

Cambios históricos frecuentes de la ortografía léxica

Análisis de los documentos castellanos del corpus CODEA

Hiroto Ueda (Universidad de Tokio)
uedahiroto@jcom.home.ne.jp

Índice

1. Introducción y objetivos	2
2. Análisis de datos.....	7
2.1. Metadatos.....	7
2.2. Formas léxicas específicas	13
2.2.1. La conjunción «y»	13
2.2.2. El artículo indefinido «un».....	16
2.2.3. El verbo «haber»	18
2.3. Formas gráficas generales	22
2.3.1. Forma unida y separada	22
2.3.2. Forma abreviada y plena	24
2.3.3. Mayúscula y minúscula.....	26
2.4. Totalidad.....	30
3. Consideraciones finales	34
3.1. Hipótesis	34
3.2. Imprenta y letra	36
Referencias	40

Resumen

El objetivo de este estudio es presentar resultados del análisis cuantitativo de la historia de las formas ortográficas léxicas utilizando los documentos de Castilla la Vieja y Castilla la Nueva del corpus CODEA («Corpus de Documentos Españoles Anteriores a 1800»). A partir de la revisión de los estudios anteriores establecemos las dos siguientes hipótesis de trabajo: (1) La historia de los cambios de las formas ortográficas léxicas obedece a una periodización progresiva dividida en tres etapas: etapa inicial estable larga → etapa de transición rápida → etapa moderna estable otra vez larga; (2) los

cambios históricos de la ortografía léxica no son incidentes aleatorios o accidentales, sino naturales y razonables, sometidos a factores lingüísticos (gráficos) y extralingüísticos (históricos, geográficos y sociales). Prestaremos especial atención a la condición de las letras, por su alta relevancia correlacional. Nuestro método es cuantitativo, a diferencia de los estudios anteriores cualitativos y presenta algunas nuevas visiones con operaciones estadísticas sencillas.

Palabras clave: cambios históricos, ortografía léxica, castellano, letra gótica, letra humanística, corpus CODEA.

1. Introducción y objetivos¹

Al tratar los rasgos históricos de la lengua española, uno se ve obligado a establecer una cronología sucesiva y de alguna manera segmentada cronológicamente. Los intentos anteriores, a la vez que respetan la división histórica más importante entre la Edad Media y la Moderna, exponen distintos rasgos grafonéticos, morfosintácticos y semántico-léxicos que caracterizan cada edad. Se supone que la división entre las dos edades históricas es tan clara e importante que no es necesario ponerla en cuestión y la mayoría de los estudios anteriores coinciden en trazar una línea divisoria alrededor del año 1500 (Sánchez Lancis 1999: 427).

En el intento general de periodización existen controversias en los métodos de investigación, en sus resultados de análisis y en las conclusiones alcanzadas. También hay manuales, cursos o ensayos sobre la historia de la lengua, donde los autores se limitan a explicar algunos ejemplos de rasgos y sucesos caracterizadores. A la hora de seguir los caminos que ha recorrido la lengua para llegar al estado actual, tales ejemplos resultan interesantes e ilustrativos, pero también echamos de menos datos cuantitativos fehacientes que garanticen la validez estadística de las observaciones ahí presentadas.

Nos planteamos cambiar el método tradicional cualitativo al nuevo cuantitativo y exhaustivo con el objeto de ofrecer una perspectiva distinta. Nuestro intento no es refutar las teorías anteriores sino más bien proponer otro

¹ Agradezco de todo corazón a Leyre Martín Aizpuru, Inmaculada Martínez, Carlos Rubio y Pedro Sánchez-Prieto Borja la ayuda que me han brindado en la preparación de este trabajo como en tantas otras ocasiones anteriores. Este trabajo ha sido subvencionado por JSPS KAKENHI Grant Number 16K02657 y el proyecto «Corpus de Documentos Españoles Anteriores a 1900 (CODEA+ 2020)» (Ref. FFI2017-82770-P).

método diferente de análisis en la creencia de que ambos estudios no son excluyentes sino complementarios.

En primer lugar, vamos a revisar los principales trabajos anteriores sobre el tema de la periodización de la historia del español. Empezamos con la observación que ofrece Sánchez Lancis (1999: 427), puesto que trata los tres puntos para nosotros cruciales: (1) etapas arbitrariamente representadas por siglos, (2) relación entre la historia interna y la externa de la lengua, y (3) tres etapas históricas establecidas por Eberenz (1991). En primer lugar, estamos de acuerdo con su crítica respecto a la división en las etapas "representadas arbitrariamente por lapsos de tiempo equivalentes a siglos", puesto que es imposible pensar que la historia de la lengua se desarrolle con el avance de un tiempo escalonado en unidades de cien años. Si se hiciera una historia de la lengua con el establecimiento cronológico previo de tales unidades, el resultado sería una confusión con lapsos vacuos y extendidos a lo largo de la linde secular. Es cierto que en la lingüística cuantitativa se usan distintas unidades de tiempo como medios para ver las relaciones mutuas y establecer divisiones estadísticamente convincentes. Sin embargo, estas unidades temporales, repetimos, no sirven más que como medios convenientes de análisis.

Por otra parte, al leer las obras citadas por Sánchez Lancis, notamos que en realidad Menéndez Pidal (1942: 47) habla de "la historia dividida en siglos" con cautela, a pesar de que utiliza el término "siglo" en el título del capítulo: "El lenguaje del siglo XVI". Tampoco Lapesa (1981) ni Cano Aguilar (1988) tratan los siglos para caracterizarlos lingüísticamente. Los dos autores exponen los siglos *a posteriori* para situar cronológicamente los hechos mencionados en los capítulos correspondientes, lo que los lectores quieren saber con exactitud de tiempo. Nosotros también dependeremos de estas unidades crónicas como escalas de medición, sin intención de ajustar los hechos lingüísticos a tales unidades de manera apriorística.

Compartimos la misma opinión de Sánchez Lancis (*loc. cit.*) de que "es muy necesario tener en cuenta también los aspectos de historia interna, propiamente lingüísticos, ya que existe una clara relación entre la historia interna y la historia externa de toda lengua". Por nuestra parte, proponemos una división bipartita de trabajo: la historia interna y la externa, en este orden, para observar la primera en relación con la segunda, pero no al revés, puesto que nuestro interés está principalmente en la lengua.

Veamos la periodización tripartita de Eberenz (1991) mencionada en Sánchez Lancis (*loc. cit.*). Eberenz (*op. cit.* 105-106), tras revisar los trabajos

anteriores encaminados a una división histórica de la lengua española, llega a la siguiente conclusión: una fase *antigua* (1200-1450), caracterizada por la "estabilidad de las estructuras esenciales de la lengua escrita", seguida de una etapa *media* (1450-1650), "marcada por una transformación más rápida" y una fase *moderna* (1650-), "que presenta de nuevo un sistema esencialmente estable". Esta división es plausible. Es de desear que se presenten los rasgos distintivos de cada período con sus respectivas relevancias estadísticas.

El modelo de Eberenz (*op. cit.*) es más convincente que una simple división bipartita habitual que se establece entre el español medieval y el moderno, puesto que sabemos a ciencia cierta que la lengua no cambia de una fase a otra de manera tajante. En los cambios sociales, siempre observamos unos pocos indicios precursores de difusión inicial lenta y, una vez llegado el momento, se produce un cambio rápido, hasta llegar finalmente a la fase otra vez larga de estabilidad de pocos movimientos. Es decir, la proporción de la nueva forma con respecto a la totalidad (antigua forma + nueva forma) empieza con 0 % para llegar a 100 %, con una línea no recta ni exponencial, sino con forma "S" plana.

Según la teoría de Labov (1994: 65), la curva-S se produce a partir de la probabilidad acumulativa de la distribución normal:

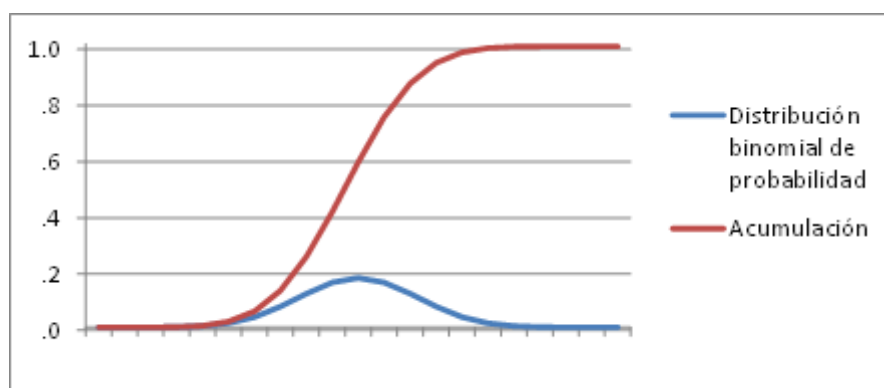


Fig. 1. Distribución normal de probabilidad y su acumulación

Denison (2003: 58) explica la razón de la curva-S por la competencia que hay entre la forma antigua y la nueva (traducción nuestra):

En las primeras etapas de un cambio, (...) la nueva forma es rara, por lo que las presiones de elección son relativamente débiles siendo lenta la subida de la tasa de cambio. En las últimas etapas del cambio, la forma antigua es rara,

por lo que el efecto selectivo de tener dos formas para comparar y elegir entre ellas es nuevamente débil, siendo lenta la tasa de cambio. Solo en el período intermedio, cuando hay un número considerable de cada forma en competencia, se acelera la tasa de cambio. De ahí la curva-S.

Para dar otra posible explicación desde el punto de vista no de formas lingüísticas competitivas, la antigua y la nueva, sino de la sociedad, de los usuarios de la lengua, partimos de la distribución normal (o binomial) de probabilidad donde los primeros precursores del cambio, por ejemplo, los primeros usuarios de un aparato nuevo (incidencia nueva), siempre son de número reducido. Al llegar a la frecuencia general de su uso, la mayoría suele obedecer a la nueva norma, al uso general del aparato, con pocas excepciones. Los que finalmente confluyen en la corriente mayoritaria son pocos. Si acumulamos la línea normal de todas las probabilidades, que presenta la forma de campana (curva azul en la Fig. 1), la línea de esta probabilidad acumulativa adopta matemáticamente la forma de "S" plana (curva roja). Todo esto se debería a la ley general de que las personas de comportamiento anormal son pocas y las personas normales son muchas.

Pensamos que ambas explicaciones, una formal y otra social, son válidas y complementarias. Efectivamente, Labov (1994: 66) explica el cambio fonético desde los dos puntos de vista, uno formal y otro social, en términos de "exposición" y "contacto" de los hablantes con las formas nuevas y antiguas (traducción nuestra):

(...) Hay dos sonidos en competencia asociados a los valores sociales característicos de los hablantes que los usan, y el progreso del cambio está asociado a la adopción de los valores de un grupo por miembros del otro. Por lo tanto, al comienzo del cambio, los hablantes de la forma más antigua rara vez se hallan expuestos a la forma más nueva, por lo que puede ocurrir poco cambio o transferencia. La tasa de cambio será mayor cuando el contacto entre los hablantes sea mayor, es decir, en el punto medio. / Podría parecer que este modelo predice un movimiento rápido hacia la finalización en la segunda mitad del cambio, ya que los hablantes de la forma conservadora estarían rodeados por hablantes de la forma más nueva, siendo su conversión la que gobierna la tasa de cambio. (...)

De esta manera, podemos aplicar el modelo de la curva-S a la teoría de

Eberenz (*loc. cit.*) que presenta la "transformación más rápida y perceptible" entre la "fase antigua" y la "fase moderna", ambas caracterizadas con una "relativa estabilidad".

Naturalmente el cambio sociolingüístico no se realiza con una fórmula de probabilidad matemática, pero creemos que la curva de forma "S" sirve como un modelo teóricamente ideal, aunque no siempre prácticamente real, del cambio sociolingüístico. Vamos a averiguarlo en la sección de análisis (sec. 3).

Un aspecto importante que se advierte en la mayoría de los estudios anteriores sobre la periodización de la lengua española es la falta del análisis cuantitativo. A excepción de un número bastante reducido de investigaciones filológicas, la mayoría se limita a exponer los rasgos gráficos, fonéticos, morfosintácticos, léxicos o semánticos de manera cualitativa, sin presentar sus frecuencias concretas. Estamos ante la cuestión de la periodización y la norma, en la cual creemos conveniente observar la situación histórica con cifras concretas y comparables, puesto que sin ellas resulta difícil evaluar su magnitud cuantitativa y, consecuentemente, su importancia cualitativa.

Como una excepción especial, nos ha llamado la atención el reciente trabajo de Granvik y Sánchez Lancis (2018), en la cual se ha aplicado el método multivariante de agrupamiento jerárquico a los datos recogidos del CORDE («*Corpus de Referencia del Español Actual*») de la RAE². Los dos autores tratan los 27 fenómenos gramaticales (morfológicos y sintácticos), que sirven de base para la periodización de la lengua española. Ambos (*ibid.*: 763-764) reconocen que "los resultados obtenidos en el presente trabajo dependen completamente de los fenómenos seleccionados para su investigación".

Otro problema no menos grave que observamos en los trabajos anteriores reside en que los investigadores no exponen clara y convincentemente con qué criterio de selección de rasgos lingüísticos se ha realizado la periodización de la historia de la lengua. El número reducido de los fenómenos parece deberse a la razón práctica del trabajo. Evidentemente, cuantos más fenómenos adecuados poseamos, más fiable será el resultado. Por nuestra parte, pensamos que incluso es posible no hacer la selección sino incluir todas las formas que se encuentran en los datos tratados. Más adelante veremos su posibilidad (sec. 3.4).

Eberenz (2009: 196) habla de "la disponibilidad de grandes corpus digitalizados" y del "acercamiento cuantitativo":

² <https://www.rae.es/recursos/banco-de-datos/crea> [27/05/2022]

En el plano filológico, hace falta volver a los textos e ir más allá del simple recuento de ocurrencias de los fenómenos. La periodización supone establecer una cronología absoluta de los cambios, tarea facilitada hoy en día por la disponibilidad de grandes corpus digitalizados. Pero el acercamiento cuantitativo es sólo la etapa inicial de la investigación, pues debe completarse por un estudio intratextual e intertextual. El proceso culminará así tanto en el análisis estructural del cambio como en la interpretación de su difusión social a partir de las obras y los documentos en que se atestigua.

