


Entre la norma ortográfica y el uso en la escritura oficial en el siglo XVIII: estudio de la variación gráfica en tres subcorpus¹

Elia Puertas Ribés
Universitat Jaume I (Castellón, España) ✉ 

<https://dx.doi.org/10.5209/dice.104719>

Fecha de envío: 03/09/2025 • Fecha de aceptación: 04/05/2026

Resumen: La lengua española del dieciocho ha sido, hasta no hace mucho, desatendida por considerarse una época de estabilización, puesto que vino de la mano de la fundación de la Real Academia Española (1713) y de su preocupación por la regularización del uso de las grafías. En este marco, el presente estudio tiene por objeto el análisis de algunos usos gráficos atestiguados en el sistema consonántico de documentos de mediados del siglo XVIII, hallados en el Archivo Municipal de Castellón e integrados en el corpus INCOM. Estos manuscritos se comparan con otros procedentes de las zonas centro y sur peninsular con la misma tipología y datación. Así, es posible comprobar en qué medida la normalización ortográfica de la Academia surtió efecto en la escritura oficial del reino de Valencia. En estos documentos, pues, es posible advertir la convivencia entre distintas tendencias gráficas aceptadas socialmente en una época determinada, lo que nos permite constatar que nos hallamos ante un momento de *estandarización en proceso* (Ramírez Luengo 2013).

Palabras clave: ortografía; siglo XVIII; corpus INCOM; Real Academia Española.

^{EN} Between spelling rules and usage in official writing in the 18th century: a study of graphic variation in three subcorpora

Abstract: Until recently, the Spanish language of the eighteenth century had been neglected because it was considered a period of stabilisation, as it coincided with the founding of the Royal Spanish Academy (1713) and its concern for the standardisation of spelling. Within this framework, the present study aims to analyse some spelling conventions attested in the consonant system of mid-18th-century documents found in the Municipal Archive of Castellón and included in the INCOM corpus. These manuscripts are compared with others from the central and southern areas of the peninsula with the same typology and dating. Thus, it is possible to verify the extent to which the Academy's orthographic standardisation had an effect on official writing in the Kingdom of Valencia. In these documents, therefore, it is possible to observe the coexistence of different spelling trends that were socially accepted at a given time, which allows us to conclude that we are dealing with a moment of *standardisation in progress* (Ramírez Luengo 2013).

Keywords: spelling; 18th century; INCOM corpus; Royal Spanish Academy.

Sumario: 1. Introducción. 2. Ortografía normativa en el primer español moderno. 3. Corpus y metodología. 3.1. Corpus base. 3.2. Corpus de control. 3.3. Metodología de análisis. 4. Análisis de los datos. 4.1. La y la <v>. 4.2. La <g>, la <j> y la <x> + <e>, <i>. 4.3. La <h>. 4.4. La <c>^{+e,i}, <ç>, <z> y <s>, <ss>. 5. Conclusiones

Como citar: Puertas Ribés, E. (2026). Entre la norma ortográfica y el uso en la escritura oficial en el siglo XVIII: estudio de la variación gráfica en tres subcorpus. *Dicenda. Estudios de lengua y literatura españolas* 44(2026) 1-12. <https://dx.doi.org/10.5209/dice.104719>

1. Introducción

¹ El presente trabajo forma parte del proyecto de investigación «Componentes socioestilísticos, idiolectales y discursivos en la variación y el cambio lingüístico en español: contribuciones desde la sociolingüística histórica» (2022-2026), financiado por el Ministerio de Ciencia e Innovación/Agencia Estatal de investigación y por FEDER Una manera de hacer Europa (Ref. PID2021-122597NB-I00) así como por el Subprograma estatal de Movilidad, del Plan Estatal de Investigación Científica, Técnica y de Innovación (PRX23/00019).

Hasta la década de los noventa, el español dieciochesco fue comúnmente considerado como una etapa de estabilización del idioma (Cano Aguilar, 2018 [2004]: 825), puesto que vino de la mano de la fundación de la Real Academia Española (RAE) (1713) y de su propósito de preservar el buen uso de la lengua española con la publicación de obras como el *Diccionario de Autoridades* (1726-1737) y la *Ortografía de la lengua castellana* (1741), entre otras. Estos sucesos históricos, que constituyen la historia externa de la lengua, sirvieron para bautizar este periodo como el inicio del *español moderno*, caracterizado —en aquel entonces— por el cese de cambios lingüísticos, de ahí que las obras cumbre de la historia de la lengua castellana, como las de Lapesa (1985) y Cano Aguilar (2018 [2004]), dedicasen a esa centuria solo unas páginas, en las que centraron su atención principalmente en la labor de la Academia y en cuestiones léxicas (Álvarez de Miranda, 2004; Calderón Campos, 2012).

Esta periodización, por tanto, no tuvo en cuenta la evolución y estructuración internas de la lengua española, pero, afortunadamente, cada vez hay más estudios basados en corpus lingüísticos diacrónicos que aúnan esfuerzos para recuperar testimonios de esa centuria y demostrar así que, lejos de ser una etapa de ‘ayuno’ en cambios (García-Godoy, 2012a: 10), el español del siglo XVIII manifiesta variación lingüística en distintos niveles (véase el volumen editado por García-Godoy, 2012b, así como los trabajos de Ramírez Luengo, 2006; Calderón Campos, 2018; González Sopeña, 2023; Blas Arroyo y Puertas Ribés, 2024a; entre otros).

Concretamente, si nos fijamos en el plano ortográfico, la publicación de tratados sobre esta materia se venía realizando ya varios siglos antes del periodo que aquí estudiamos. Y, sin embargo, a principios del siglo XVIII, la lengua española carecía todavía de un sistema ortográfico unitario (Martínez Alcalde, 2007), lo que llevó a la Academia a reconocer que donde:

[...] se experimenta notable desigualdad en la Orthographía de la Lengua Castellana es en el modo de escribir las palabras, porque cada uno ha usado el método que le ha dictado su génio y manera de hablar, ò según los primeros rudimentos que aprendió en la escuela quando niño, ú después el uso le ha ido enseñando, sin atender á otra regla que à la común y vulgar, de que siendo la Orthographía un arte de bien escribir, lo que se escribe es copia de lo que se habla, ò se concibe. Todo esto procede de no haberse establecido reglas para su verdadero uso con la autoridad necesaria, desde que se comenzó a escribir la lengua castellana (*Diccionario de Autoridades*: LXVI).

