

Soy, doy, estoy y voy: la yod desinencial en el corpus CODEA+ 2015¹

Soy, doy, estoy and voy: Desinential Yod in Corpus CODEA+ 2015

Marina Serrano Marín

GITHE-Universidad de Alcalá, España

Resumen: Desde una perspectiva que aúna diferentes disciplinas lingüísticas, se investiga la variación consustancial al cambio y la selección que afectaron a las variables de primera persona de singular de los verbos *ser*, *estar*, *dar* e *ir* en documentos notariales de los siglos XIII al XVI en el corpus CODEA+ 2015. El objeto de este análisis es determinar si las variables morfológicas verbales con yod desinencial aparecieron simultáneamente en todo el territorio de habla castellana y bilingüe de la Península o si, por el contrario, presentaron una frecuencia de aparición y una distribución cronológica, geográfica y diastrática diferentes. Se proporcionan datos empíricos cuantificables y argumentos teóricos que permiten, por una parte, establecer una periodización más ajustada del fenómeno analizado y por otra, ofrece una explicación en la que el componente geográfico y el diafásico desempeñan un papel fundamental en la reconstrucción de la variación histórica de la lengua.

Palabras clave: morfología verbal histórica, sociolingüística histórica, castellano medieval, corpus lingüísticos, CODEA+ 2015.

Abstract: From a perspective which combines different linguistic disciplines, this article researches the variation which is naturally determined by the linguistic change and the selection which have affected the 1st singular person of the present indicative variables of the verbs *ser*, *estar*, *dar* and *ir*. The key question which has given rise to this research lies in whether the verbal morphological variables with desinential yod, appeared simultaneously all over the Spanish-speaking and bilingual territory of the Peninsula, or if, on the contrary, they showed a different frequency of appearance and a different chronological, geographic and diastratic distribution. This analysis provides quantifiable empirical data and theoretical arguments which, on the one hand, let us establish a more accurate periodization of the studied phenomenon than the one provided up to now, and, on the other hand, they let us give an explanation in which the geographical and diaphasic components play an essential role in the reconstruction of the historical variation of the language.

Keywords: historical verbal morphology, Historical Sociolinguistics, castellano medieval, Linguistic corpora, CODEA+ 2015.

¹ Este trabajo es resultado de la tesis doctoral defendida en mayo de 2018 en la Universidad de Alcalá, realizada con la ayuda predoctoral MINECO BES-2013-066654, la cual se enmarca dentro del proyecto "Corpus de Documentos Españoles Anteriores a 1800 (CODEA+2015)", con referencia FFI2012-33646.

Introducción

Los hablantes de una lengua hacen uso de ella de forma variable, es decir, pueden comunicar significados diferentes y emplear elementos lingüísticos distintos. Esa manifestación variable de la lengua permite igualmente comunicar en determinadas circunstancias el mismo concepto empleando unos elementos lingüísticos en lugar de otros sin que ello produzca un cambio de significado. En dichos casos nos encontraríamos ante fenómenos de *variación lingüística* o *polimorfismo* (Radtke, 2002, p. 49).

La variación, que es consustancial al *cambio lingüístico* (Selig, 2008, p. 72), puede ser fonético-fonológica, léxica, morfosintáctica y/o discursiva. Es capaz de afectar a una parte o a la totalidad de una *variedad lingüística*, es decir, al conjunto de patrones lingüísticos homogéneo con un repertorio lo suficientemente extenso como para ser utilizado en todos los contextos de comunicación por los hablantes de una comunidad de habla. Dicha comunidad se define como el conjunto de hablantes que comparten al menos una lengua, pero también los usos y normas de carácter sociolingüístico, así como las creencias, las actitudes y el prestigio lingüísticos (Moreno Fernández, 2005, p. 23).

El cambio lingüístico, el cual ha de situarse en la dimensión del habla para que su inicio y difusión sean comprensibles, posee una doble naturaleza, la cual ya fue señalada en 1916 por Menéndez Pidal, quien anunciaba sin proponérselo una de las principales hipótesis de la sociolingüística variacionista:

Claro es que toda creación en los productos sociales o colectivos (tales como la poesía tradicional, el lenguaje o la costumbre) es obra de un individuo que en un momento de iniciativa individual se eleva sobre el nivel común de las gentes (Menéndez Pidal, 1916, p. 272).

En primer lugar, el cambio es individual. La actuación verbal de un hablante y el uso innovador que este realiza de una unidad lingüística puede desencadenar el cambio en el seno de una comunidad lingüística (Mendívil Giró, 2010, p. 118; Garatea Grau, 2016, pp. 26-28). Ahora bien, pese a que el cambio es un fenómeno resultante de las acciones humanas, no por ello tiene un carácter teleológico, ya que el cambio no es una meta de las intenciones de los hablantes (Keller, 1985, pp. 216-217).

En segundo lugar, es colectivo en el momento en el que dicha innovación experimenta una difusión en la comunidad lingüística a la que pertenece el hablante y esta adopta el nuevo uso, pues "sin comunidad lingüística no hay cambio lingüístico imaginable" (Martín Butragueño, 1997, p. 41). En la difusión del cambio entran en juego aspectos sociales como el prestigio o la influencia del hablante (Labov, 1966), por ello también puede considerarse colectivo en tanto en cuanto su inicio y no su difusión depende del uso concreto que haga un grupo social, una comunidad de habla, etc., y no un hablante particular (Romaine, 1988, p. 357).

Si atendemos a las circunstancias en las que cambia la lengua, podemos determinar dos tipos de cambio lingüístico: cambio desde arriba y cambio desde abajo (Labov, 1966). Un ejemplo de cambio desde arriba lo observamos en el sufijo *-ísimo*, el cual fue introducido por las clases letradas en la comunidad de habla. Las clases no letradas imitaron ese patrón lingüístico por el prestigio social que le otorgaron, hasta el punto de que hoy día el sufijo *-ísimo* pertenece a un registro menos formal que al que pertenecía originalmente (Gimeno Menéndez, 1990, pp. 89-

102, Moreno Fernández, 2005, pp. 177-190, De Toledo y Huerta/Pons Rodríguez, 2016, pp. 63-103). El cambio desde abajo sería aquel en el que un patrón lingüístico empleado por las clases menos letradas es aceptado y adoptado por las clases socioculturales más altas como consecuencia de la actitud positiva que experimentan hacia el nuevo fenómeno lingüístico. Por tanto, el reflejo textual de la lengua en los documentos va a estar ligado a la aceptación o rechazo que los hablantes le concedan a un fenómeno de variación concreto. Un ejemplo que ilustra un tipo de cambio de abajo a arriba es el que se observa en la difusión de las variantes de la 1ª persona de singular de presente de indicativo con partícula –y de los verbos *ser*, *estar*, *dar* e *ir*.

1. Breve revisión de las hipótesis sobre la yod desinencial

La primera persona de singular de los verbos *ser*, *estar*, *dar* e *ir* era etimológicamente monosilábica (ver Tabla 1).

