

«Hispanismo asiático», Videoconferencia en Casa Asia e Instituto Cervantes en Madrid,
19 de enero, 2021. (haber+pp.docx) rev. AMS, LMA.

Historia y geografía de «haber» + participio pasado

Contribuciones desde Japón

Hiroto Ueda (Universidad de Tokio)

1. Introducción¹

Al recibir la propuesta conjunta de Casa Asia e Instituto Cervantes en Madrid de dar una videoconferencia sobre un tema lingüístico no muy especializado y, al mismo tiempo, sobre estudios hispánicos que hay en Japón, se me ocurrió el título arriba puesto. El subtítulo, “Contribuciones desde Japón”, es un tanto raro, puesto que en estudios y ciencias no puede ni debe haber una frontera nacional. No obstante, he pensado que merece la pena aprovechar la ocasión para dar a conocer al mundo las investigaciones realizadas por los estudiosos japoneses, que mayoritariamente están publicadas en sitios poco accesibles y escritas en japonés, lo cual constituye una barrera de comunicación difícil de salvar.

El tema principal es “Historia y geografía de «haber» + participio pasado (pp en adelante)”, por ejemplo, *he escrito una carta, he llegado tarde*, etc. Cuando aprendí el uso de estas formas por primera vez en un curso universitario, pensé que no era muy difícil, ya que estaba acostumbrado a usarla en inglés, *have + pp*. Luego me di cuenta de que el verbo auxiliar «haber» no se utiliza como verbo pleno de la misma manera que *have* en inglés: *I have a book*. Posteriormente supe que en su lugar se utiliza el verbo «tener»: *Tengo un libro*. Me preguntaba: ¿A qué se debe esta diferencia?

En el curso de latín, aprendimos que la forma del pasado, por ejemplo, DIXI significa tanto el pasado indefinido ‘dije’ como el perfecto ‘he dicho’. El profesor y los libros de historia de español nos enseñaron el origen de la forma compuesta «haber» + pp (*he dicho*) en la forma supuestamente latina: PP + HABERE (*DICTUM HABEO* ‘tener’). Sin embargo, el uso auxiliar de PP + HABERE (*DICTUM*) no está explicado en el diccionario y nunca lo he visto en textos latinos. Entonces, ¿cómo se explica la combinación de PP + HABERE como el origen de la forma española «haber» + pp?

¹ Agradezco a Noritaka Fukushima, Leyre Martín Aizpuru y Antonio Moreno Sandoval por la ayuda prestada en la preparación de este trabajo.

Actualmente, al analizar los datos lingüísticos de España e Hispanoamérica, he observado que una misma forma presenta diferente uso y frecuencia dependiendo de la zona geográfica. ¿Qué causa ha producido la distribución geográfica de la misma forma en ambos lados del Atlántico? He buscado las razones de la distribución geográfica en la historia de la lengua, con los datos ofrecidos por distintos proyectos españoles de investigación, contando con los estudios anteriores de países hispanohablantes y de fuera de ellos, inclusive de Japón.

A continuación, para ofrecer las respuestas a estas preguntas, vamos a ver el origen latino de la forma compuesta «haber» + pp, su difusión en lenguas romances, el desarrollo histórico en el castellano medieval y el español moderno y, finalmente, la geografía actual de la misma forma tanto en España como en Hispanoamérica. Para no terminar la conferencia solo con las noticias de los estudios anteriores, me voy a permitir presentación de los propios míos, resultados de una reciente investigación. Concretamente voy a enfocarme en el transcurso histórico del latín > el castellano medieval > el español moderno y en la gradación geográfica del este al oeste, en paralelo.

Al abordar el mismo tema intento comprobar la validez de la hipótesis sobre la existencia de la continuidad en espacio y tiempo de propiedades lingüísticas de la forma compuesta. Me centraré en el estilo normal observado en los documentos notariales, en contraste con las obras literarias tratadas en los estudios previos. Mi propósito principal de este estudio es proponer el factor de la frecuencia de uso, como el factor más importante de la gramaticalización, es decir, el cambio del verbo pleno «haber» ‘poseer’ en el verbo auxiliar en «haber» + pp.

2. Origen y extensión

2.1. Lengua latina

Para ver la historia de la forma compuesta española, «haber» + pp, tenemos que remontarnos a la situación de la lengua latina, dividida en cuatro etapas del latín escrito: latín antiguo (LA), latín clásico (LC), latín tardío (LT) y latín medieval (LM). Kuniyama (2007: 3) delimita la cronología en el siguiente esquema (parcialmente modificado):²

² Las líneas oblicuas representan las distancias lingüísticas que hay con respecto al latín clásico (LC). Me he permitido hacer una pequeña modificación del gráfico de Kuniyama en el inicio adelantado del latín hablado, siempre apartado del latín escrito,

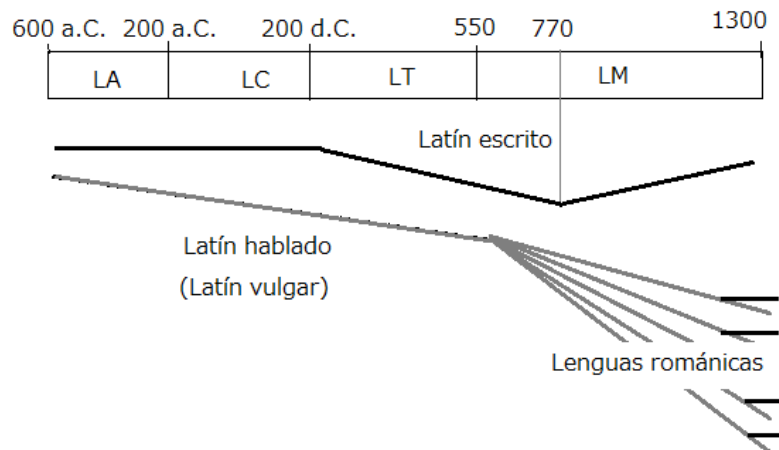


Fig. 1. Latín antiguo (LA), l. clásico (LC), l. tardío (LT), l. medieval (LM) y l. vulgar

De esta manera se distinguen entre dos modos de expresión de la lengua: latín hablado (latín vulgar y lenguas romances) y latín escrito (desde el latín antiguo hasta el latín medieval).

Por otra parte, desde la filología japonesa Nomura (2011: 1-16) llama la atención para no confundir entre la lengua hablada y el estilo coloquial, por una parte; y la lengua escrita y el estilo literario, por otra, puesto que ambos estilos, coloquial y literario, pertenecen a la lengua escrita. Creo que la clasificación de Nomura es aplicable a la situación de la lengua española y también a la mayoría de las lenguas con escritura establecida y tradición literaria en general. Se contrastan los dos modos, uno erróneo y otro correcto, en los siguientes esquemas:

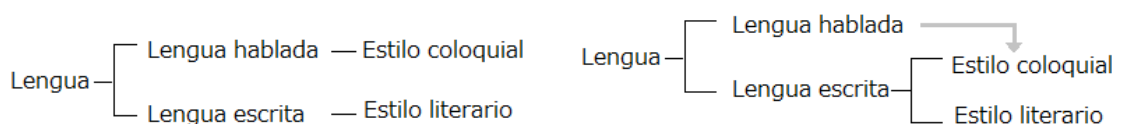


Fig. 2. Esquema erróneo / Esquema correcto

En la parte derecha de la Fig. 2, he dibujado una flecha gris que conecta la lengua hablada con el estilo coloquial perteneciente a la lengua escrita para establecer una relación o un reflejo unidireccional de la lengua hablada al estilo coloquial de la lengua escrita, pero no una igualdad de ninguna manera. El ‘estilo coloquial’, para

en lugar de la unión de las dos líneas en el gráfico original de Kunihara, puesto que es difícil suponer que el latín hablado naciera del latín escrito. He hecho la distinción entre las líneas negras que representan las obras y documentos escritos existentes y las grises, las supuestas lenguas habladas latentes bajo las manifestaciones escritas.

librarse de la igualación habitual que se hace con la ‘lengua hablada’, también es posible sustituirlo por ‘estilo normal’ en el sentido de ‘estilo no literario’. Los dos estilos, coloquial (normal) y literario, se tratan de una gradación entre los tipos caracterizantes, es decir, un texto escrito puede ser más coloquial o literario que otro en la escala estilística gradual entre los dos extremos: coloquial y literario. A mi modo de ver, los documentos notariales que trataremos en la sección 3 son más coloquiales (normales) que las obras literarias, que son tratadas en los estudios anteriores.³ Todavía más coloquiales (normales) son las transcripciones de conversaciones grabadas que veremos en la sección 4. La distinción de los dos estilos, coloquial (normal) y literario en los dos tipos de textos, documentos notariales y obras literarias respectivamente, es importante.

Tomando en consideración este esquema histórico (Fig. 1) y el estilístico (Fig. 2. derecha), vamos a abordar el tema de historia latina de PP + HABERE.

A diferencia de la situación material del castellano medieval, donde contamos con numerosos documentos recogidos en serie (sec. 3) y de la del español actual con transcripciones de las conversaciones grabadas (sec. 4), la condición del estudio de la lengua latina obliga el uso de los materiales sueltos sin continuidad ni homogeneidad.

Se sabe que la forma española «haber» + pp se remonta históricamente a la forma latina PP + HABERE ‘tener, poseer’. Esta explicación sucinta presenta dos problemas: la diferencia enorme de frecuencia entre la forma latina y la española y el cambio semántico funcional de ‘tener’ al verbo auxiliar «haber». También el participio latino (PP) poseía unas características diferentes de las que tiene el español actual: concordancia de género y número. Su posición dentro de la oración era tan libre que aparecía delante y detrás de HABERE, con la inserción de otros elementos entre HABERE y el participio o sin ella. De esta manera, a pesar de la similitud de apariencia entre la forma latina y la española, la diferencia práctica (frecuencia), morfológica (concordancia de género y número), sintáctica (posición del participio con respecto al verbo) y semántico-funcional (‘tener’ y auxiliar) entre las dos lenguas es grande. ¿A qué se deben tantas diferencias? ¿Es posible encontrar un motivo unificado, ideal para explicar la diferencia de estos comportamientos lingüísticos?

A mi modo de entender, todo se explica por la demanda de la funcionalidad del tiempo perfecto, es decir, por la utilidad del mismo. La función perfectivo era y

³ Se trata de textos medievales. En cambio, a partir del siglo XIX, con el realismo, y sobre todo en el XX el estilo de la literatura puede ser más coloquial que el de los documentos notariales, que no son nada coloquiales.

es tan fundamental que se encuentra en las lenguas del mundo de manera casi general (Bybee, Perkins and Pegliuca 1994, Oka 2013, Kume 2016). La forma latina, por ejemplo, SCRIPSI (*escribí, he escrito*), al convertirse en la forma española *escribí*, se especializó en el uso pretérito (*escribí*), a costa del perfectivo (*he escrito*). Para suplir la función perdida, el pueblo la encontró en la combinación de HABERE y participio concordado: SCRIPTAM LITTERAM HABEO ‘Tengo escrita la carta’. Esta combinación se supone que existía en la lengua hablada, a pesar de que en la escrita aparece con poca frecuencia, como veremos inmediatamente. Se trata de una formación relativamente nueva, practicada en el habla del pueblo, de modo que en los textos latinos literarios y formales, donde se evita la mezcla del estilo tradicional y el innovador, no aparecen (Hanssen 1913/1966: 229).

El esquema que presenta Fukushima (1978: 5), citado por Tsujii (1996: 65), es útil para explicar la diferencia de relación sintáctica con respecto a las dos concordancias: la de persona (pers.) entre el Sujeto y HABERE / «haber» por una parte, y la de género y número (gén. núm.) entre el participio y el Objeto por otra:

Sintaxis concordante: [Sujeto = (pers.) «haber»] - [PP = (gén. núm.) Objeto]

Sintaxis no concordante: [Sujeto = (pers.) [«haber» + pp]] - Objeto

Lo aplicamos a la sintaxis latina y a la española de la siguiente manera:

Latín: [Objeto =~ (gén. núm.) PP] ~ [Sujeto =~ (pers.) HABERE]

Español: [Sujeto -= (pers.) [«haber» + pp]] - Objeto

donde el signo de igualdad ‘=’ representa la concordancia gramatical, el guión (-) el orden normal y la tilde (~) el orden libre.

La existencia o ausencia de la concordancia entre el participio y el Objeto es tan importante que cambia totalmente la estructura sintáctica. En latín, donde la concordancia de PP y el Objeto garantiza la fuerte combinación entre ellos, mientras que en español, el participio (pp), por falta de concordancia con el Objeto, se combina con «haber» más que con el Objeto formando un tiempo compuesto [«haber» + pp]. El cambio de la posición libre de PP en latín a la fija en español en forma de [«haber» + pp] también contribuye a la formación de la función gramatical perfectiva.

Otra diferencia importante entre latín y español está en que HABERE es un verbo pleno con el significado léxico de ‘poseer, tener’⁴, mientras que «haber» es

⁴ Según el diccionario de Segura Munguía (2003: 331), HABERE posee significados de “I. [Sentido propio] tener, tener como cosa propia; poseer, ser dueño de, (...): *pecus habere*, poseer un rebaño (...) II. [Sentido figurado] tener: *h. aliquid in ore atque in*

un verbo auxiliar que forma parte del tiempo compuesto. De esta manera, HABERE es comparable con TENERE dentro del mismo plano léxico, mientras que «haber» (+ pp) es incomparable con «tener», puesto que los dos verbos pertenecen a las categorías gramaticales diferentes: verbo auxiliar y verbo pleno.

Para comprobar la frecuencia de HABERE y TENERE, junto con el uso, aunque escaso, de la combinación latina de PP + HABERE, he consultado las páginas de *Anfitrión* y *Asinaria* de Plauto (c. 254-184 BC)⁵, *Catilinarias* de Cicerón (106-43 BC)⁶, *Guerra de Galia* (I, II, III) de César (100-44 BC)⁷, *Biblia vulgata* de San Jerónimo (c. 342-420)⁸, *Itinerarium* de Aeteria (viaje de 381-384)⁹, *Etimologías, Libros XI, XII, XIII* de San Isidoro (c. 560–636)¹⁰ y documentos medievales recogidos en el corpus CORHEN («*Corpus Histórico del Español Norteamericano*», 800-1200) dirigido por María Jesús Torrens Álvarez¹¹. Vamos a ver las frecuencias de los dos verbos HABERE y TENERE que hay en estas obras y corpus¹²:

animo, tener siempre algo en los labios y en el pensamiento” mientras que el significado de TENERE es “Tener, coger; sujetar, asir. *aliquem [in] sinu t.* tener a alguno en sus brazos; *aliquid [in] manu t.*, tener algo en la mano”. En suma, HABERE corresponde básicamente a la ‘posesión’, mientras que TENERE a la ‘acción de coger’.

⁵ El texto y su traducción correspondiente son de González-Haba (1992).

⁶ Id. Campos Rodríguez (1975).

⁷ Id. García Yebra y Escolar Sobrino (2002).

⁸ Id. Gryson (1969) y Sociedades Bíblicas en América Latina (1960).

⁹ Id. Arce (1980).

¹⁰ Id. Oroz Reta y Marcos Casquero (1983).

¹¹ <https://corhen.es/> [7 de enero, 2021]

Fontán y Moure Casas (1987: 435-438) incluyen los documentos cancillerescos de Enrique I de Castilla y Toledo (1216) y de Fernando III (1219) en su *Antología del latín medieval*. Según los dos autores, estos dos documentos “corresponden a fórmulas cancillerescas tradicionales, y tienen el interés de que pocos años después, bajo el mismo Fernando III, los documentos de asuntos interiores del reino empezaron a ser redactados en castellano, quedando reservado el uso del latín para los de relaciones internacionales”. Los documentos reunidos en el corpus CORHEN incluyen los documentos fuera del ámbito cancilleresco, por lo que puede presentar rasgos populares de la lengua. En el mismo corpus he encontrado 92.250 palabras latinas en 302 documentos latinos, que se distribuye en las franjas cronológicas de la siguiente manera: año 800: 527 palabras (documentos: 1), año 900: 1384 (4), año 1000: 8375 (19), año 1100: 34638 (111), año 1200: 47326 (167).

¹² FA: frecuencia absoluta, FR: frecuencia relativa, FN: frecuencia normalizada por mil palabras.

Latin	Plauto	Cicerón	César	Vulgata	Aeteria	Isidoro	CORHEN
FA. HABERE	48	26	56	1478	57	148	519
FA. TENERE	13	13	29	185	12	27	158
Total	61	39	85	1667	69	175	677
FR. HABERE	79%	67%	66%	89%	83%	85%	77%
FR. TENERE	21%	33%	34%	11%	17%	15%	23%
Palabras	17 885	12 638	15 960	532 329	17 545	25 205	93 999
FN. HABERE	2.684	2.057	3.509	2.776	3.249	5.872	5.521
FN. TENERE	0.727	1.029	1.817	0.348	0.684	1.071	1.681

Tabla 1. HABERE y TENERE en obras y documentos latinos

Esta tabla ofrece tres tipos de frecuencia: absoluta (FA), relativa (FR) y normalizada (FN). La frecuencia absoluta representa el número de veces que la forma aparece en cada conjunto de textos. La frecuencia relativa indica la proporción de cada frecuencia que ocupa dentro de la totalidad interna de términos en comparación, derivada en forma de porcentaje (%). La frecuencia normalizada es útil para conocer su magnitud cuantitativa dentro de la totalidad de palabras que hay en el conjunto de textos. Se deriva de la frecuencia absoluta, dividida por la totalidad de palabras, multiplicada por 1000, por ej. la frecuencia normalizada de HABERE en Plauto se calcula: $48 / 17\ 885 * 1000 = 2.684$, es decir, aparece 2.7 veces en 1000 palabras.

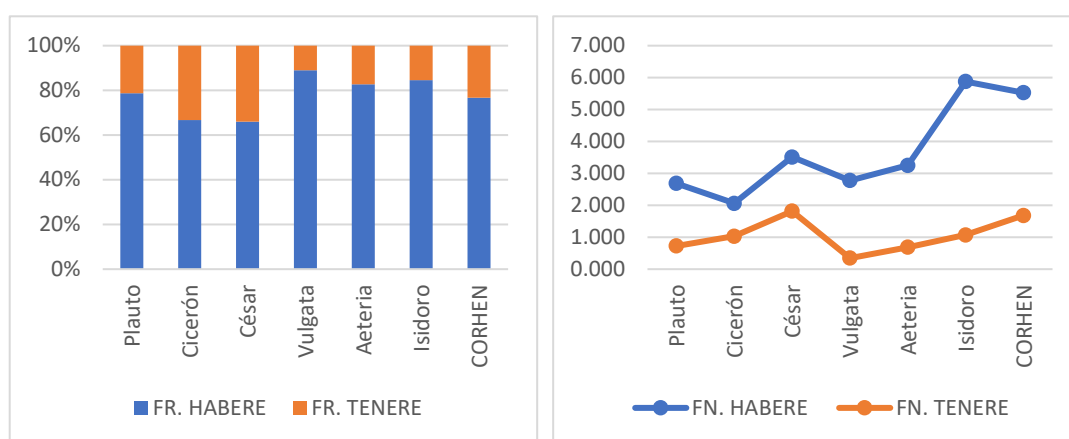


Fig. 3. HABERE y TENERE en obras y documentos latinos

Siempre la frecuencia de HABERE prevalece a la de TENERE, probablemente por su significado general de 'posesión'. Es lógico que tenga más utilidad semántica que el significado de acción de 'coger, asir'. También es notable el aumento de uso de HABERE, que se percibe a lo largo de siglos.