Por ello, la Real Academia Española inicia su propio proceso de normalización y normativización de la lengua, que convivirá con las publicaciones de otros ortógrafos, impresores y autores de la época. Prueba de ello son las distintas publicaciones de Martínez Alcalde (1990, 2007, 2010, etc.), en las que se describen detalladamente las propuestas ortográficas de esa centuria desde una perspectiva historiográfica. Además, la ortografía del XVIII cuenta en la actualidad con diversos trabajos sobre cuestiones (orto)gráficas, elaborados a partir del estudio de casos de documentos de la época (Gómez Seibane y Ramírez Luengo, 2007; Ramírez Luengo, 2012, 2013, 2015; etc.); en algunos de ellos, además, es posible constatar la convivencia de diferentes variantes en función de la zona geográfica (Company Company, 2007; Mondéjar, 2007; González Sopeña, 2023; etc.).

En nuestro caso, el presente artículo se une al estudio de fuentes documentales inéditas para examinar las tendencias gráficas en el reino de Valencia. Para ello, buceamos en un conjunto de 56 cartas administrativas, incorporadas recientemente al corpus INCOM. Esta base de datos textuales se conforma como el corpus base de trabajo, aunque, con el objetivo de verificar y contrastar los datos, manejaremos también dos subcorpus del mismo periodo y tipología textual, procedentes de dos zonas geográficas diferentes. Así pues, estos subcorpus de control se toman prestados del CODEA+2022, para la obtención de datos de la zona centro, y del CODEMA, para la recuperación de textos del sur peninsular. De este modo, se trata de comprobar si hallamos usos lingüísticos que representen ciertos rasgos gráficos dialectales, o si, más bien, esas tendencias son un reflejo general en el corpus de una *estandarización en proceso* en los usos gráficos de la época (Ramírez Luengo, 2015: 438). Para este estudio, el corpus que se maneja se ha compilado con transcripciones paleográficas de textos de carácter jurídico-administrativo que han sido redactados por escribanos, profesionales que, por lo general, solían poseer una formación sólida antes de ejercer ese cargo (Almeida Cabrejas *et al.*, 2018). De este modo, tratamos también de ahondar en la formación de los escribanos del corpus base (corpus INCOM), al tiempo que pretendemos entender hasta qué punto la normalización y normativización ortográfica de la Academia se adaptó e integró en la escritura de los amanuenses de mediados del siglo XVIII.

A continuación, el artículo se estructura como sigue. En primer lugar, se ofrece un repaso sobre las aportaciones ortográficas de la Academia (apartado 2). Seguidamente, en el tercer epígrafe se presentan los corpus lingüísticos que se manejan para la obtención de datos, así como la metodología que se ha seguido para el análisis. En el cuarto apartado, se expone el estudio de las variantes ortográficas más frecuentes en el corpus para observar la difusión, similitud y diferencias entre las distintas bases de datos textuales seleccionadas. Por último, en el 5, se extraen las conclusiones pertinentes sobre el estudio llevado a cabo en estas páginas.

2. Ortografía normativa en el primer español moderno

La fijación del sistema ortográfico español se sustenta sobre tres criterios básicos, que son: la pronunciación, el uso y la etimología (Martínez Alcalde, 2010). Por ello, la variedad de factores y criterios que se adoptaron en las obras que vieron la luz durante la primera mitad del siglo XVIII sobre los procesos de codificación

y unificación reflejan para unos el caos ortográfico que imperaba en la época (Rosenblat, 1951; Lapesa, 1985; Zamora Vicente, 1999; Arnal Purroy, 2000; etc.), mientras que, para otros, un periodo de *estandarización en proceso* (Frago, 2010; Díaz Moreno, 2014; Ramírez Luengo, 2015; etc.).

Sea como fuere, la Academia, ante esa falta de consenso en materia ortográfica, sintió la necesidad de crear sus propias normas para la creación del *Diccionario de Autoridades* (1726), de ahí que se publicase, junto a esa obra lexicográfica, el *Discurso Proemial de la Orthographía de la Lengua Castellana*, basado principalmente en los criterios etimológicos, pero también del uso que hacían autores ilustres y de la adecuación a la pronunciación (Martínez Alcalde, 2007: 112; Freixas, 2016: 144). Posteriormente, esas reflexiones ortográficas dieron lugar, en 1741, a la *Orthographía española*, un tratado mucho más elaborado que vendría a iniciar la primera de las seis ediciones que se publicarían a lo largo del siglo XVIII (Gómez Asencio, 2008: 35), aunque los procesos de oficialización llegarían todavía más tarde, en las primeras décadas del siglo XIX (Martínez Alcalde, 2007: 61).

No obstante, en el periodo que nos ocupa, si bien la Academia no era el único referente en cuestiones ortográficas, sus obras empezaron a ganar prestigio y reconocimiento, por ese motivo serán las que tendremos en cuenta en este estudio. Así pues, centrándonos ya en la obra de 1741², dado que es la que más se aproxima a los textos del corpus base (corpus INCOM), advertimos que la RAE estructura los apartados de las letras comenzando, bien con la descripción de las grafías desde el punto de vista de su pronunciación, bien con la comparación con otras letras o, incluso, con otras lenguas para facilitar la comprensión de cómo suena en la práctica. Seguidamente, cuando existen alternancias entre distintas grafías, remite normalmente a criterios etimológicos, aunque, en ocasiones, la elección de una u otra se justifica también por el uso habitual de algunas voces.