Tabla 1. Formas de presente: morfos etimológicos y morfos en –y.

VERBO	FORMA ETIMOLÓGICA	FORMA ACTUAL
<i>se(e)r</i>	Só	soy
<i>dar</i>	Dó	doy
<i>estar</i>	Estó	estoy
<i>ir</i>	Vó	voy

Entre las diferentes hipótesis que se han postulado para explicar el *incremento palatal* (Martínez-Gil, 2012, p. 935), esto es, la adición de la vocal no etimológica –y a la 1ª persona de singular de estos verbos encontramos principalmente tres². En primer lugar, autores como Ford y Spaulding (Gallego de la Puente, 2007, p. 383) consideran que –y procede del pronombre tónico de primera persona yo (*so yo > soy*, *do yo > doy*)³. De modo que formas como *soy* o *doy* procederían de la aglutinación entre los morfos verbales *só* y *dó* y el pronombre en secuencias interrogativas o en enunciativas en las que la estructura informativa se encuentra topicalizada (“e estas arras vos *dó yo* per foro de León”, CODEA-0227). En segundo lugar, la –y procedería de una forma previa con –e paragógica final (i.e. *soe*), dando lugar a un hiato que se habría resuelto posteriormente en diptongo con el cierre de la vocal final (*soe > soi*) (Lloyd, 1993, p. 565). En tercer lugar, Staaff y Corominas identifican la –y con la partícula *ý*, procedente del adverbio de lugar latino *IBĪ* (De Gorog, 1980, p. 159). Esta última hipótesis es la que considero más probable porque este mismo fenómeno se dio en otras lenguas romances como el provenzal, en el que la 3ª persona de singular del verbo *ir*, *vai*, procede de *va i < VĀDIT IBĪ*, y la 3ª persona de singular del pretérito *fon* ('fue') es la forma evolucionada de *fo'n < fo* en, formado a partir de *FUIT INDE* (Lloyd,

² Martínez-Gil (2012) plantea una cuarta hipótesis basada en la noción de minimidad prosódica, según la cual la adición de –y en *soy*, *doy*, *estoy* y *voy* vendría a satisfacer el pie métrico con esquema de troqueo moraicó (v. De Gorog, 1980; Harris, 1983; Hayes, 1989).

³ Sobre la posición del sujeto en castellano medieval, v. Neumann-Holzschuh (1996).

1993, pp. 567-568). Es más, en friulano encontramos las mismas formas verbales para estos verbos: *sòi, dòi, estòi* y *vòi* (Gregor, 1975, pp. 102-103, citado en Lloyd, 1993, p. 570).

Las otras dos hipótesis han sido desestimadas, pues a la luz de los datos que ofrece el corpus, solamente se da un caso en el que el pronombre de 1ª persona aparece pospuesto al verbo conjugado en 1ª persona de singular de presente de indicativo. Si el orden VS en el corpus es testimonial, la posible aglutinación del verbo con el pronombre es inexistente. Además, la aglutinación del pronombre con el verbo supondría la pérdida de la *-o* final de *yo*, de lo que resultaría un sonido palatal fricativo sin apoyo vocálico para su articulación. A pesar de que existen lenguas como el húngaro cuyas leyes fonotácticas permiten este tipo de sonidos en posición implosiva, o como el retorromance, en el que el pronombre de 1ª persona de singular puede presentar una forma apocopada (Müller, 1973), en español no es posible. En cuanto a la posibilidad de que las formas con *-y* provengan de unas anteriores con *-e* paragógica, el corpus corroboraría las palabras de Lloyd (1993, p. 566) cuando considera *soy* como la forma originaria y *soe* como un derivado de esta. La primera data de *soy* en el corpus es del año 1244; la de *soe*, de 1256.

2. El corpus

Para el estudio del incremento palatal no etimológico en la 1ª persona de singular de los verbos mencionados anteriormente hemos establecido un corpus que abarca desde el siglo XIII hasta el siglo XVI. Nos hemos basado en los 1641 documentos del corpus CODEA+ 2015 comprendidos en estos siglos (ver Tabla 2), lo que se traduce en un total de 8543 formas diferentes.

Tabla 2. Distribución de piezas documentales por siglo y ámbito de emisión.

ÁMBITO SIGLO	CANCIL	JUD	MUN	ECLES	PART	TOTAL
XIII	117	11	9	236	117	490
XIV	57	10	8	131	99	305
XV	57	12	15	109	66	259
XVI	78	93	52	45	319	587
TOTAL	309	126	84	521	601	1641

CANCIL (cancilleresco), JUD (judicial), MUN (municipal), ECLES (eclesiástico), PART (particular).

En la tabla que presentamos a continuación podemos observar que los cuatro siglos estudiados no cuentan con el mismo número de documentos. Aunque todo corpus aspira a ser equilibrado en sus distintos ejes, esto es, el eje diacrónico, el eje geográfico y el eje tipológico, esta es una tarea harto complicada, ya que el carácter histórico de CODEA+ 2015 y, por consiguiente, del corpus de esta investigación, dificulta el establecimiento de una correlación proporcional para cada siglo,

territorio⁴ y tipología documental (Torruella, 2016, p. 93). A pesar de ello, el corpus cumple dos premisas fundamentales para poder realizar análisis diacrónicos: es un corpus cerrado y no es manipulable (Company, 2005, pp. 144-163).

Tabla 3. Tabla resumen de las piezas documentales del corpus por siglo/provincia.

SIGLO PROVINCIA	XIII	XIV	XV	XVI	TOTAL
Álava	3	1	0	2	6
Albacete	0	0	1	3	4
Alicante	0	0	1	0	1
Almería	0	0	0	1	1
Asturias	30	12	4	0	46
Ávila	34	25	7	6	72
Badajoz	0	2	2	4	8
Barcelona	0	0	1	0	1
Burgos	41	12	8	15	76
Cáceres	3	13	4	6	26
Cádiz	0	0	0	16	16
Cantabria	20	13	10	7	50
Castellón	1	0	0	0	1
Córdoba	4	0	2	4	10
Cuenca	3	1	3	5	12
Gerona	0	0	0	1	1
Granada	0	0	2	10	12
Guadalajara	5	14	15	26	60
Guipúzcoa	1	0	1	5	7
Huelva	0	0	0	4	4
Huesca	20	16	4	1	41
Jaén	1	1	4	2	8

⁴ A pesar de que en los siglos estudiados no existía una división política del territorio por provincias, fue necesario adoptar esta distribución para que las herramientas de búsqueda del corpus pudieran recuperar los datos de los documentos y proyectarlos en mapas, tablas y gráficos.