Por nuestra parte, proponemos librarnos del "simple recuento de ocurrencias de los fenómenos" (Eberenz: *loc. cit.*), para llegar a un tratamiento exhaustivo de todas ellas en la totalidad de los fenómenos dentro de los datos de cantidad confiable, lo que no ha sido posible hasta épocas recientes. Lo veremos en la sección dedicada al análisis, donde averiguaremos la validez de nuestras hipótesis de trabajo establecidas a partir de estudios anteriores:

Hipótesis-1: La historia de los cambios de las formas ortográficas léxicas obedece a una periodización progresiva dividida en tres etapas: Etapa inicial estable larga → Etapa de transición rápida → Etapa moderna estable larga.

Hipótesis-2: Los cambios históricos de la ortografía léxica no son incidentes aleatorios o accidentales, sino naturales y razonables, dependientes de factores lingüísticos (gráficos) y extralingüísticos (históricos, geográficos y sociales).

A continuación, procedemos a indagar la validez de ambas hipótesis en la historia de la ortografía castellana.

2. Análisis de datos

2.1. Metadatos

En este estudio hemos utilizado los datos del corpus CODEA («*Corpus de documentos Españoles Anteriores a 1800*»)³ y especialmente los datos

³ GITHE (Grupo de Investigación Textos para la Historia del Español):

lematizados de documentos de la región de Castilla la Vieja (Sánchez-Prieto Borja / Ueda, 2018 y de Castilla la Nueva (Aguetas / Sánchez-Prieto Borja / Ueda, en preparación).

Tomamos en consideración las formas léxicas, junto con los siguientes metadatos: franja de año con intervalo de 50 años (1200, 1250..., 1750)⁴, región (Castilla la Vieja, Castilla la Nueva), documento (documento cancilleresco y documento particular) y letra (gótica y humanística).

Nuestro interés principal está en el cambio cronológico. Por ello, seleccionamos el intervalo más conveniente, adecuado para el análisis estadístico y la representación gráfica: 50 años. No partimos de la división preestablecida entre la Edad Media y la Moderna, puesto que nuestro objetivo es buscar los puntos cronológicos críticos de cambios ortográficos. Según nuestros experimentos, el intervalo de 100 años es demasiado amplio mientras que el de 25 años resulta poco útil para realizar la comparación.

Para el parámetro geográfico, tomamos la decisión de seleccionar la región en lugar de la provincia, puesto que no todas las provincias aparecen en todas las franjas cronológicas. En la clasificación entre Castilla la Vieja y Castilla la Nueva, excluimos los documentos cancellescos por el carácter itinerante de la cancelería real, como muestra la siguiente lista de ciudades de donde se emitieron los documentos cancellescos⁵:

Lugar de la emisión de los documentos cancellescos (año: lugar): 1223: SG, 1229: BU, 1243: VA, 1254: BU, 1255: P, S, 1256: SG, 1260: GU, 1271: P, 1272: BU, 1273: TO, 1274: BU, 1275: VA, 1276: BU, 1278: SG, 1282: VA, BU, 1283: VA, 1284: SG, TO, 1285: BU, GU, TO, 1286: P, BU, 1287: SO, 1288: VA, 1289: BU, VA, 1290: TO, 1291: BU, VA, TO, 1292: AV, BU, 1293: VA, 1303: M, 1304: BU, SO, 1307: VA, 1314: GU, 1315: BU, 1322: TO, 1326: BU, 1329: M, TO, 1338: CU, 1340: M, 1342: VA, 1347: TO, 1348: M, 1351: VA, 1352: SO, 1368: TO, 1377: BU, 1379: VA, 1381: SG, 1386: M, 1392: BU, 1401: VA, 1407: SG, 1408: GU, VA, M, 1412: VA, CU, 1417: BU, 1420: VA, 1422: TO, 1448: BU, TO, 1453: BU, 1455: AV, 1457: BU, 1458: VA, 1460: GU, 1462: VA, 1465: TO, 1470: SG, 1475: VA, AV, 1481: VA, 1488: M, 1493: SO, 1494: VA, M,

CODEA+ 2015 (Corpus de documentos españoles anteriores a 1800)

<http://www.corpuscodea.es/> [27/7/2019]

⁴ Los años de cada franja representan el año de inicio. Por ejemplo, la franja de 1200 comprende los años desde 1200 hasta 1249.

⁵ Leyenda de provincia: AV, Ávila; , BU, Burgos; CU, Cuenca; GU, Guadalajara; M, Madrid; P, Palencia; S, Santander; SG, Segovia; SO, Soria; TO, Toledo; y VA, Valladolid.

1496: VA, 1498: S, 1502: TO, 1503: M, 1505: SG, 1508: BU, 1513: VA, 1515: BU, 1518: AV, 1521: BU, VA, 1525: TO, 1532: VA, 1535: M, 1537: VA, 1546: M, 1555: VA, 1559: TO, 1560: VA, TO, 1562: VA, 1563: M, 1569: VA, 1573: M, 1592: BU, 1594: M.

La siguiente tabla muestra los números de documentos en Castilla la Vieja y Castilla la Nueva, a excepción de los documentos cancillerescos:

Región	1200	1250	1300	1350	1400	1450	1500	1550	1600	1650	1700	1750	Total
C. Vieja	7	53	34	14	17	21	44	7	24	6	9	24	260
C. Nueva	8	10	5	18	18	15	34	257	127	66	33	72	663
Total	15	63	39	32	35	36	78	264	151	72	42	96	923

Tabla 1. Número de documentos en Castilla la Vieja y Castilla la Nueva

El corpus CODEA asigna la tipología documental clasificada en cinco grupos: cancilleresco, eclesiástico, judicial, municipal, particular, distribuidos en la siguiente tabla (Tabla 2). Seleccionamos los documentos cancillerescos y los particulares por su alta frecuencia y caracterización estilística contrastiva:

Docu- mento	1200	1250	1300	1350	1400	1450	1500	1550	1600	1650	1700	1750	Total
Canciller.	5	68	28	12	15	30	30	28	3	4	1	0	224
Eclesiást.	12	34	9	23	20	16	11	8	19	3	2	3	160
Judicial	1	1	1	2	3	4	27	48	29	32	7	18	173
Municipal	0	2	0	3	3	7	18	11	6	3	2	9	64
Particular	2	26	29	4	9	9	22	197	97	34	31	66	526
Total	20	131	67	44	50	66	108	292	154	76	43	96	1147

Tabla 2. Número de documentos en cinco tipos documentales

En cuanto a las letras, el corpus CODEA presenta la siguiente distribución de números de documento, ordenada cronológicamente de manera diagonal. La siguiente tabla muestra el cambio cronológico de letras (Ueda 2017):

Letra	1200	1250	1300	1350	1400	1450	1500	1550	1600	1650	1700	1750	Total
Gótica documental	7	3	4										14
De privilegios	1	6											7
Gótica libraria	12	30	10		1								53
De albalaes		18	3										21
Gótica cursiva		57	49	18	1								125
Gótica		3	1	2	4	1							11
Gótica redonda		5		1		2	4	1					13
Gótica precortesana				13	24	1	3						41
Cortesana				1	15	52	69	2					139
Procesal						3	4	35	7				49
Humanística							1	100	51	7	1	2	162
Humanística redonda					1	2	3	36	6	3	1	10	62
Humanística cursiva							4	99	73	61	35	74	346
Bastarda							1	3		3			7
Bastardilla redonda									4	1	6	9	20
Total	20	122	67	35	46	61	89	276	141	75	43	95	*

Tabla 3. Número de documentos según la escritura. *Total: 1070

Según los estudios paleográficos (Marín Martínez y Ruiz Asensio 1996: I.283-346, II.63-83; Tamayo 2012: 365-443, 445-455), estas letras se clasifican en dos grupos fundamentales: letras góticas y letras humanísticas. Hemos trazado la línea divisoria en la misma tabla. Respetamos la misma división en los análisis subsiguientes. A continuación, presentamos algunos ejemplos de la letra gótica (Fig. 2, 3) y la humanística (Fig. 4, 5):

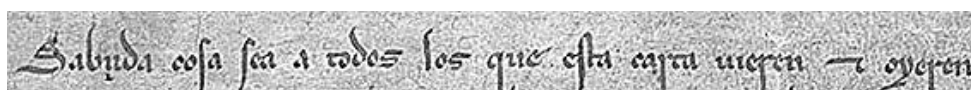


Fig. 2. CODEA-159: Burgos, 1229, documento cancilleresco, letra gótica.

Sabuda cosa sea a todos los que esta carta uieren & oyeren & oyeren

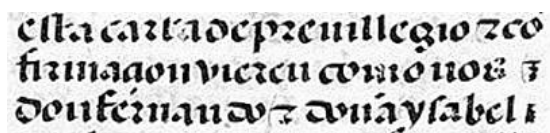


Fig. 3. CODEA-984: Toledo, 1502, documento cancilleresco, letra gótica.
 {1} *esta carta de preuilegio & con* {2} *firmacion vieren como nos*
 {3} *don fernando & donna ysabel*

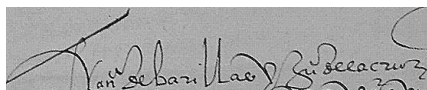


Fig. 4. CODEA-2071: Toledo, 1570, documento particular, letra humanística.
Francisco Debarillas y Juan de la Cruz

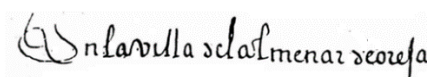


Fig. 5. CODEA-2054: Madrid, 1588, documento particular, letra humanística.
En la villa de Colmenar de Oreja

Veamos seguidamente la distribución de todas las formas en las dos regiones (CV: Castilla la Vieja, CN: Castilla la Nueva), dos tipos documentales (Can: documentos cancillerescos, Par.: documentos particulares) y dos clasificaciones de la escritura (Gót.: letra gótica, Hum.: letra humanística):

Forma	1200	1250	1300	1350	1400	1450	1500	1550	1600	1650	1700	1750
CV	2199	23097	26368	12718	15018	20650	32786	6539	8968	4298	6962	6284
CN	3308	4680	1896	19715	18306	12203	25063	79100	34927	48013	15032	22760
Can.	2804	36017	19523	13690	24608	33558	27411	30984	7998	14164	669	0
Par.	852	9684	22574	3083	10373	9120	14385	28461	13273	29566	14307	15749
Gót.	8311	60408	47787	39141	54306	57760	65445	33812	6792	815	6672	4034
Hum.	0	0	0	0	1502	2259	7354	68347	38900	64896	15991	25010

Tabla 4. Distribución de todas las formas

Para observar las correlaciones entre los seis parámetros mencionados, conviene elaborar sus respectivos gráficos de línea a lo largo de historia. A partir de un gráfico general de todos los parámetros, hemos notado la existencia de dos grupos: CV-Can.-Gót. y CN-Par.-Hum. Por ello, los hemos separado en los dos siguientes gráficos:

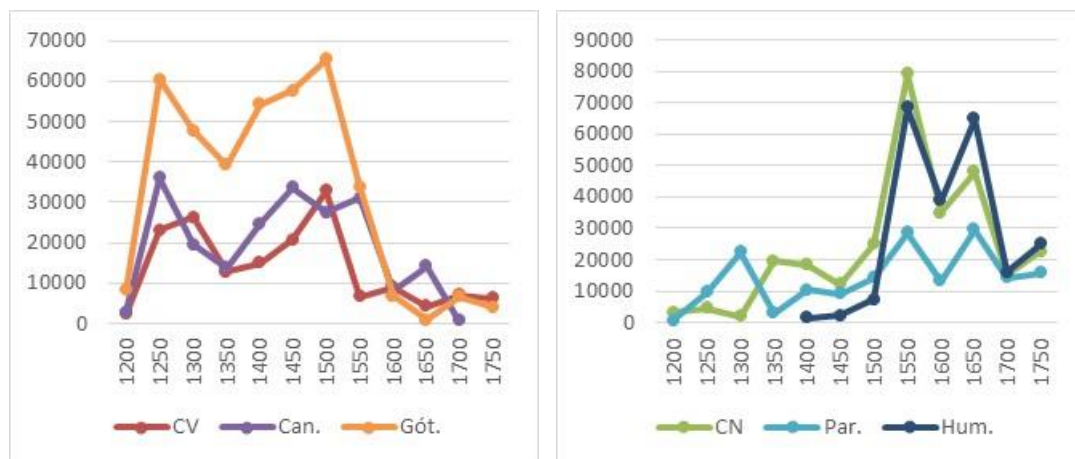


Fig. 6. Distribución de todas las formas.

(a) CV: Castilla la Vieja, Can.: Documento cancelleresco, Gót. Letra gótica.

(b) CN: Castilla la Nueva, Par.: Documento particular, Hum. Letra humanística.

Creemos importante reconocer estas correlaciones existentes dentro de cada grupo (Fig. 6a, 6b), de que trataremos a lo largo de este estudio.

También conviene confirmar las relaciones cuantitativas de formas entre las dos regiones y las dos letras. Según los siguientes gráficos, se aprecia la relación básica complementaria de las letras, es decir, fundamentalmente la letra gótica decae a medida que crece la letra humanística, aunque hay una superposición importante en la franja cronológica de 1550, época de transición, especialmente en Castilla la Nueva:

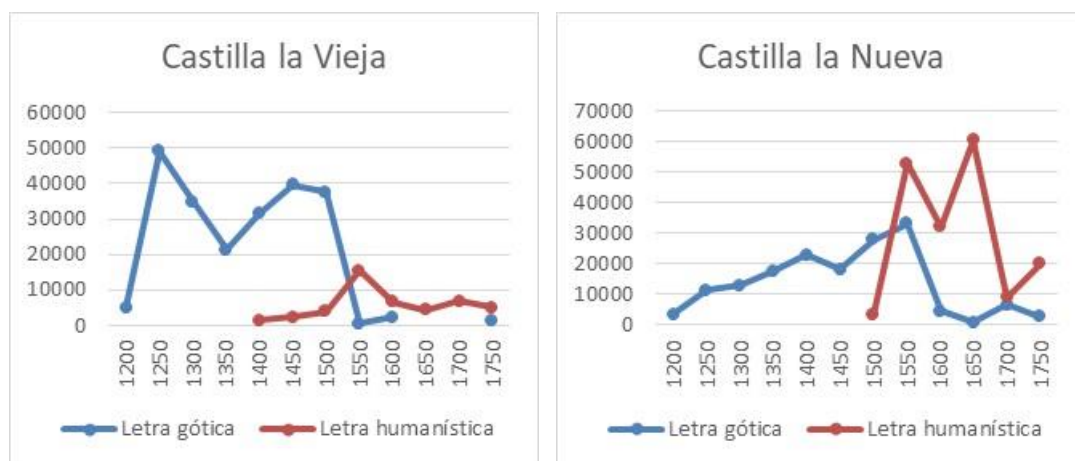


Fig. 7. Distribución de todas las formas.

(a) Castilla la Vieja. (b) Castilla la Nueva.