El primer par de letras de las que nos ocuparemos en las siguientes páginas es el formado por la y la <v>. En la entrada correspondiente a la primera de ellas se recuerda ya que no existen diferencias en su pronunciación con respecto a la segunda (<v>), por lo que se advierte que, para señalar la fijación de estas grafías, el criterio de la pronunciación no es válido, así como tampoco lo es el del uso. Por tanto, la Academia se ciñe a principios etimológicos. En el caso de la segunda grafía, la <v> se describe junto con una tercera grafía, la <u>, por tener funciones similares tanto en el sistema vocálico como en el consonántico. En nuestro caso, restringimos el estudio solamente a la <v> consonántica por ser la más frecuente en nuestro corpus base. Así pues, como se resume en la Tabla 1, la Academia sienta las reglas ortográficas de la y la <v> fundamentalmente sobre el principio etimológico, aunque, para aquellas voces de origen incierto, prescribe el empleo de la , uso habitual ya en esa época, especialmente por parte de los intelectuales.

Tabla 1. Síntesis de las reglas ortográficas establecidas para la y la <v> en la *Orthographía* (1741)

	<v>
<ul style="list-style-type: none"> • Todas las voces que en su origen se escriben con , se deben escribir con la misma grafía (p.143) • Del mismo modo, si el origen de la palabra es desconocido, se escribirá con , por ser como se pronuncia normalmente (pp.134 y 144) • Cuando el étimo latino tiene P (<i>capere</i> > <i>caber</i>) (p. 144) • Antes de las consonantes líquidas L y R (p.144). 	<ul style="list-style-type: none"> • Todas las palabras que por uso frecuente se escriben con <v> (como <i>virtud</i>, <i>virgen</i>, <i>vuestra</i>, etc.) (p. 133) • Se escribe esta grafía siempre y cuando esté presente en la raíz de la voz española (p.209).

De forma similar, los académicos inician el apartado de la <g> diferenciando los dos sonidos que se atribuyen a esta grafía en función de la vocal que la sigue. En este trabajo, centraremos la atención en el uso de la <g> seguida de las vocales anteriores (<e, i>) para la transcripción del fonema /x/. Por otro lado, este fonema puede representarse también con dos grafías más: la <j> y la <x>. En el primer caso, como ya ocurría con en la entrada de la <v>, la Academia describe la <j> en oposición a la <i> por presentar alternancias en la escritura en mayúscula, mientras que en su pronunciación suele intercambiarse con la <g> y con la <x> velares. Por su parte, la Academia confiesa las dificultades para explicar el uso de la <x>, por lo que remite al origen etimológico de la palabra para emplearla adecuadamente. Por consiguiente, se recomienda nuevamente acudir al origen de las voces, aunque reconoce que ese criterio no está exento de sus propias dificultades, puesto que el étimo no siempre presenta dichas grafías en la lengua de la que proceden. Sea como fuere, la institución dicta las reglas que se sintetizan más abajo:

² Aunque no se mencione de forma explícita el *Discurso* de 1726 en este apartado, se tiene en cuenta en el análisis de resultados.

Tabla 2. Síntesis de las reglas ortográficas para la <g>, la <x> y la <j> en la *Ortographia* (1741)

<g>, <x>	<j>
<p style="text-align: center;"><g></p> <ul style="list-style-type: none"> Las voces, cuya pronunciación con la <e,i> es 'gutural' y la contienen en su origen, se escriben con <g> (<i>geometría, gira</i>) (p.155) Aquellas que en latín conservan la <hi> (<i>hieroglyphicus > geroglífico</i>) (p.156) El resto de palabras, con pronunciación fuerte, se escribe con <j> (p. 156). <p style="text-align: center;"><x></p> <ul style="list-style-type: none"> Se usa cuando se halla en la lengua latina (p. 216) Si se pronuncia fuerte, como en su origen, se escribe con <x>; en cambio, la pronunciación suave, se marcará con el acento circunflejo en la vocal que le sigue (p.217) Cuando el étimo latino tiene S, se convierte en <x> en castellano. También se emplea en aquellas voces con pronunciación 'gutural', tanto en singular como en plural (p.217). 	<ul style="list-style-type: none"> Se emplea en nombre en cuyo origen está presente la <i>, pero su pronunciación es 'gutural' (<i>lustitia > justicia</i>) (p.177) Igualmente, las voces árabigas que se han incorporado al español (<i>jubon, jabal</i>) (p.177) Si en su origen se usa la G (y no la X) en las sílabas <i>ga, go, gu</i>, estas se escribirán con <j> (p.178) De forma similar, la sílaba latina LI o doble <l> se convierte en español en <j> (<i>consilium > consejo; malleator > majador</i>) (p.178) Ante duda de si se escribe <x> o <j>, en palabras cuyo origen no lo determina, se usa <j> (p.179).

El apartado de la <h> se abre retomando la polémica sobre su estatus como letra del abecedario hispánico, así como el papel que desempeña como aspiración en ciertas sílabas. Como se ha comentado ya en las grafías anteriores, la Academia se guía principalmente por la etimología para establecer las siguientes reglas de la tabla 3:

Tabla 3. Síntesis de las reglas ortográficas establecidas para la <h> en la *Ortographia* (1741)

<h>
<ul style="list-style-type: none"> Se conserva en las voces en cuyo origen tienen H antes de alguna vocal (p. 162) Se emplea en algunas palabras que presentan una F en el étimo (<i>facere > hacer</i>) (p. 163) Asimismo, se conserva en voces derivadas en cuya palabra simple aparece (<i>rehacer</i>); exceptuando aquellas con <i>ab-</i>, para las que se suprime la <h> del étimo latino (p. 163) Cuando la palabra empieza por <i>ue</i> (de la O latina), se añade <h> (<i>orphanus > huérfano</i>) (p. 163) Se emplea tras la <c> para marcar la pronunciación del dígrafo <ch> (p. 163) En ocasiones, si las voces vienen del griego o del hebreo, se emplea también la <h> después de la <c> para marcar el sonido velar (<i>châridad</i>) (p.164) La grafía culta latina <ph> se escribe en nombre de facultades y propios; en cambio, tras la R, se usa en aquellas voces griegas que en su origen se aspiraban. La <th>, por su parte, se emplea en palabras comunes, como <i>theatro</i>. En definitiva, el uso de la <h> después de la <r> o <t> es signo de erudición, aunque no es incorrecto omitirla, aunque sí lo será si se emplea en aquellas que no la tienen en su origen (pp. 166-167).