Se nota que la frecuencia normalizada de HABERE ‘tener, peser’ es tan alta como se iguala a «haber» + pp, que veremos más adelante. En cambio la frecuencia de TENERE ‘asir, coger’ se mantiene baja. Veamos algunos ejemplos:

- (1) a. *Non equidem ullum habeo iumentum.* ‘no tengo conmigo cabalgadura alguna.’ (Plauto. *Anfitrión*. I-I.i.)
- b. *Habemus scnatus consultum in te, Catilina, vehemens et grave* ‘Tenemos contra ti, Catilina, un decreto del Violado vigoroso y severo’ (Cicerón. *Catilinarias*, I-1)
- c. *naves habent Veneti plurimas,* ‘los vénetos tienen muchísimas naves,’ (César, *Commentarii de Bello Gallico*, III-8)
- d. *Nihil autem sagacius canibus; plus enim sensus ceteris animalibus habent.* ‘No existe animal alguno más sagaz que el perro, pues tiene los sentidos más desarrollados que todos los demás.’ (Isidoro, *Etimologías*, XII-2)

La frecuencia de TENERE es más baja, probablemente por tratarse de una acción específica (‘coger, asir’):

- (2) a. *Cur me tenes?* ‘¿Por qué me retienes?’ (Plauto. *Anfitrión* I-I.iii)
- b. *et viri qui tenebant illum includebant ei caedentes* ‘Y los hombres que custodiaban a Jesús se burlaban de él y le golpeaban;’ (*Vulgata*. Ev. sec. Lucam, 22-63)
- c. *sed modo ibi accessus Romanorum non est, totum enim illud Persae tenent.* ‘pero ahora no pueden llegar allí los romanos, pues los persas son dueños de todo aquel territorio;’ (Aeterna, 20)

A pesar de que hay en la bibliografía distintas menciones sobre la combinación de PP + HABERE a lo largo de siglos (Ernout, 1953: 216-217; Harrington, 1962: 41-42; García de Diego, 1970: 234; Grandgent, 1970: 91; Niemeyer, 1976: 478; Ernout et Thomas, 1972:223; Lloyd, 1993: 278-279; Herman, 1997: 91-94; Kuniyama, 2007: 78-79), su frecuencia es mínima¹³. En los textos

¹³ Pérez González (1985: 168) precisa “*habere* + participio apenas aparece en los documentos de los siglos IX-XI, así como en los nuestros, en los que sólo hemos hallado contados ejemplos”, mientras que según Villimer (1976: 125-126), PP + HABERE es una “perífrasis muy productiva en nuestra época [S. XIV y XV] en la Cancillería castellana, propia de todos los períodos del latín, y que sirve para expresar el estado alcanzado, pero mantenido en sí o en sus consecuencias”. Estas dos explicaciones no son contradictorias, puesto que se tratan de épocas distintas.

latinos enumerados anteriormente, he encontrado solo una vez en Cicerón, tres veces en César, ocho veces en *Vulgata*, dos veces en Aeteria, siete veces en Isidoro y ninguna en CORHEN, de los que extraigo los siguientes ejemplos:

- (3) (i) *licet et capellarum sit pressum habere nasum*. ‘No obstante, también es característico de las cabras el tener aplastadas las narices.’ (Isidoro. *Etimologiae*, XII-2)
- (ii) *Mulus autem a Graeco tractum vocabulum habet*. ‘No obstante, el vocablo «mulo» es de ascendencia griega.’ (Isidoro. *Etimologiae*, XII-1)
- (iii) *Iam ad certas res confidencias certos homines delectos ac discriptos habebat*. ‘Por lo demás, para cada misión tenía sus hombres especialmente escogidos.’ (Cicerón. *Catilinarias*, III-7)
- (iv) *cadens apertos habet oculos* ‘Caído, pero abiertos los ojos’ (Jerónimo. *Vulgata*, *Numerus*, 24)
- (v) *Tunc videntes hoc Persae averterunt ipsam aquam a civitate et fecerunt ei decursum contra ipso loco, ubi ipsi castra posita habebant*. ‘Viendo esto los persas, desviaron el agua de la ciudad y dirigieron su corriente hacia el lugar donde ellos habían fijado su campamento.’ (Aeteria. *Itinerarium*, 19-11)
- (vi) *Tunc dictum est nobis, quia in isdem diebus, qua sanctus Moyses vel filii Israhel contra illas civitates pugnaverant, castra ibi fixa habuissent* ‘en la cual: Nos dijeron, el santo Moisés o los hijos de Israel habían plantado sus campamentos mientras luchaban contra aquellas ciudades.’ (Aeteria. *Itinerarium*, 12-9)

En realidad, estamos ante cuatro tipos de relación sintáctica que hay en la combinación de PP + HABERE (Rodríguez Molina 2010: 1042-1056). En primer lugar, tipo-A, el ejemplo (3.i), se trata de las narices ‘aplastadas’, es decir, ‘chatas’. En este caso, el participio *passum* (< PREMERE) puede interpretarse como adjetivo, a pesar de la traducción ‘tener aplastadas las narices’. El ejemplo (3.ii) pertenece al segundo tipo-B. Aquí, *tractum* es participio (< TRAHERE) sin duda. Pero, el sujeto sintáctico de la oración, *Mulus* ‘mulo’, es diferente del sujeto semántico del participio *tractum* ‘traído’, puesto que *mulus* no ha traído la palabra del griego. Para llegar a la forma compuesta «haber» + pp, la condición necesaria es la identidad de sujeto de la oración y del participio, lo que observamos en los ejemplos (3.iii) y (3.iv), pertenecientes al Tipo-C. En el ejemplo (3.iii), el que tenía su hombre lo había escogido, es decir, el sujeto de *habebat* coincide con el de *delectus* (<

DELIGERE ‘escoger’). Algo parecido ocurre en el ejemplo (3.iv). Finalmente los dos ejemplos de Aeteria, (3.v) y (3.vi), que representan el Tipo-D, se consideran como orígenes de las formas analíticas de lenguas romances, inclusive el español (Eto 1998: 130-131).

Dejando aparte el Tipo-A (adjetivo), nos interesa especialmente el paso de los tres Tipos en este orden: B → C → D¹⁴:

- (4) Tipo-A: Sujeto_i Objeto_j Adj_j HABERE : Ej. (i)
 Tipo-B: Sujeto_i Objeto_j PP_j HABERE : Ej. (ii)
 Tipo-C: Sujeto_i Objeto_j PP_i HABERE : Ej. (iii), (iv)
 Tipo-D: Sujeto_i Objeto_j [PP HABERE] : Ej. (v), (vi)

La interpretación que hemos propuesto al ejemplo (3.ii) es del Tipo-B: ‘el mulo tiene su nombre traído (procedente) del griego’. Pero también es posible otra interpretación: ‘el mulo ha traído del griego su nombre’, que es del Tipo-C. De esta manera podemos suponer que el origen de la estructura de Tipo-C está en la ambivalencia semántica entre Tipo-B y Tipo-C y su desarrollo posterior en la mayor inclinación al Tipo-C. Lo mismo ocurre en el paso final de Tipo-C a Tipo-D, como en el ejemplo (3.v), citado de Aeteria, *ubi ipsi castra posita habebant*. ‘donde ellos habían fijado su campamento’ lo hemos clasificado en el Tipo-D, de acuerdo con Eto (op. cit.)¹⁵. Pero también es posible otra interpretación perteneciente al Tipo-C, ‘donde ellos tenían fijado su campamento’, por tanto, podemos pensar en el cambio gradual del significado ambivalente, Tipo-C+D, al significado más inclinado al Tipo-D¹⁶.

El proceso gradual de Tipo-C a Tipo-D se denomina ‘reanálisis’ dotado de ‘recategorización’, que explica Rodríguez Molina (2010: 1057-1061):

En el caso del verbo HABEO, su recategorización como auxiliar significó la pérdida de su red argumental, cambios en sus propiedades de selección y

¹⁴ Los subíndices i y j indican la identidad de referentes. Respetamos el orden latino usual de palabras, donde el verbo conjugado HABERE se sitúa al final de la oración.

¹⁵ Elcock (1960: 109): “The construction, (...) became increasingly employed by Christian writers, as in Egeria’s [Aeteria’s] CASTRA POSITA HABEBAT. At this stage the new compound tenses were already in being.” Terasaki (2011: 37) sitúa el cambio semántico de la forma compuesta en ‘perfecto presente’ a partir del siglo VI.

¹⁶ Sobre la ambigüedad semántica de PP + HABERE, véase Adams (2013: 615-651). Shimaoka (1970: 96-99) sitúa el cambio de *Habeo flores comparatos* a *Habeo comparatum flores* en la época de dinastía merovingia (481-751), cuando la forma HABERE COMPARATUM se igualaba con COMPARAVI.

rección sintáctica, así como la adquisición de rasgos funcionales. Así, HABEO ya no selecciona más un complemento nominal, sino uno clausal, comparte un único sujeto con el participio y pierde la capacidad léxica como verbo de posesión para expresar solo rasgos funcionales flexivos, esto es, la anterioridad temporal (...) y la relevancia actual.

En el mismo cambio, interviene un factor importante: la frecuencia. Con el uso frecuente de la construcción de Tipo-C, es normal que se pierda cada vez más el significado de HABERE ‘poseer’ y, en su lugar, el centro de peso semántico empieza a caer en el participio pasado (PP) formando una unidad verbal compuesta de «haber» + pp. El participio pasado (PP), que antes jugaba un papel secundario pasivo, se convierte ahora en el elemento principal activo, dejando «haber» al lado como un mero elemento funcional, no semántico. Se trata de una simplificación estructural de cuatro elementos, Sujeto + Verbo + Objeto + Complemento de Objeto (PP), relativamente poco frecuente, a la combinación de tres elementos: Sujeto + Verbo + Objeto, frecuentísima en la lengua en general.

Por otra parte, ahora que acabamos de reconocer la importancia de la frecuencia, nos llama la atención la frecuencia sumamente reducida de PP + HABERE en textos latinos, puesto que la forma española correspondiente, «haber» + pp, posee una frecuencia notablemente alta. Incluso, como veremos más adelante, los primeros documentos castellanos ofrecían ejemplos de «haber» + pp con cierta frecuencia significativa. ¿A qué se debe esta diferencia numérica? La razón se puede buscar en estos mismos ejemplos que acabamos de ver. Se trata de usos fortuitos de la forma HABERE (‘tener’) + PP (Tipo-C), que expresaba el estado resultante de la acción anterior, que corresponde a la forma española «tener» + pp (*Tengo escrita la carta*), que tampoco presenta la frecuencia tan alta como la forma «haber» + pp (*He escrito la carta*), dotada de la función perfectiva. En los ejemplos citados hemos visto que PP + HABERE es todavía menos frecuente que HABERE sin PP. Esto se explica por tratarse de un uso léxico especial de HABERE ‘tener’. En cambio, la forma «haber» + pp no se trata de un uso léxico especial de «haber», sino de un uso funcional general del mismo.

El cambio del uso especial de la forma latina PP + HABERE en el uso general de la forma española «haber» + pp no ha sido repentino, sino paulatino y gradual. En las obras y documentos latinos tratados, hemos visto algunos indicios, escasos pero importantes. A mi modo de ver, su frecuencia aumentaba en la lengua hablada por su utilidad comunicativa que el pueblo reconocía, a pesar de que en la

lengua escrita aparecen pocas veces, por no aceptar fácilmente las nuevas formas coloquiales nacidas y propias de la lengua hablada del pueblo. La lengua escrita suele tardar en aceptar las formas coloquiales que hay en la lengua hablada, hasta que finalmente las utilizan cuando las formas dejan de ser coloquiales y llegan a ser normales.

2.2. Lenguas romances

Antes de abordar el tema de «haber» + pp en la historia de español, conviene observar la perspectiva románica general para situar el español dentro del contexto de tiempo y espacio.

Según Kuniyara (2007: 2), el latín hablado cambiaba continuamente de generación a generación y se diversificaba en distintas provincias romanas y finalmente entre el siglo VI y el VIII se convertían en las distintas lenguas romances (Fig. 1).

Kataoka (1982a: 293-295) compara las formas compuestas del tiempo perfecto de las seis lenguas romances, con respecto al verbo auxiliar: el español y el rumano utilizan PP + HABERE, mientras que el portugués, TENERE o HABERE, siendo el primero más usual. El italiano, el francés, el retorromance utiliza HABERE para todos los verbos transitivos y la mayoría de los intransitivos, reservando ESSE para una parte de los intransitivos. Ejemplos (primera persona singular y plural):

- (5) Portugués: *tenho amado, temos amado; hei amado, havemos amado*
Español: *he amado, hemos amado; he arribado, hemos arribado*
Francés: *ai aimé, avon aimé; sui arrivé[vée], sommes arrivés[vées]*
Italiano: *ho amato, abbiamo amato; sono arrivato[a], siamo arrivati[e]*
Retorromance: *hai amau, havein amau; sun arrivaus[vada], essan arrivai [vadas]*
Rumano: *am sperat, am sperat*

En esta visión románica, nos interesa la situación de español, que se encuentra entre el portugués con TENERE y el francés con ESSE.



Fig. 4. Lenguas romances en Europa (*apud* Posner 1996: 19)

Por otra parte, el mismo autor explica el origen de PP + HABERE en el uso del PP como un “adjetivo verbal”, por ejemplo, HABEO CULTELLUM COMPARATUM ‘Tengo un cuchillo comprado’ (Plauto), que se convierte posteriormente en el tiempo perfecto ‘He comprado un cuchillo’, nacido de la innovación del pueblo (Kataoka 1982b: 250). Al mismo tiempo, apunta que el verbo auxiliar se pone delante del participio en todas las lenguas romances, lo que impide su conversión en una forma sintética como ocurre en la forma futura o condicional (*compraré, compraría*) y favorece la forma perifrástica del tiempo compuesto. Encontramos la razón por la que el verbo auxiliar precede al participio en la ubicación del Objeto CULTELLUM (‘cuchillo’), que suele venir detrás del verbo, junto con su complemento, que es el participio COMPARATUM (‘comprado’).

De la explicación de Kataoka, confirmamos la división principal de las lenguas romances entre el portugués mayoritariamente con TENERE y el resto (español, francés, italiano, retorromance, rumano) con HABERE.

Las lenguas romances dentro de la Península Ibérica se distribuyen aproximadamente de la siguiente manera (Bando y Asaka 2005: xvi; Andrés Díaz 2013: 18):



Fig. 5. Lenguas romances en la Península Ibérica

Según Andrés Díaz (2013: 598), las lenguas iberorromances se dividen en dos grupos con respecto al tiempo compuesto: lenguas con «tener»: portugués (*tenho cantado*), gallego (*teño cantado*), asturiano (*tengo ~ teo cantado*), mirandés (*tengo cantado*), y lenguas con «haber»: castellano (*he cantado*), aragonés (*he cantau ~ cantato*), aranés (*è cantat*) y catalán (*he cantat*). El mismo autor (op. cit.: 595, 598) precisa:

Los tiempos compuestos con «haber» y con «tener» no son equivalentes. Los primeros expresan aspecto perfectivo con indicación del término de la acción; (...) Los tiempos compuestos con «tener», además del aspecto perfectivo, conllevan reiteración de la acción. Así, por ejemplo, gall-port. *falei* y ast. *falé* equivalen a 'hablé' y 'he hablado', mientras que port. *tenho falado*, gall. *teño falado* y ast. *tengo falao* significan 'he hablado en repetidas ocasiones, he venido hablando reiteradamente'. (595) / En gallegoportugués y asturleonés no existe la oposición «canté» / «he cantado», que se expresa siempre en el pretérito indefinido «canté»: así, mientras que «he» + participio expresa un pasado relacionado con el presente, «tengo» + participio añade la idea de reiteración o continuidad indefinida de la acción (gall. *Teño lavado roupa*. port. *Tenho lavado roupa*. ast. *Tengo llavao ropa*). (598)

De esta descripción, la división principal de las lenguas de TENERE se expande no solamente al portugués sino a otras lenguas occidentales: gallego, asturiano y mirandés. La continuidad geográfica de estas cuatro lenguas no debe ser accidental y, por tanto, merece la pena buscar sus razones históricas.

Terasaki (1987: 73-74) ordena las funciones semánticas del perfecto compuesto (PC, *he escrito*) en relación con el pretérito simple (PS, *escribí*) que varían en diferentes lenguas romances con la siguiente Tabla-2, sumamente útil para observar la gradación de uso panrománica:

funciones \ tipos		I	II	III	IV	V
aoristo		PS	PS	PS	PC'	PC
pf	inmed.	PS	PS	PC	PC	PC
	indef.	PS	PS	PC	PC	PC
	cont.	PS	PC	PC	PC	PC
	result.	PC	PC	PC	PC	PC

Tabla 2. Pretérito simple (PS) y perfecto compuesto (PC). Terasaki (1987)

Los cuatro usos del perfecto (pf) están explicados en Alarcos Llorach (1947: 46), en el orden de la evolución histórica. Los cito acompañados de la abreviación de Terasaki (loc.cit.):

1. result(ativo). “Expresión de la duración presente del resultado de una acción anterior (*pagado vos he*)”.
2. cont(inuativo). “Expresión de la acción continuada que ha producido un estado presente (*tanto avemos fecho*)”.
3. inmed(iato). “Expresión de una acción momentánea inmediatamente anterior al presente gramatical (*tan gran palabra havéis oído*)”.
4. indef(inido). “Expresión de una acción momentánea no inmediatamente anterior, pero sentida en relación con el presente (*según después me ha dicho*)”.

Los números romanos indican grupo de lenguas (Terasaki *loc.cit.*):

Al primer tipo [I] pertenecen los dialectos italianos meridionales, en los cuales el empleo del PC está limitado al tipo resultativo. En la antigüedad el latín vulgar sería del mismo tipo. El segundo grupo [II] se compone del español mexicano, otros dialectos americanos del español¹⁷, el gallego y el portugués, aunque en los últimos dos idiomas la forma de PC es “*ter + participo* (TENEOS SCRIPTUM)”. (...) Ejemplos del español mexicano: *¿Ya comiste?* / *¿Pensaste*

¹⁷ Terasaki (2011: 253) incluye en [II] Galicia, Asturias, León, Canarias e Hispanoamérica.

bien lo que te dije? / (...) El tercer grupo [III] consta de español castellano y occitano, cuyo PC se extiende por todos los terrenos del perfecto. Al cuarto tipo [IV] corresponde el catalán, que emplea, para el valor de aoristo, una forma compuesta peculiar, llamado pretérito perfecto perifrástico: *anar* + infinitivo (EO SCRIBERE). (...) / En el quinto grupo [V] caben el francés, el italiano septentrional y el rumano. Presentan la última etapa del desarrollo del PC, que prevalece casi por completo sobre el PS, particularmente en la lengua popular. (...)

El orden histórico que propone el cuadro de Terasaki (1, 2, 4, 3) es diferente del de Alarcos Llorach (1, 2, 3, 4), probablemente por tratar los usos 3 y 4 de manera conjunta¹⁸. Propongo volver al orden de Alarcos Llorach (1, 2, 3, 4), puesto que el uso 4 (indefinido, “presente ampliado”) se aproxima al uso aorístico. Por mi parte, para ver la continuidad histórica, me permito incluir el latín hablado coetáneo del latín clásico en la primera columna [1] con la distribución unitaria del tiempo PS y otra columna [4] correspondiente al español arcaizante mencionado en Alarcos Llorach (1947: 44)¹⁹ y hacer confluir el catalán con francés, italiano septentrional y rumano en la última columna, por considerar la forma compuesta *anar* + infinitivo en catalán como una especie de perfecto compuesto (PC). En la división [6] he incluido el español andino, donde “todo pasado perfecto tiende a expresarse mediante el perfecto compuesto” (Moreno de Alba, 2006: 2119)²⁰:

¹⁸ Terasaki (1987: 72): “Como observa Alarcos Llorach, creemos que en las primitivas etapas del español el PC tomó la función resultativa y luego la continuativa. En la etapa posterior conseguiría, en nuestro criterio, los usos del pasado indefinido y del pasado inmediato.”

¹⁹ Alarcos Llorach (loc.cit.): “Desde el siglo XVII, cuantas veces la forma simple aparece en lugar de la compuesta, se debe a arcaísmo afectado, a latinismo o a necesidades poéticas. (...) *hoy perdí, cielos, la esperanza que tenía* (Calderón, El Príncipe Constante, 3, 118).” Creo que los ejemplos ahí citados son del presente ampliado, puesto que se encuentra los adverbios idóneos para este uso: *en nuestros días, hoy, hasta ahora*. Supongo que los arcaístas afectados no utilizarían el pretérito simple para los usos 1. resultativo, 2. continuativo e 3. inmediato.

²⁰ [1] latín hablado coetáneo del clásico; [2] latín vulgar, italiano meridional; [3] español americano, español canario, gallego, portugués; [4] español arcaizante; [5] español castellano, occitano; [6] español andino, catalán, francés, italiano septentrional, rumano.

funciones \ tipos		[1]	[2]	[3]	[4]	[5]	[6]
pf	1. result.	PS	PC	PC	PC	PC	PC
	2. cont.	PS	PS	PC	PC	PC	PC
	3. inmed.	PS	PS	PS	PC	PC	PC
	4. indef.	PS	PS	PS	PS	PC	PC
5. aoristo		PS	PS	PS	PS	PS	PC

Tabla 3. Pretérito simple (PS) y perfecto compuesto (PC).

Tabla modificada de Terasaki (1987).

Este cuadro es útil para comprender la situación castellana medieval y la española moderna dentro de la perspectiva tanto histórica como geográfica de las lenguas romances. Es curioso observar la gradación formal del tiempo compuesto desde la periferia occidental, gallego, portugués, español americano hasta la zona oriental, catalán, francés, italiano septentrional y rumano. El español castellano y el occitano se encuentran en la zona intermedia. Esta visión panrománica se comprueba también en la Península Ibérica y, como veremos inmediatamente, en la comparación de los tres dialectos peninsulares: leonés, castellano y aragonés.