Con respecto a los tipos documentales, los siguientes gráficos muestran

el cambio general de la letra gótica a la letra humanística tanto en los documentos cancillerescos como en los particulares. Nos llama la atención el crecimiento destacado de la letra humanística en los documentos particulares, que contrasta con los documentos cancillerescos donde se observa la disminución general incluso en la letra humanística:



Fig. 8.. Distribución de todas las formas.

(a) Documento cancilleresco. (b) Documento particular.

Tomamos en cuenta estas correlaciones básicas a la hora de analizar las vicisitudes de las formas léxicas, que veremos a continuación.

2.2. Formas léxicas específicas

Vamos a indagar los casos léxicos específicos de alta frecuencia: conjunción «y», artículo indefinido «un» y verbo auxiliar «haber», que creemos apropiados para observar los grandes rasgos ortográficos históricos.

Veamos la proporción de casos concordantes entre las formas léxicas en la transcripción paleográfica y las de la forma moderna (Con.), con respecto a los casos discordantes entre las dos (Dis.), que se representa en la ratio de formas concordantes dentro de la Suma (Con. + Dis.). La llamamos 'ratio de concordancia' (RC):

$$RC \text{ (ratio de concordancia)} = \text{Concordancia} / \text{Suma}$$

2.2.1. La conjunción «y»

La conjunción «y» es peculiar en poseer el fonema /i/ a pesar de su

carácter átono, al lado de la grafía '&', representante del signo tironiano procedente de la escritura latina y la coexistencia con la forma latina 'et' y su sucesora española medieval 'e', y la grafía moderna 'y', una variabilidad nada fácil de explicar. Para abordar estos temas, vamos a recorrer rápidamente la historia, con el mismo marco geográfico, social y escritural.

Primero, vamos a observar las vicisitudes de frecuencia de las variantes mencionadas en total, sin hacer distinción de parámetros. Según el siguiente gráfico, la frecuencia de la forma actual 'y' experimenta una subida repentina a partir de 1500. Veamos inmediatamente la causa de tan drástica subida⁶:



Fig. 9. Variantes históricas de «y». Total.
Frecuencia normalizada por 10 000 palabras.

Según los siguientes gráficos de ratios de concordancia, parece que la forma moderna 'y' llega a la mayoría numérica (0.5) en 1550, un tanto antes en Castilla la Vieja que la Nueva (Fig. 14b), y en los documentos cancillerescos antes que en los particulares (Fig. 15a).

⁶ El representante 'i' incluye 'j', y 'e' incluye 'et'.

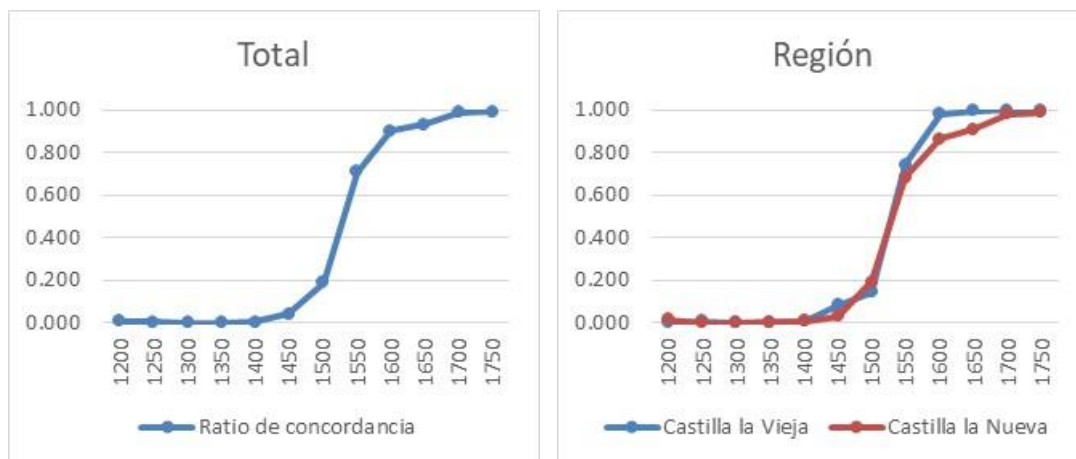


Fig. 10. Ratio de la forma 'y'. (a) Total. (b) Región.

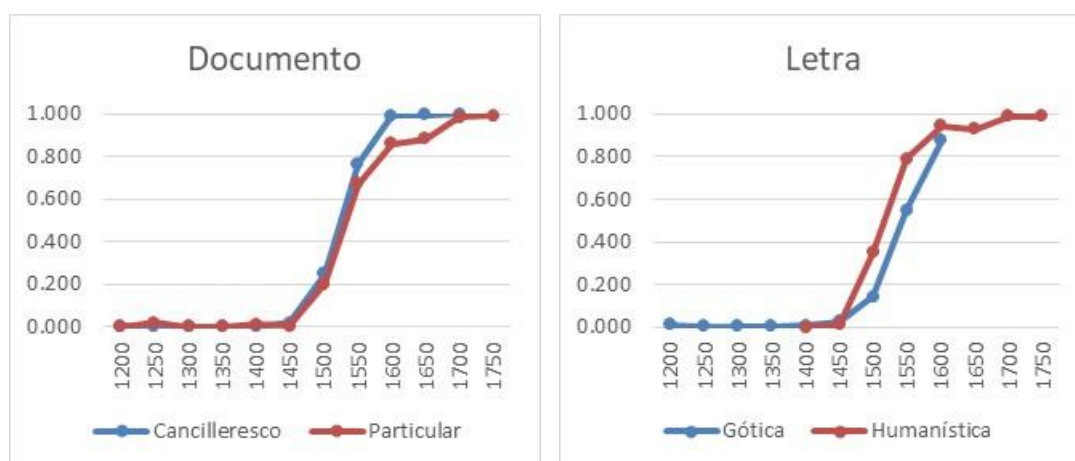


Fig. 11. Ratio de la forma 'y'. (a) Documento. (b) Letra.

En la escritura gótica, hasta 1500, se utilizaba mayoritariamente el signo tironiano '&' y en 1550 de repente se cambió la situación con preferencia de 'y', debido probablemente a la influencia de la letra humanística, que en 1550 prefirió la 'y' a costa de '&'. Su subida posterior es rápida desde 1450 hasta 1550, tal como observamos en la Fig. 11b. Las curvas de subida presentan la forma "S" plana, discutida en la sección 1.

A nuestro modo de ver, la repentina aparición de 'y' como conjunción no es explicable del todo por el aspecto gráfico parecido al signo tironiano en forma de '7' (Salvador y Lodaes 2001: 352; Fig. 2 y 3). Sin negar la posibilidad de atribuir la causa a la semejanza física entre el signo tironiano parecido a '7' y la letra 'y', proponemos considerar otra causa principal, puesto que la forma de la letra 'y' existía mucho tiempo antes, siendo difícil suponerla como causa inmediata de la formación ortográfica tan repentina y radical de la conjunción

justo en 1550. Suponemos que la repentina incidencia de la 'y' como conjunción se debe a un factor externo: la llegada de Italia de la nueva escritura humanística⁷. De esto trataremos más adelante (sec. 4).

2.2.2. El artículo indefinido «un»

Para la vocal /ú en la posición inicial del artículo indefinido «un» ha habido dos grupos de variantes principales: uno con 'v*' (frecuencia): 'vn' (1250), 'vn<a>' (15), 'vna' (1405), 'vnas' (96), 'vnos' (125); y otro con 'u*' (frecuencia): 'un' (584), 'una' (575), 'unas' (44), 'unos' (38). Esta variación se ha resuelto finalmente en 'u*' tras una larga historia de competición, como muestra el siguiente gráfico. En primer lugar, nos llama la atención la forma de letra "U" en la curva de Total (Fig.12a). Vamos a buscar su razón de esta forma peculiar:

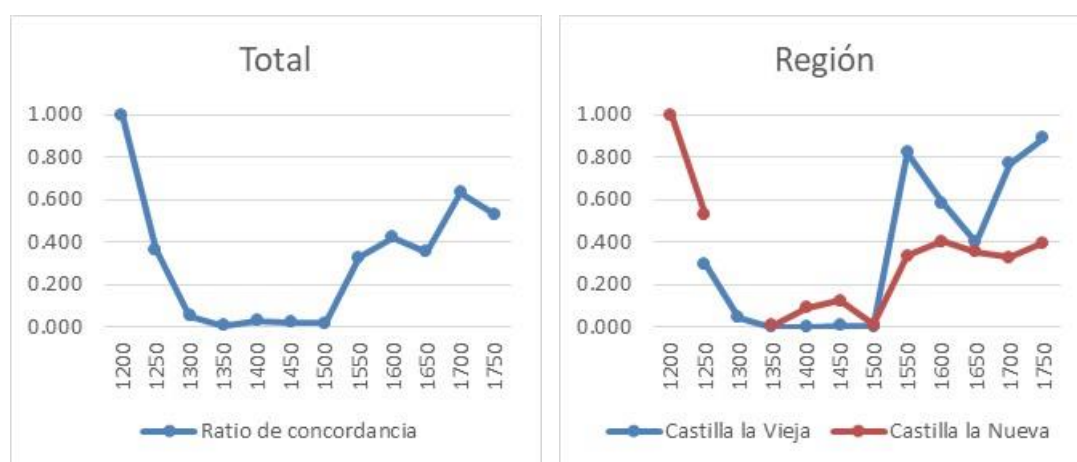


Fig. 12. Ratio de la forma 'un*' en «un». (a) Total. (b) Región.

⁷ Para la explicación del fonema /i/ en lugar de /e/, hemos pensado en la analogía con el frecuente diptongo creciente (Ueda, Sánchez-Prieto Borja y Moreno Sandoval, 2020).

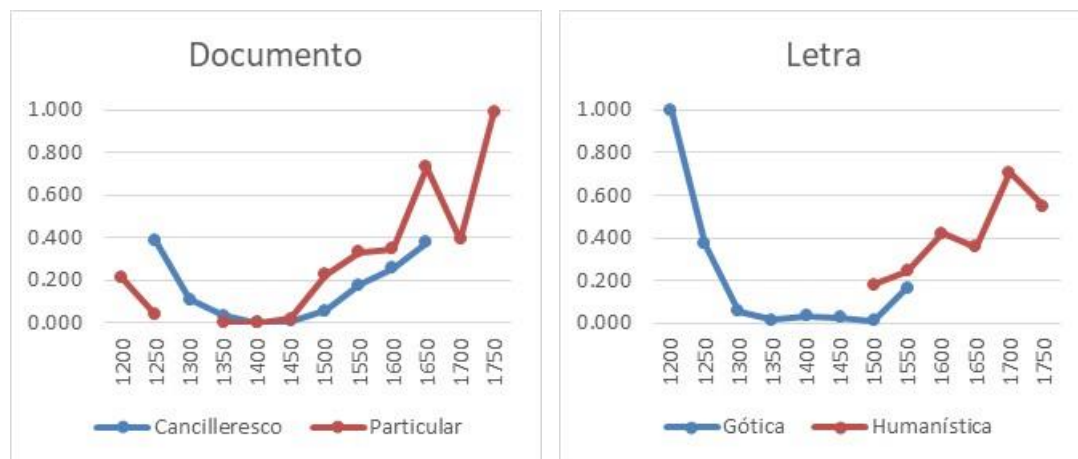


Fig. 13. Ratio de la forma 'un*' en «un». (a) Documento. (b) Letra.

En la distinción geográfica, observamos que tanto en Castilla la Vieja como en la Castilla la Nueva se nota la subida repentina en 1550 (Fig. 12b). En cuanto a la distinción documental, el documento particular avanza más que el documento cancilleresco, que sigue indeciso en adoptar la 'u*' inicial (con menos de 50 %), manteniendo la antigua forma con la 'v*' inicial⁸:

Nos interesa especialmente la competencia de las dos formas, la antigua con 'v*' y la nueva con 'u*' en la letra gótica en contraste con la humanística. En la gótica, desciende el uso de 'u*' inicial casi constantemente, mientras que en la letra humanística, la nueva forma con 'u*' inicial avanza cada vez más hasta llegar a la mayoría absoluta en 1700 (71 %). Encontramos aquí la razón por la que se presenta la forma "U" en la distribución cronológica de la ratio (Fig. 12a). La forma "U" que hemos observado en la curva de la totalidad parece deberse a la disminución casi constante de 'u*' en la letra gótica, que ocurre en la Edad Media y a su aumento general en la letra humanística en la Edad Moderna.

Es bien sabido que en latín no existía la distinción entre 'u' y 'v', utilizándose solo la mayúscula 'V'. Y en la Edad Media, eran dos variantes de una misma letra. La variante medieval 'v' en 'vno' puede haber servido para darle mayor presencia al inicio de palabra, lo mismo que 'y', en lugar de 'i' en 'yglesia', 'ynfante', etc⁹. De ahí, la larga vacilación entre las dos grafías representantes de

⁸ Sobre los casos peculiares de 'un*' en 1200 de documentos cancillerescos, Sánchez-Prieto Borja (2004:437) contrasta el reinado de Fernando III (1217-1252) con 'u' y Alfonso X (1252-1284), con 'v'. Según Martín Aizpuru (2020: 223), el infante Sancho (1277-84, rey 1284-95) mandó expedir documentos con las formas de los numerales con 'v-'.

⁹ Juan de Valdés en su *Diálogo de la lengua* (hacia 1535) comentaba que él usaba 'v-' inicial "por ornamento de la escritura" (edición de Lope Blanch, 1969:

la vocal /ú hasta la última decisión definitiva adoptada por la Real Academia Española en el *Diccionario de Autoridades* (1726-1739) con la 'u' para la vocal y 'v' para la consonante (Torrens Álvarez 2007: 177). Esta solución no sobrevino repentinamente en el siglo XVIII, sino que había ido adoptándose paulatinamente a través de tiempo en la letra humanística, libre de la larga tradición gótica, donde mantenía el uso tradicional destacado de 'v' al inicio de palabra.

2.2.3. El verbo «haber»

En esta sección, nos enfocamos en la posición inicial de múltiples formas del verbo «haber». La siguiente lista de las frecuencias de cada grafía en orden descendente muestra que históricamente ha sido más frecuente el inicio sin 'h' ('V*-', frecuencia total: 4541) que con ella ('h*-', 1360):

'V*-' (frecuencia): a (554), aya (346), an (342), auer (338), auemos (168), auia (137), ay (137), ayan (121), aver (112), av<er> (103), e (82), auiendo (57), aber (56), auja (54), oujere (50), oviere (45), auian (44), auedes (44), avia (41), ayades (41), aviendo (40), ouiere (40), ya (39), au<er> (38), aujdos (37), oujere<n> (32), avemos (30), ayad<e>s (29), avidos (28), ouiesse (28), abia (27), aued<e>s (24), auie (24), ayamos (22), ai (21), obiere (21), ouiessen (21), ye (21), auie<n> (20), aveys (20), aue<r> (20), aujan (20), abra (20), ouiere<n> (20), ..., Total (4541).