La Rioja	27	6	9	4	46
León	49	7	9	12	77
Lugo	2	0	0	0	2
Madrid	2	12	9	261	284
Málaga	0	0	0	4	4
Murcia	1	1	1	3	6
Navarra	35	34	4	5	78
Orense	1	0	0	0	1
Palencia	46	0	12	2	60
Salamanca	12	26	19	6	63
Segovia	15	3	8	7	33
Sevilla	29	9	3	29	70
Soria	1	3	2	3	9
Teruel	2	15	25	2	44
Toledo	30	12	18	40	100
Valencia	0	0	1	0	1
Valladolid	43	22	23	38	126
Vizcaya	0	1	0	5	6
Zamora	10	8	10	4	32
Zaragoza	8	30	29	9	76
Europa	0	0	0	16	16
s.l. (sin lugar)	20	3	4	21	48
TOTAL	499	292	251	599	1641

3. Resultados

3.1. Distribución cronológica y geográfica

A la luz de los datos obtenidos en el corpus, se corroboraría la afirmación de Pensado acerca de la data cronológica, ya que el primer morfo de la 1ª persona de singular de presente de indicativo se registra en un documento del año 1227, presumiblemente de Teruel (1).

[...] e a todo el concello, que a frere Roderici o a Pero Ferrer, mi fillo, o Toda Petri o Teresa Petri poder no les doy en el Castiello de Santa María ni en nul logar de los otros (CODEA-1223, Teruel, 1227).



Mapa 1. Distribución geográfica de las variantes etimológicas y de las variantes con -y en el siglo XIII.

En cuanto a la primera data en el corpus de los otros verbos, soy⁵ aparece por primera vez en un documento de Salamanca de 1244 (2). En 1492 se registra estoy en un documento de Cuenca (3). Por último, voy no se documenta hasta 1581 en un documento emitido en Nápoles (4).

(2.) soy fiador e manero de defender el cabildo con esta heredade (CODEA-0509, Salamanca, 1244).

(3.) yo estoy presta & aparejada de fa<z>er & complir en sastifacion de mis pecados toda la penitencia que por vuestras reverencias me fuere puesta (CODEA-2498, Cuenca, 1492).

(4.) y aunque lo que voy a decir sé que no lleva mucho camino toda vía lo diré como a señora mía (CODEA-2276, Nápoles, 1581).

Tabla 4. Distribución por siglos de doy, soy, estoy y voy.

SIGLO	Nº DE FORMAS	Nº DOCUMENTOS
XIII	12	9
XIV	9	5
XV	41	23
XVI	208	122
XVII	266	122

⁵ La 2ª persona de plural sois no se habría formado sobre la 1ª persona de singular soy, sino que sería resultado de la evolución fonética sodes > sois (v. Rini, 1996).

El gráfico 1 muestra cómo a partir de 1327, aumenta la frecuencia de uso de los morfos con *-y*, lo que supone la frecuencia absoluta más alta en el segundo cuarto del siglo XVI. De 12 formas registradas en 9 documentos en el siglo XIII, pasamos a 9 casos recogidos en 5 documentos en el siglo XIV. Sin embargo, en el siglo XV quedan documentadas 41 en 23 documentos, 208 en 122 piezas en el siglo XVI y 266 en el XVII igualmente en 122 piezas documentales.

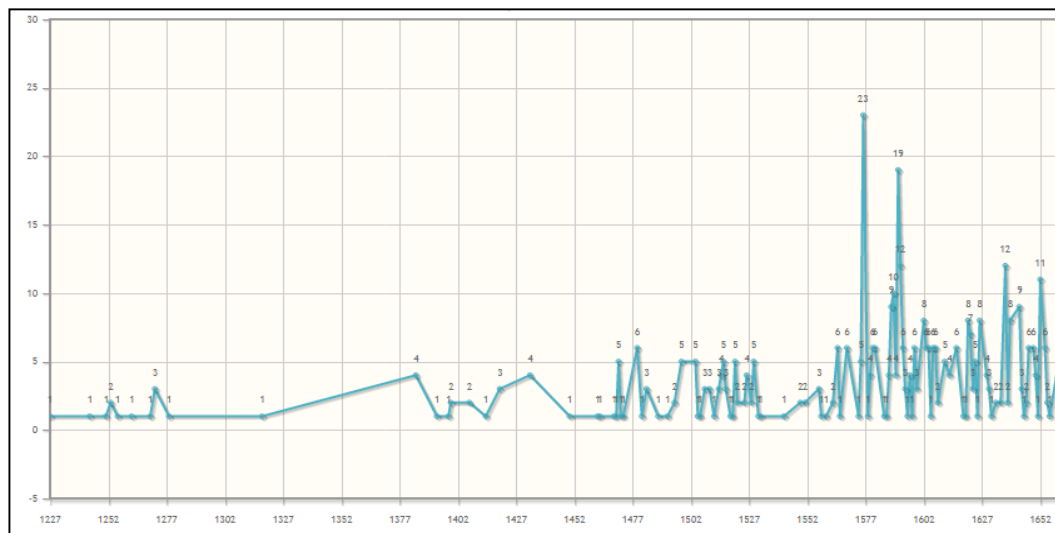


Gráfico 1. Distribución por años de *doy*, *soy*, *estoy* y *voy*.

Esta progresión, especialmente exponencial a partir del siglo XV, ha de contrastarse con la frecuencia de uso de los morfos etimológicos *só*, *estó*, *dó* y *vó*.



Mapa 2. Distribución geográfica de las variantes etimológicas y de las variantes con *-y* en el siglo XIV.

El corpus recupera 262 formas de *só* desde el año 1208 hasta el año 1520. En este año se registra *só* por última vez en un documento toledano de 1520 (5), pero este se trata de una copia de un documento del siglo XIII. En 1502 *só* aparece al final de una

carta de compraventa en la parte correspondiente a la fórmula del escribano (6). Hemos de remontarnos a una carta de compraventa oscense datada en 1500 para hallar un uso pleno del verbo. El único ejemplo en el texto se encuentra en una parte formulaica del documento (7).

(5.) e otorgo que só pagado d'ellos (CODEA-0291, Toledo, 1520).

(6.) la fiz escribir y fiz aquí mio signo e só testigo (CODEA-1006, Cádiz, 1502).

(7.) E yo dito Joán de Tafalla, qui present só, la dita fiancería fago e atorgo siquiere tal fiança de salvo de la dita donación (CODEA-0943, Huesca, 1500).

Tal y como se observa en el gráfico 2, es en el paso del siglo XIII al XIV cuando se registran un mayor número de casos. A partir del siglo XV decrecen los ejemplos hasta contabilizar un ejemplo en tres documentos distintos en el siglo XVI. No se documenta ningún caso en el siglo XVII.