Terasaki (2011: 254) resume la situación panrománica de la siguiente manera (traducción mía):

Al observar el conjunto de las lenguas romances, encontramos dos polaridades: el grupo de francés e italiano coloquiales con la prevalencia del tiempo compuesto donde el perfecto presente casi ha sustituido al pretérito simple y el otro grupo de portugués y occitano donde prevalece el pretérito simple y casi no se utiliza el perfecto presente. El español está situado en la posición intermedia, en la que el español americano se aproxima al segundo grupo.

Según la lista de Alarcos Llorach (loc.cit.) y el cuadro de Terasaki (loc.cit.), el uso inicial del perfecto compuesto es 'resultativo'. Supongo que la gramaticalización de «haber» + pp inicia en la ambigüedad semántica que hay con «tener» + pp de sentido resultativo, de que habla Nakaoka (1993: 175; traducción mía)²¹:

La forma de 'aver + pp' competía semánticamente con la forma de 'tener + pp'. Especialmente en el español medieval cuando *aver* se utilizaba como verbo

²¹ Véase también Terasaki (2011: 159).

léxico [con sentido de ‘tener, poseer’], la diferencia semántica y funcional se hacía ambigua frecuentemente: *Avié el mesquiniello los braços encorvados, / teniélos endurecidos.*

Suzuki (1999) ha realizado un estudio comparativo de las distintas funciones presentadas en Alarcos Llorach (1980: 46) en el presente perfecto (*he cantado*) y pretérito simple (*canté*) en las cuatro obras de los siglos XVI y VII: *La Celestina* de Fernando de Rojas (CEL: 1499-1514), 2. *Diálogo de la lengua* de Juan de Valdés (DIA: 1535), *Pasos* de Lope de Rueda (PAS: “El deleitoso”, 1565; “Del registro de representantes”, 1570) y *Entremeses* de Miguel de Cervantes (ENT: 1615). Ha analizado los diálogos que hay en estas obras. Me permito citar las frecuencias absolutas y relativas y elaborar dos gráficos con las relativas (Tabla 5, 6 en Apéndice):

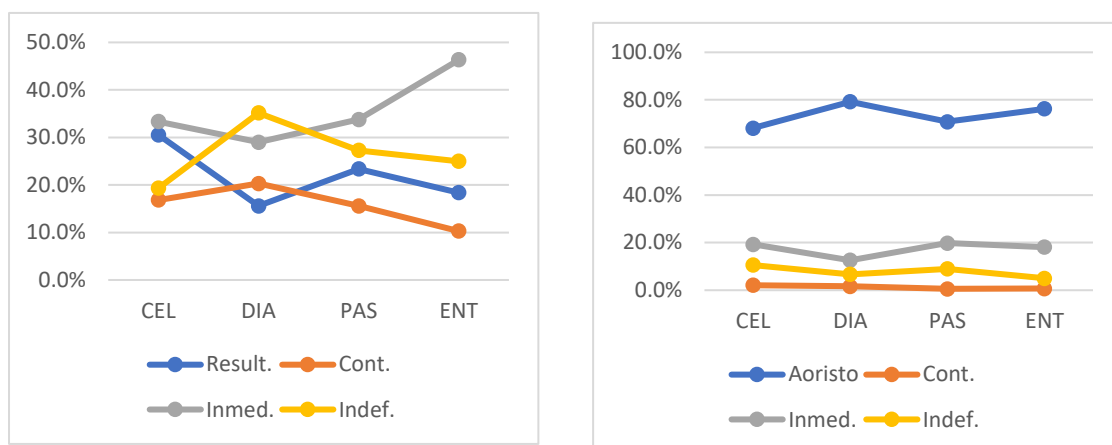


Fig. 6. Perfecto compuesto (*he cantado*) / Pretérito simple (*canté*). Suzuki (1999: 81).

De esta manera podemos apreciar las tendencias de cuatro usos del perfecto compuesto, que compiten en las cuatro obras tratadas. Nos interesa la tendencia descendente del uso resultativo, el origen de «haber» + pp, probablemente en relación con la nueva forma «tener» + pp. ***

¿Por qué causa se presenta este proceso progresivo del perfecto compuesto? Al respecto, Posner (1996: 180) explica de la siguiente manera:

(...) en muchas variedades romances modernas, las formas del compuesto han reemplazado a las formas del pasado simple en el habla coloquial. La explicación podría estar en el deseo de evitar las complejidades morfológicas del pasado simple (...), en la tendencia general a las construcciones analíticas o en el énfasis que recae en el aquí y ahora en el uso coloquial. (...)

Además de la causa morfológica (simplicidad formal de «haber» + pp), la sintáctica (preferencia por la construcción analítica) y la pragmática (énfasis en el aquí y ahora del hablante) mencionados por Posner, merece la pena pensar en la gradación semántica que hemos observado en el cuadro de Terasaki (op. cit.), donde podemos apreciar distintas etapas continuas y coherentes que explican la invasión cada vez más avanzada de la forma compuesta (*he comido, he tenido*) en el territorio de la forma simple (*comí, tuve*), especialmente en las lenguas de oriente, en contraste con las de occidente a las que no llega la misma invasión.

3. Historia de español

3. 1. Corpus

En esta sección, para conocer la historia de la forma compuesta «haber» + pp, voy a utilizar los documentos notariales recogidos en el corpus lingüístico CODEA («*Corpus de Documentos Españoles Anteriores a 1800*», dirigido por Pedro Sánchez-Prieto)²², que están integrados en nuestra plataforma LYNEAL («*Letras y Números en Análisis Lingüísticos*»)²³ que hemos venido desarrollando en colaboración con Antonio Moreno Sandoval. Los documentos notariales son idóneos para nuestro propósito de analizar los procesos históricos de la lengua, por presentar la fecha y el lugar de manera más inmediata que las obras literarias, que inevitablemente manifiestan el estilo tradicional precedente y con los cambios posteriores causados por la transferencia de copias no coetáneas (Jacob 2001: 155)²⁴.

Otra característica importante de los documentos notariales es su cantidad enorme, continuidad histórica, extensión geográfica y homogeneidad clasificada. Del corpus total, vamos a seleccionar los documentos procedentes de las cuatro regiones centrales: León, Castilla la Vieja, Castilla la Nueva y Aragón para situar la variedad castellana en comparación con las regiones colindantes:²⁵

²² <http://www.corpuscodea.es/> [7 de enero de 2021]

²³ <https://lecture.ecc.u-tokyo.ac.jp/~cueda/lyneal/> [7 de enero de 2021]

²⁴ Menéndez Pidal (1976: 360): “Parece que la ley de la concordancia [de participio con el objeto] se generalizó más en el siglo XIII que en el XII; lo cierto es que era más propia de la lengua de Per Abbat que de la del autor de El Cid”

²⁵ Utilizo la antigua zonificación regional y no la división administrativa actual, por considerar más idónea para estudios históricos y geográficos de la lengua española. León (LE: León, Zamora, Salamanca), Castilla la Vieja (CV: Cantabria, Palencia, Burgos, La Rioja, Valladolid, Ávila, Segovia, Soria), Castilla la Nueva (CN: Madrid, Guadalajara, Toledo, Cuenca, Ciudad Real, Albacete) y Aragón (AR: Huesca, Zaragoza, Teruel).

La tercera condición no menos importante especialmente para nuestro propósito de investigación es su marcado rasgo distintivo estilístico, normal, no literario. Aprovechando esta condición propia de los documentos notariales, podemos comparar dos tipos de texto muy diferente uno de otro, obras literarias tratadas en los estudios previos y los documentos notariales que analizaremos en esta sección (sec. 3).

La siguiente tabla muestra la frecuencia de la forma «haber» + pp, junto con la totalidad de palabras, frecuencia normalizada por mil palabras (FN)²⁶ y división establecida por el orden descendente de FN (Tabla 6 en Apéndice):

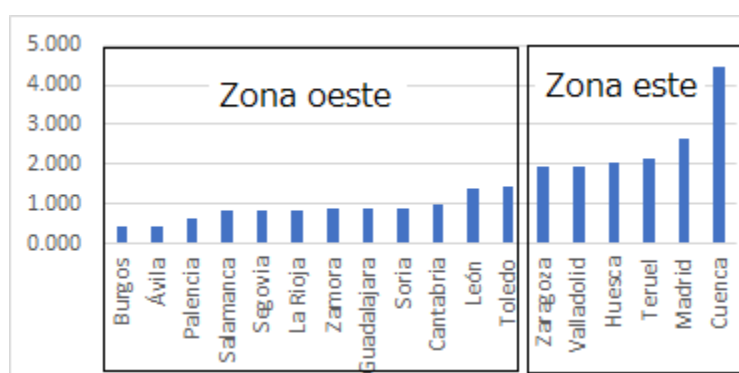


Fig. 7. «haber»+pp en provincias y división Oeste-Este. CODEA.

Este método que llamo ‘poscategorización’ difiere del método habitual de dialectología española. He hecho la división entre la Zona Oeste y la Zona Este, de acuerdo con la frecuencia normalizada de «haber» + pp en un punto entre Toledo y Zaragoza que corresponde aproximadamente a la división regional de Castilla - León y Aragón²⁷. En esta división en la Zona Oeste están incluidas no solamente provincias de Aragón (Huesca, Zaragoza, Teruel), sino también Valladolid, Madrid y Cuenca por ofrecer las altas cifras de frecuencia de uso de la forma gramatical «haber» + pp²⁸. No se trata de una división establecida *a priori* entre las regiones

²⁶ Recordemos que la frecuencia normalizada por mil palabras se deriva de la siguiente manera: por ejemplo, la frecuencia absoluta de «haber» + pp en Burgos 19 dividida por 44 007 multiplicada por 1000 igual a 0.432. La frecuencia absoluta no sirve para comparar con otras frecuencias, debido a la diferencia de cantidad de palabras que hay en cada provincia. En cambio, la frecuencia normalizada es comparable, puesto que está calculada sobre la base común de mil palabras.

²⁷ En realidad, es difícil trazar la división histórica de Castilla - León y Aragón, puesto que la frontera de los tres reinos se movía continuamente a lo largo de siglos, como visualizan Vives (1978), López-Davalillo Larrea (2000) y Induráin Pons (2018).

²⁸ No se encuentran datos en Ciudad Real y Albacete.

administrativas, sino de una zonificación lingüística realizada *a posteriori* en busca de un agrupamiento apropiado para nuestro propósito específico. En el siguiente mapa están visualizadas las localizaciones geográficas de las provincias concernientes:



Fig. 8. Zona Oeste y Zona Este de «haber»+pp. CODEA.

En cuanto a la cronología, el año de inicio es 1200, puesto que antes de esta fecha no contamos con suficientes datos. Extendemos la cronología hasta 1600, es decir, hasta 1649, cuando se determina la tendencia numérica de la forma en cuestión. Dentro de los cinco tipos documentales: cancilleresco, eclesiástico, judicial, municipal y particular, vamos a excluir el tipo cancilleresco por su carácter itinerante, es decir, la cancillería real se movía de un sitio a otro continuamente, lo que impide la identificación geográfica de la lengua, puesto que el lugar de emisión del documento no corresponde necesariamente al habla local. El número total de palabras en cada franja cronológica es el siguiente:

Año	1200	1250	1300	1350	1400	1450	1500	1550	1600
Oeste	11 376	50 631	52 993	40 829	63 396	70 574	66 605	25 220	20 404
Este	3 340	25 985	19 358	44 883	61 986	37 709	45 178	85 627	35 228
Total	14 716	76 616	72 351	85 712	125 382	108 283	111 783	110 847	55 632

Tabla 7. Cantidad de palabras en la Zona Oeste y Zona Este . CODEA.

El año de cada fila representa el inicio de la franja cronológica, de modo que, por ejemplo, 1200 corresponde al periodo desde el año 1200 hasta 1249. Estas cifras son importantes y necesarias para conocer la magnitud numérica comparable, la frecuencia normalizada (FN), de fenómenos gramaticales que vamos a ver a continuación.

3. 2. Factor de frecuencia

En la sección 2.2, hemos observado un gradiente geográfico de la forma compuesta «haber» + pp en distintas lenguas romances tanto en la Romania como en la Península Ibérica, según el cual las lenguas de oriente (francés, catalán, aranés, aragonés) favorecen la forma «haber» + pp, mientras que las de occidente (portugués, gallego, asturiano, mirandés) son reacias a ella. El castellano medieval y el español moderno se sitúan en medio de las dos extremas. Ahora vamos a ver la misma tendencia dentro de las cuatro regiones tratadas en el corpus CODEA, divididas en las dos zonas, Este y Oeste (sec. 3.1, Tabla 8 en Apéndice):

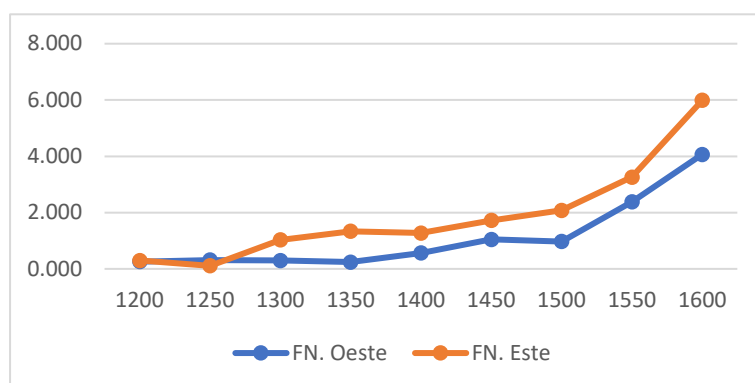


Fig. 9. «haber» + pp. Zona Oeste y Zona Este. CODEA.

Ejemplos: «**haber**» + pp. *lo que les yo he mandado e ordenado segund sobredicho es*, 1314, Segovia / *la qual dicha viña yo ove comprado e compré de Ferrant García de Camargo*, 1369, Burgos.

En el gráfico de la frecuencia normalizada (Fig. 9), podemos observar la

distribución numérica en la Zona Oeste y Zona Este, casi siempre en la este más frecuente que en la Oeste. La frecuencia reducida de «haber» + pp en la Zona Oeste puede explicarse en el contexto colindante de las lenguas romances occidentales que hemos visto en la sección 2.2. Por otra parte, la alta frecuencia de la Zona Este se extiende hacia oriente en la lengua catalana, que se relaciona con el francés, la lengua más avanzada de la perífrasis en toda Romania (sec. 2.2, 3.1).

Los datos que presenta Rodríguez Molina (2010: 2014) confirma la tendencia parecida de la distribución en espacio y tiempo (frecuencia normalizada por 50 documentos notariales) (Tabla 9 en Apéndice):

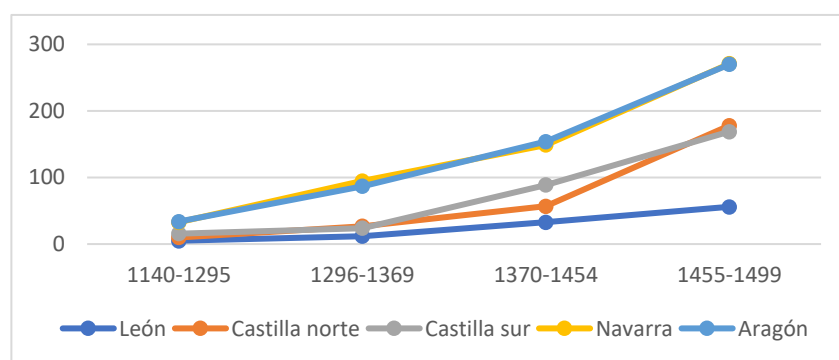


Fig. 10. «haber» + pp documentos notariales. Rodríguez Molina (2010)²⁹.

Según el mismo investigador, la mayor frecuencia de uso de «haber» + pp en la zona oriental (Castilla norte y Aragón) puede deberse a la influencia de la población franca (op.cit.: 1225):

Esta hipótesis puede apoyarse en el hecho de que los documentos del Alto Aragón o Navarra muestran más ejemplos innovadores que los de la zona del valle del Ebro; ya que los documentos de Jaca y su jurisdicción, Huesca, Pamplona y Estella son los que mayor frecuencia de los usos innovadores registran.

Rodríguez Molina (2010.: 1218) también ubica la zona primitiva de expansión de la perífrasis en el oriente de la Península:

Mi hipótesis es que tanto el foco de la innovación como la zona primitiva de expansión de los tiempos compuestos se encuentran en el oriente de la Península Ibérica, y que desde este núcleo inicial la perífrasis se fue

²⁹ En el gráfico se ven solapadas las líneas de Navarra y Aragón.

extendiendo de norte a sur, y, sobre todo, de este a oeste.

Según Barrio de la Rosa (2016: 268-269), el significado primigenio del verbo «haber» es de carácter incoativo: ‘entrar en posesión de algo’ y ha encontrado quince ejemplos de uso incoativo de «tener» en la zona oriental de la Península (op.cit.: 271):

Los primeros ejemplos de *tener* en contextos innovadores, propios de la semántica de *haber* y asociados con él, se distribuyen, principalmente, antes de 1500, en enclaves orientales pertenecientes a las actuales provincias de Teruel, Zaragoza, Huesca y Navarra.

Seguidamente, vamos a analizar la frecuencia de «haber» con participio («haber» + pp) en comparación con «haber» sin participio («haber»-pp) en el corpus CODEA (Tabla 10 en Apéndice):

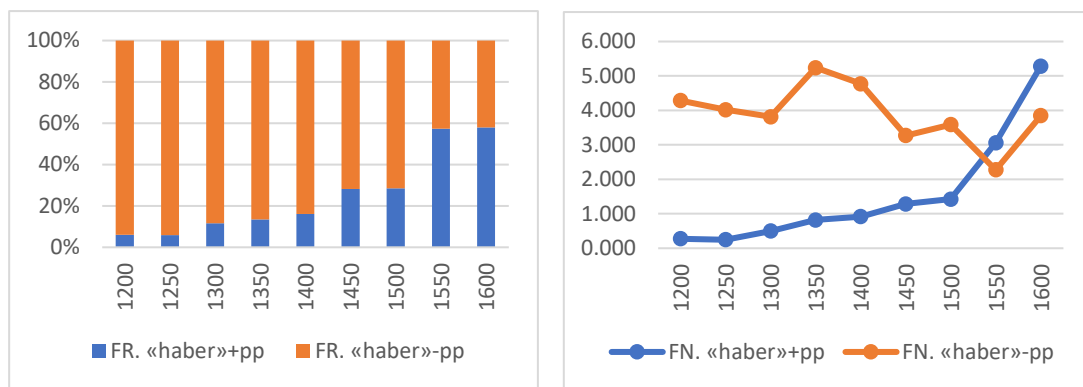


Fig. 11. «haber» + pp / «haber» sin pp. CODEA.

Ejemplos: (1) **«haber» + pp**: *lo cual dixo que avíades usado de tanto tiempo*, 1387, Madrid / *en los sus registros e escrituras avía quedado una carta de fuero*, 1507, León; (2) **«haber» sin pp**: *las nuestras casas mayores que avemos en esta dicha villa*, 1409, Guadalajara / *los lugares donde an e tienen sus testigos*, Madrid, 1532, Madrid.

En estas cifras y gráficos, podemos observar el crecimiento constante de «haber» con participio desde 1200, cuando había casos reducidos (6%), seguido de las etapas con cifras relativas (%) y normalizadas cada vez más grandes. En cambio, la frecuencia de «haber» sin participio disminuye con grandes altibajos. Por lo tanto, no estamos precisamente ante los datos que muestren la formación gramatical de «haber» + pp en correlación inversa con el uso pleno de «haber» sin pp.

Las cifras de frecuencia relativa (FR, %) pueden ser poco fiables, puesto que están condicionadas dentro de un conjunto uniforme de 100%. Especialmente en la comparación entre dos casos, cuando uno sube o baja, el otro necesariamente tiene que bajar o subir, como observamos en el gráfico izquierdo de la Fig. 10. Llama la atención el desarrollo parado de «haber» + pp entre 1550 y 1600 que se observa en el mismo gráfico. Esto está causado no por el poco desarrollo de «haber» + pp, sino más bien por la subida repentina de «haber» - pp en 1600, lo que se aprecia en el gráfico de la frecuencia normalizada (FN), donde precisamente se observa la subida importante de «haber» + pp en 1550 y 1600. En este sentido, la frecuencia normalizada dentro de todos los textos concernientes, representada por el enorme número de palabras, es más de fiar que la simple frecuencia relativa (FR), limitada forzosamente dentro de la totalidad de 100%. Por lo tanto, la frecuencia relativa es engañosa (Fig. 11, izquierda). En realidad, al observar el gráfico de la frecuencia normalizada (Fig. 11, derecha) notamos algunos años cuando ambos suben, concretamente en 1350, 1500, 1600, lo que demuestra la falsa correlación inversa entre «haber» + pp y «haber» - pp.