'h*-' (frecuencia): ha (345), he (209), han (180), hauer (122), hauia (58), hauiendo (42), ha<n> (39), huuiere (37), hemos (35), hauian (20), haber (20), habia (19), haver (15), huuiese (15), h<e> (12), havia (10), haviendo (10), hubiere (8), habra (7), haia (7), hay (7), habiendo (6), huuieren (6), haya (6), hà (6), haidos (6), huuiesen (6), haido (5), hubiese (5), hayan (5), haueis (5), huuiesse (4), havian (4), hubo (4), havra (3), habidos (3), habria (3), huviese (3), huviesen (3), hauemos (3), haura (3), ..., Total (1360).

El cambio definitivo se produjo en 1750 como muestra el siguiente gráfico (Fig. 14a):

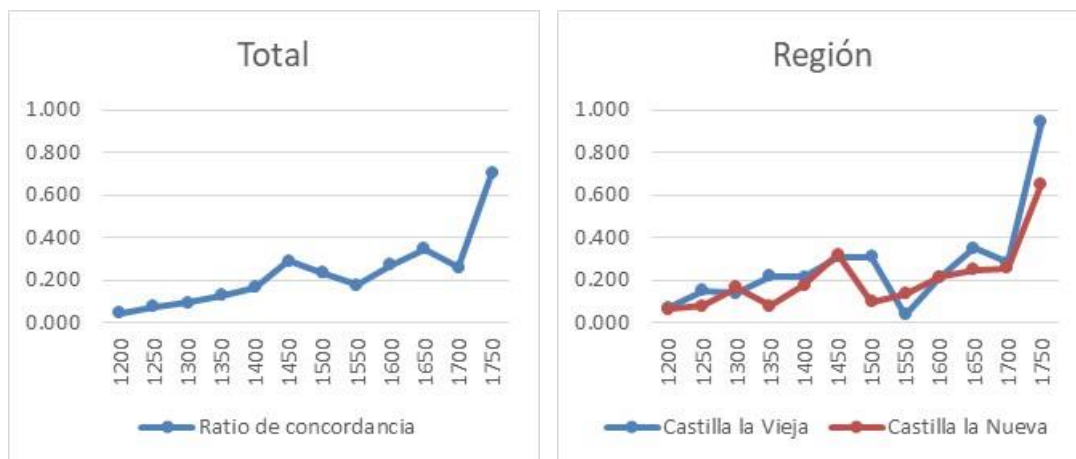


Fig. 14. Ratio de la forma 'h*' en «haber». (a) Total. (b) Región.

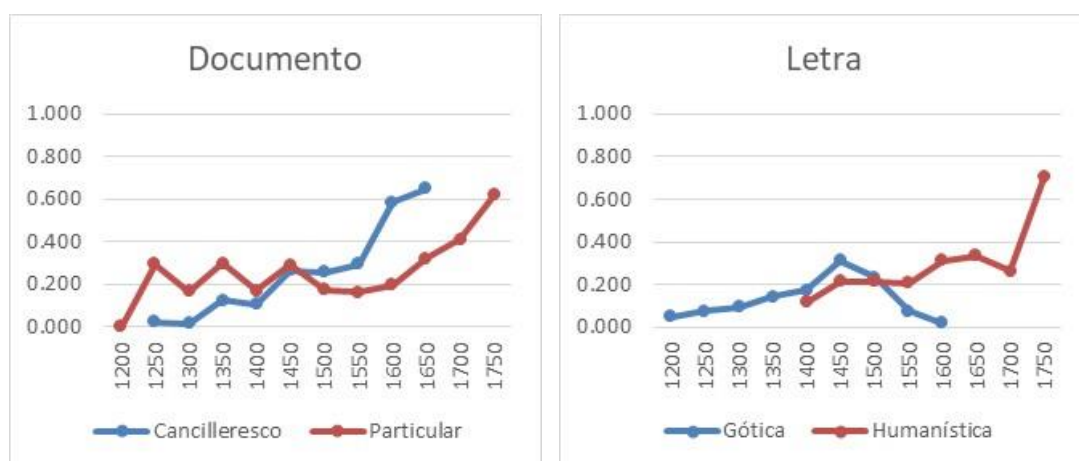


Fig. 15. Ratio de la forma 'h*' en «haber». (a) Documento. (b) Letra.

En los gráficos de las dos regiones (Fig. 14b), destacamos la coincidencia temporal del cambio de supremacía en 1750.

En cuanto al tipo de documentos (Fig. 15a), ambos tipos, el cancilleresco y el particular, ascienden, aunque con subidas y bajadas en la Edad Media. A partir de 1500, la subida es constante en los dos tipos documentales.

El factor Letra (Fig. 15b) parece coincidir con la tendencia general (Fig. 14a). La letra gótica sigue indecisa siempre, mientras que la letra humanística se decanta en favor de la hache inicial. La subida repentina destacada en 1750 se debe, probablemente, a la ortografía normativa de la Real Academia Española.

Por otra parte, en las dos listas ('V*' y 'h*') observamos una cierta heterogeneidad dentro del mismo paradigma verbal de «haber». Concretamente, advertimos una alta frecuencia de 'ha' y 'he' en la Edad Media, cuando el uso general del paradigma se presentaba sin 'h-' inicial, por ejemplo, 'auer' o 'aya'.

Al respecto, Juan de Valdés, al responder a la pregunta del porqué de la 'h' inicial de 'ha', y posteriormente de 'he' (ibid., 97), escribía en su libro, *Diálogo de la lengua* (hacia 1535) lo siguiente (Lope Blanch 1969: 75-76):

Marcio: (...) Agora me dezid por qué unas vezes escrivís *a* con *h* y otras sin ella.

Valdés: Por hazer diferencia de quando es verbo o quando es preposición, y assí siempre que es verbo la escrivo con *h* y digo: *Quien ha buen vezino, ha buen maitino*, (...); y quando es preposición escrivola sin *h*, diziendo *A buen callar llaman Sancho*, (...) (75-76)
(...)

Valdés: (...) haziéndose enemigos de la *h*, ninguna diferencia hazen entre *e* quando es conjunción, y *he* quando es verbo, porque siempre la esriven sin *h*, en lo qual, como os he dicho del *ha*, yerran grandemente. (97)

Efectivamente, esta distinción entre *he*, *ha* del verbo, por una parte, y la *e* (conjunción) y *a* (preposición), por otra, por la que aboga Valdés personalmente, parece manifestar cierta tendencia general en la Edad Media, como demuestran los siguientes gráficos:

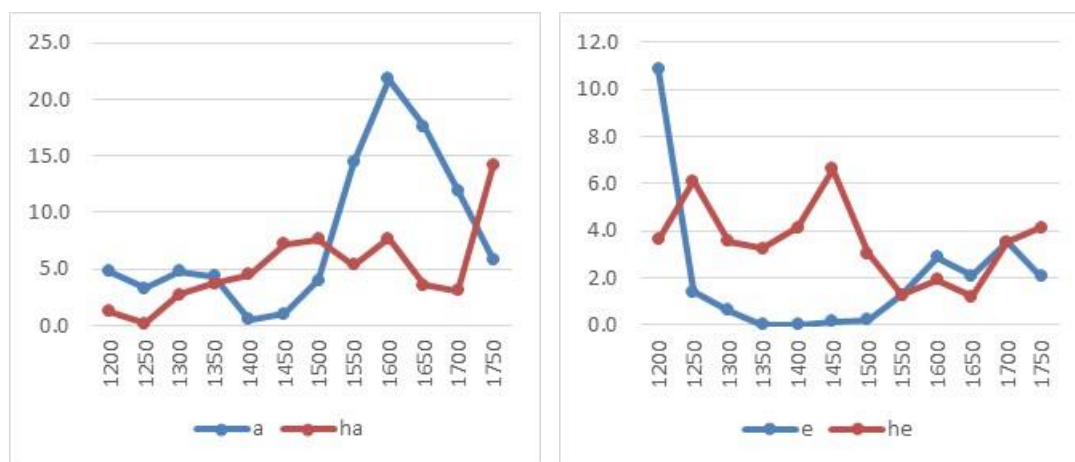


Fig. 16. Formas verbales del verbo «haber». (a) *a* (prep.) ~ *ha*. (b) *e* (conj.) ~ *he*. Frecuencia normalizada por 10 000 palabras.

Es interesante el uso mayoritario de 'he' en la Edad Media, con su extensión a 1500 (1500-1549) por la que Valdés manifestaba su preferencia, cuando la forma 'e' (conjunción) presentaba una frecuencia importante (sec. 2.2.1). En cambio, el caso de 'ha' es un tanto diferente, pues compite con la

preposición 'a' de altísima frecuencia a lo largo de historia.

En este punto, conviene comparar estos gráficos con los de *an* ~ *han* y *a[uv]er* ~ *ha[uv]er*:

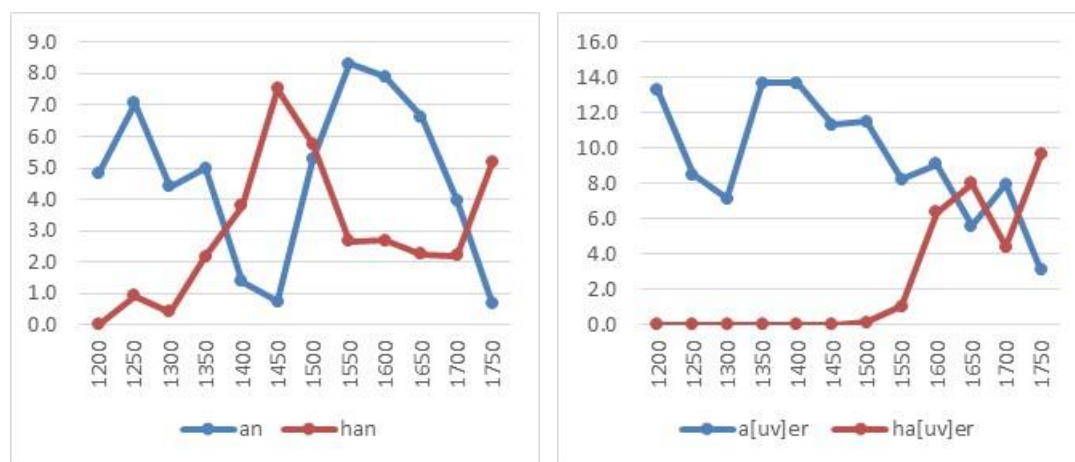


Fig. 17. Formas verbales del verbo «haber».

(a) *an* ~ *han*. (b) *a[uv]er* ~ *ha[uv]er*.

Frecuencia normalizada por 10 000 palabras.

A nuestro modo de ver, los comportamientos numéricos de 'an' y 'han' (Fig. 17a) se parecen a los de 'a' y 'ha' (Fig. 16a) a lo largo de tiempo, probablemente debido a que '(h)an' pertenece al paradigma verbal de presente de indicativo en tercera persona lo mismo que la forma '(h)a', con única diferencia formal de la '-n' final y la funcional de número (singular - plural). Parece bastante natural que haya una analogía entre la forma singular y la plural, ambas dotadas de alta frecuencia.

Para explicar el porqué de la 'h-' inicial para el verbo «haber», además de su etimología latina (lat. HABERE), hay que tener en cuenta su articulación muda, que no cambia la fonética, y la posición inicial, el lugar donde se manifiesta la función demarcativa por excelencia, igual que en 'y' inicial ('yglesia', 'ynfante') y 'v-' inicial ('vno', 'vso'), tratadas en las secciones anteriores (2.2.1, 2.2.2). La Real Academia Española y la Asociación de Academias de la Lengua Española (2010: 142) explican que "la mayoría de las haches que aparecen en palabras españolas se deben a la presencia de esta letra en su étimo latino". Según Marcos García (2012: 151), la reconstrucción de la 'h-', perdida en la Edad Media, se atribuye a la "reacción erudita de los renacentistas, amantes del latín". Por nuestra parte, creemos que también es importante reconocer, como una de las causas de la aparición de la hache muda,

el uso distintivo de 'ha' y 'he' del verbo frecuente «haber», frente a 'a' (preposición) y 'e' (conjunción), como afirmaba Juan de Valdés en el siglo XVI, con añadidura de la forma analógica 'han' (< 'ha'), seguidos de la posterior generalización de otras formas del mismo paradigma verbal. Creemos que la decisión académica en el siglo XVIII fue adoptada no solo por la etimología latina, sino también por la larga competencia castellana entre 'h*-' y 'V*-' librada en los siglos anteriores.

2.3. Formas gráficas generales

En esta sección, vamos a analizar los aspectos paleográficos generales en formas unidas, formas abreviadas y formas con mayúscula y minúscula. Aprovechamos la distinción establecida en el corpus entre transcripción paleográfica y presentación crítica para observar sus correspondientes variaciones gráficas.

2.3.1. Forma unida y separada

Para analizar la ortografía léxica, es fundamental establecer la unidad léxica como una forma separada. La realidad, sin embargo, es que en paleografía aparecen numerosas formas unidas, que se hallan separadas en la presentación crítica y también en la ortografía moderna. Nos interesa saber cuándo y cómo se separaron léxicamente formas paleográficas que con anterioridad existían unidas.

Sánchez-Prieto (1998: 158-170) trata ampliamente las formas paleográficas unidas y separadas a partir de la clasificación de palabras en elementos gramaticales y elementos léxicos: (a) dos elementos gramaticales, por ej. *empós (en pos)*, *porque (por que)*, *porqué (por qué)*; (b) elemento gramatical + elemento léxico, por ej. *apenas (a penas)*; (c) dos elementos léxicos, por ej. *seiscientos (seis cientos)*. Por nuestra parte, para llevar a cabo un tratamiento cuantitativo, hemos llevado a cabo establecido una distinción simple entre la forma unida discordante (es decir, la forma paleográfica diferente de la moderna) y la forma separada concordante (la forma paleográfica con separación idéntica de la moderna). La siguiente lista pertenece a las formas unidas más frecuentes:

Formas unidas más frecuentes: de-la (3126), de-los (1401), d<e>-la (988), de-las (893), en-la (881), a-los (816), a-la (736), d-esta (490), q<ue>-lo (476),

En-el (455), de-lo (428), d<e>-los (415), en-<e>l (388), d-<e>l (316), q<ue>-los (301), d-el- (288), de-la- (278), a-uos (275), d-ella (266), q<ue>-la (248), en-las (244), d<e>-las (240), q<ue>-les (235), a-las (225), d-ellos (203), en-los (197), a-lo (181), en-esta (165), d-ello (164), d-este (151), de-su (150), a-cada (145), e<n>-la (143), d-ellas (139), de-uos (138), A-todos (138), so-la (128), a-mj (126), En-ella (124), d<e>-lo (116), q<ue>-las (108), a-vos (108), so-pena (107), en-la- (107), d-esto (107), se-le (106), de-el (104), ..., Total (48 065).