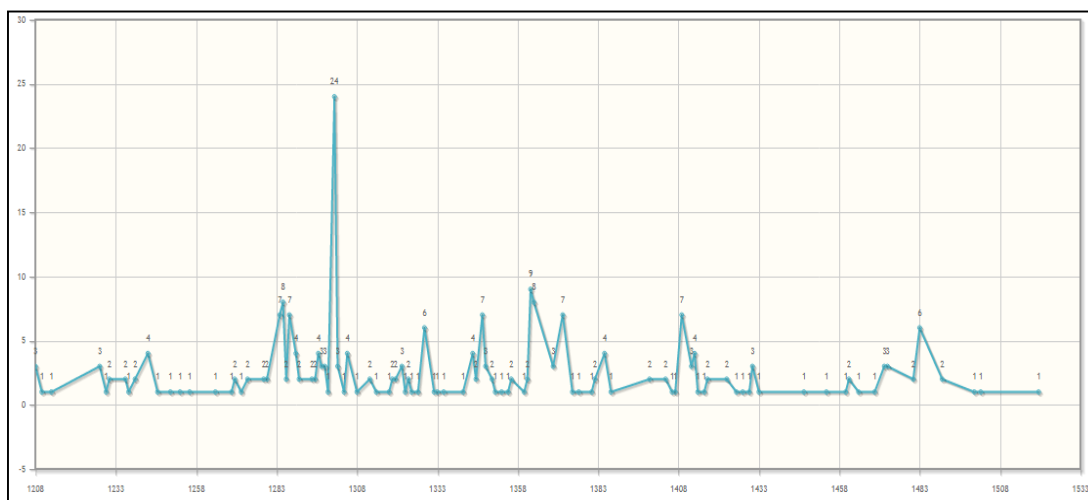


Gráfico 2. Distribución por años de só.

En segundo lugar, del morfo *estó* se documentan cuatro casos: uno, en 1312 (8); uno, en 1419 (9); uno, en 1525 (10) y uno, en 1569 (11).

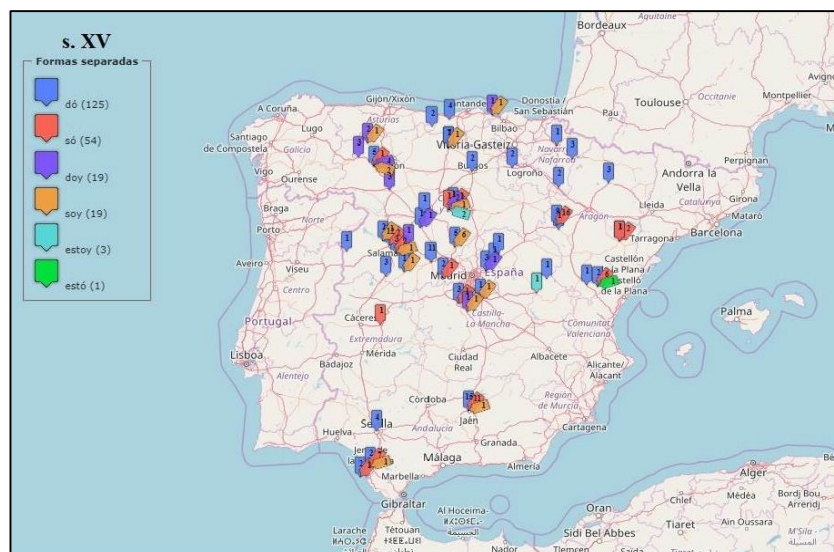
(8.) Sía coneçuda cosa a todos omnes cómo yo Guillerma Davicenda, que *estó* en Osca (CODEA-0946, Huesca, 1312).

(9.) e sobre unas casas mías que yo é, sitiadas en el dito lugar de Ruvihuelos, en las cuales yo agora *estó* y habito (CODEA-0794, Teruel, 1419).

(10.) sobre lo cuala pido complimiento de justicia, & para ello vuestro noble <...> imploro, & *estó* presto de dar información de lo sobredicho (CODEA-2497, Cuenca, 1525).

(11.) E yo el dicho Aparicio de Buendía, que presente *estó* a todo lo que dicho es (CODEA-0176, Guadalajara, 1569).

Pese a que su frecuencia de aparición es una de las más bajas de sus morfos homólogos, su desaparición y sustitución por *estoy* tiene lugar a partir del segundo cuarto del siglo XVI.



Mapa 3. Distribución geográfica de las variantes etimológicas y de las variantes con -y en el siglo XV.

En tercer lugar, se registran 417 casos de *dó* distribuidos en 230 documentos a lo largo del período comprendido entre los años 1213 (12) y 1515 (13). Su mayor frecuencia de uso se da a lo largo del siglo XIII.

(12.) E estas arras vos *dó* yo per foro de León (CODEA-0227, Valladolid, 1213).

(13.) ca para todo lo susodicho e cada cosa e parte d'ello *dó* e otorgo el dicho mi poder cumplido a vós (CODEA-0174, Guadalajara, 1515).

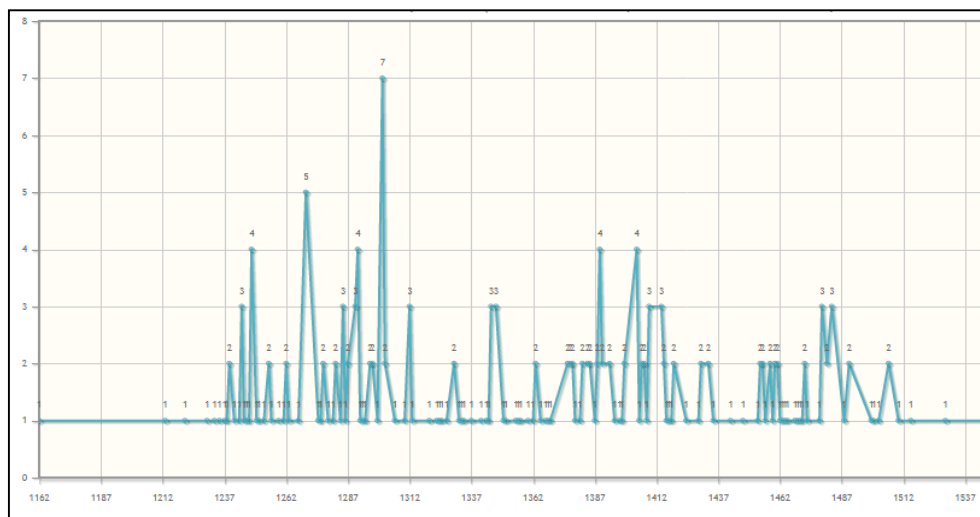
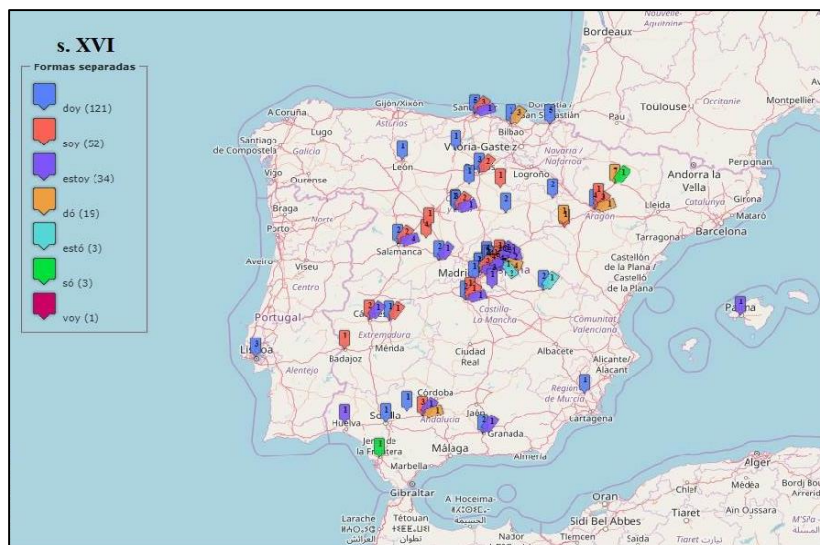


Gráfico 3. Distribución por años de *dó*.