Por último, en este trabajo describiremos las sibilantes, para las cuales la Academia considera pertinente comenzar su descripción aclarando su pronunciación en función de la vocal que la sigue, aunque, esto ya se venía haciendo en las entradas de grafías anteriores, advierte de que tanto la pronunciación como el uso de las voces con <s> líquida varían con el tiempo. Por tanto, no se ofrece mayor detalle sobre la <s>. No obstante, es interesante recordar que la Academia dedica un apartado a la duplicación de consonantes; en esa sección, si bien fija como regla general no duplicar las grafías, hace una excepción para algunas, como la doble ese (<ss>), en casos como los que se citan en la tabla 4. Véanse, a continuación, las recomendaciones para las grafías <c>, <z>, <ç> y <ss>:

Tabla 4. Síntesis de las reglas ortográficas para <c>, <z>, <ç> y <ss> en la *Ortographia* (1741)

<c>	<ss>
<ul style="list-style-type: none"> Las sílabas <i>ce</i>, <i>ci</i> se escriben siempre con <c>, mientras que la <z> se mantiene solo en aquellas que la tienen en su origen (p. 146) Los nombres que terminan el singular con <z>, hacen el plural -ces con <c> (p.147). 	<ul style="list-style-type: none"> Se duplica en los superlativos y en los pretéritos de subjuntivo (p.231) Ocurre lo mismo en los derivados que se elide la última letra del prefijo por una segunda <s> (<i>assentar</i>) (p.232) Se duplica en el adverbio <i>assí</i> para distinguirlo del pretérito de <i>asir</i> (p.232).
<z>, <ç>	
<ul style="list-style-type: none"> La Academia advierte que la <ç> tiene la misma pronunciación que la <z>, por lo que se considera superflua (p.150) La <z> se usa si está en el origen ante las vocales <i>a</i>, <i>o</i>, <i>u</i>, o al final de algunas palabras en singular en las que el latín tenía <i>x</i> (p.322). 	

3. Corpus y metodología

Para el estudio de la fijación ortográfica se requiere una muestra textual lo suficientemente representativa como para obtener datos que nos permitan llevar a cabo una descripción fiable. Por eso, de acuerdo con la metodología propuesta por Ramírez Luengo (2020), no solo nos hemos servido de los testimonios extraídos del corpus INCOM, sino que además se han seleccionado otras fuentes semejantes “desde el punto de vista tipológico y sociolingüístico, de manera que los resultados obtenidos en los análisis sean comparables” (p. 37), tanto cualitativa como cuantitativamente. Así pues, para la obtención de datos manejaremos textos procedentes de tres corpus lingüísticos: corpus INCOM, CODEA+2022 y CODEMA. El primero se usa como corpus base (§ 3.1.) para el estudio, mientras que los otros dos funcionan como corpus de control (§ 3.2.) para ampliar, comparar y contrastar la información.

3.1. Corpus base

El corpus INCOM, compilado por el grupo de investigación *Sociolingüística* (O93), de la Universitat Jaume I, reúne una cifra superior a los catorce millones de palabras, escritas por más de seis mil individuos de diferente extracción social y dialectal, entre mediados del siglo XVI y la primera mitad del XX (Blas Arroyo y Puertas Ribés, 2024b). Además, a estos materiales se han incorporado recientemente otros documentos, de carácter inédito, procedentes del Archivo Municipal de Castellón, que dan voz a más de veinte individuos de diversas localidades de las provincias de Castellón y Valencia, y cuya temática gira en torno al asentamiento forzoso y a la persecución que sufrió la comunidad gitana a mediados del siglo XVIII (1745-1750) (Puertas Ribés, 2026). De esos textos, se han seleccionado 56 cartas administrativas escritas por escribanos que suman un total de 10.122 palabras³.

3.2. Corpus de control

Por su parte, el corpus de control se conforma por dos subcorpus: por un lado, uno recupera testimonios de la región central de la península a través del corpus CODEA+2022; por otro lado, el segundo recoge textos del sur, concretamente, de la zona occidental de Andalucía, gracias al corpus CODEMA.

El CODEA+2022 (*Corpus de Documentos Españoles Anteriores a 1900*), compilado por el grupo de investigación GITHE, de la Universidad de Alcalá, reúne textos desde el siglo XVI hasta el XIX, aunque presenta ciertas carencias muestrales para la centuria que nos ocupa. Por tanto, para equilibrar esa muestra, a los 13 manuscritos del CODEA+2022 se han añadido cuatro documentos del corpus INCOM, redactados y emitidos por escribanos de la corte madrileña, y que habían sido previamente descartados como fuente documental para el corpus base por no pertenecer al reino de Valencia. De este modo, el subcorpus CODEA+2022 (e INCOM) asciende a 10.372 palabras.

Por su parte, el segundo corpus del que nos servimos aquí es el del proyecto CODEMA (*Corpus Diacrónico de Documentación Malagueña*), del grupo ARINTA, de la Universidad de Málaga, que reúne documentos municipales, eclesiásticos y privados de esa ciudad entre los siglos XVI y XIX. En este caso, se han seleccionado 12 documentos, que alcanzan una extensión similar a los dos anteriores (10.504 palabras).

A continuación, se ofrece una tabla con los datos que se manejan para este estudio. Como puede observarse, el número de palabras de los textos es similar en los tres corpus (aproximadamente unas 10.000).

³ La selección de las cartas no responde al orden numérico habitual, sino más bien al que se establece cronológicamente dentro del corpus INCOM. Por tanto, se recuperan fragmentos que van más allá de la carta 56 (por ejemplo, el de la carta 65), mientras que algunos manuscritos, como las cartas de la 35 a la 38, se excluyen del corpus base, por haber sido redactadas en zonas geográficas que no pertenecen al reino de Valencia.

Además, han sido redactados por escribanos a mediados del siglo XVIII y pertenecen a tipologías textuales similares (actas y declaraciones, textos legislativos, testimonios y declaraciones, etc.).

