de Investigación de la Lengua Española.
- Sánchez González de Herrero, María Nieves. 2008. «De Proprietatibus Rerum: versiones castellanas». *Cahiers de recherches médiévales* 16:349-66. doi: <https://doi.org/10.4000/crm.11072>.
- Sánchez Vicente, Andrea. 2020. «Cuestiones pendientes sobre la historia de *connusco*, *convusco* y su paradigma». Pp. 235-52 en *Studies in Hispano-Romance Historical Linguistics and Lexicography: A Tribute to John J. Nitti*, editado por S. Kania y C. Kauffeld. New York: Hispanic Seminary of Medieval Studies.
- Serrano Marín, Marina. 2018. *Estudio de la morfología verbal del español en fuentes documentales de los siglos XIII-XVI*. Tesis doctoral: Universidad de Alcalá.
- Siewierska, Anna. 2013. «Gender Distinctions in Independent Personal Pronouns». en *The World Atlas of Language Structures Online*, editado por M. Dryer y M. Haspelmath. Leipzig: Max Planck Institute for Evolutionary Anthropology.
- Silva-Corvalán, Carmen. 1994. *Language contact and change: Spanish in Los Angeles*. Oxford: Oxford University Press.
- Silva-Corvalán, Carmen, y Andrés Enrique-Arias. 2017. *Sociolingüística y pragmática del español*. 2ª edición. Washington, DC: Georgetown University Press.
- Spitzer, Leo. 1947. «Vosotros». *Revista de Filología Española* 21:170-75.
- Szpiech, Ryan. 2016. «Translating between the Lines: Medieval Polemic, Romance Bibles, and the Castilian Works of Abner of Burgos/Alfonso of Valladolid». *Medieval encounters* 22:113-39.
- Taberner, Cristina, Ricardo Pichel, y Andrés Enrique-Arias. 2020. «El español en contacto con otras lenguas peninsulares: los corpus diacrónicos de Galicia, Navarra y Mallorca». *Estudios de Lingüística del Español* 42:177-217.
- Thomason, Sarah G. 2001. *Language contact: an introduction*. Edinburgh: Edinburgh University Press.
- Thomason, Sarah G., y Daniel L. Everett. 2001. «Pronoun Borrowing». *Proceedings of the Annual Meeting of the Berkeley Linguistics Society* 27:301-16. doi:

- 10.3765/bls.v27i1.1107.
- Torrens Álvarez, María Jesús. 2007. *Evolución e historia de la lengua española*. Madrid: Arco/Libros.
- Torruella Casañas, Joan. 2009. «Bases científicas en la investigación a partir de corpus: el caso del Corpus Informatitzat del Català Antic». Pp. 95-116 en *Diacronía de las lenguas iberorrománicas: nuevas aportaciones desde la lingüística de corpus*, editado por A. Enrique-Arias. Madrid / Frankfurt am Main: Iberoamericana / Vervuert.
- Trudgill, Peter. 2011. *Sociolinguistic typology. Social determinants of linguistic complexity*. Oxford: Oxford University Press.
- Vázquez Cuesta, Pilar, y María Albertina Mendes Da Luz. 1987. *Gramática portuguesa*. I. 3ª. Madrid: Gredos.
- Weinreich, Uriel, William Labov, y Marvin I. Herzog. 1968. «Empirical foundations for a theory of language change». Pp. 95-195 en *Directions for historical linguistics. A Symposium*, editado por W. P. Lehmann y Y. Malkiel. Austin: University of Texas Press.
- Wickham, Hadley, Mara Averick, Jennifer Bryan, Winston Chang, Lucy D'Agostino McGowan, Romain François, Garrett Golemund, Alex Hayes, Lionel Henry, Jim Hester, Max Kuhn, Thomas Lin Pedersen, Evan Miller, Stephan Milton Bache, Kirill Müller, Jeroen Ooms, David Robinson, Dana Paige Seidel, Vitalie Spinu, Kobske Takahashi, Davis Vaughan, Claus Wilke, Kara Woo, y Hiroaki Yutani. 2019. «Welcome to the tidyverse». *Journal of Open Source Software* 4(43):1686. doi: 10.21105/joss.01686.
- Zamora Vicente, Alonso. 1970. *Dialectología española*. Madrid: Gredos.