En cuarto y último lugar, no se documenta ningún caso de *vó*.



Mapa 4. Distribución geográfica de las variantes etimológicas y de las variantes con -y en el siglo XVI.

3.1.1. Otras variantes

Junto a las variantes mencionadas hasta el momento, hemos de prestar igualmente atención a otras marcadas diatópicamente. Tal es el caso de las variantes asturleoneras *estoe/estou*, *soe/sou*, *doe/dou* y *vou* (Gallego de la Puente, 2007, p. 384). De todas ellas solo se documentan en el corpus las formas *soe* (14-16) y *dou* (17-21).

(14.) de que yo *soe* ben pagada (CODEA-0440, Asturias, 1256).

(15.) por XV morabedís de moneda leonés, de los cuales morabedís yo *soe* bien pagada (CODEA-0569, Asturias, 1263).

(16.) de aquellos a que yo *soe* tenido de rogar (CODEA-0663, Asturias, 1395).

(17.) yo don Pedro Ponz por atal pleito vos *dou* estas arras e vos vendo esta heredat (CODEA-0415, León, 1235).

(18a.) you García Martínez, taullero de Colombranos, *dou* e outorgo quanto derecho ey enos mulinus (CODEA-0428, León, 1274).

(18b.) Así estos mulinus vos *dou* e vos entergo (CODEA-0428, León, 1274).

(19.) yo Diego Martinez, prior d'este dicho monasterio *dou* a vós el dicho Arias Ferrández con esta renta otra casa mía (CODEA-0461, León, 1400).

(20.) e tirando una viña que *dou* al convento del dicho monasterio (CODEA-0462, León, 1399).

(21a.) cómo eu don Alfonso Meéndez de Bornes, *dou* e outorgo a don Pedro, abat de Morerola, e al convento d'es mismo lugar quanto ei e devo aver (CODEA-1234, Zamora, 1255).

(21b.) todo lo *dou* e outorgo livremiente e quitamente al abat (CODEA-1234, Zamora, 1255).

(21c.) e la *dou* e la entrego a Morerola como desuso é dicho (CODEA-1234, Zamora, 1255).

La concentración geográfica de estos morfos se da principalmente en León, provincia en la que se recogen 5 casos (17-20). En Zamora encontramos 3 casos (21a-21c), igual que en Asturias (14-16).



Mapa 5. Distribución geográfica de los morfos *soe* y *dou*.

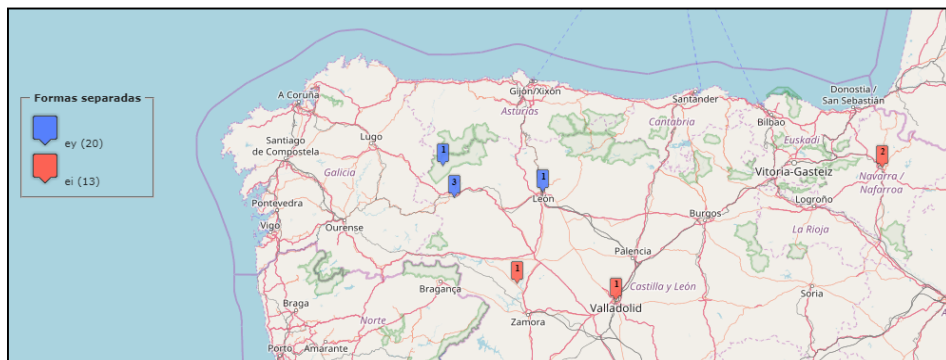
El ejemplo 14 de *soe* se documenta en una carta de compraventa de 1256 emitida presumiblemente en Asturias, al igual que el ejemplo 15 del año 1263. Ambos forman parte de la construcción de carácter formulaico [SER + bien pagado]⁶. El ejemplo 16, por el contrario, pertenece a una estructura intransitiva de obligación que no ha tenido continuación en la lengua.

En lo que al morfo *dou* se refiere, los ocho casos se reparten entre una carta de arras, una carta de arrendamiento y tres cartas de donación. Siempre aparece en la parte formulaica del documento en la que se manifiesta quién realiza la donación y/o otorgamiento y quién va a ser el receptor de la misma. Su carácter formulaico puede ser una de las razones de su mayor pervivencia en los documentos, ya que el último caso registrado data de 1400 (19).

A pesar de considerar que no existe ninguna relación entre las variantes con la partícula *-y* y el morfo *hey* de la 1ª persona de singular de presente del verbo haber, creemos oportuno dedicarle unas líneas.

Del mismo modo que *soe* y *dou* están marcadas diatópicamente para la zona leonesa, *hey* y sus variantes grafemáticas *ey* y *ei* se adscriben igualmente a áreas de influjo asturleonés. Se han encontrado dos ejemplos en Navarra.

⁶ *Ser bien pagado* es uno de los diferentes sintagmas libres que nos podemos encontrar en las piezas documentales, cuyo uso recurrente ha determinado el orden de las palabras que lo forman, pero sin poder considerarlo colocación o unidad fraseológica en sentido pleno (Cuadros Muñoz, 2011, p. 205).



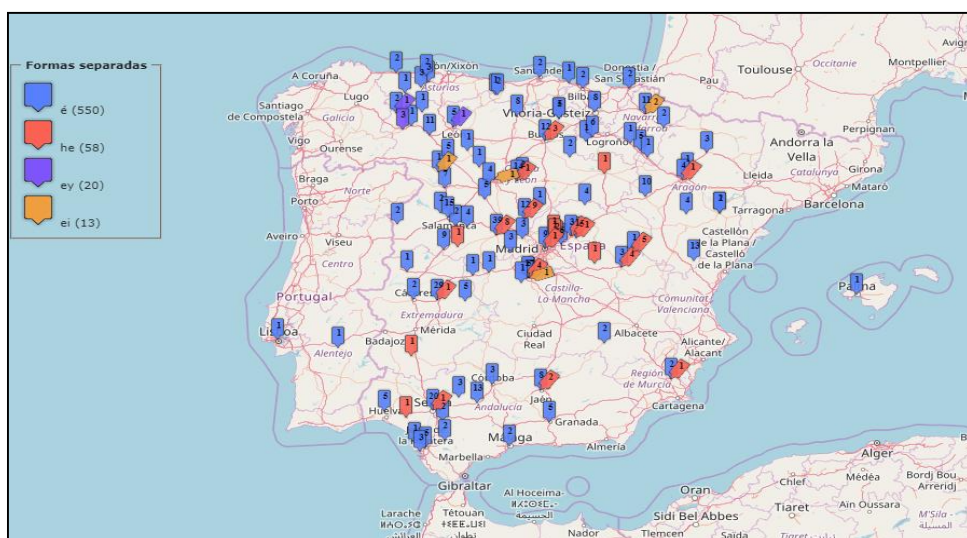
Mapa 6. Distribución geográfica de los morfos diptongados de 1ª pers. de sing. del verbo haber.