¿A qué causa se debe la subida constante de «haber» + pp en la frecuencia normalizada (FN)? Para responder a esta importante cuestión, hay que observar el gráfico derecho de la Fig. 10, donde apreciamos la subida constante de «haber» + pp en contraste con la curva descendente irregular de «haber» - pp. La ascendencia constante de «haber» + pp, sin solución de continuidad, representa el modo típico de la gramaticalización (cambio del elemento léxico en el gramático funcional, concretamente la auxiliarización de «haber»), necesariamente unidireccional. Nos llama la atención la última parte de la curva ascendente, especialmente pronunciada en 1550 y 1600. Estamos ante la parte inicial de la curva de forma “S”, que se caracteriza por el movimiento lento en el inicio, el rápido en el intermedio y, al finalizar la curva, otra vez el lento. Esta secuencia de lento - rápido - lento es causada por la curva normal de forma de campana constituida con dos partes de inicio y final de baja frecuencia y la parte central de alta frecuencia. El cambio lingüístico se supone que aparece con baja frecuencia y si el cambio se trata de una gramaticalización eficiente la velocidad de desarrollo aumenta cada vez más y en la etapa final para concluir el desarrollo disminuye la velocidad y construye la norma estable. La frecuencia acumulativa de la distribución de forma de campana se dibuja con la forma de “S”. Creo que una gramaticalización tan eficiente como la formación de «haber» + pp no necesita una causa externa, sino que se desarrolla por sí solo con impulso interno, lo que se observa en el gráfico de la frecuencia

normalizada (Fig. 11 derecha). Ahí se aprecia la parte inicial de la curva ascendente típica de forma “S” de «haber» + pp, independientemente de la curva de «haber» - pp.

El impulso interno de la gramaticalización no se encuentra independiente del entorno social geográfico, puesto que nadie vive solitariamente sin relación ni comunicación con la sociedad. De ahí que se presente la continuidad histórico-geográfica de la distribución de frecuencia de uso que acabamos de ver (Tabla 10, Fig. 10). En este sentido, no veo la influencia negativa de «haber» sin pp en la formación gramatical de «haber» + pp, es decir, el cada vez menor uso de «haber» sin pp no causa necesariamente la subida del grado de gramaticalidad de «haber» + pp. Vamos a ver otros factores gramaticales supuestos para explicar el proceso de gramaticalización de la forma compuesta.

3. 3. Factor semántico de «haber» y «tener»

Para estudiar el uso de verbos posesivos, «haber» y «tener», Ueno (1988, 1989, 1990) escogió “cinco obras literarias representativas de cada siglo” (1988: 121): *Poema de Mio Cid* (CID, hacia 1140); *Milagros de Nuestra Señora* (MLG, hacia 1246); *El Conde Lucanor* (LUC, 1335); *La Celestina* (CEL, 1490) y *Lazarillo de Tormes* (LAZ, publicada en 1554). Utilizando estos textos, ha comparado dos verbos «haber» y «tener» en su función de 1. Verbo principal ‘posesivo transitivo’ (1988), 2. Verbo auxiliar del perfecto compuesto (1989) y 3. Verbo auxiliar modal de perífrasis obligativa (1990), de los cuales nos interesa especialmente 1 (1988: 122) y 2 (1989: 68)³⁰ (Tabla 11 en Apéndice).

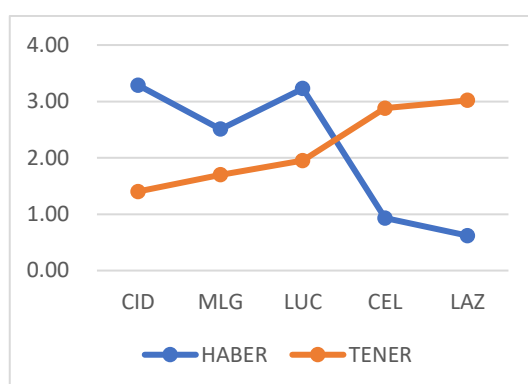


Fig. 12. Verbo principal ‘posesivo’. Ueno (1988).

La conclusión que ha sacado Ueno (1990: 137) es la siguiente:

³⁰ Utilizo solo las frecuencias absolutas para derivar las frecuencias relativas (%) y elaborar sus gráficos.

En resumen, *tener* ha venido poco a poco invadiendo las funciones de *haber* y le ha quitado casi toda la función de posesión y quedándose con ella. (...) en cambio, *haber* va limitándose en su uso a cada vez menos funciones, salvo la de verbo auxiliar para perfecto compuesto de los intransitivos y reflexivos, función que le ha quitado a *ser*, su rival en el español antiguo.

Para observar la extensión cronológica posterior vamos a aprovechar los datos ofrecidos por Pountain (1985: 342), donde apreciamos la vicisitud de los porcentajes desde *Poema de mio Cid* hasta Santa Teresa de Ávila (el gráfico es elaboración mía) (Tabla 12 en Apéndice):³¹

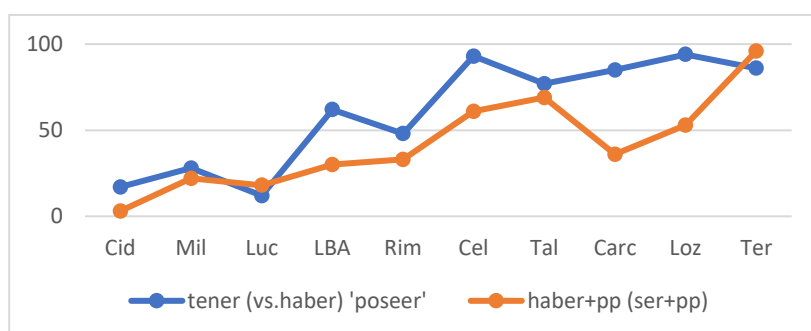


Fig. 13. Relación entre *tener* y *haber*. Pountain (1985).

Cito el comentario del mismo autor (loc. cit. traducción mía):

(...) el paso de *tener* hacia la expresión no marcada de posesión a expensas de *haber* es un proceso anterior y se acerca a su finalización más rápidamente que la sustitución de *ser* por *haber* como auxiliar perfecto.

Aquí la “expresión no marcada de posesión” se refiere a la posesión ‘literal’ (*tengo mil libras*), ‘inalienable’ (*tiene los ojos cerrados*) y ‘floja’ (*tengo un trabajo en una fábrica*). Se considera que la “sustitución de *ser* por *haber* como auxiliar perfecto” indica el progreso de auxiliarización de «haber»: *son llegados* > *han llegado*. Se trata de verbos intransitivos y reflexivos.

Como acabamos de ver, una de las causas importantes de la

³¹ Cid: *Poema de mio Cid*; Mil: Gonzalo de Berceo, *Milagros de Nuestra Señora*; Luc: Don Juan Manuel, *El Conde Lucanor*; LBA: Juan Ruiz, Arcipreste de Hita, *Libro de buen amor*; Rim: Pero López de Ayala, *Libro de poemas o Rimado de Palacio*; Cel: Fernando de Rojas, *La Celestina*; Tal: Alfonso Martínez de Toledo, *Arcupreste de Talavera o Corbacho*; Carc: Diego de San Pedro, *La cárcel de amor*; Loz: Francisco Delicado, *La lozana andaluza*; Ter: Santa Teresa de Ávila, *Epistolario*. He cambiado el orden de Tal - Cel en Cel-Tal, de acuerdo con la cronología literaria.

gramaticalización de «haber» + pp se considera el uso cada vez más frecuente del verbo «tener», que invade el campo semántico de ‘poseer’ del verbo «haber» (Olbertz 1992: 256). Veamos las frecuencias de «haber» y «tener» sin participio en el corpus CODEA, donde efectivamente «tener» crece continuamente, a medida que «haber» sin participio se aminora, pero con grandes vacilaciones (Tabla 13 en Apéndice):

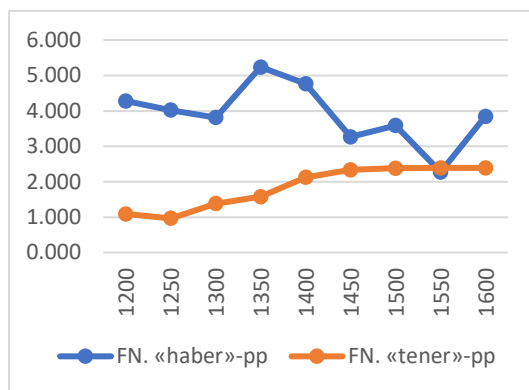


Fig. 14. «haber» sin participio y «tener». CODEA.

Ejemplos: **«haber»-pp.** *nós la dicha compannía avemos cerca de la dicha iglesia, 1405, Guadalajara / unas casas que yo é e tengo en esta dicha cibdat de Toledo, 1447, Toledo. **«tener»** *el que toviere el logar del arçobispo reciba so diezmo, 1226, Toledo / E esta casa tiene comunales puertas sin cerradura. 1287. Ávila.**

Ahora bien, nos interesa especialmente la relación que hay entre la gramaticalización de «haber» + pp y el progreso de uso de «tener» (a exclusión de «tener» + pp) (Tabla 14 en Apéndice):

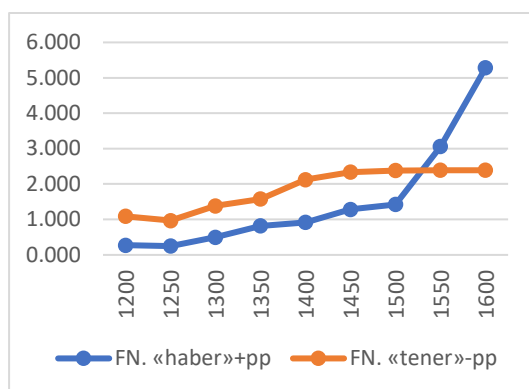


Fig. 15. «haber»+PP y«tener». CODEA.

Según estos datos, «haber» + pp parece correlacionar con «tener» hasta 1500.

Después de este año, «haber» + pp se independiza de «tener», supuestamente por causa de la finalización de su gramaticalización. Sin embargo, es difícil dar una conclusión sobre la relación causa - efecto de «haber» + pp y «tener», puesto que hay dos modos de pensar: causa «tener» → efecto «haber» + pp (por cadena de propulsión) o causa «haber» + pp → efecto «tener» (por cadena de tracción). La supremacía numérica de «tener» por encima de «haber» + pp antes de 1500 no constituye necesariamente un argumento para apoyar la iniciativa de «tener» como causa del crecimiento de «haber» + pp, puesto que la frecuencia de «haber» + pp podría subir independientemente de la de «tener». Pienso que la gramaticalización de «haber» + pp se ha realizado no necesariamente por causa de la invasión de «tener» en el campo semántico del verbo pleno «haber» ‘poseer’, sino por su propio desarrollo natural de gramaticalización unidireccional, en relación con la subida progresiva de frecuencia de uso como hemos visto en la sección anterior.

Vamos a ver la distribución de frecuencia de «tener» en espacio y tiempo (Tabla 15 en Apéndice):

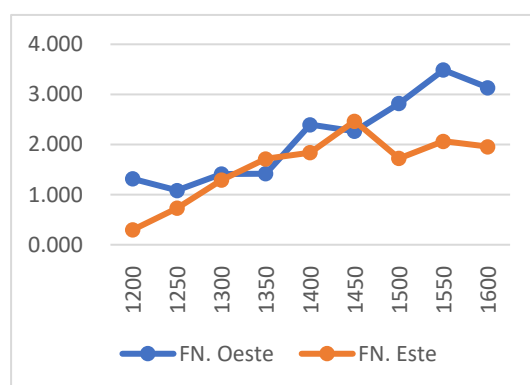


Fig. 16. «tener» - pp. Frecuencia normalizada. CODEA.

Es menos notable la tendencia ascendente general en la cronología de «tener» que la de «haber» + pp que hemos visto en la curva de la totalidad de regiones (Fig. 9). Es cierto que al sumar todas las frecuencias, aparece cierta curva característica de tendencia ascendente en «tener» (Fig. 15), pero al observar los detalles separados de zonas, las líneas se presentan de manera bastante heterogénea (Fig. 16). Esto nos obliga a poner cierta reserva en afirmar la hipótesis consistente en el crecimiento de «tener» como causa de la gramaticalización de la forma compuesta «haber» + pp. Es cierto que «tener» va ganando el campo semántico de «haber» ‘poseer’ (Fig. 14), pero esto no implica siempre la gramaticalización de «haber» + pp, que podría crecer ciertamente con pérdida de «haber» ‘poseer’, pero también sin ella, lo que ocurre por ejemplo en francés e

incluso en inglés, donde el verbo *avoir* y *have* mantienen el significado léxico de ‘poseer’ al lado de su uso auxiliar. El verbo pleno español «haber» con el significado léxico ‘poseer, tener’ no es necesariamente incompatible con su uso auxiliar «haber» + pp, como hemos observado en los siglos medievales, cuando las dos formas, el verbo pleno y el auxiliar, coexistían sin problema (Fig. 11). El crecimiento de «tener» y la gramaticalización de «haber» + pp pueden presentar la coincidencia sin correlación, cada uno con su propio desarrollo independiente del otro.

La tesis sobre la correlación entre «tener» y «haber» + pp se apoya en la supuesta correlación inversa entre «haber» - pp y «haber» + pp (Fig. 11), lo que he refutado anteriormente (sec. 3.2).

3. 4. Factores gramaticales

Esta subsección es poco productiva en el sentido de que los factores gramaticales que voy a tratar aquí, a mi modo de ver, no construyen indicios adecuados de la gramaticalización, concretamente de la auxiliarización de «haber» en documentos no literarios. Partimos de las cuatro normas gramaticales presentadas por Company (1983: 237-238):

El hecho fundamental que caracteriza al español medieval, en lo que respecta al funcionamiento de los tiempos compuestos, es la posibilidad de varias normas en una misma gramática: (i) Concordancia o no del participio de verbos transitivos con el Objeto directo de su oración (...) (ii) Interposición o no de constituyentes entre auxiliar y participio (...) (iii) Orden auxiliar + participio o participio + auxiliar (...) (iv) Un sistema de doble auxiliaridad con ser o haber según se tratara de verbos transitivos o intransitivos (...) / Esta gama de posibilidades en un mismo estado de lengua es lo peculiar del español medieval con respecto al español moderno, en el cual estas dos formas, auxiliar y participio, han sufrido un claro proceso de cohesión, indicado fundamentalmente por la gramaticalización del participio, con la consecuente pérdida de concordancia, el orden fijo de los formativos de la construcción: auxiliar + participio y la imposibilidad de interponer constituyentes entre ambos elementos.

Es importante reconocer el orden de proceso: la iniciativa de la gramaticalización del participio y su consecuencia en la no concordancia, el orden fijo y la no intercalación. Este esquema es válido para los textos literarios, pero plantea problemas para los documentos notariales como veremos a continuación.

3. 3. 1. Concordancia del participio

Como hemos visto en la sección 2.1, la gramaticalización de «haber» + pp consiste en el reanálisis de la estructura oracional. Reproduzco el mismo esquema:

Latín: [Objeto =~ (gén. núm.) PP] ~ [Sujeto =~ (pers.) HABERE]
Español: [Sujeto -= (pers.) [«haber» + pp]] - Objeto

En la estructura latina, la combinación de PP con Objeto está fortalecida por la concordancia de género y número entre los dos, mientras que la de pp con Objeto en la estructura española es débil, por falta de la concordancia gramatical y por alto grado de la gramaticalización de «haber» + pp. Esta reestructuración gramatical nace del mayor uso de la misma forma y, al mismo tiempo, por su parte la forma gramaticalizada produce todavía mayor uso de la misma por ser precisamente gramatical. La desaparición de la concordancia consiste en la unificación de la forma del participio. ¿En qué forma se unifica? Naturalmente en la forma más frecuente, menos especificada: la masculina singular, por ser más habitual. ¿Por qué razón se unifica la forma del participio? Por falta de necesidad de concordancia entre Objeto y participio en la forma compuesta gramaticalizada de «haber» + pp.

Yamamura (1982: 402) ha realizado un recuento de 172 casos de la concordancia de Objeto y participio en la forma «haber» + pp de *Poema de mio Cid*. Se dividen en tres grupos: (1) 62 casos del participio concordante (*nos han abiltados ifantes de Carrión*, 2942); (2) 24 casos del participio discordante (*desta batalla que avemos arrancado*, 814); (3) 86 casos de masculino singular, donde no sabemos si hay concordancia o no. La misma autora destaca el alto porcentaje de la concordancia (1) ($62 / [62 + 24] = 72\%$) y lo interpreta como una fijación formal incompleta de «haber» + pp.

Ito (1978: 108) destaca la opinión de Miguel de Unamuno (1977: 264) sobre la concordancia del participio en el *Poema de mio Cid*:

El uso frecuentísimo del *lo* representando una oración entera y la extensión grandísima de las oraciones en que el participio atributivo se refirió a una oración entera fue lo que generalizó el uso del participio atributivo neutro, que fue más tarde prevaleciendo en todos los casos. El verso 198: *atorgar nos hedes esto que auemos parado* nos presenta un ejemplo típico de la clase de oraciones cuya construcción, extendida analógicamente ha pasado a todos los casos.

Junto con la alta frecuencia de *lo* neutro como representante de una oración entera, también la del Objeto masculino singular en general es importante por ser el término preponderante, no marcado, como veremos inmediatamente³².

Sato (1991: 57) considera que la forma del participio, concordante o discordante, es el factor determinante de la formación gramatical de la combinación «haber» + pp, dejando al lado el orden de «haber» y pp, y la intercalación de elementos entre ellos, puesto que estos dos fenómenos gramaticales se deben a razones estilísticas de enunciación y *variatio*. Su cálculo sobre la concordancia del participio en las cuatro obras literarias presenta las siguientes cifras: *Cantar de mio Cid*: concordancia 51 (72.9%) / discordancia 19 (27.1%); *Milagros de Nuestra Señora*: c. 36 (92.3%) / d. 3 (7.7%); *El Conde Lucanor*: c. 6 (11.5%) / d. 46 (88.5%); *La Celestina*: c. 1 (0.3%) / d. 99.7%³³. Según la autora, se ha alcanzado la fijación de la forma compuesta en *El Conde Lucanor*.

Tsujii (1989: 157) ofrece la estadística de los casos de concordancia y discordancia del participio en «haber» + pp en tres obras literarias: *Lapidario* (1250) 16 (70%) : 7 (30%); *Conde Lucanor* (1335) 5 (11%) : 42 (89%); *Arcipreste de Talavera o Corbacho* (1438) 3 (11%) : 24 (89%), según lo cual parece que en el siglo XIV hubo el cambio de prevalencia en la discordancia. Por otra parte nos llaman la atención las frecuencias importantes de concordancia que hay en las obras literarias a lo largo de estos siglos: 70%, 11%, 11%, lo que se contrasta a la situación de los documentos notariales que veremos inmediatamente..

En cuanto a la forma del participio según género (masculino, femenino) y número (singular, plural), los estudios anteriores han indicado su unificación sucesiva en forma neutra (masculina singular). Nuestros datos de documentos notariales (CODEA) también lo confirman por las siguientes cifras y gráficos (Tabla 16 en Apéndice)³⁴:

³² Nakaoka (1993: 173) anota que desde tiempo temprano hay casos del “mantenimiento del participio en «haber» + pp en la forma neutra”. Por lo tanto, la aparente discordancia de género y número del participio es más bien una unificación morfológica en la forma neutra, no en la forma masculina singular, puesto que en el participio de «haber» + pp unificado, no existe la distinción de género y número, sino que se mantiene la forma neutra..

³³ No disponemos de la frecuencia absoluta.

³⁴ mas: masculino, fem: femenino, sg: singular, pl: plural.

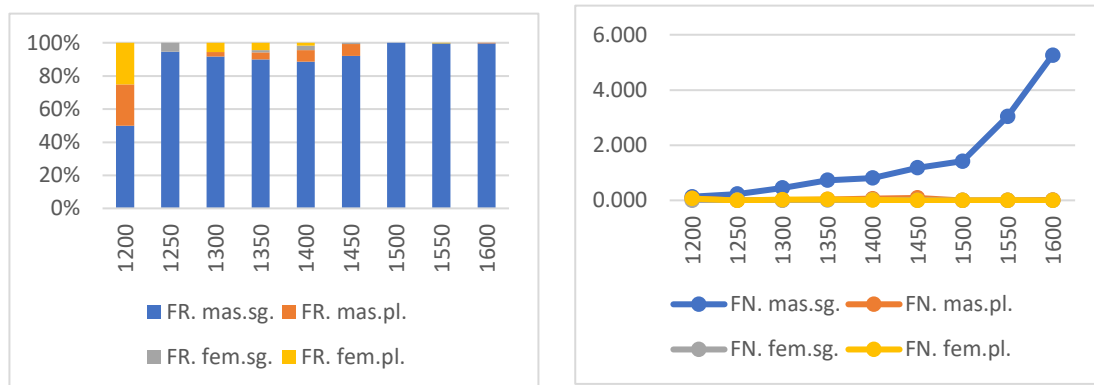


Fig. 17. Género y número del participio. CODEA.