La supremacía numérica de las formas unidas de preposición + artículo es notable. También es importante la conjunción 'que' + artículo. Ambas pertenecen al grupo (a) de Sánchez-Prieto (1998: 160-163).

Al realizar la búsqueda exhaustiva se observa el predominio de frecuencia de las formas separadas por encima de las unidas, como muestra la Fig. 18a con la ratio de concordancia siempre superior a 0.5, es decir, históricamente las formas concordantes (formas separadas) siempre son mucho más frecuentes que las discordantes. En su cronología advertimos la curva descendente hasta 1550 y a partir de 1600, la curva asciende, lo que significa que en la Edad Media predomina la tendencia cada vez más grande de la forma unida, mientras que en la Moderna, se prefiere cada vez más la forma separada:

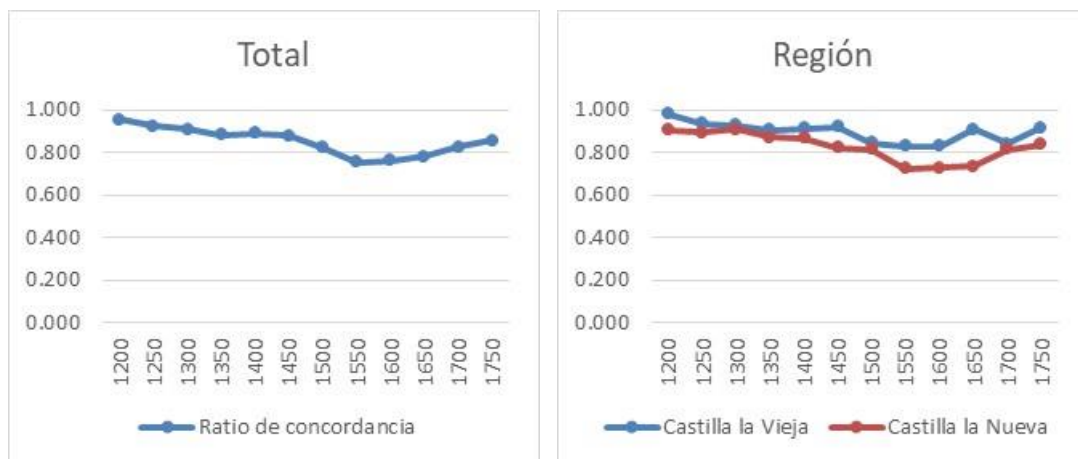


Fig. 18. Ratio de la forma separada. (a) Total. (b) Región.

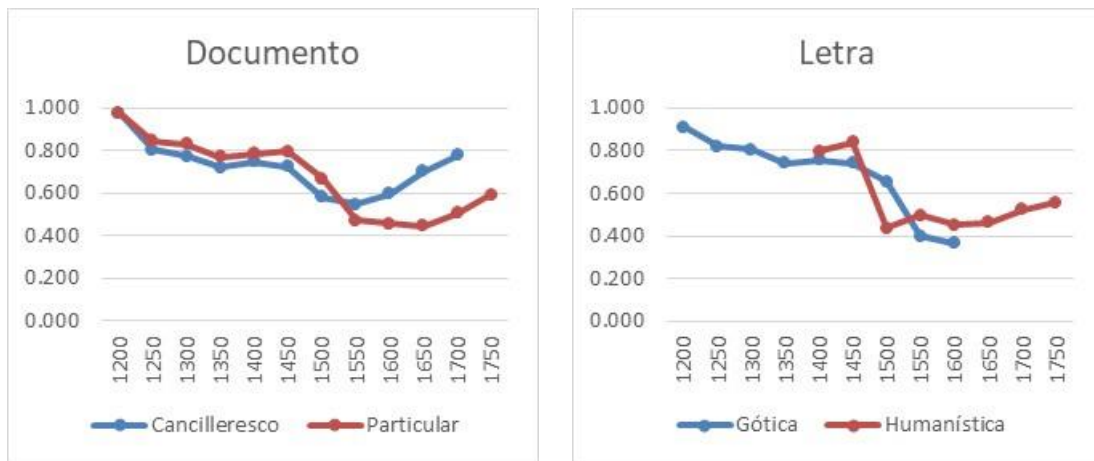


Fig. 19. Ratio de la forma separada. (a) Documento. (b) Letra

En el gráfico de la Región (Fig. 18b), observamos que la forma separada (forma concordante) es más frecuente en Castilla la Vieja que en Castilla la Nueva. En cuanto al Documento, a partir de 1600 se nota vacilación de la forma separada en los documentos cancillerescos. Se presenta la tendencia histórica de manera más destacable en el factor de la letra (Fig. 19b), donde apreciamos con claridad la curva descendente en la letra gótica y la ascendente en la letra humanística a partir de 1600. Por consiguiente, el factor de Letra parece ser causante importante de la forma moderna separada.

2.3.2. Forma abreviada y plena

Al comparar los documentos del corpus a través del tiempo, se nota la abundancia de las formas abreviadas (formas discordantes) en la Edad Media (Sánchez-Prieto 1998: 91-95), frente al aumento de las formas plenas (formas concordantes) en la Edad Moderna, como muestran las siguientes listas y el gráfico (Fig.11a):

Forma abreviada (frecuencia): q<ue> (15669), d<ic>ho (3488), d<ic>ha (2540), d<e> (1993), no<n> (1666), nj<n> (1598), p<ar>a (1541), co<n> (1384), n<uest>ro (1216), <e>l (1211), d<ic>hos (1092), n<uest>ra (912), sen<n>or (861), an<n>o (804), v<uestra> (802), an<n>os (732), c<art>a (699), p<ar>te (605), d<ic>has (578), e<n> (566), dich<o> (501), porq<ue> (473), q<ua>l (467), m<a>r<avedi>s (464), aq<ui> (456), man<er>a (454), bie<n> (450), n<uest>ros (446), R<eales> (418), d<e>l (417), d<on> (402), joh<a>n (386), q<ua>nto (385), egl<es>ia (380), ..., Total: 126 183.

Forma plena (frecuencia): de (45616), & (24688), y (19586), el (16626), la (16619), en (15656), a (13262), por (10398), los (9429), que (9209), E (7082), d (6330), lo (6062), las (5460), se (5195), esta (4229), su (4181), o (3581), con (3251), como (2976), del (2608), dicho (2526), yo (2433), mj (2431), es (2412), nos (2284), don (2239), al (2084), no (2076), para (2041), todo (1891), vos (1872), dicha (1828), et (1776), carta (1736), ..., Total: 536 133.

Entre las dos regiones, se observa un movimiento numérico parecido de la ratio de concordancia a lo largo de tiempo (Fig.11b):

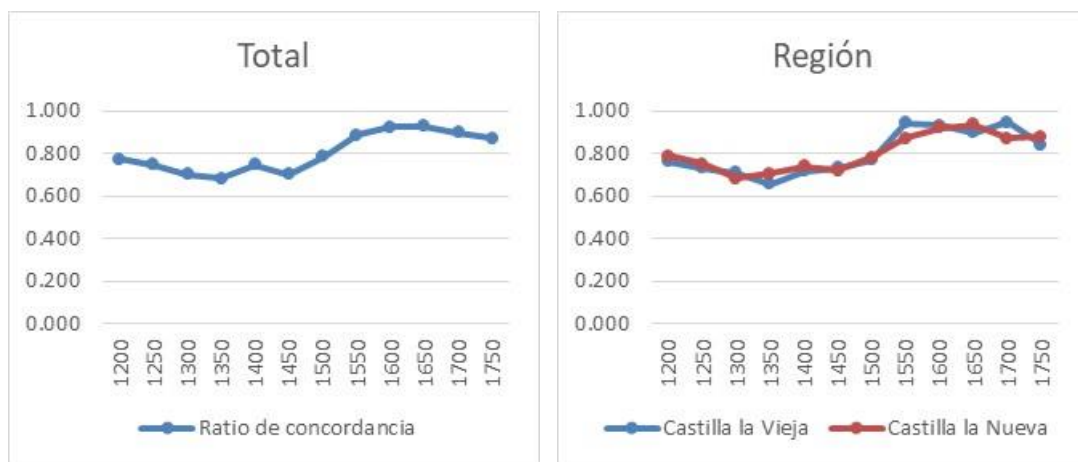


Fig. 20. Ratio de la forma plena. (a) Total. (b) Forma plena. Región.

Tampoco se presenta un gran contraste entre los dos tipos documentales, el cancilleresco y el particular (Fig. 21a). En cambio, observamos una gran diferencia en las curvas cronológicas de las letras (gótica y humanística), puesto que la subida de la forma plena (concordante) en 1450 en la letra humanística es destacable, y todavía más en contraste con la bajada de la letra gótica en la misma franja cronológica (Fig. 21b):

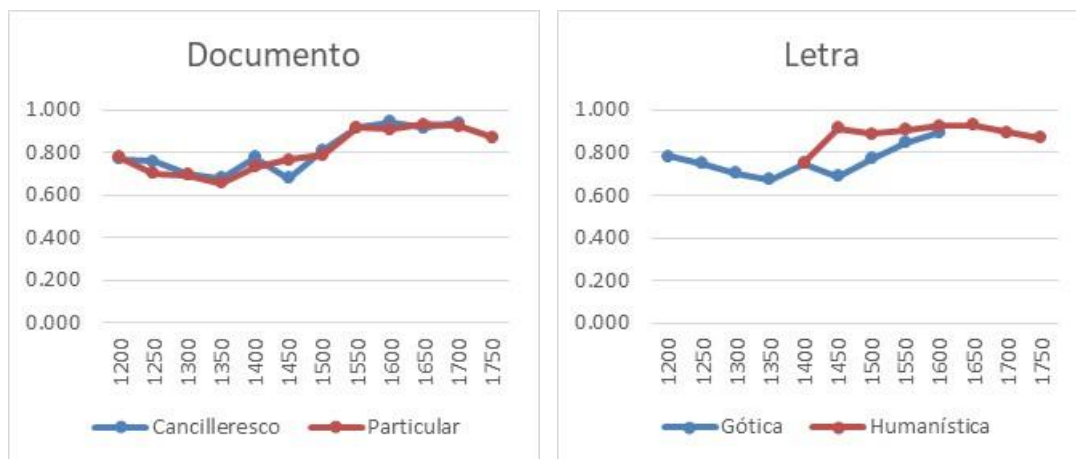


Fig. 21. Ratio de la forma plena. (a) Documento. (b) Letra.

Por lo tanto, la subida de la ratio de concordancia observada en el gráfico de la totalidad (Fig. 11a) parece deberse al uso de la letra humanística en 1450 (Fig. 12b). Las curvas de la ratio de concordancia muestran la forma de "S" plana (sec. 1).

2.3.3. Mayúscula y minúscula

La cuestión del uso de la mayúscula se divide en los nombres propios y en el inicio de una oración (Sánchez Prieto 1998: 170-174). En primer lugar, veamos los casos del nombre propio. La siguiente lista muestra los nombres propios más frecuentes con minúscula y con mayúscula, donde comprobamos más casos con minúscula que mayúscula en la historia:

Nombres propios más frecuentes con minúscula (frecuencia): perez (446), dios (381), johan (230), domingo (230), martin (211), martinez (211), pero (206), maria (194), juan (179), alfonso (147), garçia (141), diego (133), toledo (130), pedro (128), iohan (120), ferrandez (115), ..., Total (29 504).

Nombres propios más frecuentes con mayúscula (frecuencia): Juan (759), Toledo (298), Pedro (247), Maria (228), Leon (195), Castiella (190), Martin (188), Johan (182), Dios (169), Sancho (165), Cordoua (147), Seuilla (144), Alonso (136), Francisco (134), Jahen (133), Antonio (129), Domingo (121), Manuel (117), Pero (116), Gallizia (109), Rodrigo (106), ..., Total (13 148).

Tras comprobar la mayoría general de la minúscula en la historia,

veamos el movimiento concreto de la ratio de la mayúscula a lo largo de los siglos (Fig. 22a.), donde observamos la bajada casi general en la Edad Media y a partir de 1550 con su posterior aumento progresivo hasta llegar al cambio de mayoría en 1700. En cuanto a la región, tanto Castilla la Nueva como la Vieja presentan la misma curva en forma de letra "U" (Fig. 22b):

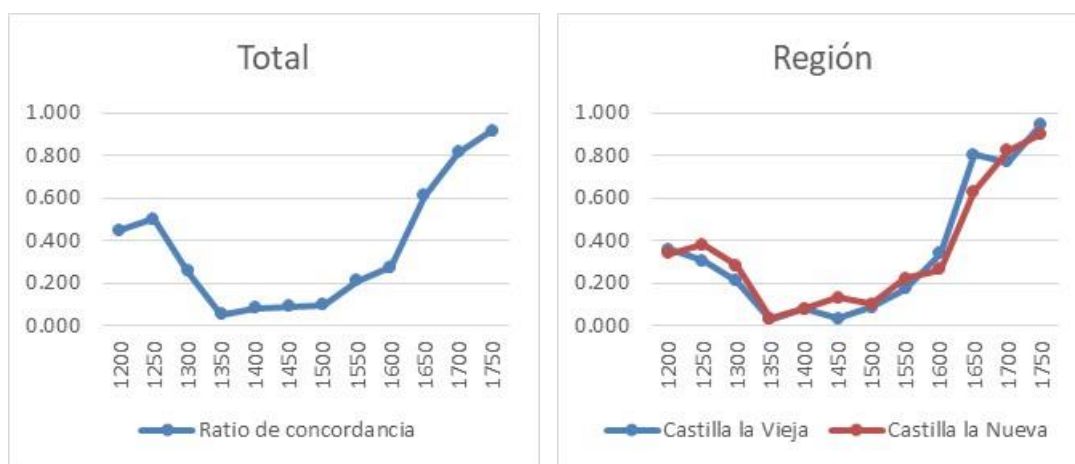


Fig. 22. Ratio de la mayúscula. Nombre propio. (a) Total. (b) Región.