Las 33 ocurrencias de *ei/ey* encontradas en 21 documentos diferentes abarcan el período comprendido entre 1233 y 1399. Tras ese año solo se da un caso en un documento autógrafo escrito en Valladolid en 1565 por Guiomar de Vilhena, mujer de habla portuguesa:

(22.) De mi señora la condesa ei sabido que vosa señoría estaua ya fora de todo los piligros de la mar y camino (CODEA-2343, Valladolid, 1565).

En este caso concreto el morfo *ei* no se trataría de un arcaísmo, sino de la 1ª persona de singular de presente de indicativo del verbo portugués *aver*.

En contraste con la forma *hey*, el morfo monoptongo *he* se documenta ampliamente en todo el territorio peninsular. Igualmente, en el área geográfica del dominio lingüístico asturleonés tal y como expone Müller (1973: 246). Este señala la pronta monoptongación que sufrió *hey* en leonés (*hey* > *he*), pese a ser la forma que más tarde se conservaría. En el resto de la Península se difundió la forma *he* (Mapa 7).



Mapa 7. Distribución geográfica de los morfos verbales de la variante *hey*.

Los ejemplos de *hey* en Navarra pueden tener relación con las primeras dataciones de *ai* y *hay*⁷. La primera de ellas se encuentra en una carta de compraventa presumiblemente de Cantabria fechada en 1202 (CODEA-0216), mientras que la segunda se localiza en una carta de donación navarra de 1337 (CODEA-0985). De las variantes analíticas *á y* e *ý á* no hemos encontrado ningún caso en el corpus.

3.2. Distribución diastrática y diafásica

Para poder explicar con mayor claridad la distribución diastrática y diafásica de los dos grupos de variantes, consideramos oportuno comenzar observando el siguiente gráfico:

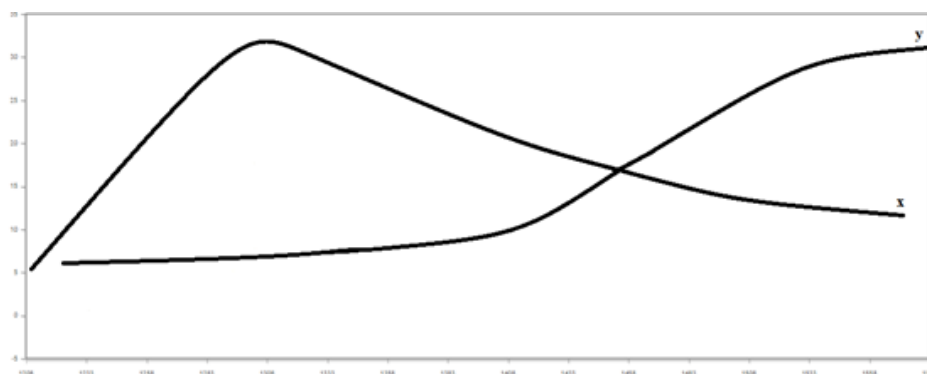


Gráfico 4. Distribución temporal de las variantes etimológicas (x) y las variantes con la partícula -y (y).

A medida que disminuye la frecuencia de uso de las variantes etimológicas (*só*, *estó*, *dó*), aumenta progresivamente la de las variantes etimológicas con -y (*soy*, *estoy*, *doy*, *voy*). La difusión temporal de las variantes innovadoras (y) corresponde al modelo de la curva en S (Conde Silvestre, 2007, p. 143), según el cual los hablantes de las clases bajas son los que participan más activamente en la adopción y difusión de las variantes innovadoras. Si extrapolamos la hipótesis del patrón social curvilíneo (Chambers y Trudgill, 1988; Labov, 1994, 2001) al uso de las variantes de la 1ª persona de singular de presente de indicativo que presentan -y, la difusión de su uso se habría iniciado en aquellos hablantes cuyo nivel de lengua se corresponde estilísticamente con un registro medio bajo. Ello se vería refrendado por los resultados obtenidos al realizar búsquedas en las que se han cruzado la variable siglo y la variable ámbito de emisión.

⁷ Véase, Schmidely (1988).

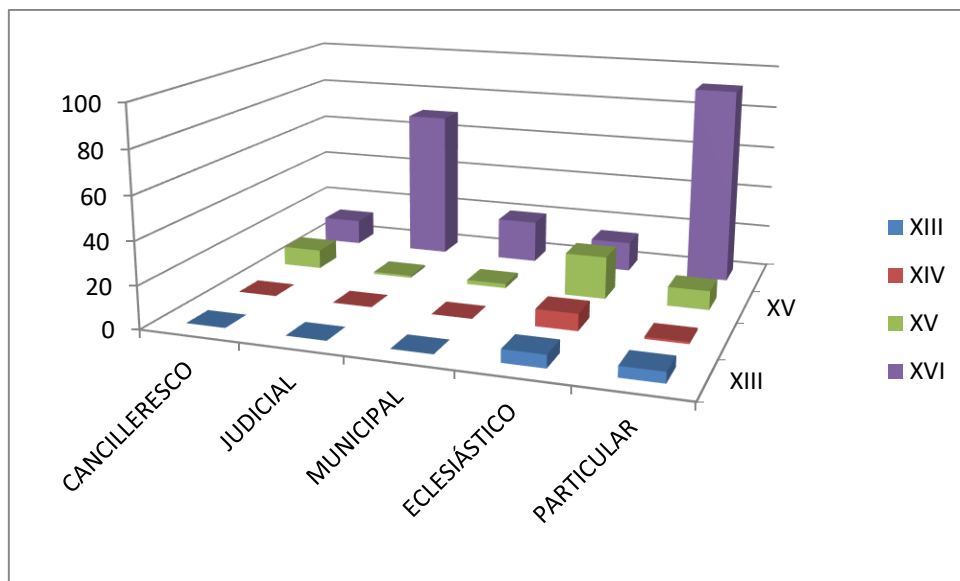


Gráfico 5. Distribución de *doy*, *estoy*, *soy* y *voy* en relación con el ámbito de emisión.