Tabla 5. Dimensiones del corpus analizado

	INCOM	CODEA (e INCOM)	CODEMA
N.º de palabras	10.122	10.372	10.504
N.º de textos	56	17	12
Tipología textual	Cartas administrativas; notificaciones.	Actas y declaraciones; Textos legislativos; Testimonios y declaraciones; certificaciones	Actas y declaraciones; Textos legislativos.

En definitiva, la selección de estos corpus responde a la posibilidad de acceder a fuentes documentales de las tres zonas geográficas mencionadas y trabajar, además, con la edición paleográfica de los manuscritos, transcripción que, en los tres casos, sigue la presentación propuesta por la red internacional CHARTA.

3.3. Metodología de análisis

A continuación, los textos elegidos se han compilado en tres subcorpus en la plataforma de Sketch Engine con el objetivo de automatizar las búsquedas, tanto por subgrupo como conjuntamente. Para el estudio ortográfico, se han seleccionado aquellas variantes gráficas que presentan mayor inestabilidad en el plano escrito, como son la y <v>; la <g>, la <j> y la <x>; la <h>; y la <s>, la <ss>, la <c>, la <ç> y la <z>.

Las búsquedas se han realizado por medio de la funcionalidad de *Wordlist*, que permite obtener listados de palabras con su frecuencia de aparición. Dentro de esta función, nos hemos servido de la opción *Advanced* con el propósito de realizar las búsquedas por medio de expresiones regulares (*regex*), es decir, patrones que permiten encontrar distintas cadenas de texto. Por ejemplo, para las palabras cuya inicial era la , se ha empleado la expresión 'b.*'. Así, se han encontrado las voces que empiezan por esa letra seguida de cualquier número de caracteres. Ahora bien, para poder trabajar con mayor precisión, las ediciones paleográficas se han revisado manualmente antes de su compilación con el objetivo de separar aquellas palabras que se hubiesen transcrito unidas y se han eliminado también las comillas angulares para reconocer con mayor facilidad las palabras. No obstante, las abreviaturas se han dejado sin desarrollar, para estudiar solamente las grafías empleadas por los escribanos. Además, se han excluido del estudio los nombres propios, dado que sus grafías no se fijarán hasta el siglo XIX (Almeida Cabrejas, 2014a: 186), así como palabras vacías (*stopwords*) o aquellas que pudieran presentar frecuencias extremas, por ejemplo, el posesivo *vuestra* en el estudio de las grafías y <v> (y de la <s>).

Por último, huelga decir que los datos se analizan tanto desde una perspectiva cualitativa como cuantitativa. En el primer caso, nos basaremos en el estudio de los textos en función de las reglas ortográficas propuestas por la Academia en 1726 y 1741. Seguidamente, los datos cuantitativos se presentan en la tabla con los valores absolutos (*N*), mientras que el valor relativo se ofrece solo en aquellos que cumplen (*Cumple*) las normas propuestas por la Academia.

4. Análisis de los datos

Como se ha mencionado más arriba, el estudio que se presenta a continuación parte de la transcripción paleográfica de los textos, dado que respeta las características gráficas de los manuscritos. La elección de esta edición se justifica, pues, por el valor que adquiere para el estudio (orto)gráfico —y, también, fonético— de los documentos (Sánchez-Prieto, 2011: 16). En los siguientes apartados analizaremos las tendencias generales de los tres subcorpus para describir las diferencias entre los usos de la y la <v> (§ 4.1.); de la <g>, <j> y <x> (§ 4.2.); la <h> (§ 4.3.); y <s>, <ss>, <c>, <ç> y <z> (§ 4.4.).

4.1. La y la <v>

Las grafías y <v> —cuya distinción fonética en español es inexistente— representan en la escritura el fonema /b/. Para su uso, como se muestra en la tabla 1 (§ 2), la Academia se rige principalmente por criterios etimológicos para indicar el uso de una u otra consonante, aunque, en ocasiones, acude a la pronunciación y uso frecuente de algunas voces. En el caso de la labiodental (<v>), en la práctica puede alternarse con la <u>, así se muestra en el trabajo de Arnal Purroy (2000); sin embargo, en el corpus INCOM no aparece documentada con ese valor, por lo que centraremos nuestra atención en las grafías y <v>, esta última solo como consonante.

Así pues, si buceamos en los subcorpus es posible advertir alternancias en la escritura de algunas palabras. Ahora bien, aunque estas variaciones se producen normalmente en palabras con la misma raíz léxica (CODEMA: *venta/benta*, *vendedor/bendedor*), conviene señalar que los testimonios que no siguen la *Orthographía* (1741) al inicio de palabra son aquellos que, en realidad, mantienen la recomendación extendida de que “se debe usar de la B quando no se sabe ciertamente, ò se duda del origen de las dicciones” (*Diccionario de Autoridades*: LXXII). Es, posiblemente, el caso de todos los testimonios que no se han clasificado en la tabla 6 como que ‘Cumplen’.

En posición interna, la tendencia mayoritaria es confundir preferentemente verbos —en distintos tiempos verbales— como *haber/haver* (e, incluso, *hauer* en el corpus CODEMA); *deber/dever*; *recibir/recivir*. En el corpus base de este estudio, por ejemplo, hallamos la alternancia en distintos documentos redactados por el mismo escribano, el del corregidor de Castellón, Cristóbal Pascual:

- (1) los referidos quince Gitanos para conducirles a esta Villa; Y ordeno, y mando a las justicias por donde transitaren les **recivan** de noche en sus carseles, y les guarden con veinte hombres armados de que quedaran responsables; [...]. (INCOM, *carta* 40)
- (2) quietud del Reyno asegurar esta clase de Jentes segun lo encarga su Magestad prevengo â Vuestras mercedes que luego que **reciban** esta orden, y reconoscan con el mayor cuidado el pueblo y termino a fin del logro de los expresados gitanos y caso de [...]. (INCOM, *carta* 42)

Posiblemente, la confluencia de los antiguos fonemas oclusivo y fricativo en /b/ en épocas anteriores dificultó la escritura de una u otra grafía, lo que no fue una tarea fácil para la Academia, ya que no siempre fue coherente. Esto se observa en la escritura del verbo *haber* (con <v> en los primeros tres tomos del diccionario, mientras que aparece con en el cuarto) (Almeida Cabrejas, 2014a: 12); de ahí las alternancias que hallamos en nuestro corpus, que, pese a no responder a las reglas establecidas por aquel entonces, siguen los usos habituales que predominaban hasta entonces.