También en el gráfico del Documento (Fig. 23a) presentan la misma tendencia el documento cancilleresco y el particular, pues en ambos casos se observa la curva en forma de "U". ¿A qué se debe la curva en forma de "U" que hemos observado en los tres gráficos? La respuesta a esta cuestión parece estar en el siguiente gráfico de la Letra (Fig. 23b), donde observamos el descenso general de la ratio en la letra gótica en contraste con el ascenso uniforme de la letra humanística. Es probable que la curva de forma "U" se haya presentado debido a la distribución cronológica de la mayúscula cada vez más disminuida en letra gótica mayoritariamente en la Edad Media, y el aumento de la misma en la letra humanística en la Edad Moderna. Esta diferencia es importante como factor que explica la tendencia general de la mayúscula en el nombre propio.

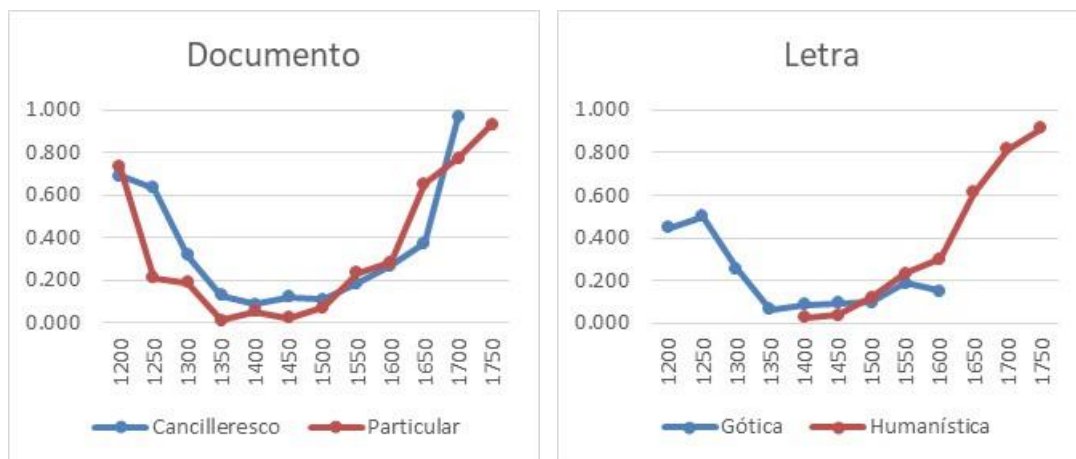


Fig. 23. Ratio de la mayúscula. Nombre propio. (a) Documento. (b) Letra.

Ahora bien, veamos otro uso de mayúscula, pero no correspondiente a nombres propios, en posición inicial de oración. Nos concentramos en las formas paleográficas correspondientes a la forma crítica con mayúscula. Los ejemplos más frecuentes son los siguientes:

Otras formas con minúscula (frecuencia): & (519), y (494), yo (335), nuestro (291), yten (278), sant (266), santa (245), testigos (193), sennor (144), san (120), a- (117), por (99), señor (93), ccc (92), en (85), dada (84), sancta (82), y- (81), nuestra (76), que (75), lo (74), fecha (73), e (73), ffecha (72), santo (70), de (69), el (63), sepades (58), esta (58), la (58), ..., Total (9605).

Otras formas con mayúscula (frecuencia): E (1083), Et (1052), Don (666), Sepan (269), En (242), Y (214), Dada (170), El (152), Por (132), Yo (111), E- (94), Rey (93), EN- (90), Mas (88), Remato (83), ego (81), La (81), Yten (77), San (76), Y- (68), Señor (68), Vuestra (61), Sant (55), Dos (55), Real (55), A (53), Magestad (52), Sepades (52), Otrossi (52), ..., Total (9592).

En el gráfico de Total (Fig. 24a) se observa la frecuencia de mayúscula al inicio y al final, con una curva de forma "U", parecida a la de los nombres propios (Fig. 22a). Las dos regiones presentan las curvas de la ratio de la mayúscula un tanto diferentes, pero la tendencia general es la misma, en forma de "U" (Fig. 11b):

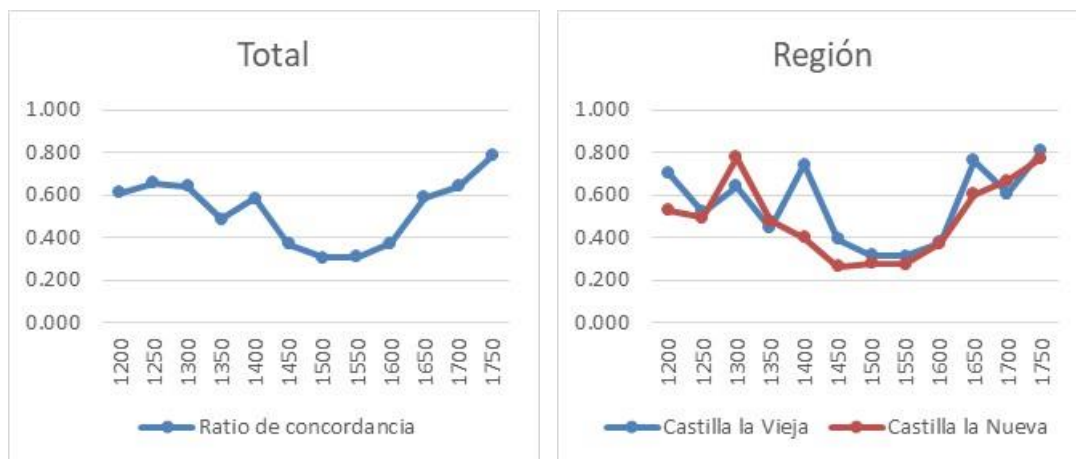


Fig. 24. Ratio de la mayúscula. Otras categorías. (a) Total. (b) Región.

Respecto al Documento, también se observa la misma tendencia (Fig. 25a). En la distinción de Letra, la humanística toma la iniciativa con respecto a la mayúscula en 1500, mientras que en la gótica descende paulatinamente hasta desaparecer casi por completo (Fig. 12b):

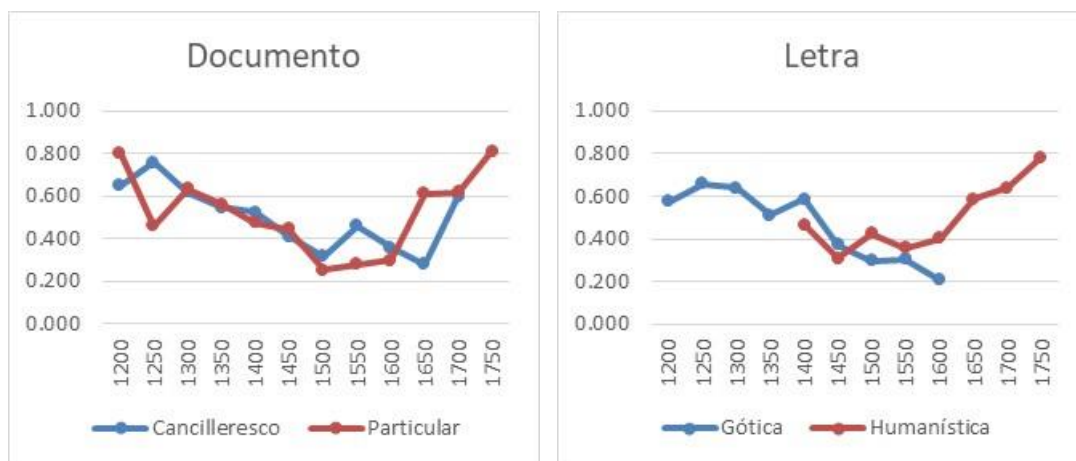


Fig. 25. Ratio de la mayúscula. Otras categorías. (a) Documento. (b) Letra.

Al comparar los gráficos del nombre propio (Fig. 22, 23) con los de otras categorías (Fig. 24, 25), notamos una similitud grande en las curvas, lo que puede indicar la no distinción del nombre propio como categoría caracterizadora de la mayúscula, a lo largo de historia. También es interesante observar las curvas en forma de letra "U", con altas frecuencias al inicio y al final de la cronología tratada y la baja frecuencia en el intermedio sin hacer distinción entre el nombre propio y otras categorías.

Es difícil explicar la curva en forma de "U" desde el punto de vista de la

región y del documento, puesto que no encontramos las razones de tal cambio ni en la región ni en el documento. Es más, no se han presentado diferencias de la ratio de mayúscula entre los miembros del mismo factor. En cambio, desde el punto de vista de la letra, todo es plausible dada la tendencia descendente en la gótica y la ascendente en la humanística. La letra gótica se caracteriza por la indistinción de formas mayúsculas y minúsculas manteniendo la misma presentación gráfica con el criterio tradicional de la escritura. La humanística, por el contrario, buscaba la forma idónea prestando atención a la lectura.

Aparentemente, la curva total en forma de "U" destruiría la hipótesis de periodización progresiva (Hipótesis-1, sec. 1). En realidad, desde el punto de vista de las dos letras, sin embargo, la historia de la mayúscula se vuelve progresiva tanto en la gótica como en la humanística: en la letra gótica, progresiva negativa, y en la letra humanística, progresiva positiva, lo cual históricamente es natural y razonable.

2.4. Totalidad

En las secciones anteriores, hemos observado cambios cronológicos de la ortografía léxica divididos en formas léxicas específicas (sec. 3.2) y en formas gráficas generales (sec. 3.3). En este punto conviene ver la totalidad de los cambios históricos de la ortografía castellana. Para ello, hemos preparado los siguientes gráficos, en los cuales se observa el cambio de la ratio de concordancia exhaustiva de todas las formas léxicas, sin hacer distinción de las formas léxicas específicas y las formas gráficas generales:

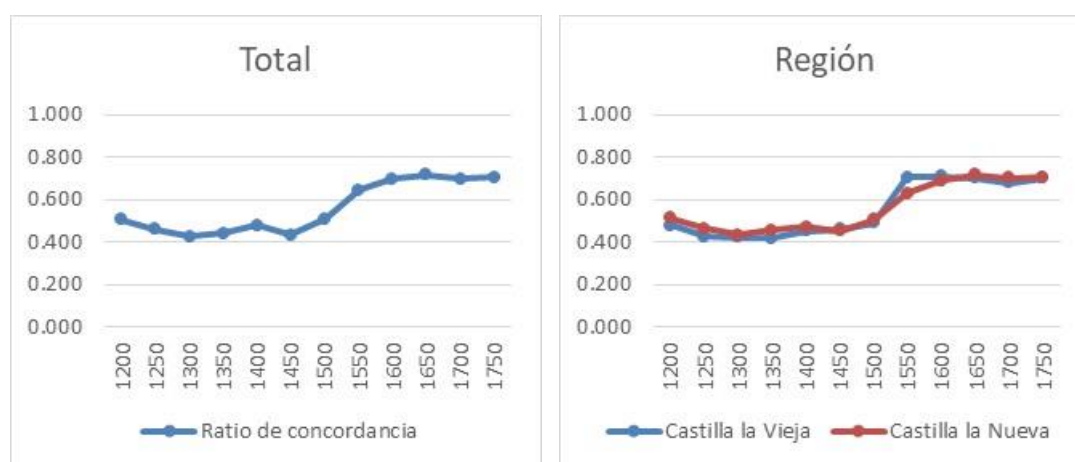


Fig. 26. Ratio de concordancia exhaustiva. (a) Total. (b) Región.

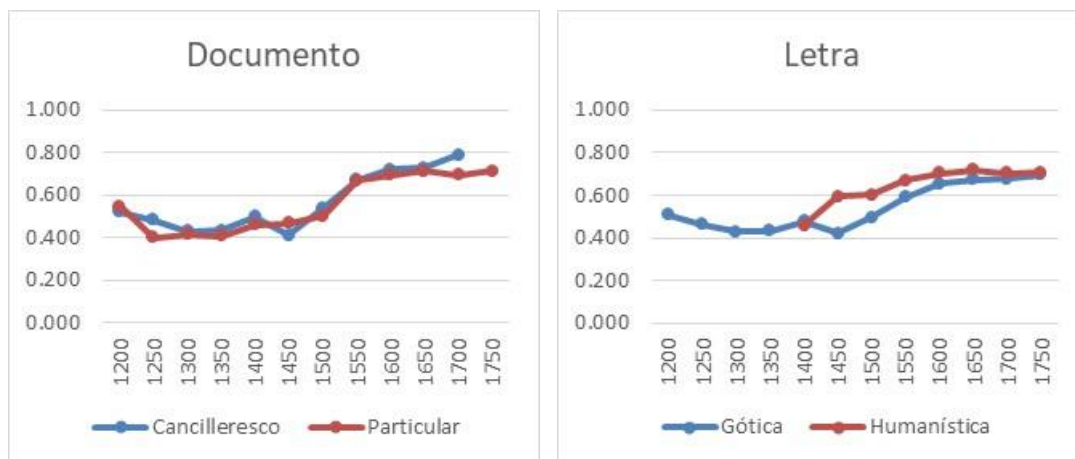


Fig. 27. Ratio de concordancia exhaustiva. (a) Documento. (b) Letra.

En los tres primeros gráficos, se observa la curva típica de "S" plana, con una relativa estabilidad en la parte inicial baja (1200-1450), con el segmento intermedio de subida rápida (1500-1550) y con la parte final otra vez estable (1600-1750). De esta manera se comprueba la validez de la teoría del cambio histórico general (sec. 1), aunque hemos visto casos concretos que no muestran la misma curva de "S" plana (sec. 2.2.2, 2.2.3, 2.3.1, 2.3.3).

Por otra parte, es importante destacar la peculiaridad de las curvas de la letra, donde encontramos un notable contraste entre la letra gótica y la humanística (Fig. 27b). Según el gráfico, la letra humanística avanza más en la modernización que la gótica desde 1450 en adelante. Esto es muy significativo históricamente.

La siguiente lista muestra las frecuencias de las formas concordantes con las formas modernas:

Forma concordante (categoría gramatical) frecuencia: <el> (art) 48077, <de> (prep) 47760, <y> (conj) 19302, <en> (prep) 15626, <a> (prep) 12294, <por> (prep) 10372, <que> (conj) 9110, <lo> (pro) 8834, <ser> (v) 7214, <este> (dem) 7060, <decir> (v) 6400, <su> (pos) 5921, <se> (pro) 3819, <o> (conj) 3483, <él> (pro) 3280, <con> (prep) 3248, <dar> (v) 2755, <yo> (pro) 2427, <todo> (indef) 2306, <como> (conj) 2305, ..., Total: 369 851.