Mientras que en los siglos XIII y XIV solo se dan estas variantes en los ámbitos eclesiástico y particular, esto es, en aquellos ámbitos identificados con los registros menos formales de la lengua, a partir del siglo XV estas variantes comienzan a aparecer en los documentos emitidos por la cancillería, cuyo registro es culto. De modo que se puede concluir que la difusión del uso de las variantes con *-y* en la dimensión social se trata de un cambio de abajo a arriba. Un cambio que se originó en los estratos inferiores de la sociedad y que de forma paulatina fue adoptado por los estratos más altos hasta convertirse finalmente en la norma.

Conclusiones

Una vez analizados los datos obtenidos en el corpus CODEA+ 2015 acerca de la adición de la yod desinencial a la 1ª persona de singular de presente de indicativo de los verbos *ser*, *estar*, *dar* e *ir*, podemos esbozar las tendencias de la progresión de este cambio y su difusión.

Las cronologías de los morfos etimológicos y de los morfos con *-y* han permitido limar la cronología establecida para las variantes innovadoras de este fenómeno, adelantando su primera data al siglo XIII tal y como afirma Cano Aguilar (1998). Desde este siglo se documenta la convivencia de las formas *soy*, *doy*, *estoy* y *voy* y las formas etimológicas *só*, *dó* y *estó*. La forma *vó* no aparece recogida en el corpus.

El ascenso de las variantes con la partícula *-y* a partir de la segunda mitad del siglo XVI coincide con el descenso de las variantes etimológicas. Según los datos del corpus, su consolidación se produce mayoritariamente a finales del XVI, contrariamente a lo que afirman otros autores como Cano Aguilar (1998), que la retrasan hasta el siglo XVII en la totalidad del territorio peninsular.

En cuanto a la distribución geográfica de los dos tipos de variantes, se observan diferencias entre las zonas occidental y oriental de la Península. Las primeras manifestaciones de las variantes con *-y* se localizan en el oeste de la Península, concretamente en León y Salamanca. Se podría aventurar, por tanto, que las

variantes sin –y final se adscribirían principalmente a la zona oriental. Sin embargo, en el siglo XIII las variantes etimológicas continuaban siendo las más frecuentes.

A partir del siglo XIV los morfos con –y continúan su difusión hacia el este. Su uso se amplía hacia el sur y se extiende durante el siglo XVI a la zona centro. A medida que avanzan las variantes *soy*, *doy*, *estoy* y *voy*, retroceden las etimológicas. Las últimas documentaciones de las formas *só*, *dó* y *estó* se dan en la zona oriental peninsular que coincide con el área de dominio aragonés. Esta, a excepción de una forma documentada en Navarra en el siglo XIII, no conoció de forma generalizada las formas innovadoras con yod desinencial hasta el siglo XVII, por lo que para este fenómeno esta área se caracterizaría por ser más conservadora.

Las variantes con –y en documentación navarroaragonesa fueron introducidas a través de documentos emitidos por la cancillería en la zona central. Así, mientras que en Aragón la adopción de estas formas podría describirse como un cambio de arriba a abajo, en el resto de la Península su aparición y difusión se originó en los estratos inferiores de la sociedad. En los siglos XIII y XIV solo se dan estas variantes en los ámbitos eclesiástico y particular, identificados con los registros menos formales de la lengua. Es a partir del siglo XV cuando estas variantes comienzan a aparecer en los documentos emitidos por la cancillería, por lo que de forma paulatina fueron adoptadas por los estratos más altos hasta convertirse finalmente en la norma mediante un cambio de abajo a arriba.

Si entre Castilla y Aragón se puede dibujar una isoglosa que indique la conservación de las variantes etimológicas *só*, *dó* y *estó* frente a su sustitución por *soy*, *doy* y *estoy*, igualmente podemos trazar otra con las variantes *soe* y *dou*. Dentro de la zona occidental estas se han documentado únicamente en la zona más occidental de León, Zamora y Asturias, área de influencia del asturleonés.

Bibliografía

- ABAD MERINO, M. (2009). Uso de corpus documental en sociolingüística histórica y retos para su elaboración en el sureste peninsular (ss. XV-XVII). In ENRIQUE-ARIAS, A. (ed.), *Diacronía de las lenguas románicas. Nuevas aportaciones desde la lingüística de corpus*. Madrid, Frankfurt am Main: Iberoamericana Vervuert, pp. 251-266.
- CANO AGUILAR, R. (1998). *El español a través de los tiempos*. Madrid: Arco Libros.
- CHAMBERS, J. & TRUDGILL, P. (1988). *Dialectology*. Cambridge: Cambridge University Press.
- COMPANY COMPANY, C. (2005). Una paradoja de la lingüística histórica romance: el florecimiento de la sintaxis histórica románica. *La Corónica*, 34 (1), pp. 144-163.
- CONDE SILVESTRE, J. C. (2007). *Sociolingüística histórica*. Madrid: Gredos.
- CUADROS MUÑOZ, R. (2011). *Sepan quantos este privilegio vieren: observaciones sobre el paradigma demostrativo de primera persona en la documentación diplomática medieval*. *Cahiers d'études hispaniques médiévales*, 34, pp. 203-233.
- DE GOROG, R. P. (1980). L'origine des formes espagnoles *doy*, *estoy*, *soy*, *voy*. *Cahiers de Linguistique Hispanique Médiévale*, 5, pp. 157-162.
- DE TOLEDO Y HUERTA, Á. O. & PONS RODRÍGUEZ, L. (2016). *Textos para la Historia del Español X. Queja política y escritura epistolar durante la Guerra de Independencia*:

documentación de la Junta Suprema Central en el AHN. Alcalá de Henares: Servicio de Publicaciones de la Universidad de Alcalá.