Ejemplos: (1) **Masculino singular.** *a esa dicha villa de Peñafiel, adonde había estado preso cinco o seis días en la cárcel*, 1574, Valladolid / (2) **Masculino plural.** *E vós abat don Peidro aviedes dados CC moravedís a mio padre Día Sánchez*, 1244, Burgos / (3) **Femenino singular.** *nós la avemos determinada por ciertas afrontaciones de cadaúna*, 1257, Zaragoza / (4) **Femenino plural.** *e avan recomendadas aquéllas en compensación*, 1396, Huesca

Los datos de 1200, a pesar de presentar el porcentaje en la misma escala que otros, no son muy fiables por lo escasas que son las cifras absolutas. De todos modos, la curva ascendente pronunciada de la frecuencia normalizada (FN) a partir de 1500 es notable. También hemos de destacar su preponderancia casi total en la frecuencia relativa (FR. %). Las otras formas (masculina plural, femenina singular y plural) son indicadoras de la discordancia de género y número del participio con el Objeto. La forma masculina singular incluye tanto la forma concordante con el Objeto masculino singular como la forma discordante con el Objeto masculino plural, femenino singular o plural. Supongo que la preponderancia numérica de la forma masculina singular, sea concordante o discordante, es tanto causa como efecto de la formación gramatical de «haber» + pp, al mismo tiempo.

Según Rodríguez Molina (2016: 456), la misma concordancia está condicionada por los siguientes parámetros lingüísticos: “(...) la categoría gramatical del Objeto, su posición sintáctica, los rasgos morfológicos de los controladores (género, número) y su posición relativa en la escala de animación ampliada (definitud, animación)”. Aquí, “los rasgos morfológicos de los controladores (género, número)” son de pronombres concordantes con el participio. Sus datos son obras literarias y fueros del siglo XIII a XV, donde existen numerosos casos de discordancia que permiten análisis cuantitativos para llegar a su conclusión

fidedigna. En cambio, los documentos notariales recogidos en el corpus CODEA presentan otro aspecto muy diferente, donde contamos con solo 39 casos dentro de 1174 en total (Tabla 16). Tampoco estos 39 casos arrojan una tendencia numérica de los rasgos morfológicos (género y número). Los pocos casos de concordancia de género y número del participio con el Objeto indican la fase final de la auxiliarización de «haber» (+ pp) ya en el inicio de la cronología tratada (1200). De modo que no sirve mucho para conocer el grado de gramaticalización en los siglos posteriores.

3. 3. 2. Orden e intercalación

Yamamura (1982) ha analizado cuantitativamente el uso del pretérito compuesto (*ha dicho*) en *Cantar de mio Cid*. Para estudiar la formación histórica de «haber» + pp, son importantes dos tablas de frecuencia que proporcionan la autora: de la intercalación del pronombre y de la concordancia de Objeto y participio (sec. 3.3.1). Aquí nos interesa la intercalación del pronombre (Yamamura, 1982: 400):

Orden / posición	Entre «haber» y pp	Delante de «haber»+pp	Detrás de «haber»+pp
Inicial, pp+«haber»	9 (Ej. 1)	0	0
No inicial, id.	19 (Ej. 2)	0	0
Inicial, «haber»+pp	0	0	1 (Ej. 3)
No inicial, id.	3 (Ej. 4a)	51 (Ej. 4b)	0

Tabla 17. Posición de pronombre átono en *Cantar de mio Cid*. Yamamura (1982)

Ejemplos: Ej. 1. Inicial de oración, pp precedente: el pronombre entre «haber» y pp. *esposos los an ifantes de Carrión*. (3219) / Ej. 2. No inicial de oración, pp precedente: el pronombre entre «haber» y pp. *Estas apreciaduras mio Cid presas los an*. (3250). / Ej. 3. Inicial de oración, «haber» precedente: el pronombre, detrás de «haber»+pp. *a Saragoça metuda l'a en paria* (914).³⁵ / Ej. 4a. No inicial de oración, «haber» precedente: el pronombre, entre «haber» y pp. *honores e tierras avellas condonadas* (887)³⁶. / Ej. 4b. No inicial de oración, «haber» precedente: el pronombre, delante de «haber» y pp. *Assí commo lo ha dicho, todos adobados son*. (383).

En lugar de observar la posición del pronombre con los tres elementos:

³⁵ Yamamura trata la *a* inicial como una forma verbal. En cambio, la interpreto como una preposición y, en su lugar, el verbo lo encuentro en *l'a*.

³⁶ Se trata de una forma imperativa *aved*, detrás de la cual es normal que venga el pronombre.

«haber», pp, pronombre, propongo una clasificación más simple con dos elementos: [pro(nombre)-«haber»] y [participio]. Al sustraer los dos casos dudosos (Ej. 3 y 4a), quedamos básicamente con dos estructuras: (1) pp + pro-«haber» (en posición inicial y no inicial: Ej. 1, 2) y (2) pro-«haber» + pp (en posición no inicial: Ej. 4b). De esta manera, confirmamos el orden establecido del pronombre delante del verbo conjugado³⁷ y la variación doble de la posición de pp con respecto a «haber» que se unifica posteriormente en una posición fija de pp, que constituye un cambio significativo de la formación de «haber» + pp. Llama la atención la supremacía numérica de pro-«haber» + pp (Ej. 4b) en posición no inicial, que posteriormente prevalecería en todas las posiciones.

Veamos la variación del orden de «haber» y participio que se observa en el corpus CODEA (Tabla 18 en Apéndice):



Fig. 18. «haber» + pp / pp + «haber». CODEA.

Ejemplos: (1) «haber» + pp. *Iten he comido carne de judíos & moros*, 1492, Cuenca / *E diz que non le avéis querido admitir ni recibir*, 1513, Madrid. (2) pp + «haber». *Premisas las devidas recomendaciones, recibido é una vuestra carta de requisición (...)*, 1473, Toledo / (...), *finestras o forados que feito avredes*, (...), 1417, Teruel (...), *e segund esto que dicho é(...)*, 1364, Valladolid.

Estas cifras y gráficos muestran que los casos del orden marcado pp + «haber» son escasos en general, sin manifestar ninguna tendencia cronológica. Por lo tanto, no podemos usar la frecuencia del orden de «haber» + pp, con respecto al de pp + «haber» como indicio de la gramaticalización.

Estoy de acuerdo con la explicación que da Rodríguez Molina (2010: 1352) sobre el orden de «haber» y participio:

³⁷ Efectivamente, Yamamura (op.cit.: 400) menciona: “desde la antigüedad hubo muy poca posibilidad de que se fusionaran el participio y el pronombre [en este orden]”.

(...), hay que señalar que desde los primeros textos romances el orden más frecuente es ya el que se ha impuesto en español actual, como han puesto de relieve los trabajos que se han ocupado de esta cuestión. Así, a partir del siglo XIII hay ya un claro predominio del orden en el que el auxiliar precede al participio, y en el siglo XV la inversión del participio puede considerarse como residual.

Y seguidamente pone una nota: “Por ello, es difícil precisar con exactitud la cronología del cambio V-Aux > Aux-V, pues este debió producirse en el período previo al nacimiento de la escritura romance.”

Creo que de ahí viene la distribución escasa y esporádica de pp + «haber» y de intercalación en CODEA que acabamos de ver en la Tabla 18 y la Fig. 17, probablemente debida al carácter terminado y residual de que habla Rodríguez Molina (loc.cit.). Se trata de documentos notariales, que reflejan el habla de la cronología de manera más inmediata que las obras literarias y judiciales, donde pueden aparecer rasgos arcaicos ya desaparecidos en el habla del pueblo³⁸.

Otro tema relacionado con el del orden de «haber» y participio es la posibilidad de intercalación de pronombres átonos y elementos léxicos. Sin embargo, la frecuencia de la intercalación es escasísima en CODEA (seis casos en total):

Ejemplos: «**haber**» + [incalación] + pp. (...) *los ditos VI soldos II dineros qu' é yo ordenado con esta condición*, (...), 1379, Teruel / (...) *e oficiales o cualquier d'ellos lo oviesen todo así juzgado, e sentenciado, e dado e mandado*, 1455, Zamora / (...) *o cualquier d'ellas lo oviesen todo así juzgado, e sentenciado, e dado e mandado*, 1459, Zamora / (...) *e les preguntó si avían bien visto e mirado las viñas* (...), 1462, Salamanca / (...) *o qualquier d'ellos lo oviesen así visto e julgado e dado por su sentencia* (...), 1472, Salamanca / (...) *y falta de la dentadura sin remedio porque avía ya mudado*, (...), 1595, Cuenca.

Romani (2006: 289) presenta el porcentaje de casos de «haber» y pp adyacentes, es decir, sin intercalación, frente a los casos con intercalación: siglo XII 83% - 17%, XIII 88% - 12%, XIV 93% - 7%, XV 99% - 1%. A mi modo de ver, el porcentaje que ocupan los casos de intercalación (17%, 12%, 7%, 1%) es todavía

³⁸ Precisamente Company (1983: 252) anota: “Es posible pensar que la interpolación de constituyentes entre auxiliar y participio fuera característica de normas más elaboradas como la literaria, dado que en los *Documentos lingüísticos de Castilla*, de carácter notarial, esta variable ofrece una frecuencia de uso excesivamente baja (0.04%), en comparación con los demás textos analizados, todos ellos literarios.”

bastante alto, en comparación con los datos de CODEA, debido a la diferencia estilística que hay entre los materiales de Romani (2006) que incluyen obras literarias en buena medida y los documentos exclusivamente notariales (CODEA).

3. 3. 3. Transitivo e intransitivo

Anteriormente (sec. 2.1) hemos visto que la forma compuesta «haber» + pp nació de la combinación latina de PP + HABERE, donde el verbo HABERE es transitivo con significado de ‘tener, poseer’, que exige un Objeto directo: SCRIPTAM LITTERAM HABEO ‘Tengo escrita la carta’. En cambio, en el español actual, la misma forma compuesta «haber» + pp se utiliza tanto con el Objeto como sin él: *He escrito la carta / He caminado mucho*. Es de suponer que en las épocas tempranas, el castellano utilizaba «haber» + pp solo con los verbos transitivos casi a exclusión de los intransitivos (cf. Yllera 1980: 230-243; 276-281)³⁹. Veamos sus frecuencias cronológicas (Tabla 19 en Apéndice):

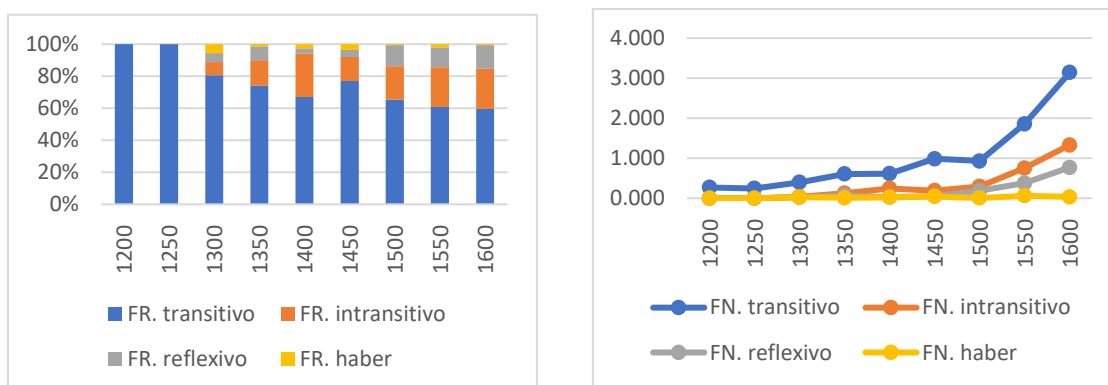


Fig. 19. «haber» + pp (transitivo / intransitivo / reflexivo / haber). CODEA.

Ejemplos: (1) **pp (transitivo)**. *otros ganaderos an metido sus ganados en los dichos términos*. 1492, Toledo / *habían cometido grave y atroz delito y desacato nuestro*, 1574, Valladolid. (2) **pp (intransitivo)**. *Sancho Pérez avía venido e entrado en la iglesia*, 1317, Zaragoza / *por ruego d'ellos avían ido a ver las dichas viñas*, 1462, Salamanca. (3) **pp (reflexivo)**. *se á encargado de fazer la dicha fiesta perpetuamente*, 1588, Madrid / *nunca por cartas nos avíamos*

³⁹ Penny (2006: 193) considera más bien los casos sin el Objeto que los propios verbos intransitivos: “cuando encontramos ejemplos de HABEO + participio sin Objeto directo patente (es decir, sin nada que exprese «poseído»), como en HABEO INTELLECTUM en lugar de INTELLEXI, «he entendido», podemos concluir que HABEO se ha gramaticalizado como morfema que expresa aspecto perfecto y que el concepto de posesión se ha perdido.”

comunicado, 1642, Valladolid. (4) **pp (haber)** en *nuestra villa de Alcalá de Henares á avido y ay desorden*, 1596, Madrid / *poner remedio en algunos descuidos que á avido en la administración de la hazienda*, 1626, León.

La tabla y el gráfico de las frecuencias relativas (FR. %) confirman que la forma «haber» + pp intransitivo no existe en los primeros años (1200, 1250), nace en 1300 y la proporción con respecto a la forma competidora «haber» + pp transitivo crece en épocas posteriores, lo que no implica que «haber» + pp transitivo decrezca. Efectivamente la tabla y el gráfico de la frecuencia normalizada (FN) muestran el crecimiento numérico constante tanto de la forma transitiva como la intransitiva y la reflexiva. El crecimiento de uso de «haber» + pp (intransitivo) no puede ser indicio de la auxiliarización de «haber». A mi modo de ver, los dos crecen de manera independiente⁴⁰.

Por su parte, Rodríguez Molina (2010: 1074) considera que “el retroceso de la sintaxis medieval (pérdida de la concordancia, de la interpolación y del orden V-Aux) es un desarrollo independiente y no coincidente con la generalización semántica de HABER + PTCP [«haber» + pp]”. En su lugar, para medir el grado de gramaticalización, propone utilizar los nuevos ocho parámetros: A. Animación del sujeto, B. Transitividad, C. Diátesis, D. Clase de objeto, E. Combinatoria léxica I: Estado, F. Combinatoria léxica II: Modales, G. Integración paradigmática, H. Frecuencia de uso. Entre ellos, son especialmente destacables B y H, que poseen importancia numérica, progresiva a lo largo de siglos. Como hemos tratado el factor de frecuencia anteriormente (sec. 3.2.), nos enfocaremos exclusivamente en B. Transitividad. A continuación, cito su explicación (op.cit: 1138):

B. TRANSITIVIDAD. El modelo predice que habrá diferencias cronológicas, diatópicas y textuales de naturaleza cuantitativa en la proporción de esquemas transitivos e intransitivos. Cuanto más antiguos sean los textos, menor número de predicados transitivos aparecerán, pues *haber* + PTCP [PP] se habrá extendido a predicados intransitivos. (...). ***

El mismo autor presenta la siguiente distribución de «haber» + pp con participio intransitivo en tiempo y espacio (op.cit: 1147). Se trata del crecimiento de

⁴⁰ Aquí otra vez la frecuencia relativa engaña, presentado la menor cifra del participio transitivo en 1600 (60%), a pesar de su más alta frecuencia normalizada (3.146).

su uso a lo largo de siglos: *es venido* > *ha venido* (Tabla 20 en Apéndice):⁴¹

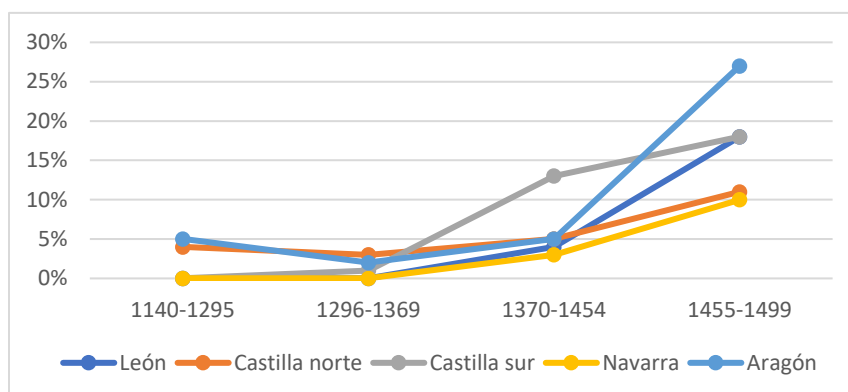


Fig. 20. «haber» + pp con participio intransitivo. Rodríguez Molina (2010).

Se trata del crecimiento general de la nueva forma «haber» + pp (intransitivo), confluida con la forma precedente «haber» + pp (transitivo). La línea ascendente de Aragón en la segunda mitad del siglo XV es notable, lo que es acorde con la geografía general de «haber» + pp, como hemos visto anteriormente en CODEA (sec. 3.2).

El motivo gramatical del cambio de «ser» + pp (intransitivo) a «haber» + pp (intransitivo) está explicado por Nakaoka (1993: 175, traducción mía):

‘ser + pp’ era la forma para la voz pasiva y, al mismo tiempo, para la forma que expresaba el resultado de la acción perfectiva (correspondiente a la forma moderna «estar» + pp). De esta manera, el rendimiento funcional de una forma se hizo excesivo y, por lo tanto, la misma forma [«ser» + pp] no se mantuvo como una forma del tiempo compuesto. De esta manera, la duplicación funcional producía la duplicación semántica de la misma forma, lo que causaba la ambigüedad. *alli esta muy desmayada, / a un Requexo es assentada* (‘en un rincón se sentó / está sentada’) - *Pedíla por mujer, ssó con ella casado* (‘me casé con ella / estoy casado con ella’)

Ahora bien, veamos la frecuencia normalizada de «haber» + pp (intransitivo) en el corpus CODEA con parámetros cronogeográficos (Tabla 21 en

⁴¹ León (Galicia, Asturias, León, Extremadura leonesa); Castilla norte (Tierra de Campos, Burgos, Extremadura castellana); Castilla sur (Toledo, Cuenca y Guadalajara, Andalucía, Murcia); Navarra (Navarra, Valle del Ebro); Aragón (Alto Aragón, Bajo Aragón).

Apéndice):

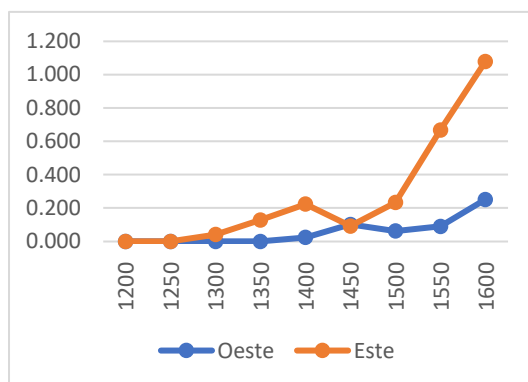


Fig. 21. «haber» + pp (intransitivo). CODEA.

En el mismo gráfico, observamos la baja frecuencia de «haber» + pp (intransitivo) en general hasta 1500. En 1550 y 1600 se diversifica cuantitativamente entre la Zona Este y la Oeste. Se confirma la supremacía de la Zona Este. En cambio, en la Zona Oeste no se observa una ascendencia tan notable como en la Zona Este. De modo que el criterio de intransitividad no representa completamente el grado de gramaticalización de la forma compuesta. Es útil para la Zona Este, pero no lo es tanto en la Zona Oeste en documentos notariales.

El valor del segundo parámetro de la gramaticalización, la frecuencia de uso, es innegable, como hemos visto en la sección 3.2.