En contraste, disponemos de las formas discordantes con las modernas, que han experimentado cambios ortográficos desde el estado inicial en la franja correspondiente hasta la forma final moderna, incluidas la abreviatura, los acentos y la distinción entre mayúsculas y minúsculas. Consideramos que las

frecuencias de las formas discordantes representan la magnitud de los cambios históricos de manera cuantitativa:

Forma discordante (categoría gramatical) frecuencia, & (conj) 24672, q<ue> (conj) 15599, e (conj) 6959, d (prep) 6323, d<ic>ho (v) 3471, d<ic>ha (v) 2527, d<e> (prep) 1987, et (conj) 1775, mj (pos) 1726, no<n> (adv) 1666, p<ar>a (prep) 1540, nos (pro) 1471, vna (art) 1403, co<n> (prep) 1382, mas (adv) 1362, l (art) 1343, vn (art) 1248, n<uest>ro (pos) 1216, <e>l (art) 1143, nj<n> (adv) 1098, d<ic>hos (v) 1080, mill (num) 1046, uos (pro) 1031, ..., Total: 292 465.

En más de la mitad de los casos de la siguiente tabla se observa coincidencia entre la transcripción paleográfica y la forma moderna (53 %), aunque los casos discordantes tampoco son desdeñables (47 %).

*	Con.	Dis.	Suma	RC: Con./Suma
Total	358 470	315 310	673 780	0.532

Tabla 5. Casos concordantes (Con.), discordantes (Dis.) y ratio de concordancia (RC).

Ahora nos interesa la ratio de concordancia (RC) en relación con los factores de la cronología (Edad Media, Edad Moderna)¹⁰, la región (Castilla la Vieja y Castilla la Nueva), el tipo de documento (documento cancilleresco y documento particular) y la letra (gótica y humanística) (Tabla 6). Para ver el grado de influencia del factor, reordenamos la tabla en la columna de RC. Según la misma tabla, el factor que más favorece la discordancia es la letra gótica, y el factor que más prefiere la concordancia es la letra humanística. Entre estos dos factores extremos, se encuentran distintos factores intermedios con sus correspondientes valores de RC:

¹⁰ Tras comprobar la relativa estabilidad de RC desde 1200 (1200-1249) hasta 1450 (1450-1499), agrupamos este período en la denominación aproximada conveniente de "Edad Media" y desde 1500 (1500-1549) hasta 1750 (1750-1799) en la de "Edad Moderna".

Factor	Con.	Dis.	Suma	RC: Con./Suma
L. gótica	192 570	180 901	373 471	0.516
E. media	201 822	187 609	389 431	0.518
C. la Vieja	153 895	137 925	291 820	0.527
C. la Nueva	200 138	173 209	373 347	0.536
D. cancilleresco	95 577	78 926	174 503	0.548
E. moderna	116 243	95 260	211 503	0.550
D. particular	156 648	127 701	284 349	0.551
L. humanística	136 677	107 704	244 381	0.559

Tabla 6. Casos concordantes (Con.), discordantes (Dis.). Factor separado

Al comparar los factores dentro de la misma categoría (región, tipo documental, letra), se reconocen dos grupos fundamentales, como muestra la Tabla 7. Según la tabla anterior (Tabla 6), la diferencia entre la letra gótica y la humanística es extrema, mientras que la existente entre las dos regiones es escasa, lo que explica el poder explicativo que posee el factor Letra y la presentación contradictoria en el factor Región en algunas ocasiones que hemos observado:

*	Grupo-1. Discordante	→ Grupo-2. Concordante
Cronología	Edad Media	→ Edad Moderna
Región	Castilla la Vieja	→ Castilla la Nueva
Documento	Cancilleresco	→ Particular
Letra	Gótica	→ Humanística

Tabla 7. Grupo-1. Discordante; Grupo-2. Concordante.

En vez de calcular la ratio de concordancia en cada factor separadamente (Tabla 7), también es posible calcularla con los factores combinados, como muestra la siguiente tabla (Tabla 8)¹¹:

¹¹ No se han presentado casos en las siguientes combinaciones:
E. media / C. la Vieja / D. cancilleresco / L. humanística
E. media / C. la Nueva / D. cancilleresco / L. humanística
E. media / C. la Nueva / D. particular / L. humanística

Cronología	Región	Documento	Letra	Con.	Dis.	Suma	RC: C/S
E. media	+C. la Nueva	+D. cancilleresco	+L. gótica	17 765	19 329	37 094	0.479
E. moderna	+C. la Nueva	+D. particular	+L. gótica	4 710	4 543	9 253	0.509
E. media	+C. la Vieja	+D. particular	+L. gótica	22 037	19 190	41 227	0.535
E. media	+C. la Nueva	+D. particular	+L. gótica	4 556	3 947	8 503	0.536
E. moderna	+C. la Vieja	+D. cancilleresco	+L. gótica	6 943	5 878	12 821	0.542
E. moderna	+C. la Nueva	+D. particular	+L. humanística	40 610	33 921	74 531	0.545
E. media	+C. la Vieja	+D. cancilleresco	+L. gótica	49 188	40 499	89 687	0.548
E. moderna	+C. la Vieja	+D. particular	+L. gótica	5 184	4 019	9 203	0.563
E. moderna	+C. la Vieja	+D. cancilleresco	+L. humanística	6 000	4 425	10 425	0.576
E. moderna	+C. la Vieja	+D. particular	+L. humanística	7 916	5 691	13 607	0.582
E. moderna	+C. la Nueva	+D. cancilleresco	+L. gótica	7 900	5 455	13 355	0.592
E. moderna	+C. la Nueva	+D. cancilleresco	+L. humanística	24 266	15 868	40 134	0.605
E. media	+C. la Vieja	+D. particular	+L. humanística	1 630	592	2 222	0.734

Tabla 8. Casos concordantes (Con.), discordantes (Dis.). Factor combinado.

En esta tabla se presentan las cifras de RC (ratio de concordancia) de manera más sensible, es decir, con el rango más amplio [0.479-0.734]. Entre los cuatro factores de variación, el de Letra (gótica, humanística) es más coherente en su distribución de la ratio de concordancia, entre altas cifras de RC y la letra humanística, por una parte, y entre bajas cifras de RC y la letra gótica, por otra. Otros factores (cronología, región, documento) son menos coherentes, en el sentido de que no se observa una concentración destacable.

Por consiguiente, dentro de los cuatro factores tratados, la letra es más importante para determinar la ratio de concordancia, índice de la modernización, lo que también hemos comprobado en las secciones anteriores (2.2, 2.3).

3. Consideraciones finales

3.1. Hipótesis

Hemos llegado al momento de presentar conclusiones sobre las dos hipótesis establecidas en la sección 2:

Hipótesis-1: La historia de los cambios de las formas ortográficas léxicas obedece a una periodización progresiva dividida en tres etapas: Etapa inicial estable larga → Etapa de transición rápida → Etapa moderna estable larga.

Creemos haber demostrado *grosso modo* la validez de la hipótesis. Podemos establecer una periodización de cambios históricos de la ortografía española: etapa antigua (1200-1500), etapa transitoria (1500-1599), etapa moderna (1600-), de acuerdo con los cambios que presentan las variantes ortográficas numéricamente importantes. Nuestra periodización, basada en todas las formas ortográficas de palabras de CODEA (en Castilla), con una visión retrospectiva desde las formas modernas, muestra una peculiaridad que facilita una división principalmente bifásica, Edad Media y Edad Moderna con una etapa de transición de 100 años.

Esta observación general del cambio de totalidad no garantiza una homogeneidad en todos los aspectos interiores. Es decir, al lado de la curva de forma “S” (Fig. 10, 11, 20, 21, 26, 27), hemos observado también la curva de forma “U” (Fig. 12, 13, 22, 23, 24, 25) e incluso una curva descendente (Fig. 18, 19). Las curvas ascendentes monótonas en Fig. 14 y 15 pueden ser correspondientes a la parte inicial de la curva de forma “S”. Esta heterogeneidad de los cambios cronológicos es explicable por la razón de que los aspectos tratados en este estudio no guardan unas relaciones mutuas, sino que son independientes y peculiares. Encontramos la única tendencia común en la modernización estandarizada dirigida al estado actual.

Es interesante observar la coincidencia de la división tradicional entre la Edad Media y la Moderna y nuestra división tripartita con una etapa de transición: 1500-1599. El cambio radical que hemos observado en el análisis masivo exhaustivo se debe a la subida de la frecuencia total de las formas concordantes y/o a la bajada de la de las formas discordantes.

¿A qué se debe este cambio masivo que se presenta en la etapa transitoria de la centuria de 1500? Responder a esta pregunta fundamental equivale a comprobar la validez de nuestra segunda hipótesis:

Hipótesis-2: Los cambios históricos de la ortografía no son incidentes aleatorios o accidentales, sino naturales y razonables, dependientes de factores lingüísticos (gráficos) y extralingüísticos (históricos, geográficos y sociales).

Hemos considerado los factores lingüísticos (gráficos) de cada caso: las condiciones léxicas en la totalidad de las formas abreviadas, la condición del signo tironiano ('&') y de la fonética histórica de los diptongos crecientes en la

conjunción «y», el origen de 'u' latina y la distinción inicial propia de 'v-' en el artículo «un», así como la colisión homonímica y la analogía dentro del paradigma del verbo «haber». Estas condiciones lingüísticas son peculiares de cada caso.

Por otra parte, los factores extralingüísticos (históricos, geográficos y sociales) son comunes en todos los casos tratados. Desde el principio, los textos tratados guardan altas correlaciones dentro de cada uno de los grupos de factores: Edad Media, Castilla la Vieja, documento cancilleresco y letra gótica, por una parte (Grupo-1); y Edad Moderna, Castilla la Nueva, documento particular y letra humanística, por otra (Grupo-2), naturalmente con superposiciones intermedias transitorias. En efecto, hemos observado que, dentro de cada grupo, se manifiesta una cierta similitud interesante en las líneas de cambios cronológicos.

3.2. Imprenta y letra

Para aproximarnos a la estandarización de la lengua, los investigadores destacan la importancia histórica que supone la introducción de la imprenta en Europa en general (Febvre et Martin 1971); y en España en particular, en sus aspectos históricos y sociales: la pronta llegada de la imprenta en la Península¹², el elevado número de libros impresos, superior a Inglaterra con gran diferencia, la difusión de la imprenta en numerosas ciudades, antes de finalizar el siglo XV, que son "un justo índice de una creciente actividad intelectual" (Trevor Davies 1973: 44-45). La "creciente actividad intelectual" se observa precisamente en la tasa de alfabetización que expone Nalle (1989: 72), basándose en los documentos inquisitoriales en la diócesis de Cuenca clasificada por año de nacimiento (de hombres): 1463-1510: 9%, 1511-1530: 27%, 1531-1550: 36%, 1551-1570: 46%, 1571-1590: 54%, 1591-1637: 52%.

Goez (2007: 126) anota que "[l]os impresores y tipógrafos podían contribuir indirecta e informalmente a la estandarización de una variante, sobre todo si revisaban las obras más publicadas y más leídas con el fin de eliminar las variantes lingüísticas anticuadas o regionales". Sobre la comparación de la versión impresa de la *Gran conquista de Ultramar* (1503) con el manuscrito del s. XIII, estudiada por Harris-Northall (1996), Tuten y Tejedero-Herrero (2011: 297) comentan que "la versión impresa revela una amplia intervención editorial que redujo la variación en todos los aspectos, reflejando la ideología lingüística

¹² Según Odriozola (1982: 51), la primera época incunable hispana comienza hacia los años 1471-1472 en Segovia y en 1473-1474 en Barcelona y Valencia.

de humanistas influyentes como Nebrija" (traducción nuestra). Efectivamente en Harris-Northall (1996: 129-133) encontramos casos de estandarización en las formas plenas con respecto a las apocopadas originales: 'diz' > 'dize', 'tod' > 'todo'; formas no ensordecidas en la consonante en posición final de palabra: 'cipdat' > 'dibdad'; formas separadas de preposición y artículo: 'poral' > 'para el', 'entrel' > 'entre el', etc., todas ellas observadas en el castellano, que a principios del siglo XVI estaba en plena estandarización nacional de un Estado recién unificado (Harris-Northall, 1996: 144).

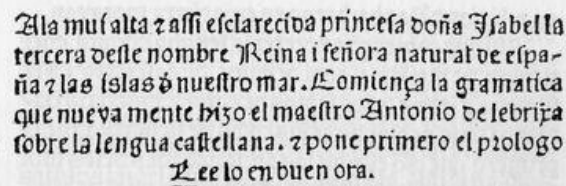
Por nuestra parte, proponemos relacionar la introducción de la imprenta con la selección de la letra humanística, puesto que la mayoría de los libros impresos son de la letra humanística a lo largo de historia, a excepción de la época inicial de la imprenta en la que encontramos libros impresos en la letra gótica¹³. Ong (1987: 122) contrasta la cultura del manuscrito, orientada hacia el productor, con abundantes abreviaciones que favorecen al copista, y la del impreso, orientada hacia el consumidor con un texto más legible. En este sentido, el cambio del manuscrito al impreso implica, fundamentalmente, el cambio de la letra al servicio del lector. Consideramos que también el traspaso de la letra gótica a la humanística está más orientado hacia el lector que el del manuscrito al impreso, lo que podemos comprobar en los dos impresos, uno en la letra gótica (Antonio de Nebrija, *Gramática de la lengua castellana*¹⁴, Salamanca, 1492) y otro en la humanística (Miguel de Cervantes, *El ingenioso hidalgo don Quixote de la Mancha*¹⁵, Madrid, 1605)¹⁶. Efectivamente, en el libro de Nebrija (Fig. 28) encontramos el signo tironiano en forma de '7' para la conjunción, propio de la letra gótica manuscrita, mientras que en el de Cervantes aparece la 'y' (Fig. 29), igual que el contraste observado anteriormente (sec. 2.2.1):

¹³ Según Ruiz (2000: 187), en los libros incunables (1472?-1500) también hubo cambio de la letra gótica a la letra humanística en la etapa de expansión de la imprenta (1491-1500). Para cotejar gráficamente los manuscritos y los libros desarrollados en Europa, veáanse Knight (1998) y Knight (2012).

¹⁴ Facsímil de Librerías París-Valencia (Valencia, sin fecha).

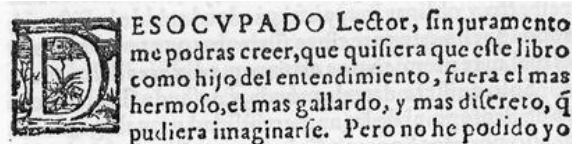
¹⁵ Facsímil de Junta de Castilla y León, Fundación Siglo, The Hispanic Society of America, Valladolid, 2001.