Índice de figuras

Figura 1. Representación idealizada de una curva en S. Adaptada de Denison (2003:53)	32
Figura 2. Volumen de palabras entre los siglos XIII-XVI en Castilla central y oriental (CCO), León y Castilla occidental (LCO) y Navarra y Aragón (NyA) en el CODEA+.	65
Figura 3. Volumen de palabras por cada medio siglo en el CICA.	86
Figura 4. Volumen de palabras por tipología textual en el CICA.	87
Figura 5. Volumen de palabras en cada uno de los siglos y zonas geográficas en textos documentales (CODEA+) y literarios.	88
Figura 6. Volumen de palabras por dialecto en los textos documentales del CICA.	101
Figura 7. Volumen de ocurrencias por dialecto en la muestra aleatoria de los textos documentales del CICA.	101
Figura 8. Evolución de las variantes <i>nosotros</i> y <i>vosotros</i> en la muestra de los textos literarios y en el CODEA+.	105
Figura 9. Porcentaje de aparición de las formas reforzadas de primera y segunda persona del plural <i>nosotros</i> y <i>vosotros</i> en la muestra de textos literarios.	106
Figura 10. Porcentaje de aparición de las formas reforzadas de primera y segunda persona del plural <i>nosotros</i> y <i>vosotros</i> en el CODEA+.	107
Figura 11. Porcentaje de uso de las formas simples <i>nós</i> y <i>vós</i> con referente singular frente a las formas correspondientes con referente plural en el CODEA+.	108
Figura 12. Frecuencia de aparición y número de ocurrencias de las variantes simples y compuestas en posición de sujeto, en término de preposición o en otros contextos (textos literarios).	109
Figura 13. Frecuencia de aparición y número de ocurrencias de las variantes simples y compuestas en posición de sujeto, en término de preposición o en otros contextos (textos documentales).	109
Figura 14. Presencia de formas simples y compuestas según zona geográfica (textos literarios).	113
Figura 15. Difusión de los pronombres <i>nosotros</i> y <i>vosotros</i> desde el siglo XIII al XVI en los textos documentales del CODEA+ 2015, el CICA y el TMILG.	116
Figura 16. Porcentaje de aparición de las formas reforzadas <i>nosotros</i> y <i>vosotros</i> en los textos cancillerescos, eclesiásticos y particulares del CODEA+ por mitad de siglo.	119
Figura 17. Porcentaje de aparición de las formas reforzadas <i>nosotros</i> y <i>vosotros</i> en los textos cancillerescos, eclesiásticos y particulares del CODEA+ por siglo y por zona geográfica.	120
Figura 18. Frecuencia por cada 100.000 palabras de cada uno de los modificadores en el CODEA+ (siglos XIII-XVI).	130
Figura 19. Frecuencia por cada 100.000 palabras de cada uno de los modificadores en el corpus de textos literarios (siglos XIII-XVI).	130
Figura 20. Frecuencia por cada 100.000 palabras de cada uno de los modificadores junto con <i>nós</i> , <i>vós</i> , <i>nosotros</i> , <i>vosotros</i> en el corpus de textos literarios (siglos XIII-XVI)	132
Figura 21. Aparición de los pronombres con modificador en primera (<i>nós</i>) y segunda (<i>vós</i>) persona en el CODEA+.	134
Figura 22. Aparición de los pronombres con modificador en primera (<i>nós</i>) y segunda (<i>vós</i>) persona en el corpus de textos literarios.	134