Otro tema interesante relacionado con la transitividad del participio es la selección de «haber» o «ser» de acuerdo con el rasgo gramatical del participio⁴². Company

⁴² En realidad, tenemos que atender a la ‘hipótesis inacusativa’, que está explicada por Rodríguez Molina (2010: 1022): “La existencia de la doble auxiliaridad en español antiguo ha sido notada desde antiguo y, generalmente, se ha asociado con el carácter transitivo / intransitivo de los verbos. De este modo, el verbo *ser* quedaría como auxiliar de los verbos intransitivos, *haber* de los transitivos. Esta es la descripción que recoge la mayoría de los tratados de gramática histórica y las monografías existentes sobre la lengua de textos literarios antiguos. No obstante, actualmente se piensa que la doble auxiliaridad romance es un fenómeno vinculado a la intransitividad escindida y a la hipótesis inacusativa (...). Esta hipótesis predice que los verbos intransitivos inergativos y los verbos transitivos seleccionan HABER como auxiliar, mientras que solo los verbos intransitivos inacusativos se auxilian con SER. Si bien esta hipótesis no está exenta de problemas, resulta mucho más satisfactoria que la hipótesis tradicional” y pone la siguiente nota: “La hipótesis inacusativa resulta mucho más satisfactoria que la hipótesis tradicional porque permite explicar el hecho de que algunos verbos intransitivos, como *andar* o *correr* nunca se auxiliaron con el verbo *ser*, puesto que estos verbos son inergativos y no inacusativos: tienen sujetos agentivos, no permiten la formación de construcciones absolutas (*andado el niño, *corrido Juan), etc.” Pero, “la aplicación de los criterios de inacusatividad (...) a los datos del español antiguo no está exenta de problemas, (...) verbos que en principio pueden considerarse inacusativos

(1983: 241) lo explica de la siguiente manera:

A diferencia del español actual en que *haber* está generalizado como el único auxiliar de los tiempos compuestos del verbo, en los textos medievales en estudio coexisten, (...), dos sistemas parciales para indicar la auxiliaridad. / a) Por un lado *haber* parece especializarse como auxiliar de los tiempos compuestos de verbos TRANSITIVOS: (...) / b) Por otro, *ser* no sólo es auxiliar de las perífrasis de significado pasivo como en el español de hoy: (...) sino que se especializa como auxiliar de los tiempos compuestos de verbos INTRANSITIVOS, fundamentalmente de movimiento —*venir, ir, salir, llegar*. . . etc.— o los desinentes *nacer* y *morir*, y también es *ser* el auxiliar de construcciones que hoy interpretaríamos como reflexivas (...)

Vamos a comparar la frecuencia de «haber» + pp y la de «ser» + pp, ambas del participio intransitivo en el corpus CODEA (Tabla 22 en Apéndice):

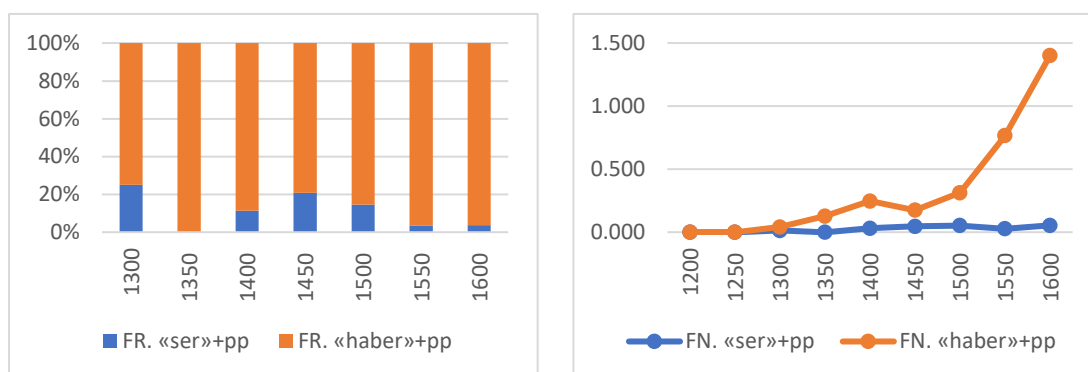


Fig. 22. «ser» / «haber» + pp (intransitivo). CODEA.

Ejemplos. «ser»+pp (intransitivo). *era cosa notoria que era fallecido e pasado d'esta presente vida*, 1507, León / *Pues el marqués de Tarazona es ya venido*, 1615, Madrid. «haber»+pp. *por mi indisposición nunca emos venido en declaración*, 1506, Zaragoza / *no á llegado la carta a mis manos.*, 1591, Madrid.

Ciertamente se observa el uso de «ser» a lo largo de siglos, pero siempre con

aparecen en algunas ocasiones auxiliados con el verbo *haber* (...). en español medieval, los verbos meteorológicos y los verbos *ser* y *estar* se auxilian por norma general con *haber*, así como la pasiva compuesta (ha sido dicho) y los verbos reflexivos puros (yo me lavo) (...) Sin embargo, en otras lenguas romances, como es el caso del italiano, tanto los verbos meteorológicos como el verbo SER, las pasivas compuestas y los verbos reflexivos puros se auxilian con el verbo SER”. De manera que es una “hipótesis que presenta numerosos escollos para su defensa”.

mucha menos frecuencia que el de «haber». Tampoco se aprecia su uso descendente sino con altibajos irregulares, lo que no permite tomar como criterio coherente de gramaticalización de «haber» + pp. También hay que tomar en consideración la diferencia estilística que existe entre obras literarias de estilo escrito tradicional y documentos notariales de estilo más cercano a la tendencia innovadora.

3. 3. 4. Conjunto

En lo que se refiere a las propiedades que posee la forma «haber» + pp, Azofra Sierra (2009: 95) las resume de la siguiente manera:

La forma más frecuente del nuevo perfecto compuesto, en español medieval, debió de ser la que presentaba el auxiliar pospuesto y concordancia de participio con complemento: *escrita he una carta*. Posteriormente, como consecuencia del proceso de gramaticalización, se producen cambios en la estructura sintáctica: / 1. En primer lugar, al desvincularse el participio y el CD [complemento directo, Objeto], deja de haber concordancia en género y número entre ellos, por lo que el participio se inmoviliza en la forma de masculino singular hacia finales del siglo XV: *escrito he una carta*. / 2. Por otro lado, de acuerdo con la tipología del español, el orden de los elementos de la perífrasis tiende a estabilizarse como auxiliar + participio + CD: *he escrito una carta*. / 3. Además, al estrecharse la relación entre auxiliar y participio, cada vez será más difícil encontrar elementos intercalados entre ellos: esp. mod. *la he escrito* (esp. med. *escrito la he*) / 4. Por último, debido a su desemantización total, *haber* deja de sentirse como transitivo y puede así convertirse en auxiliar también con verbos que no son transitivos, reemplazando a *ser*.

Notamos que el orden que presenta la misma autora, 1. No concordancia del participio pasado (pp), 2. Orden auxiliar + pp, 3. Contigüidad de auxiliar y pp, 4. «haber» + pp (intransitivo) es estructural, no necesariamente cronológico, puesto que empieza con 1. No concordancia de pp, correspondiente al siglo tan tardío como XV.

A continuación, por completar el estudio cuantitativo de la forma compuesta, cito el cuadro de Mendivil Giró (2015: 270), adoptado de Olbertz (1993:256) (Tabla 23 en Apéndice)⁴³:

⁴³ Para ver la tendencia total del movimiento histórico, he cambiado el orden de las propiedades lingüísticas y elaborar el gráfico correspondiente: 1. Orden canónico

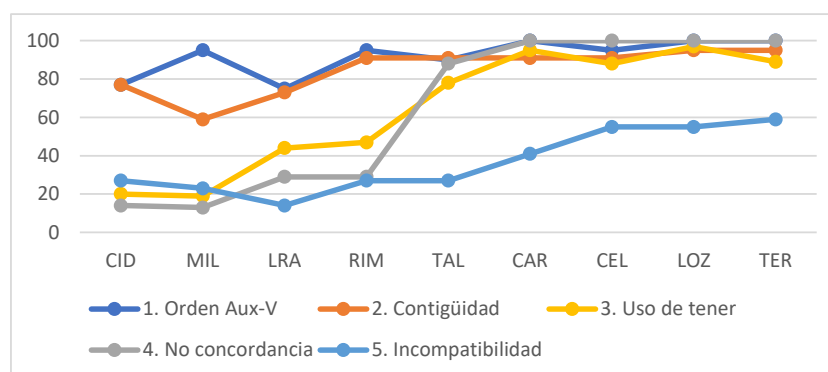


Fig. 23. Evolución de las propiedades de «haber» + pp (%). Olbertz (1993)

Por estas cifras y el gráfico, podemos observar el transcurso cronológico de 1. Orden de «haber» + pp, 2. Contigüidad de «haber» y pp → 3. Uso de *tener*, 4. No concordancia del participio → 5. Incompatibilidad semántica entre «haber» ‘tener’ + pp. Llama la atención la alta correlación que hay entre las últimas tres propiedades lingüísticas, 3, 4, 5. No obstante, no me atrevo a proponer que los primeros elementos en el orden cronológico sean la causa de la gramaticalización de «haber» + pp, puesto que también es posible que sean efectos de ella. Expondré mi conclusión sobre la causa más razonable en la sección 5.

Los datos reunidos de extensas fuentes de Romani (2006: 284, 287, 289, 325) confirman la tendencia numérica similar (Tabla 24 en Apéndice)⁴⁴:

entre auxiliar y participio fijado; 2. Contigüidad entre el auxiliar y pp ; 3. Grado en el que se expresa la posesión con el verbo «tener»; 4. Ausencia de concordancia; 5. Incompatibilidad semántica entre «haber» ‘tener’ y pp / CID: *Poema de mio Cid* (1140-1180); MIL: *Milagros de Nuestra Señora* (1252?) de Berceo; LBA: *Libro de buen amor* (1343) del Arcipreste de Hita; RIM: *Rimado de palacio* (1386) de López de Ayala; TAL: *Arcipreste de Talavera o Corbacho* (1438) de Martínez de Toledo; CAR: *Cárcel de amor* (1492) de Diego de San Pedro; CEL: *Celestina* (1599) de Rojas; LOZ: *La lozana andaluza* (1528) de Francisco Delicado; TER: *Epistolario* (1562) de Teresa de Jesús.

⁴⁴ Todas las cifras son frecuencias relativas (%) con respecto al término opuesto. “Haber/ser + pp” se refiere a la proporción de «haber» + pp con respecto a «ser» + pp (intransitivo). CORPUS: Siglo XII: *Documentos lingüísticos de España* [DLE], *La fazienda de ultramar*, *Cantar de mio Cid*, *Crestomatía del español medieval* [CEM] ; Siglo XIII: [DLE], *Primera crónica general de España*, [CEM], *Fuero de Alcalá de Henares*, *Colección diplomática de Sepúlveda* [CDS], *Libro de los animales que cazan*; Siglo XIV: [DLE], [CEM], [CDS], *Textos para la historia del español* [THE], *Libro de la montería*, *Crónica del rey don Pedro*; Siglo XV: [DLE], [CEM], [CDS], [THE], *Obras completas del marqués de Santillana*, *Generaciones y semblanzas*, *Claros varones de Castilla*, *Cárcel de amor*.

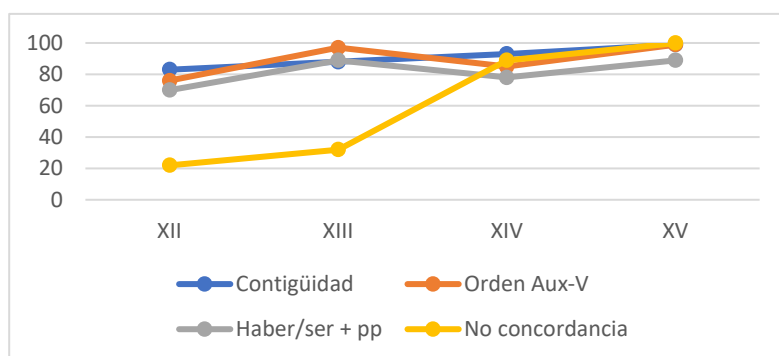


Fig. 24. Evolución de las propiedades de «haber» + pp (%). Romani (2006)

Como acabamos de ver, los estudios de Olbetz (1993) y Romani (2006) con los textos mayoritariamente literarios coinciden en la misma tendencia ascendente marcada de ‘no concordancia’ y las dos propiedades, ‘contigüidad’ y ‘orden de «haber» + pp’ son cronológicamente estables en alto grado de gramaticalización. En cambio, los materiales de CODEA, de documentos exclusivamente notariales, muestran una muy reducida presencia de estas propiedades de «haber» + pp (sec. 3.3.1, 3.3.2, 3.3.3), lo que es debido al estilo innovador no literario donde la gramaticalización estaba bastante avanzada y casi finalizada. La diferencia estilística entre los dos tipos de texto es grande. Por esta razón, estos indicios propuestos son válidos para estudiar los textos literarios, mientras que no lo son para los documentos notariales. En su lugar, la frecuencia de uso de «haber» + pp sirve como indicio fiable de la gramaticalización en general, tanto para el estilo tradicional como el innovador (sec. 3.2). La frecuencia normalizada de uso tiene mérito de ser comparable para todos los textos y para todos los propósitos de la investigación lingüística.

4. Geografía del español actual

4.1. España

Para observar la situación actual de «haber» + pp en España, contamos con los datos del proyecto de investigación COSER («*Corpus Oral y Sonoro del Español Rural*») dirigido por Inés Fernández Ordóñez⁴⁵. Gracias a la ayuda prestada por la directora del proyecto, hemos instalado los datos ofrecidos en nuestro sistema LYNEAL⁴⁶. Son transcripciones de conversación en dialectos locales en estilo tradicional pertenecientes a personas mayores de edad.

⁴⁵ <http://www.corpusrural.es/> [7 de enero de 2021]

⁴⁶ <https://lecture.ecc.u-tokyo.ac.jp/~cueda/lyneal/coser.htm> [7 de enero de 2021]

La siguiente lista muestra la frecuencia de la forma compuesta, normalizada por mil palabras (FN.1000), junto con la probabilidad binomial acumulativa superior (Binom.sup.)⁴⁷. Las cifras en negritas son de la probabilidad menor de 0.05, estadísticamente significativa:

N	Provincia	haber+PP	Palabras	FN.1000	Binom.sup.
30	Islas Baleares	88	11 289	7.795	0.000
26	Teruel	234	42 532	5.502	0.000
20	Tarragona	77	14 071	5.472	0.000
23	Ávila	174	32 036	5.431	0.000
14	Soria	239	44 193	5.408	0.000
9	La Rioja	335	65 945	5.080	0.000
7	Navarra	194	39 254	4.942	0.000
3	Vizcaya	237	47 994	4.938	0.000
31	Valencia	359	73 433	4.889	0.000
11	Burgos	343	72 396	4.738	0.000
22	Segovia	181	39 814	4.546	0.000
16	Zaragoza	155	34 869	4.445	0.002
27	Cuenca	216	49 341	4.378	0.001
6	Álava	181	42 272	4.282	0.003
15	Valladolid	172	40 282	4.270	0.005
17	Lérida	40	9 544	4.191	0.135
2	Cantabria	159	38 313	4.150	0.015
4	Guipúzcoa	167	40 288	4.145	0.014
12	Huesca	220	54 345	4.048	0.014
28	Castellón	174	43 519	3.998	0.036
24	Guadalajara	168	44 719	3.757	0.162
40	Sevilla	242	64 683	3.741	0.130
13	Palencia	191	52 645	3.628	0.281
21	Salamanca	210	58 127	3.613	0.291
25	Madrid	209	61 249	3.412	0.608
36	Alicante	92	27 291	3.371	0.625
32	Cáceres	217	65 049	3.336	0.731

⁴⁷ He utilizado función de Excel: =1-BINOMDIST(E2-1,F2,\$E\$49/\$F\$49,1) donde E2 es la frecuencia absoluta de forma, en el caso de Islas Baleares, 88; F2 es totalidad de palabras que hay en el mismo lugar, 11 289, \$E\$49 es la frecuencia total de la forma, 7 536 y \$F\$49 es la frecuencia total de todos los lugares, 2 170 574.

19	Barcelona	63	19 153	3.289	0.683
35	Badajoz	195	61 300	3.181	0.897
41	Huelva	160	50 853	3.146	0.902
34	Ciudad Real	141	45 225	3.118	0.908
42	Granada	156	50 287	3.102	0.928
33	Albacete	254	83 285	3.050	0.984
38	Córdoba	156	54 541	2.860	0.994
43	Almería	141	49 928	2.824	0.995
29	Toledo	158	57 456	2.750	0.999
44	Málaga	200	80 176	2.495	1.000
39	Jaén	89	38 592	2.306	1.000
37	Murcia	179	80 680	2.219	1.000
45	Cádiz	85	39 903	2.130	1.000
18	Zamora	117	57 286	2.042	1.000
46	Santa Cruz de Tenerife	51	29 476	1.730	1.000
8	León	65	59 053	1.101	1.000
5	Lugo	7	6 884	1.017	1.000
47	Las Palmas de Gran Canaria	33	42 757	0.772	1.000
10	Pontevedra	6	11 874	0.505	1.000
1	Asturias	6	423 72	0.142	1.000
*	Total	7 536	2 170 574	3.472	0.502

Tabla 25. Frecuencia normalizada de «haber» + pp en España. COSER.

El número de la primera columna corresponde al número de cada provincia que aparece en el siguiente mapa, donde podemos apreciar la frecuencia normalizada en forma de barra vertical. El sistema dibuja un área sombreada que abarca las provincias de frecuencia significativamente alta, con la frecuencia normalizada igual o superior a 3.998, dotada de la probabilidad binomial menor de 0.05.

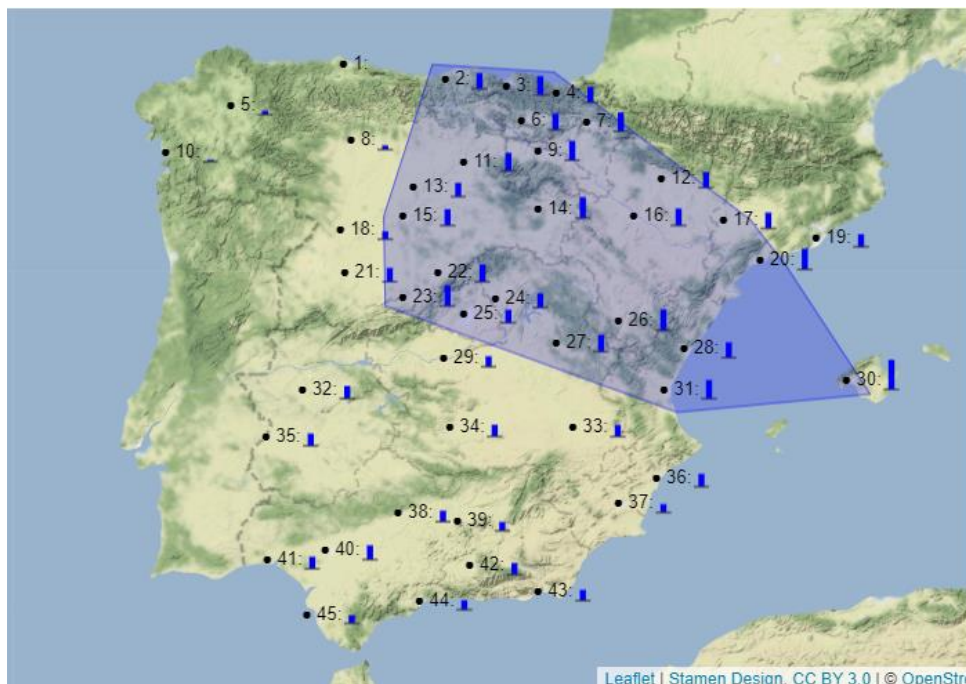


Fig. 25. «haber» + pp en la Península Ibérica. Frecuencia normalizada por mil palabras. COSER.

Ejemplos: *Ha dicho mi madre si puede amasar mañana, La Rioja / hemos compactao muy bien y le digo la verdad, hemos hecho un trozo de pan, eh, sí, sí, sí. Valencia*

De esta manera, se configura un área coherente de alta frecuencia normalizada en la Península. El área tan bien definida no puede ser producida de manera accidental, sino debe de reflejar el uso preferente de la forma compuesta en la zona continua. Se trata de la zona oriental, concretamente Cataluña y Aragón con extensión a Castilla la Vieja y norte de la Nueva. Como hemos visto anteriormente el carácter oriental de la forma compuesta dentro de la Romania e Iberorromania (sección 2.2) y la cronología de la misma forma centrada en Aragón, en contraste con León, región reacia a la misma forma. Es interesante comprobar que la situación actual es acorde con los datos presentados en las investigaciones diacrónicas anteriores y nuestras (sec. 3.2).

4.2. Hispanoamérica

Los estudios dialectales de Hispanoamérica indican que ahí la forma presente de «haber» + pp, *he comido*, se utiliza con frecuencia reducida y, en su lugar,

se utiliza la forma de pretérito simple, *comí*. De la diferencia entre el español de España y el de México, Lope Blanch (1991: 142-143) precisa lo siguiente:

(...) Contrariamente a lo que sucede en España, en México no se emplea la forma compuesta para expresar las acciones verificadas en el antepresente, en el 'presente ampliado', sino siempre la forma simple ("*Llegó hace un momento*"). Por otra parte, el significado aspectual del pretérito compuesto no es, en México, el perfectivo, sino el de acción reiterada o continuada que llega hasta el presente, hasta el ahora. Pues bien, estas peculiaridades del uso mexicano, tan distinto del castellano actual, están muy próximas a las formas castellanas medievales y prerrenacentistas, (...)

Suzuki (2004) cuestiona la hipótesis del arcaísmo supuesto para el tiempo pasado en América y tras investigar documentos coloniales escritos en ciudades americanas, Santo Domingo, México, Lima, Santiago de Chile, Tucumán y Buenos Aires, del siglo XVI al XVIII, llega a la siguiente conclusión (351: traducción mía):

(...) el uso frecuente del pretérito y el menor uso del presente perfecto en el español americano moderno no representa un estado como resultado de la conservación constante del uso antiguo observado en la época colonial. Más bien se supone que desde la época colonial, en la que hubo un estado transitorio de distinción, el español europeo y el español latinoamericano han establecido su propio criterio de diferenciación respectivamente.