¹⁶ Don Quijote advierte a Sancho Panza: "no se la [carta dirigida a Dulcinea] des a trasladar a ningún escribano, que hacen letra procesada [letra procesal, variante tardía de la gótica], que no la entenderá Satanás." (cap. XXV), lo que demuestra lo ilegible que era la letra gótica para un intelectual de aquel entonces.



Ala muy alta z assi esclarecida princesa doña Isabella
tercera deste nombre Reyna i señora natural de espa-
ña z las islas de nuestro mar. Comiença la gramatica
que nueva mente hizo el maestro Antonio de lebrija
sobre la lengua castellana. z pone primero el prologo
Lee lo en buen ora.

Fig. 28. Nebrija, 1492, letra gótica.



ESOCVPADO Lector, sin juramento
me podras creer, que quisiera que este libro
como hijo del entendimiento, fuera el mas
hermoso, el mas gallardo, y mas discreto, q
pudiera imaginarse. Pero no he podido yo

Fig. 29. Cervantes, 1605, letra humanística

En relación con los factores, tanto lingüísticos como extralingüísticos, creemos que la introducción de la letra humanística en la sociedad castellana constituye uno de los aspectos importantes que produjo las transformaciones ortográficas. En cambio, resulta difícil establecer otro factor como productor principal de los cambios ortográficos, puesto que la Edad Moderna por sí sola no puede ser una causa de ellos, ni la condición geográfica, ni la tipología del documento. Tampoco el cambio del medio de comunicación escrita, del manuscrito al impreso, puede ser causa principal por sí solo. Estas condiciones, cronológicas, geográficas y sociales, necesitan otra condición determinante para impulsar los efectos ortográficos. Precisamente, encontramos la causa decisiva en la letra. Estamos estudiando los grafemas históricos, de modo que es natural y lógico tratar el tema desde el punto de vista de la letra. Pensamos que otras condiciones, de orden cronológico, geográfico y social, pueden ser causas coadyuvantes al lado de la causa principal, que es el cambio de la letra gótica a la humanística, libre de la tradición de escritura medieval.

La letra humanística nació, o más bien renació, en Florencia (Italia) por influjo de Francesco Petrarca (1304-1374), a partir de la letra carolina de los siglos X y XI, en rechazo de la letra gótica, usual en aquel entonces. Citamos la explicación histórica dada por Romero Tallafigo (2022):

Petrarca maldijo a esta gótica cursiva de libros como *vaga, luxurians* o lujuriosa, propia de pictores, de trazos artificiosos (*artificiosis tractibus litteris*) e inventada para cualquier cosa menos para leer (*ad alium quam ad legendum inventa*). El gran poeta de Italia anheló una letra cuidada y clara (*castigata et clara*), que fuera bien captada por los ojos (*seque ultro oculis ingerens*).

Según Marín Martínez y Ruiz Asencio (1996), en la segunda mitad del siglo XV, desde Florencia, la letra humanística se difundió a toda Italia, incluido el Reino de Nápoles que formaba parte de la Corona de Aragón. Posteriormente la introducción de la misma letra en Castilla se inicia con los Reyes Católicos,

como consecuencia, en parte, del movimiento de unidad nacional (...), pero a pasos muy lentos y compartiendo el campo con las góticas tradicionales castellanas que mantiene su hegemonía. (...) si durante los primeros años imperiales [de Carlos I, 1516-1556] se empieza a salir de ese predominio [de la letra gótica], es por la vía de la enseñanza popular que se sistematizó entonces de algún modo en materias de lectura y escritura, recayendo dicha enseñanza, como era natural, en favor de la humanística. (*ibid.* II. 71)

Todo ello se comprueba en los números de documentos recogidos en el corpus CODEA que muestra la siguiente tabla:

Región	Letra	1200	1250	1300	1350	1400	1450	1500	1550	1600	1650	1700	1750
Aragón	Gótica	2	8	15	37	39	7	2					
	Human.						7	8	2	4	3	1	24
C. Vieja	Gótica	13	98	49	23	26	38	45	1	1			
	Human.					1	2	4	14	23	6	9	24
C. Nueva	Gótica	10	25	18	12	20	21	36	33	6			
	Human.							5	228	112	69	34	72

Tabla 9. Documentos de la letra gótica (Gótica) y la letra humanística (Human.) Aragón, Castilla la Vieja (C. Vieja), Castilla la Nueva (C. Nueva).

Efectivamente la introducción de la letra humanística en Aragón es 50 años más temprana que en Castilla. Dentro de Castilla, nos llama la atención la supremacía casi explosiva de la letra humanística en Castilla la Nueva en 1550 y sus altas frecuencias posteriores.

Entre los tipos de documento (cancilleresco y particular) y las dos letras gótica y humanística) se observa la cronología en la siguiente tabla (Tabla 10). Según la misma, parece que el documento particular toma la iniciativa en la introducción de la letra humanística más fuertemente y temprano que el

documento canclleresco:

Docu- mento	Letra	1200	1250	1300	1350	1400	1450	1500	1550	1600	1650	1700	1750
Cancille- resco	Gótica	5	64	28	12	15	29	23	2				
	Huma.							4	24	2	4	1	
Particular	Gótica	4	25	29	3	7	6	17	6	3			
	Huma.						2	3	183	86	34	31	66

Tabla 10. Documentos de la letra gótica (Gótica) y la letra humanística (Huma.) Documento canclleresco, Documento particular.

Se conoce la inmensa influencia cultural que recibió España de Italia en la centuria de 1500 (Kamen 1973, 1976: 81-82), en la pintura con El Greco (1541-1614), en la arquitectura con Juan de Herrera (1530-1597) y en la literatura con Juan de Valdés (c.1490-1541), por citar personalidades representativas. Creemos que dentro de la influencia cultural italiana está bien situada la letra humanística, aunque no manifestada en individuos históricos sino en el conjunto de personas de la Corona y del pueblo, es decir, en la sociedad.

Según Moreno Fernández (2005: 135), "el siglo XVI colocó al castellano en la vanguardia de la «estandarización» de las lenguas de Europa, gracias a las obras de un puñado de hombres de letras excepcionales: Elio Antonio de Nebrija, Sebastián de Covarrubias, Bernardo de Aldrete, Gonzalo Correas." Creemos que la letra humanística constituye otro capítulo importante de la «estandarización» de la lengua, «estandarización» masiva de comunicación escrita, necesaria y oportuna en la época del gran crecimiento demográfico y deslizamiento masivo de la población desde el Norte hacia el Sur, de que trata el mismo autor (141).

También nos interesa la continuación de los periodos tratados. En los siglos posteriores, no sabemos si la modernización sube en línea recta o exponencial u otra vez en forma de "S". En un futuro cercano plantearemos la indagación de los siglos XIX y XX para aproximarnos a la historia completa de la ortografía española.

Referencias

Abad, Francisco, 1998a, «Problemas de periodización y caracterización en

- historia de la lengua literaria española (I)», *Revista de filología románica*, 15, 13-33.
- Abad, Francisco, 1998b, «Problemas de periodización y caracterización en historia de la lengua literaria española (II)», *Epos*, 14, 493-513.
- Agujetas, María / Sánchez-Prieto Borja, Pedro / Ueda, Hiroto. En preparación. "Inventario léxico del corpus CODEA. II. Castilla la Nueva" en <https://lecture.ecc.u-tokyo.ac.jp/~cuedályneal/ilc-cn.htm> [27/3/2022]
- Cano Aguilar, Rafael. 1988. *El español a través de los tiempos*. Madrid: Arco Libros.
- Danison, David. 2003. "Log(ist)ic and simplistic S-curves", in R. Hickey (ed.) *Motives for language change*, Cambridge University Press, pp. 54-70.
- Eberenz, Rolf. 1991. "Castellano antiguo y español moderno: reflexiones sobre la periodización en la historia de la lengua", *Revista de Filología Española*, 71, 79-106.
- Eberenz, Rolf. 2009. "La periodización de la historia morfosintáctica del español: propuestas y aportaciones recientes", *Cahiers d'Études Hispaniques Médiévales*, 32, 181-201.
- Febvre, Lucien et Martin, Henri-Jean. 1971. *L'Apparition du livre*. 2eme ed. Paris: Albin Michel. 1998. Traducción japonesa (Sekine, Motoko; Hasegawa, Teruo; Miyashita Shiro; Tatsuo Tsukimura). Shomotsu no shutsugen. Tokio: Chikkuma shobo.
- Harris-Northall, Ray. 1996. "Printed Books and Linguistic Standardization in Spain: The 1503 *Gran Conquista de Ultramar*", *Romance Philology*, 50, 2, 123-146.
- Kamen, Henry. 1973, 1976. *A Concise History of Spain*. London: Thames and Hudson. Japanese translation Niwa Mitsuo. *Supein. Rekisi to bunka*. Tokyo: Tokai University Press.
- Knight, Stan. 1998. *Historical Scripts: From Classical Times to the Renaissance*. New Castle, Delaware: Oak Knoll Press. Japanese translation by Toshiyuki Takamiya, 2001. Tokyo: Keio University Press.
- Knight, Stan. 2012. *Historical Types from Gutenberg to Ashendene*. New Castle, Delaware: Oak Knoll Press, Japanese translation by Toshiyuki Takamiya and Mari Agata, 2014. Tokyo: Keio University Press.
- Labov, William. 1994. *Principles of linguistic change. Vol. 1. Internal factors*. Oxford: Blackwell.
- Lapesa, Rafael. 1980. *Historia de la lengua española*. 9a ed. Madrid: Gredos.
- Lathrop, Thomas A. 2002. *Curso de gramática histórica española*. Barcelona:

- Ariel.
- Lope Blanch, Juan M. 1969. *Edición, introducción y notas de Juan de Valdés, Diálogo de la lengua*. Madrid: Castalia.
- Marcos García, Juan José. 2012. *Historia del alfabeto y de las letras del abecedario romano*. Madrid: Editorial Dykinson.
- Marín Martínez, Tomás / Ruiz Asencio, José Manuel. 1996. *Paleografía y diplomática. I, II*. Madrid: Universidad Nacional de Educación a Distancia.
- Martín Aizpuru, Leyre. 2022. *La escritura cancilleresca de Fernando III, Alfonso X, Sancho IV y Fernando IV*. Bern: Peter Lang.
- Menéndez Pidal, Ramón. 1942. "El lenguaje del siglo XVI", en Ramón Menéndez Pidal, *La lengua de Cristóbal Colón, el estilo de Santa Teresa y otros estudios sobre el siglo XVI*, Madrid, Espasa-Calpe, 47-84.
- Moreno Fernández, Francisco. 2005. *Historia social de las lenguas de España*. Barcelona: Ariel.
- Odriozola, Antonio. 1982. "La imprenta en Castilla en el siglo XV", en *Historia de la imprenta hispana*. Madrid: Editora Nacional, 91-220.
- Ong, Walter J. 1987. *Oralidad y escritura. Tecnología de la palabra*. Traducción de Angélica Scherp. México DF: Fondo de Cultura Económica.
- Quilis, Antonio. 1980. *Estudio y edición de Antonio Nebrija, Gramática de la lengua castellana*. Madrid: Editora Nacional.
- Real Academia Española y Asociación de Academias de la Lengua Española. 2010. *Ortografía de la lengua española*. Madrid. Espasa Libros.
- Romero Tallafigo, Manuel. 2022. "La escritura humanística", en <https://personal.us.es/tallafigóhumanística%20.htm> [2 de mayo, 2022]
- Ruiz García, Elisa. 2000. "La imprenta en Castilla durante el siglo XV", en Riesco Terrero, Ángel (ed.). *Introducción a la paleografía y la diplomática general*. Madrid: Editorial Síntesis, 177-187.
- Salvador, Gregorio / y Lodaes, Juan R. 2001. *Historia de las letras*. Madrid: Espasa Calpe.
- Sánchez Lancis, Carlos. 1999. "Sintaxis histórica, informática y periodización del español", en J. M. Bleca / G. Clavería / C. Sánchez / J. Torruella (eds.) *Filología e Informática. Nuevas tecnologías en los estudios filológicos*, Universidad Autónoma de Barcelona, 415-446.
- Sánchez Lancis, Carlos. 2001. "Historia de la lengua, gramática histórica y

- periodización en español", *Estudi General*, 21, 395-412.
- Sánchez Lancis, Carlos. 2009. "Corpus diacrónicos y periodización del español", *Cahiers d'Études Hispaniques Médiévales*, 32, 159-180.
- Sánchez-Prieto Borja, Pedro. 1998. *Cómo editar los textos medievales. Criterios para su presentación gráfica*. Madrid: ArcóLibros.
- Sánchez-Prieto Borja, Pedro. 2004. "La normalización del castellano escrito en el siglo XIII. Los caracteres de la lengua: grafía y fonema", en Rafael Cano Aguilar (coord.), *Historia de la lengua española*. Barcelona: Ariel, 423-448.
- Sánchez-Prieto Borja, Pedro / Ueda, Hiroto (2018). "Inventario léxico del corpus CODEA. I. Castilla la Vieja" en <https://lecture.ecc.u-tokyo.ac.jp/~cueda/lyneal/ilc-cv.htm> [27/7/2019]
- Nalle, Sara T. 1989. "Literacy and Culture in Early Modern Castile", *Past and Present*, 125, 65-96.
- Tamayo, Alberto. 2012. *Historia de la escritura latina e hispánica*. Gijón: Ediciones Trea.
- Torrens Álvarez, María Jesús. 2007. *Evolución e historia de la lengua española*. 2a edición. Madrid: ArcóLibros.
- Trevor Davies, R. 1973. *El gran siglo de España, 1501-1621*. Madrid. Akal Editor.
- Tuten, Donald N. / Tejedo-Herrero, Fernando. 2011. "The Relationship between Historical Linguistics and Sociolinguistics" in Manuel Díaz-Campos (ed.) *The Handbook of Hispanic Sociolinguistics*, Oxford: Blackwell Publishing, 283-302.
- Ueda, Hiroto. 2017. "Unilateral Correspondence Analysis, applied to Spanish linguistic data in time and space", *Sixteenth International Conference on Methods in Dialectology (METHODS XVI)*, National Institute for Japanese Language and Linguistics, Tokyo, 10 August, 2017.
<https://lecture.ecc.u-tokyo.ac.jp/~cueda/kenkyuchiri/correspondence/correspondence2017.pdf>
- Ueda, Hiroto / Sánchez-Prieto Borja, Pedro / Moreno Sandoval, Antonio. 2020. "Lematización y visualización cartográfica del corpus CODEA. Formas de la conjunción 'y' en el norte de Castilla medieval", *Estudios de Lingüística del Español*, 42 (2020), 245-261.
<https://infoling.org/elies/42/elies42-10.pdf>

[Fin]