Cuantitativamente, en el conjunto del corpus se han encontrado 253 palabras que empiezan con la , de las cuales menos de la mitad sigue las reglas ortográficas de la RAE (N=119; 47 %). En interior de palabra, en cambio, la tiende a emplearse conforme a lo establecido por la RAE en el 88,6 % de los casos (N=788). De forma similar, la <v> obtiene un total de 518 casos en posición inicial, de los cuales el 93,8 % se adecúa a la ortografía de la Academia; además, de forma similar en los tres subcorpus (INCOM: 97,5 %; CODEA e INCOM: 97,5 %; CODEMA: 87,7 %). Sin embargo, parece que en posición medial genera más vacilaciones entre los escribanos, especialmente en aquellos del corpus base (49,6 %) y del CODEA+2022 (e INCOM) (59,6 %). De todos modos, como muestran los datos de la tabla 6, la tendencia general demuestra que los escribanos conocían los usos a la hora de emplear estas grafías, aunque, como ocurre con la inicial, se mantienen tradiciones anteriores en caso de duda.

Tabla 6. Resultados de la distribución y <v> en el corpus

	B-	Cumple	-B-	Cumple	V-	Cumple	-V-	Cumple	Total
INCOM	72	66 (91,7%)	273	247 (90,5%)	160	156 (97,5%)	369	183 (49,6%)	874 (74,6%)
CODEA	131	24 (18,3%)	330	287 (87%)	163	159 (97,5%)	108	64 (59,2%)	732 (73%)
CODEMA	50	29 (58%)	286	254 (88,8%)	195	171 (87,7%)	390	357 (91,5%)	921 (88%)
Total	253	119 (47%)	889	788 (88,6%)	518	486 (93,8%)	867	604 (69,7%)	2.527

4.2. La <g>, la <j> y la <x> + <e>, <i>

Asimismo, es interesante examinar el uso de <g>, <j> e <x> seguidas de las vocales <e>, <i> para la representación del fonema fricativo velar sordo, dado que es el que origina más confusiones. Recordemos que la Academia restringe el empleo de la <g>^{+e,i} para cuando se halle en su origen; la <j>, además de seguir la escritura de las palabras de las que derivan, se usa también en otros contextos, como en las voces que presentan la sílaba LI latina o la doble <l>. Por último, la <x> se escribe en aquellos términos que pronuncian esta grafía como [cs] o [gs] (*examen*), así como en las palabras que ofrecían <s> o <x> en su origen.

En posición inicial, hallamos un desequilibrio entre el uso de la <g> (N=246), la <j> (N=5) y la <x> (N=2), esta última recogida solamente en el corpus base (INCOM) para la palabra *xitano* en dos documentos del escribano del capitán general de Valencia. En este sentido, cabe señalar, que el sustantivo *gitano* (N=65) se recoge también en este corpus como *jitanos* (N=1) e, incluso, como *guitano* (N=5). En este último caso, cuatro de los testimonios aparecen en documentos redactados por el escribano del corregidor de Castellón, mientras que el otro restante, en uno de los alcaldes, posiblemente por imitación a la carta recibida por el dirigente castellanense. A continuación, se ofrecen cuatro fragmentos que ilustran las vacilaciones que presenta el sustantivo *gitano*:

- (3) carseles publicas de esta villa, devo dezir que con la noticia de que Silvestre Castellon, llamado comunmente Marco el **Gitano** vezino de la Villa de Borriana, seria en esta practicado algunas ventas, y trueques de vagajes, contra lo prevenido, en (INCOM, *carta* 61)
- (4) Mui señor mio. Habiendo ayer, restituidos, al respecto el ofisial, que con la Partida, salio a conducir, los Onze **Jitanos**, que de esas carseles, se ynbiaron â Alicante; Pass â Partisiparselo â Vuestra merced. Yncluyendo le el rezibo, (INCOM, *carta* 63)

- (5) las diligencias de embargo de bienes contra Joseph Herrandi vezino de esta villa por los 85p que esta deviendo al **Guitano**: he embargado un rozin propio de dicho Herrando que tengo augurado en poder del Repositorio lo que participo a Vuestra (INCOM, *carta 64*)
- (6) Del duque de Caylus señor don Joseph Bermudo Aviso a Vuestra Señoría el rezívo del rexistro de familias de **xitanos** establezidas en esa Villa correspondiente al proximo pasado Mes, que me remíte en Carta de 7. del corriente: Y en quanto (INCOM, *carta 11*)

En posición medial, sin embargo, no se usa tampoco de forma homogénea dentro de los subcorpus. En concreto, en el corpus principal (INCOM) la alternancia se advierte mayoritariamente en la distribución de <x> y <g>, o <x> y <j>, en determinadas palabras como, por ejemplo, *regimiento/reximiento*; *registro/rexistro*; *ejecutar/executar*, etc. Del mismo modo, tanto en el corpus INCOM como en el CODEA+2022 hallamos fluctuaciones en el empleo de <g> y <j> (*margen/marjen*; *gente/jente*; *muger/mujer*, etc.), aunque se documentan en textos de distintos escribanos. En definitiva, los datos de nuestro corpus revelan un uso mucho más desequilibrado. La <g>, por ejemplo, no parece suscitar dudas a la hora de emplearse como se establece en la ortografía de la época. En cambio, no ocurre lo mismo con la <j>, que, tanto en posición inicial como interna, parece que los porcentajes en los que se ciñe a las reglas son nulos (en posición inicial) o mucho más bajos (en posición interna > INCOM: 17,6 %; CODEMA: 26,8 %); exceptuando el segundo subcorpus (CODEA+2022 e INCOM), que alcanza el 50 %. No obstante, dado el bajo número de casos obtenidos para esta grafía, estos datos no permiten llegar a ninguna conclusión. La <x> inicial solo se advierte en el corpus base, mientras que en posición medial reúne un total de 133 testimonios, de los cuales solo el 59,4 % se ajusta a las prescripciones normativas.