- GALLEGU DE LA PUENTE, I. (2007). Las primeras personas de los presentes de ser y dar en documentos notariales riojanos de los siglos XIII y XIV. *Interlingüística*, 17, pp. 381-387.
- GARATEA GRAU, C. (2016). El habla y no la lengua; la diacronía y no la sincronía: la inversión de dos ideas saussureanas. *Signo y Seña*, 30, pp. 22-36.
- GIMENO MENÉNDEZ, F. (1990). De Sociolingüística histórica: en torno a los orígenes del español. In MORENO FERNÁNDEZ, F. (ed.), *Estudios sobre variación lingüística*. Alcalá de Henares: Servicio de Publicaciones de la Universidad de Alcalá, pp. 89-102.
- GREGOR, D. B. (1975). *Friulan: Language and literature*. New York: Oleander Press.
- HARRIS, J. (1983). *Syllable Structure and Stress in Spanish: A Nonlinear Analysis*. Cambridge, Massachusetts: MIT Press.
- HAYES, B. (1989). Compensatory lengthening in moraic phonology. *Linguistic Inquiry*, 20, pp. 253-306.
- KELLER, R. (1985). Towards a Theory of Linguistic Change. In BALLMER, T. T. (2019), *Linguistic Dynamics*. Berlin, Munich, Boston: De Gruyter, pp. 211-237.
- LABOV, W. (1994). *Principles of Linguistic Change*. Vol. 1: *Internal Factors*. Oxford: Blackwell.
- LABOV, W. (2001). *Principles of Linguistic Change*. Vol. 2: *External Factors*. Oxford: Blackwell.
- LABOV, W. (2006). *The Social Stratification of English in New York City*. Cambridge: Cambridge University Press.
- LLOYD, P. M. (1993). *Del latín al español. Fonología y morfología históricas*, 1 vol. Madrid: Gredos.
- MARTÍN BUTRAGUEÑO, P. (1997). Aproximación sociolingüística al estudio de la variación y el cambio sintáctico. Esbozo de algunos problemas generales. In MORENO FERNÁNDEZ, F. (ed.), *Trabajos de sociolingüística hispánica*. Alcalá de Henares: Servicio de publicaciones de la Universidad de Alcalá, pp. 37-67.
- MARTÍNEZ-GIL, F. (2012). Sobre la eclosión histórica de soy, voy, doy, estoy y hay: una solución prosódica. In MONTERO CARTELLE, E. & MANZANO ROVIRA, C. (ed.), *Actas del VIII Congreso Internacional de Historia de la Lengua Española*. Santiago de Compostela, 14-18 de septiembre de 2009. Santiago de Compostela: Meubook, pp. 935-946.
- MENDÍVIL GIRÓ, J. L. (2010). Coseriu, Saussure y el problema del cambio lingüístico. *BSEHL*, 7, pp. 109-127.
- MENÉNDEZ PIDAL, R. (1916). Poesía popular y romancero. *RFE*, III, pp. 233-289.
- MORENO FERNÁNDEZ, F. (2005). *Principios de sociolingüística y sociología del lenguaje*. Barcelona: Ariel.
- MÜLLER, C. (1973). *Estadística lingüística*. Madrid: Gredos.
- NEUMANN-HOLZSCHUH, I. (1996). Reflexiones acerca de una descripción funcional de la posición del sujeto en el español medieval. In ALONSO GONZÁLEZ, A., CASTRO RAMOS, L., GUTIÉRREZ RODILLA, B. & PASCUAL RODRÍGUEZ, J. A. (coords.), *Actas del III Congreso Internacional de Historia de la Lengua Española*. Salamanca, 22-27 de noviembre de 1993. Madrid: Arco Libros, pp. 471-488.
- PENNY, R. (2010). *Gramática histórica del español*. Barcelona: Ariel.

- PENSADO, C. (1988). Soy, estoy, doy, voy como solución de una dificultad fonotáctica. In Zamora, A. (ed.), *Homenaje a Alonso Zamora Vicente*, 1 vol. Madrid: Castalia, pp. 207-218.
- RADTKE, E. (2002). Der Polymorphismus in der romanischen Dialektologie und Sprachgeschichte – eine historiographische Skizzierung. In Wesch, A., Weidenbusch, W., Kailuweit, R. & Laca, B. (ed.), *Sprachgeschichte als Varietätengeschichte. Anlässlich des 60. Geburtstages von Jens Lüdtkke*. Tübingen: Stauffenburg Verlag, pp. 45-51.
- RINI, J. (1996). The 'Clinching Factor' in the Addition of -y in Spanish *doy, estoy, soy, voy*. *Journal of Hispanic Research*, 4, pp. 1-11.
- ROMAINE, S. (1988). Contribution from Sociolinguistics to Historical Linguistics. In Joly, A. (ed.), *La linguistique génétique*. Lille: Presses Universitaires de Lille, pp. 343-368.
- SÁNCHEZ-PRIETO BORJA, P. (2015). Español antiguo. In ILIESCU, M. & EUGEN, R. (ed.), *Manuel des Anthologies, corpus et textes romans*. Berlín, Boston: De Gruyter, pp. 113-146.
- SCHMIDELY, J. (1988). La -y de *doy, estoy, soy, doy*. In ARIZA, M., SALVADOR A. & VIUDAS, A. (ed.), *Actas del I Congreso Internacional de Historia de la lengua española (Vol. 1)*, pp. 611-619.
- SCHMIDELY, J. (1996). La -y de *hay*. In ALONSO GONZÁLEZ, A., CASTRO RAMOS, L., GUTIÉRREZ RODILLA, B. & PASCUAL RODRÍGUEZ, J. A. (coord.), *Actas del III Congreso Internacional de Historia de la Lengua Española*. Salamanca, 22-27 de noviembre de 1993. Madrid: Arco Libros, pp. 195-204.
- SELIG, M. (2008). Geschichte, Variation, Wandel. Spachwandel und historische Corpora. In STARK, E., SCHMIDT-RIESE, R. & STOLL, E. (coord.), *Romanische Syntax im Wandel*. Tübingen: Gunter Narr Verlag Tübingen, pp. 67-88.
- TORRUELLA, J. (2016). Tres propuestas en el ámbito de la lingüística de corpus. In KABATEK, J. (ed.), *Lingüística de corpus y Lingüística histórica iberorrománica*. Berlín, Boston: De Gruyter, pp. 90-112.

Sitografía

- GITHE (Grupo de Investigación Textos para la Historia del Español): CODEA+ 2015 (Corpus de documentos españoles anteriores a 1800). <http://www.corpuscodea.es> [01/2017-05/2020].

Relación de tablas, gráficos y mapas

Tabla 1. Formas de presente: morfos etimológicos y morfos en -y.

Tabla 2. Distribución de piezas documentales por siglo y ámbito de emisión.

Tabla 3. Tabla resumen de las piezas documentales del corpus por siglo/provincia.

Tabla 4. Distribución por siglos de *doy, soy, estoy y voy*.

Gráfico 1. Distribución por años de *doy, soy, estoy y voy*.

Gráfico 2. Distribución por años de *só*.

Gráfico 3. Distribución por años de *dó*.

Gráfico 4. Distribución temporal de las variantes etimológicas (x) y las variantes con la partícula -y (y).

Gráfico 5. Distribución de *doy, estoy, soy y voy* en relación con el ámbito de emisión.

- Mapa 1. Distribución geográfica de las variantes etimológicas y de las variantes con -y en el siglo XIII.
- Mapa 2. Distribución geográfica de las variantes etimológicas y de las variantes con -y en el siglo XIV.
- Mapa 3. Distribución geográfica de las variantes etimológicas y de las variantes con -y en el siglo XV.
- Mapa 4. Distribución geográfica de las variantes etimológicas y de las variantes con -y en el siglo XVI.
- Mapa 5. Distribución geográfica de los morfos *soe* y *dou*.
- Mapa 6. Distribución geográfica de los morfos diptongados de 1ª pers. de sing. del verbo *haber*.
- Mapa 7. Distribución geográfica de los morfos verbales de la variante *hey*.