En cuanto al uso de la anterioridad inmediata (*Llegó hace un momento*), según su cálculo del presente perfecto (*he comido*), México presenta la menor frecuencia relativa (%) dentro de las ciudades investigadas (op.cit.: 287)⁴⁸, lo que me hace pensar en su continuidad histórica hasta la actualidad. Naturalmente no se trata del mantenimiento del uso medieval, puesto que desde finales del siglo XV hasta el XVII siempre las dos formas, pretérito simple y perfecto compuesto coexistían para el uso de anterioridad inmediata (Suzuki op. cit.: 313). Razonablemente Suzuki explica lo siguiente (op. cit. traducción mía)⁴⁹:

(...) se puede pensar en la probabilidad de que la etapa transitoria observada en el español peninsular también esté reflejada en los documentos coloniales. Es

⁴⁸ Santo Domingo (frecuencia 59, 35.3%), México (21, 17.6%), Lima (48, 55.8%), Santiago de Chile (100, 28.2%), Tucumán (14, 33.3%), Buenos Aires (10, 19.6%).

⁴⁹ Efectivamente se observa el proceso ascendente histórico en el porcentaje de pretérito simple con respecto a perfecto compuesto que ofrece Moreno de Alba (2006: 2113): siglo XVI 61%, XVII 74%, XVIII 80%, XIX 85% (textos no literarios).

necesario pensar en la probabilidad de que la supremacía del pretérito simple, tal como se observa en el español mexicano actual, no se presentara claramente desde el inicio de la transposición, sino que, a partir de esta etapa transitoria, se formara en posteridad.

Conviene averiguar las distintas modalidades dentro del español europeo y el americano en su visión conjunta total, tomando en consideración sus condiciones históricas y geográficas. Para conocer la situación actual del mundo hispanohablante en su estilo coloquial, los datos que ofrece el proyecto PRESEEA («*Proyecto para el Estudio Sociolingüístico del Español de España y de América*») dirigido por Francisco Moreno Fernández, son útiles por su homogeneidad de constitución y su variabilidad sociolingüística (sexo, edad, nivel de educación)⁵⁰. La misma variedad sirve no solamente para ver la relación que hay entre los factores lingüísticos y los extralingüísticos, sino también para garantizar la representatividad del habla local conjunta. En esta ocasión no hacemos la distinción por factores sociolingüísticos y utilizamos los textos transcritos reunidos de 18 informantes de cada localidad, constituidos de dos variables de sexo, tres grupos de edad y tres niveles de educación⁵¹:

N	Lugar	«haber»+pp	Palabras	FN.1000	Binom.inf.
17	ES-Alcalá	1 407	127 780	11.011	1.000
12	PE-Lima	1 246	127 913	9.741	1.000
10	CO-Cali	976	101 771	9.590	1.000
15	ES-Santander	1 575	164 390	9.581	1.000
20	ES-Sevilla	1 318	140 379	9.389	1.000
18	ES-Madrid	1 425	162 503	8.769	1.000
22	ES-Málaga	1 435	166 669	8.610	1.000
21	ES-Granada	996	120 135	8.291	1.000
9	CO-Pereira	801	105 679	7.580	1.000
13	CH-Santiago de Chile	1 269	172 909	7.339	1.000

⁵⁰ <https://preseea.linguas.net/>

⁵¹ La constitución de la tabla es la misma de la del corpus COSER (Tabla 26), excepto la columna de Binom.inf. que se trata de la probabilidad binomial acumulativa inferior. Para su cálculo, he utilizado la función de Excel, BINOMDIST de la siguiente manera:

$$=BINOMDIST(E2,F2,E24/F24,1)$$

donde E2 es la frecuencia absoluta de forma, en el caso de ES-Alcalá, 1 407; F2 es totalidad de palabras que hay en el mismo lugar, 127 780, \$E\$24 es la frecuencia total de la forma, 20 366; \$F\$24 es la frecuencia total de todos los lugares, 3 244 086.

8	CO-Medellín	893	124 253	7.187	1.000
1	CU-La Habana	910	132 887	6.848	0.993
19	ES-Valencia	725	108 705	6.669	0.930
6	GU-Guatemala	739	122 466	6.034	0.109
4	MX-Guadalajara	755	144 942	5.209	0.000
7	CO-Barranquilla	815	159 520	5.109	0.000
2	MX-Mexicali	442	110 031	4.017	0.000
11	VE-Caracas	658	164 499	4.000	0.000
5	MX-México, D. F.	755	210 060	3.594	0.000
14	UR-Montevideo	393	127 522	3.082	0.000
3	MX-Monterrey	607	213 280	2.846	0.000
16	ES-Santiago de Compostela	226	215 793	1.047	0.000
*	Total	20 366	3 224 086	6.317	0.502

Tabla 26. Frecuencia normalizada de «haber» + pp en Hispanoamérica. PRESEEA.

Aquí nos interesa la baja frecuencia de «haber» + pp que se presenta en algunas ciudades hispanohablantes. Vamos a destacar las frecuencias bajas de probabilidad significativamente reducida, menos de 0.05, indicadas en negrita en la tabla, que corresponden a los círculos rojos en el mapa:



Fig. 26. Frecuencia normalizada de «haber» + pp en Hispanoamérica. PRESEEA.

Ejemplos: *lo malo es que hoy en día // se ha acabado bastante eso / no era como antes // VE-Caracas : me han robado cosas cuando te dije me han robado / claro / caminando / con mochila CH-Santiago de Chile.*

Efectivamente las cuatro ciudades mexicanas presentan la frecuencia significativamente baja, lo que es explicable por razones históricas que acabamos de ver. En Barranquilla de Colombia, Caracas de Venezuela y Montevideo de Uruguay también el uso de «haber» + pp es limitado. De momento no sabemos sus razones ni causas. En cambio, la poca frecuencia de «haber» + pp en Santiago de Compostela puede ser producida por la influencia de la lengua gallega, tratada en las secciones 2.2 y 3.2.

En la Tabla 26 y el mapa (Fig. 26) podemos observar la distribución numérica de «haber» + pp en la amplia geografía hispanohablante, donde una vez más confirmamos el contraste cuantitativo entre el oriente con más casos de «haber» + pp y el occidente con menos casos, tanto en la Península (ciudades de centro y oriente versus Santiago de Compostela) como en el mundo hispanohablante total (ciudades de España versus las de Hispanoamérica mencionadas en el párrafo anterior). Si la situación de México se interpreta como una representación occidental que se extiende desde Galicia, León y Andalucía en España (Fig. 25) hasta el amplio territorio mexicano y norte de Sudamérica y la ciudad de la Plata, Montevideo en Hispanoamérica (Fig. 26), la continuidad de la distribución de la forma compuesta es explicable desde el punto de vista histórico y geográfico al mismo tiempo (Rosario 1970: 50, Sánchez Méndez 2003: 339-340, Terasaki 2011: 254).

Históricamente, según Lipski (1994, 1996: 60-61), los hablantes de las lenguas de oriente (catalán, valenciano y aragonés) eran escasos en los primeros emigrantes:

Entre los primeros emigrantes a las Américas, los hablantes de catalán, valenciano y aragonés fueron escasos en proporción, pues preferían emigrar a zonas mediterráneas como Cerdeña y Sicilia. Durante cierto tiempo, la política oficial española prohibió la emigración catalana al Nuevo Mundo. Esto nos deja aún un gran número de lenguas y dialectos regionales del centro y del oeste que hay que tener en cuenta, dada la mezcolanza de emigrantes que se dirigieron a la América española.

Y según Frago García (1999: 27) “los navarros y aragoneses en la hora de

Descubrimiento ya estaban mayoritariamente castellanizados”. De esta manera se confirma el contraste cuantitativo fundamental entre el oriente minoritario y el occidente mayoritario en los primeros emigrantes españoles a América.

La situación lingüística de Bolivia y la zona andina colindante constituye una excepción importante, donde se utiliza la forma «haber» + pp con alta frecuencia (Kany 1970: 200, Moreno de Alba 2006: 2117-2118). Pérez Sáez (2003: 81) sugiere que “se trata de usos conectados con el quichua”.

5. Conclusión

En esta conferencia, he hablado de la historia y geografía de la forma compuesta española «haber» + participio pasado, empezando con la situación latina, penumbra anterior al nacimiento de la misma forma con las primeras gestaciones de PP + HABERE, junto con algunos ejemplos posibles del uso perfectivo desarrollado del participio transitivo del verbo HABERE ‘poseer’. En el dominio de las lenguas romances, hemos visto que el español ocupa un sitio intermedio entre el uso preferente en las lenguas de oriente (catalán y francés) y el escaso uso en las lenguas de occidente (gallego y portugués). Dentro de la Península hemos observado la misma gradación de este (aragonés y aranés) a oeste (asturiano y mirandés). Al investigar los documentos notariales emitidos en Aragón, Castilla y León, también hemos comprobado la inclinación descendente de la Zona Este a la Zona Oeste en la frecuencia de uso de la forma compuesta. Tanto en la Romania como en Iberorromania, el castellano se sitúa en la zona de transición entre las dos polaridades de variación. Y dentro del amplio dominio hispanohablante de ambos lados del océano Atlántico, existe la misma tendencia de este a oeste, probablemente debida a la escasa emigración desde el oriente de la Península.

Se trata de una nueva forma que apareció en la larga historia de la lengua latina. No sabemos exactamente ni cuándo ni dónde, pero parece datar en la época antes de la diversificación en las lenguas romances en una zona céntrica de la Romania, puesto que la misma forma existe en casi todo el territorio romano. Para averiguarlo, desgraciadamente no contamos con los documentos suficientes. Simplemente disponemos de algunos textos latinos donde la forma compuesta aparece con muy baja frecuencia, lo que no significa necesariamente que no se utilizara mucho en el habla del pueblo. La poca frecuencia de uso en textos latinos antiguos puede deberse al estilo literario que caracteriza las obras literarias escritas y copiadas a lo largo de historia, que no quieren utilizar el nuevo estilo coloquial en la superficie de pergamino y papel.

En el siglo XIII en adelante, observamos su uso cada vez más habitual en distintos tipos documentales. El cambio grande entre el texto latino y el castellano parece indicar la conversión del estilo coloquial al normal de la misma forma compuesta. Una vez convertida en el estilo normal, considerablemente antes del siglo XIII, su uso se hace gradualmente más frecuente para llegar a establecer un tiempo gramatical importante de alta utilidad. La velocidad de este cambio gramatical varía en el ámbito románico e iberorrománico de Europa de sur y en la tierra circumatlántico hispanohablante, siempre respetando la continuidad de tiempo y espacio, con más casos en el oriente, menos casos en el occidente con crecimiento cronológico general. En este estudio creo haber demostrado la validez de la hipótesis consistente en la continuidad histórica y geográfica de la forma compuesta «haber» + pp.

¿A qué se debe la formación gramatical constante, cada vez más acelerada de «haber» + pp, que hemos observado en la historia de español? Estamos ante la gramaticalización siempre acompañada del crecimiento de frecuencia. Ahora nos preguntamos si la frecuencia es la causa o el efecto de la gramaticalización, pregunta que se hace Bybee (2007: 17-18) en un capítulo de su libro que trata la cuestión de frecuencia en la organización del lenguaje. Su respuesta es la siguiente (traducción mía):

(...), la respuesta a la pregunta de si la frecuencia es una causa o un efecto es compleja. Por un lado, la frecuencia es solo un recuento, un patrón observable en los textos, que por supuesto es un efecto. Por otro lado, la frecuencia o repetición de experiencias tiene un impacto en las representaciones cognitivas y de esta manera se convierte en una causa de los efectos discutidos en este libro.

Pues bien, la cuestión no se trata de si la frecuencia es una causa o un efecto de procesos gramaticales, sino más bien de si la frecuencia es una causa y efecto de la gramaticalización al mismo tiempo. Efectivamente, la misma autora (Bybee, 2003: 602) menciona lo siguiente (traducción mía):

La frecuencia no es solo el resultado de la gramaticalización, también es un contribuyente principal al proceso, una fuerza activa que instiga el cambio que ocurre en la gramaticalización.

Sin embargo, no estoy tan seguro para afirmar que la frecuencia puede ser producto de la gramaticalización, puesto que se encuentran palabras gramaticales

poco frecuentes, por ejemplo, la preposición *salvo* (*Llegaron todos salvo Lucas*) y la conjunción *apenas* (*Apenas salí, se puso a nevar*), al lado de las sumamente frecuentes (*a, de, en, con, por; que, cuando, como*). Por otra parte, en cambio, si la frecuencia puede ser causa de la gramaticalización, tampoco es explicable la existencia de sustantivos y verbos no gramaticalizados, pero dotados de sumamente alta frecuencia, por ejemplo, *casa, decir, grande, etc.*

La correlación que hay entre la gramaticalización y la frecuencia puede ser espuria, producida por un factor común, puesto que no sabemos a ciencia cierta cuál de las dos toma la iniciativa en el proceso del cambio histórico. La respuesta a esta cuestión la encuentro en el uso mismo de la lengua de los hablantes. Si el hablante de la lengua encuentra la 'utilidad' de la forma recién nacida, la transforma en un elemento gramatical, con una forma gramatical adecuada para su comodidad de uso, sin concordancia, en orden fijo de «haber» y participio, con uniformidad del verbo auxiliar en «haber»; y, al mismo tiempo, la utiliza con alta frecuencia, precisamente porque es útil y ahora todavía más con la forma cómoda.

De modo que los cambios gramaticales coincidentes, disminución de concordancia, fijación del orden de «haber» y participio y unificación del verbo auxiliar en «haber», son factores coadyuvantes pero no principales. El factor principal de causa y efecto de la gramaticalización, concretamente de la auxiliarización de «haber», es el uso frecuente de la forma compuesta verbal, debido a su alta utilidad y comodidad. Por esta razón y por los datos numéricos que hemos visto, considero que el factor de la frecuencia es el más importante.

Por la 'ley de los grandes números', sabemos que la media de una muestra grande, homogénea y aleatoria (en nuestro caso, materiales de investigación) tiende a estar cerca de la media de la población completa, que corresponde en nuestro estudio a la evidencia numérica de fenómenos lingüísticos. Por ejemplo, al aumentar el número de tiradas al azar de una moneda, la media de las veces de que salga 'cara' se aproxima cada vez más a la ratio de mitad (0.5), cosa que no ocurre necesariamente si tiramos la moneda pocas veces. Por esta razón en nuestro estudio de la lengua conviene utilizar los materiales constituidos de numerosas palabras, seleccionados al azar con estratificación adecuada y no sesgada. La condición aleatoria y estratificada es difícil de satisfacer en las obras literarias, puesto que ahí se domina el estilo particular bastante marcado, lo que se salva en relativamente mayor medida en los documentos lingüísticos. En nuestro estudio de historia y geografía de la lengua, se desea utilizar los numerosos materiales menos sesgados posibles.

Para abordar el tema diacrónico y diatópico de la lengua, los corpus de documentos históricos y geográficos son sumamente útiles por ofrecer los datos diacrónicos y diatópicos sin solución de continuidad. La situación actual de la lengua se investiga con los inmensos datos de habla coloquial, transcritos y estructurados adecuadamente. A diferencia de las obras latinas y castellanas literarias, la cantidad de palabras que hay en el corpus lingüístico es enorme y homogéneo, lo que garantiza la representatividad relativa local y cronológica.

Por último, la tecnología informática facilita un cálculo inmenso como el recuento de los casos en cuestión y el de totalidad de palabras, ambos imprescindibles para derivar la frecuencia normalizada, que es importante para conocer su magnitud comparable de ocurrencia. Se trata de un cálculo sencillo, de recuento de casos, pero muy difícil de realizar manualmente por la cantidad enorme de unidades lingüísticas en un corpus grande. Ahora es factible por el último desarrollo informático.

Gracias a la ayuda prestada por los grandes proyectos españoles de investigación, que ofrecen buenos e inmensos datos en forma de corpus lingüísticos, estamos en condiciones de realizar estudios históricos y geográficos de español con evidencias cuantitativas fiables, lo que ha sido sumamente difícil, casi imposible, hasta hace muy poco. Otra ventaja en la actualidad es la relación estrecha que existe entre los investigadores sin barrera de frontera, que se ha realizado en los últimos años por el maravilloso desarrollo tecnológico de información. En esta situación favorable, esperamos seguir en comunicación con colegas nacionales e internacionales con confianza y amistad de siempre.

Es todo. Muchas gracias.

Referencias

- Adams, J. N. 2013. *Social Variation and the Latin Language*. Cambridge University Press.
- Alarcos Llorach, Emilio. 1947/1970. “Perfecto simple y compuesto en español”, en *Estudios de gramática funcional del español*. Madrid: Gredos, 13-49.
- Andrés Díaz, Ramón. 2013. *Gramática comparada de las lenguas ibéricas*. Gijón: Ediciones Trea.
- Arce, Agustín. 1980. *Itinerario de la virgen Egeria*. Madrid. Biblioteca de Autores Cristianos.
- Azofra Sierra, María Elena. 2009. *Morfosintaxis histórica del español: de la teoría a la práctica*. Madrid: Universidad Nacional de Educación a Distancia.

- Barrio de la Rosa, Florencio del. 2016. “De *haber* a *tener*. La difusión de *tener* como verbo de posesión en la historia del español: Contextos y focos”, en Benito Moreno y Toledo y Huerta (eds.), 239-280.
- Benito Moreno, Carlota de y Octavio de Toledo y Huerta (eds.). 2016. *En torno a 'haber'. Construcciones, usos y variación desde el latín hasta la actualidad*. Berlin. Peter Lang.
- Bybee, Joan L. 2003. “Mechanisms of Change in Grammaticization: the Role of Frequency”, en Brian D. Joseph y Richard D. Janda (eds.): *The Handbook of Historical Linguistics*, Oxford, Blackwell, págs. 602-623.
- Bybee, Joan L. 2007. *Frequency of Use and the Organization of Language*. Oxford University Press.
- Bybee, Joan L. and Osten Dahl. 1989. “The Creation of Tense and Aspect Systems in the Languages of the World”, *Studies in Language* 13-1. 51-103.
- Bybee, Joan, Revere Perkins and William Pegliuca. 1994. *The Evolution of Grammar. Tense, Aspect, and Modality in the Languages of the World*. Chicago: The University of Chicago Press.
- Campos Rodríguez, Francisco. 1975. M. T. *Cicerón. Catilinarias. Texto latino con traducción literal y literaria*. Tercera ed. Madrid: Gredos.
- Carter, Diana M. 2003. “A Tense Distortion: Variation in the Use of the Present Perfect in Spain”, *Toronto Working Papers in Linguistics*, 21: 15-26
- Company, Concepción (dir.) 2006. *Sintaxis histórica de la lengua española. Primera parte: La frase verbal*. México: Universidad Autónoma de México y Fondo de Cultura Económica.
- Company, Concepción. 1983. “Sintaxis y valores de los tiempos compuestos en el español medieval”, *Nueva Revista de Filología Hispánica*, 32, 235-257.
- Elcock, W. D. 1960. *The Romance Languages*. London. Faber & Faber.
- Ernout, Alfred et François Thomas. 1972. *Syntaxe latine*, 2^e edition. Paris: Librairie Klincksieck.
- Ernout, Alfred. 1953: *Morphologie historique du latin*. Paris: Librairie Klincksieck.
- Eto, Ichiro. 1998. “Estudio de usos de HABERE en el Itinerario de la Virgen Egeria”, (en japonés) *Kanda gaigo daigaku kiyoo*, 10, 121-150.
- Fontán, Antonio y Ana Moure Casas. 1987. *Antología del latín medieval. Introducción y textos*. Madrid: Gredos.
- Fukushima, Noritaka. 1978. “Historia breve del aspecto en español”, (en japonés) *Gaikokugo, gaikoku bungaku kenkyu*, Universidad de Estudios Extranjeros de Osaka.