Tabla 7. Resultados de la distribución <g>, <j> y <x> + <e>, <i> en el corpus

	G-	Cumple	-G-	Cumple	J-	Cumple	-J-	Cumple	-X-	Cumple	Total
INCOM	94	89 (94,7%)	117	117 (100%)	2	0 (0%)	17	3 (17,6%)	49	44 (89,8%)	279 (90,6%)
CODEA	84	84 (100%)	38	38 (100%)	3	0 (0%)	8	4 (50%)	33	21 (63,6%)	166 (88,5%)
CODEMA	68	68 (100%)	29	29 (100%)	0	0 (0%)	56	15 (26,8%)	51	14 (27,5%)	204 (61,7%)
Total	246	241 (98%)	184	184 (100%)	5	0 (0%)	81	22 (27,2%)	133	79 (59,4%)	649

4.3. La <h>

El apartado dedicado a la grafía <h> en la *Ortografía* (1741) se abre reflexionando sobre si su identidad se corresponde con una letra o con una mera marca de aspiración. Pese a ello, la Academia recomienda su uso tras la <c> en dos casos: por un lado, para marcar la pronunciación del dígrafo <ch> y, por otro, cuando así lo hagan las lenguas clásicas, por ejemplo, en palabras como *charidad*. Asimismo, se prescribe en las voces que en su origen presentan esta letra (como *hoy*, *honra*, etc.) y en aquellas que antes tenían F y se ha suavizado su pronunciación (como *facier*); también se mantiene en la derivación de palabras simples que presentan *h* (*deshacer*), exceptuando las que usan el prefijo *ab-* (*abhorre*>*aborrecer*); la *o* latina que diptonga en *ue* en español requiere de esta grafía para marcar que la *u* es vocal. Por último, la grafía culta *ph* se conserva para facultades y nombres propios, así como tras las voces del griego con R que marcaban aspiración y después de T en palabras comunes (pero solo si están en su origen: *Teresa* y no *Theresa*), aunque se recuerda que su omisión no se considera falta ortográfica.

Así, los datos del corpus muestran una clara tendencia a las normas prescritas por la Academia, que, a diferencia de trabajos como el de Ramírez Luengo (2013: 29), en el que se advierte mayor inestabilidad en el uso de esta grafía, en el presente corpus vemos que se aplican las reglas en distintos contextos. No obstante, se documentan casos que no se ciñen a la norma, aunque son escasos. En el corpus INCOM, en el que se obtiene un porcentaje menor, las alternancias en la presencia o ausencia de la <h> se advierten en el adverbio *así/hasí*; sustantivos como *edad/hedad*; verbos como *avisar/havisar*; *abonar/havonar* y, la tercera persona del imperfecto de indicativo del verbo *ser*, que se recoge en los tres subcorpus tanto con la forma *hera* (5 casos en total) como *era* (6 casos), y que constituye uno de los rasgos lingüísticos más característicos de la escritura procesal (Almeida Cabrejas, 2014b: 10). Ahora bien, como ya ocurría en el apartado anterior (§ 4.2.), estas alternancias se registran principalmente en documentos compuestos por diferentes escribanos, por lo que hallamos pocos testimonios en los que un mismo amanuense emplea las dos formas. Por ejemplo, en los siguientes fragmentos, redactados por el escribano del capitán general de Valencia, podemos ver la alternancia entre *edad/hedad*:

- (7) Y respecto ala Familia de Joseph Castellon de **hedad** de 40 años casado con María Jiménez de 27 y la Familia de Miguel Hernandez de 65 años y Antonia de Bargas su muger de 47: que Vuestra Señoría havona les mandaré poner en libertad. (INCOM, *carta 17*)
- (8) Y prevengo a Vuestra Señoría que en estando juntos los delos referidos parages, me remíta testimonio, con expresión de sus nombres, **edades** , si son solteros ô casados y de donde venían [...]. (INCOM, *carta 22*)

Es interesante destacar, además, que los escribanos del corpus base (INCOM) presentan una clara tendencia a la ultracorrección, puesto que escriben con hache voces como *higualmente* (carta 46), *horden* (carta 48), *hayer* (carta 57), entre otras. Por otro lado, cabe señalar también la notable presencia de las grafías cultas latinizantes en los tres corpus: <th> (N=54), <ph> (N=159) y <ch> (N=29). En los dos primeros casos, <th> y <ph>, se documentan principalmente en nombres propios, como *Mathias* (INCOM, CODEA+2022), *Jospeh* (INCOM, CODEA+2022, CODEMA), etc., aunque hay sustantivos comunes en los que también se recogen (*theniente*, *thenor*, etc.). En cambio, el dígrafo <ch>, si bien se documenta esporádicamente en antropónimos –como *Christoval* (INCOM)–, su uso más habitual se corresponde con nombres comunes del corpus de control, como *chancelados* (CODEMA), entre otros.

Tabla 8. Resultados de la distribución de la <h> en el corpus

	H-	Cumple	-H-	Cumple	Total
INCOM	206	177 (85,9%)	257	257 (100%)	463 (93,7%)
CODEA	201	195 (97%)	353	353 (100%)	554 (98,9%)
CODEMA	244	226 (92,6%)	355	355 (100%)	599 (96,9%)
Total	651	598 (91,9%)	965	965 (100%)	1.616

4.4. La <c>^{+e,i}, <ç>, <z> y <s>, <ss>

La desaparición del rasgo distintivo de sonoridad y el cambio en el lugar de articulación conllevó una modificación en el sistema gráfico. Así pues, para la representación del fonema interdental fricativo sordo (/θ/) la Academia prescribió la grafía <c>, tanto ante las vocales <e, i> como para el plural de las palabras que, por su origen, terminaban en <z> (*pez>peces*). En cambio, considera que la <ç> es superflua, por lo que recomienda el uso de la <z>, siempre que se halle en el étimo de las voces, y en aquellas que, en latín, las vocales <a, o, u> iban precedidas por la consonante <x>.