Figura 23. Distribución de las variantes pronominales reforzadas por zona en los textos literarios (siglos XIII-XVI).....	136
Figura 24. Frecuencia por cada 100.000 palabras de cada uno de los modificadores en los textos literarios de León y Castilla occidental (siglos XIII-XVI).....	136
Figura 25. Aparición de las formas pronominales reforzadas en el TMILG y el CICA por mitad de siglo.....	138
Figura 26. Evolución de las tres variantes con referente plural en textos documentales desde el siglo XIII al XVI.....	151
Figura 27. Evolución de las tres variantes con referente plural en textos literarios desde el siglo XIII al XVI.....	152
Figura 28. Porcentaje de aparición y ocurrencias de cada una de las variantes de primera y segunda persona en textos documentales de los siglos XIII al XVI.....	154
Figura 29. Porcentaje de aparición y ocurrencias de cada una de las variantes de primera y segunda persona en textos literarios de los siglos XIII al XVI.....	154
Figura 30. Ocurrencias y porcentajes de aparición de cada una de las variantes en los verbos más frecuentes del corpus de textos literarios (siglos XIII-XVI).	156
Figura 31. Ocurrencias y porcentaje de aparición de las variantes tras <i>con</i> en cada una de las zonas geográficas en textos literarios (siglos XIV-XVI).	158

Índice de tablas

Tabla 1. Resumen de obras y fuentes del corpus de textos literarios.	85
Tabla 2. Número de palabras, peso en el corpus y muestra aleatoria para cada texto literario.	100
Tabla 3. Ejemplos extraídos aleatoriamente de los textos documentales del CICA y número de ocurrencias de pronombres con referente plural (<i>nós, vós, nosaltres, vosaltres</i>).	102
Tabla 4. Frecuencias de <i>nós, vós, nosotros</i> y <i>vosotros</i> desde el siglo XIII al XVI en la muestra de textos literarios (N = 1933).	104
Tabla 5. Frecuencias de <i>nós, vós, nosotros</i> y <i>vosotros</i> desde 1251 a 1550 en CODEA+ (N = 5005).	104
Tabla 6. Frecuencia de las variantes simples y compuestas en posición de sujeto y en término de preposición en textos literarios.	110
Tabla 7. Recuentos de las variantes simples, de las reforzadas con <i>otros</i> y de las reforzadas con otros modificadores en la muestra del corpus literario.	126
Tabla 8. Recuento de las variantes reforzadas en el corpus completo de textos literarios.	127
Tabla 9. Ilustración de la regla de tres utilizada para estimar el número de formas simples del corpus literario completo a partir de la muestra.	127
Tabla 10. Estimación de las variantes <i>nós</i> y <i>vós</i> con referente plural del corpus de textos literarios a partir de la muestra.	127
Tabla 11. Ocurrencias y porcentajes de las variantes con y sin modificador en el CODEA+.	129
Tabla 12. Ocurrencias y porcentajes de las variantes con y sin modificador en el corpus de textos literarios.	129
Tabla 13. Porcentaje de aparición de <i>nós todos</i> y <i>vós todos</i> en comparación con el resto de formas reforzadas en el TMILG, el CODEA+ y el CICA.	138
Tabla 14. Resumen de ejemplos con referente plural extraídos del corpus ampliado de textos documentales según fuente.	148
Tabla 15. Resumen de porcentajes y ocurrencias de cada una de las variantes en los testimonios del corpus de textos literarios.	149
Tabla 16. Frecuencias de cada una de las variantes en primera y segunda persona (textos documentales).	155
Tabla 17. Frecuencias de cada una de las variantes en primera y segunda persona (textos literarios).	155

Índice de mapas

Mapa 1. Propuesta de categorización del <i>continuum</i> norte peninsular en 7 franjas de norte a sur.....	62
Mapa 2. Localización de las formas compuestas <i>nosotros</i> y <i>vosotros</i> en los siglos XIII-XVI en los textos documentales del CODEA+ .	114
Mapa 3. Presencia de formas compuestas en los documentos notariales del CODEA+, el CICA y el TMILG desde el siglo XIII al XVI.....	117
Mapa 4. Localización de cada una de las variantes pronominales reforzadas en el CODEA+ (siglos XIII-XV).	135
Mapa 5. Distribución de las variantes <i>connusco/convusco</i> , <i>con nós/con vós</i> y <i>con nosotros/con vosotros</i> en los siglos XIII-XV (textos documentales).	157
Mapa 6. Realizaciones de las variantes de <i>connigo</i> en el ALEANR.....	160
Mapa 7. Documentaciones de las variantes <i>connusco</i> y <i>convusco</i> con referente singular en el corpus de textos documentales.....	160