- García de Diego, Vicente. 1970. *Gramática histórica española. Tercera edición corregida*, Madrid, Gredos.
- García Yebra, Valentín e Hipólito Escolar Sobrino. 2002. *César. Guerra de las Galias. Libro I-II-III*. Segunda ed. Madrid: Gredos.
- González-Haba, Mercedes. 1992. *Plauto. Comedias, I, Anfitrión – La Comedia De Los Asnos – La Comedia De La Olla - Las Dos Báquides – Los Cautivos – Cásina*. Madrid: Gredos.
- Grandgent, Charles H. 1970. *Introducción al latín vulgar*, traducción y anotación por Francisco de B. Moll, 4ª edición. Madrid: CSIC
- Gryson, Weber. 1969. *Biblia sacra vulgata*. Editio quinta. Stuttgart. Deutsche Bibelgesellschaft.
- Hanssen, Federico. 1913/1966. *Gramática histórica de la lengua castellana*. París: Ediciones Hispano-Americanas.
- Harrington, K. P. 1962. *Medieval Latin*, second edition. Chicago: The University of Chicago Press.
- Herman, József. 1997. *El latín vulgar*, edición española reelaborada y ampliada con la colaboración de Carmen Arias Abellán. Barcelona: Ariel.
- Induráin Pons, Jordi (dir.). 2018. *Atlas histórico de España*. Barcelona: Larousse.
- Ito, Taigo. 1978. “Unos problemas de la sintaxis histórica del español medieval”, *Lingüística Hispánica*. 1, 95-116.
- Jacob, Daniel. 2001. “¿Representatividad lingüística o autonomía pragmática del texto antiguo? El ejemplo del pasado compuesto”, en Daniel Jacob y Johanes Kabatek (eds.): *Lengua medieval y tradiciones discursivas en la Península Ibérica*, Madrid / Frankfurt am Main, Iberoamericana / Vervuert Verlag, 153-176.
- Kany, Charles E. 1970. *Sintaxis hispanoamericana*. Madrid: Gredos.
- Kataoka, Kozaburo. 1982a. *Gramática comparativa de las lenguas romances*, en japonés. Tokio: Asahishuppansha.
- Kataoka, Kozaburo. 1982b. *Gramática histórica de las lenguas romances*, en japonés. Tokio: Asahishuppansha.
- Keniston, Hayward. 1937. *The Syntax of Castilian Prose. The Sixteenth Century*. Chicago: The University of Chicago Press.
- Kume, Yuusuke. 2016. “Una reflexión sobre el desarrollo diatópico de la forma japonesa V+te+V”, (en japonés) en Ogawa Yoshiki, Nagano Akiko y Kikuchi Akira (eds.) *Cambio y variación lingüística y teoría lingüística desde el corpus*. (en japonés) Tokio: Kaitakusha, 249-265.

- Kunihara, Kichinosuke. 2007. *Rudimenta mediae latinitatis*. (en japonés) Tokio: Daigakushorin.
- Lipski, John M. 1994, 1996. *El español de América*. Madrid: Cátedra. Traduc. de *Latin American Spanish*, 1994. London, Longman.
- Lloyd, Paul M. 1993. *Del latín al español. I. Fonología y morfología históricas de la lengua española*, Madrid, Gredos.
- Lope Blanch, Juan Miguel. 1961, 1991. “Sobre el uso del pretérito en el español de México”, en *Studia Philologica, Homenaje a Dámaso Alonso*, Madrid, Gredos, vol. II, págs. 376-386, recogido en *Estudios sobre el español de México*, 1991, 131-143.
- López-Davalillo Larrea, Julio. 2000. *Atlas histórico de España y Portugal. Desde el paleolítico hasta el siglo XX*. Madrid: Síntesis.
- Mendivil Giró, José Luis. 2015. *El cambio lingüístico. Sus causas, mecanismos y consecuencias*. Madrid: Editorial Síntesis.
- Menéndez Pidal, Ramón. 1976. *Cantar de Mio Cid, Texto, gramática y vocabulario. Primera parte. Crítica del texto - gramática*. 5a ed. Madrid: Espasa Calpe.
- Moreno de Alba, José G. 2006. “Evolución diacrónica y diatópica de los valores del pretérito perfecto”, en José Jesús de Bustos Tovar y José Luis Girón Alconchel (eds.): *Actas del VI Congreso Internacional de Historia de la Lengua Española* (Madrid, UCM, 29 de septiembre-3 de octubre de 2003). Madrid: Arco Libros, vol. III, 2105-2121.
- Nakaoka, Shoji. 1993. *Introducción al español medieval*. (en japonés) Tokio: Daigakushorin.
- Niemeyer, J. F. 1976. *Mediae latinitatis lexicon minus*. Leiden. E. J. Brill.
- Nomura, Takeshi. 2011. *Historia de la lengua japonesa hablada*. (en japonés). Tokio. Yoshikawa Koobunkan.
- Oka, Tomoyuki. 2013. “Gramaticalización y tipología de tiempo y aspecto. Categorización lingüística de existencia y tiempo”, (en japonés) en Kanasugi Takao, Oka Tomoyuki, Yonekura Yoko (eds.) *Lingüística Histórica Cognitiva* (en japonés). Tokio: Kuroshio Shuppan. 39-75.
- Olbertz, Hella. 1993. “The Grammaticalization of Spanish *haber* plus participle”, en Jaap van Marle (ed.): *Historical Linguistics 1991. Papers from the 10th International Conference on Historical Linguistics (Amsterdam, 12-16 august 1991)*, Amsterdam / Philadelphia, John Benjamins, págs. 243-263.
- Oroz Reta, José y Manuel A. Marcos Casquero. 1983. *San Isidoro de Sevilla. Etimologías*. Edición bilingüe. II. Libros XI-XX. Madrid: Biblioteca de

Autores Cristianos.

- Penny, Ralph. 2006. *Gramática histórica del español*. Barcelona: Ariel.
- Pérez González, Maurisio. 1985. *El latín de la cancillería castellana (1158-1214)*. Ediciones Universidad Salamanca.
- Pérez Sáez, Vicente J. 2003. *Sobre la lengua en el noroeste argentino. Capítulos para su descripción e historia*. Universidad Nacional de Salta.
- Posner, Rebecca. 1996. *Las lenguas romances*. Madrid: Cátedra.
- Pountain, Christopher J. 1985. "Copulas, Verbs of Possession and Auxiliaries in Old Spanish: the Evidence for Structurally Interdependent Changes", *Bulletin of Hispanic Studies* 62, 337–355.
- Rodríguez Molina, Javier. 2010. *La gramaticalización de los tiempos compuestos en español antiguo: cinco cambios diacrónicos*. Tesis doctoral. Universidad Autónoma de Madrid.
- Rodríguez Molina, Javier. 2016. "Patrones de variación en la concordancia del participio en español antiguo", en Benito Moreno y Toledo y Huerta (eds.), 417-467.
- Romani, Patrizia. 2006. "Tiempos de formación romance I. Los tiempos compuestos", en Concepción Company (dir.): *Sintaxis histórica de la lengua española. Primera parte: La frase verbal*. México: Universidad Autónoma de México y Fondo de Cultura Económica, 243-346.
- Rosario, Rubén del. 1970. *El español de América*. Sharon, Conn.: Troutman Press.
- Sánchez Méndez, Juan. 2003. *Historia de la lengua española en América*. Valencia: Tirant lo Blanch.
- Segura Munguía, Santiago. 2003. *Nuevo diccionario etimológico latín-español y de las voces derivadas*. Bilbao. Universidad de Deusto.
- Shimaoka, Shigeru. 1970. *Sobre lenguas romances*, en japonés. Tokio: Daigakushorin.
- Sociedades Bíblicas en América Latina. 1960. *La santa biblia. Antiguo y nuevo testamento*. Antigua versión de Casiodoro de Reina (1569) revisada por Cipriano de Valera (1602). Otras revisiones: 1862, 1909 y 1960. Asunción.
- Suzuki, Emiko. 1999. "Evolución de la distinción funcional entre el pretérito perfecto compuesto y el pretérito simple", *Estudios Lingüísticos Hispánicos*, núm. 14, 77-99.
- Suzuki, Emiko. 2004. *Estudio diacrónico del pretérito y el presente perfecto en español*. (en japonés) Tesis doctoral. Universidad de Estudios Extranjeros de Tokio.

- Terasaki, Hideki. 1987. “Perfecto compuesto español, en comparación con otras lenguas románicas”, *Estudios Lingüísticos Hispánicos*, núm. 2, 65-80.
- Terasaki, Hideki. 2011. *Historia de la lengua española*. (en japonés) Tokio: Daigakushorin.
- Tsujii, Muneaki. 1989. “La subjuntivación de la forma en *-ra* y el movimiento de *haber, tener* + p.p. en el español medieval. La correlación entre ambos en la transición sistemática”, (en japonés) *The Review of Inquiry and Research*. The Kansai University of Foreign Studies, Osaka, 49, 143-165.
- Tsujii, Muneaki. 1996. *Un estudio sobre las formas en -ra y -se, y ‘había + participio’ en el español medieval: dentro del marco de la transición desde la forma sintética hacia la analítica*. Tesis doctoral. Kansai Gaidai University.
- Unamuno, Miguel de. 1977. *Gramática y glosario del Poema del Cid. Contribuciones al estudio de los orígenes de la lengua española*. Madrid: Espasa-Calpe.
- Villimer, Santiago. 1976. *Estudios de latín medieval. Documentos de la cancillería castellana. SS. XIV y XV*. Vitoria: Colegio Universitario de Álava.
- Vives, J. Vicens. 1978. *Atlas de historia de España*. Décima edición. Barcelona: Teide.
- Yamamura, Hiromi. 1982. “El uso de la forma ‘pretérito perfecto compuesto’ en el Cantar de mio Cid”, (en japonés) en *Miyagi Noboru kyozyu kanreki kinen ronbunshu*, 393-409.
- Yllera, Alicia. 1980. *Sintaxis histórica del verbo español: Las perífrasis medievales*, Zaragoza, Universidad de Zaragoza.

Apéndice. Tablas

func. / obra	CEL	DIA	PAS	ENT	CEL	DIA	PAS	ENT
Result.	98	43	18	25	30.5%	15.6%	23.4%	18.4%
Cont.	54	56	12	14	16.8%	20.3%	15.6%	10.3%
Inmed.	107	80	26	63	33.3%	29.0%	33.8%	46.3%
Indef.	62	97	21	34	19.3%	35.1%	27.3%	25.0%
Total	321	276	77	136	100.0%	100.0%	100.0%	100.0%

Tabla 4. Perfecto compuesto (*he cantado*). Suzuki (1999: 81)

func. / obra	CEL	DIA	PAS	ENT	CEL	DIA	PAS	ENT
Aoristo	619	202	143	122	68.1%	79.2%	70.8%	76.3%
Cont.	19	4	1	1	2.1%	1.6%	0.5%	0.6%
Inmed.	175	32	40	29	19.3%	12.5%	19.8%	18.1%

Indef.	96	17	18	8	10.6%	6.7%	8.9%	5.0%
Total	909	255	202	160	100.0%	100.0%	100.0%	100.0%

Tabla 5. Pretérito simple (*canté*), Suzuki (1999: 81).

Provincia	«haber» + pp	Palabras	FN	División
Burgos	19	44 007	0.432	Oeste
Ávila	19	43 709	0.435	Oeste
Palencia	8	13 252	0.604	Oeste
Salamanca	46	56 719	0.811	Oeste
Segovia	19	22 984	0.827	Oeste
La Rioja	13	15 333	0.848	Oeste
Zamora	20	22 574	0.886	Oeste
Guadalajara	49	55 146	0.889	Oeste
Soria	5	5 619	0.890	Oeste
Cantabria	21	20 931	1.003	Oeste
León	44	31 937	1.378	Oeste
Toledo	100	69 817	1.432	Oeste
Zaragoza	128	66 666	1.920	Este
Valladolid	197	102 460	1.923	Este
Huesca	34	16 717	2.034	Este
Teruel	82	38 620	2.123	Este
Madrid	334	126 547	2.639	Este
Cuenca	37	8 284	4.466	Este

Tabla 6. «haber»+pp en provincias y división Oeste-Este. CODEA.

Año	1200	1250	1300	1350	1400	1450	1500	1550	1600
FA. Oeste	3	16	16	10	36	74	65	60	83
FA. Este	1	3	20	60	79	65	94	279	211
FN. Oeste	0.264	0.316	0.302	0.245	0.568	1.049	0.976	2.379	4.068
FN. Este	0.299	0.115	1.033	1.337	1.274	1.724	2.081	3.258	5.990

Tabla 8. «haber» + pp. Zona Oeste y Zona Este. CODEA.

Etapa	León	Castilla norte	Castilla sur	Navarra	Aragón
1140-1295	5	10	16	33	34
1296-1369	12	27	24	95	87
1370-1454	33	57	89	149	154
1455-1499	56	178	169	271	270

Tabla 9. «haber» + pp documentos notariales. Rodríguez Molina (2010).

Año	1200	1250	1300	1350	1400	1450	1500	1550	1600
FA.«haber»+pp	4	19	36	70	115	139	159	339	294
FA.«haber»-pp	63	308	276	449	598	354	401	252	214
FR.«haber»+pp	6%	6%	12%	13%	16%	28%	28%	57%	58%
FR.«haber»-pp	94%	94%	88%	87%	84%	72%	72%	43%	42%
FN.«haber»+pp	0.272	0.248	0.498	0.817	0.917	1.284	1.422	3.058	5.285
FN.«haber»-pp	4.281	4.020	3.815	5.238	4.769	3.269	3.587	2.273	3.847

Tabla 10. «haber» + pp / «haber» sin pp. CODEA.

Verbo	CID	MLG	LUC	CEL	LAZ	CID	MLG	LUC	CEL	LAZ
HABER	205	148	511	127	24	3.29	2.51	3.23	0.93	0.62
TENER	87	100	309	392	117	1.40	1.70	1.95	2.88	3.02

Tabla 11. Verbo principal 'posesivo transitivo'⁵² Ueno (1988).

<i>tener y haber</i> (%)	Cid	Mil	Luc	LBA	Rim	Cel	Tal	Carc	Loz	Ter
<i>tener</i> (vs. <i>haber</i>) 'poseer'	17	28	12	62	48	93	77	85	94	86
<i>haber</i> +pp (vs. <i>ser</i> +pp)	3	22	18	30	33	61	69	36	53	96

Tabla 12. Relación entre *tener y haber*. Pountain (1985).

Año	1200	1250	1300	1350	1400	1450	1500	1550	1600
FA.«haber»-pp	63	308	276	449	598	354	401	252	214
FA.«tener»-pp	16	74	100	135	266	253	266	265	133
FN.«haber»-pp	4.281	4.020	3.815	5.238	4.769	3.269	3.587	2.273	3.847
FN.«tener»-pp	1.087	0.966	1.382	1.575	2.122	2.336	2.380	2.391	2.391

Tabla 13. «haber» sin participio y «tener». CODEA.

Año	1200	1250	1300	1350	1400	1450	1500	1550	1600
FA.«haber»+pp	4	19	36	70	115	139	159	339	294
FA.«tener»-pp	16	74	100	135	266	253	266	265	133
FN.«haber»+pp	0.272	0.248	0.498	0.817	0.917	1.284	1.422	3.058	5.285
FN.«tener»-pp	1.087	0.966	1.382	1.575	2.122	2.336	2.380	2.391	2.391

Tabla 14. «haber»+pp y «tener»-pp. CODEA.

⁵² La parte derecha de la tabla se trata de frecuencia normalizada por 100 verbos.

«tener» - pp	1200	1250	1300	1350	1400	1450	1500	1550	1600
FA. Oeste	15	55	75	58	152	160	188	88	64
FA. Este	1	19	25	77	114	93	78	177	69
FN. Oeste	1.319	1.086	1.415	1.421	2.398	2.267	2.823	3.489	3.137
FN. Este	0.299	0.731	1.291	1.716	1.839	2.466	1.727	2.067	1.959

Tabla 15. «tener». Frecuencia absoluta / Frecuencia normalizada. CODEA.

Año	1200	1250	1300	1350	1400	1450	1500	1550	1600
FA. mas.sg.	2	18	33	63	102	128	159	337	293
FA. mas.pl.	1	0	1	3	8	10	0	0	1
FA. fem.sg.	0	1	0	1	3	1	0	0	0
FA. fem.pl.	1	0	2	3	2	0	0	1	0
FR. mas.sg.	50%	95%	92%	90%	89%	92%	100%	100%	100%
FR. mas.pl.	25%	0%	3%	4%	7%	7%	0%	0%	0%
FR. fem.sg.	0%	5%	0%	1%	3%	1%	0%	0%	0%
FR. fem.pl.	25%	0%	6%	4%	2%	0%	0%	0%	0%
FN. mas.sg.	0.136	0.235	0.456	0.735	0.814	1.182	1.422	3.040	5.267
FN. mas.pl.	0.068	0.000	0.014	0.035	0.064	0.092	0.000	0.000	0.018
FN. fem.sg.	0.000	0.013	0.000	0.012	0.024	0.009	0.000	0.000	0.000
FN. fem.pl.	0.068	0.000	0.028	0.035	0.016	0.000	0.000	0.009	0.000

Tabla 16. Género y número del participio. CODEA.

Año	1200	1250	1300	1350	1400	1450	1500	1550	1600
FA. «haber»+pp	4	18	35	62	112	132	157	338	293
FA. pp+«haber»	0	1	1	8	3	7	2	1	1
FR. «haber»+pp	100%	95%	97%	89%	97%	95%	99%	100%	100%
FR. pp+«haber»	0%	5%	3%	11%	3%	5%	1%	0%	0%
FN. «haber»+pp	0.272	0.235	0.484	0.723	0.893	1.219	1.405	3.049	5.267
FN. pp+«haber»	0.000	0.013	0.014	0.093	0.024	0.065	0.018	0.009	0.018

Tabla 18. «haber» + pp / pp + «haber». CODEA.

Año	1200	1250	1300	1350	1400	1450	1500	1550	1600
FA. transitivo	4	19	29	52	77	107	104	206	175
FA. intransitivo	0	0	3	11	31	21	33	84	74
FA. reflexivo	0	0	2	6	4	6	21	42	43
FA. haber	0	0	2	1	3	5	1	7	2
FR. transitivo	100%	100%	81%	74%	67%	77%	65%	61%	60%
FR. intransitivo	0%	0%	8%	16%	27%	15%	21%	25%	25%
FR. reflexivo	0%	0%	6%	9%	3%	4%	13%	12%	15%
FR. haber	0%	0%	6%	1%	3%	4%	1%	2%	1%
FN. transitivo	0.272	0.248	0.401	0.607	0.614	0.988	0.930	1.858	3.146
FN. intransitivo	0.000	0.000	0.041	0.128	0.247	0.194	0.295	0.758	1.330
FN. reflexivo	0.000	0.000	0.028	0.070	0.032	0.055	0.188	0.379	0.773
FN. haber	0.000	0.000	0.028	0.012	0.024	0.046	0.009	0.063	0.036

Tabla 19. «haber» + pp (transitivo / intransitivo / reflexivo / haber). CODEA.

Etapa	León	Castilla norte	Castilla sur	Navarra	Aragón
1140-1295	0%	4%	0%	0%	5%
1296-1369	0%	3%	1%	0%	2%
1370-1454	4%	5%	13%	3%	5%
1455-1499	18%	11%	18%	10%	27%

Tabla 20. «haber» + pp con predicados intransitivos (vs. transitivo).

Rodríguez Molina (2010).

Año	1200	1250	1300	1350	1400	1450	1500	1550	1600
FA. Oeste	0	0	0	0	3	11	7	10	14
FA. Este	0	0	3	11	28	10	26	74	60
FN. Oeste	0.000	0.000	0.000	0.000	0.024	0.102	0.063	0.090	0.252
FN. Este	0.000	0.000	0.041	0.128	0.223	0.092	0.233	0.668	1.079

Tabla 21. «haber» + pp (intransitivo). CODEA.

Año	1200	1250	1300	1350	1400	1450	1500	1550	1600
FA. «ser»+pp	0	0	1	0	4	5	6	3	3
FA. «haber»+pp	0	0	3	11	31	19	35	85	78
FR. «ser»+pp	*	*	25%	0%	11%	21%	15%	3%	4%
FR. «haber»+pp	*	*	75%	100%	89%	79%	85%	97%	96%
FN. «ser»+pp	0.000	0.000	0.014	0.000	0.032	0.046	0.054	0.027	0.054
FN. «haber»+pp	0.000	0.000	0.041	0.128	0.247	0.175	0.313	0.767	1.402

Tabla 22. «ser» / «haber» + pp (intransitivo). CODEA.

Propiedad	CID	MIL	LRA	RIM	TAL	CAR	CEL	LOZ	TER
1. Orden de aux.-pp	77	95	75	95	90	100	95	100	100
2. Contigüidad de aux.-pp	77	59	73	91	91	91	91	95	95
3. Uso de «tener»	20	19	44	47	78	95	88	97	89
4. No concordancia de pp	14	13	29	29	88	100	100	100	100
5. Incompatibilidad	27	23	14	27	27	41	55	55	59

Tabla 23. Evolución de las propiedades de «haber» + pp (%). Olbertz (1993)

Propiedad	XII	XIII	XIV	XV
1. Contigüidad de aux.-pp	83	88	93	99
2. Orden de aux.-pp	76	97	85	99
3. Haber/ser + pp	70	89	78	89
4. No concordancia de pp	22	32	89	100

Tabla 24. Evolución de las propiedades de «haber» + pp (%). Romani (2006)

[FIN]