En nuestro corpus, la escritura de la <c> para representar el fonema (/θ/) no supone grandes diferencias de cómo se hacía hasta el momento, de ahí que los porcentajes tanto al inicio (100 %) como en posición interna (95,5 %) sean representativos de la continuación ortográfica. No obstante, observamos algunos casos esporádicos en los que se confunden las grafías <c> y <s> (INCOM: *acignados*, *pricion*; CODEA+2022: *anciosa*, *ocaciones*; CODEMA: *posecion*, *comicion*, etc.). Conviene señalar que estas formas no responden estrictamente a una vacilación del sistema gráfico, sino que constituyen una alternancia mediatizada por la pronunciación seseante de los autores.

De forma similar, la <z> tiende a emplearse en aquellos casos en los que se debería escribir la <c>. De hecho, los testimonios documentados en el conjunto del corpus al inicio de palabra son aquellos en los que la grafía aparece seguida de <e> o <i>, en el corpus INCOM, por ejemplo, encontramos *zevada* y *zercanía*; en el CODEA+2022, *ziento* y *zedulas*; mientras que, en el CODEMA, *ziudad* y *celebrar*; entre otras. En posición medial, se recoge un número elevado de testimonios en los que ocurre lo mismo, especialmente en el corpus INCOM, que solo alcanza el 22,5 % de casos que siguen las normas ortográficas y que son, mayoritariamente, cuando aparece con las vocales <a, o, u>; en cambio, el 77,5 % responde a usos antiguos de la grafía, recogidos en muchos casos en obras lexicográficas de inicios del siglo dieciocho, así se atestigua en el *Nuevo tesoro lexicográfico de la lengua española* (NTLLE): *quinze* (Stevens, 1706); *vezino*, *onze* (Bluteau, 1721); etc. Además, en los corpus de control (CODEA+2022 y CODEMA), encontramos algunas ocurrencias con <ç>: en el primero de ellos (CODEA+2022) se halla tanto al inicio (*çerrado*, *çircunstançia*, *çelebro*, etc.) como en interior de palabra (*veçino*, *ofiçio*, etc.), mientras que en el segundo (CODEMA), únicamente se registran voces que la contienen en posición intermedia: *notiçia*, *venefiçio*, etc.

Tabla 9. Resultados de la distribución <c>, <ç> y <z> en el corpus

	C-	Cumple	-C-	Cumple	Z-	Cumple	-Z-	Cumple	-Ç-	Cumple	Total
INCOM	32	32 (100%)	303	290 (95,7%)	5	0 (0%)	133	30 (22,5%)	0	0 (0%)	473 (74,4%)
CODEA	35	35 (100%)	192	186 (96,8%)	66	0 (0%)	220	89 (40,5%)	164	0 (0%)	677 (45,8%)
CODEMA	30	30 (100%)	252	237 (94%)	43	3 (6,9%)	216	90 (41,6%)	7	0 (0%)	548 (65,7%)
Total	97	97 (100%)	747	713 (95,5%)	114	66 (5,3%)	569	209 (37,7%)	171	0 (0%)	1.698

Por otro lado, la transcripción del fonema palatal fricativo sordo (/s/) se prescribe por medio de la grafía <s>, por lo que se reserva la <ss> para contextos muy específicos (superlativos, pretéritos de subjuntivo, palabras derivadas y el adverbio *así* para evitar confusiones con el verbo *asir*).

Así pues, cuando aparece al inicio de la palabra, no presenta alternancias ni prácticamente vacilaciones. De hecho, la mayoría pertenece al subcorpus CODEMA, como *sinco*, *sitado*, *celebrar*, que corroboran la

confusión entre los fonemas /s/ y /θ/; aunque en el de INCOM también hay algún caso similar, como *sircunstancia*, *serrado*; y, más aisladamente, en el CODEA+2022: *selebra*, *selosa*. En estos casos, la escritura actúa como un reflejo de la realidad fonético-fonológica del escribiente, en la que la neutralización de la oposición entre sibilantes se traslada al plano escrito.

Por lo que se refiere a la doble ese, únicamente encontramos testimonios esporádicos en el CODEMA (*sseñoría*, *sscivano*). En posición medial, es interesante recuperar también los datos del CODEMA, en los que se advierte nuevamente la misma alternancia mediatizada por la pronunciación (*hise*, *presiso*, *aseyte*, etc.), así como la representación del grupo consonántico latino KS, que, en ocasiones, se escribe con esta grafía (*esepto*). Vemos que, de forma similar, en el corpus INCOM se hallan testimonios redactados en algunas alcaldías de municipios castellanenses (Villarreal y Almazora), en los que se confunde <c> por <s>:

- (9) servidor Joseph Alvaro Alcalde Señor Don Joseph Bermudo Mui señor mio. Habiendo ayer, restituidos, al respecto, el **ofisial**, que con la Partida, salio a **condusir**, los Onze Jitanos, que de esas **carseles**, se ynbiaron â Alicante [...]. (INCOM, *carta* 63)
- (10) [...] vezina de esta Villa de Villareal ha venido a mi casa, y me ha dicho, como Diego Castellón de la Villa de Morella, que al **pareser** se halla en la ciudad de cartagena, la deve su honor con palabra de casamiento que la dio. (INCOM, *carta* 65)

Por último, la duplicación de esta grafía (<ss>) alterna en muchas formas que, o bien no la contienen en su étimo, o bien se ha suprimido, por lo que es posible advertir la convivencia de tradiciones antiguas con las que siguen la nueva propuesta ortográfica: *cassa*, *esscribano* (INCOM); *desseo*, *caussa* (CODEA+2022); *usso*, *perssonas* (CODEMA), entre otras.

En definitiva, los datos del corpus permiten constatar que las sibilantes habían dado ya por esta centuria dos sistemas diferentes (el meridional y el castellano). De hecho, como se analiza en el estudio de González Sopena (2023), vemos el seseo como fenómeno presente en los documentos del CODEMA, representado ortográficamente tanto con la <s> como con la <ss>. Sin embargo, pensamos que los datos obtenidos en el INCOM podrían ser fruto del contacto de lenguas entre el castellano y el valenciano. Este fenómeno de transferencia fonética, que requeriría de un análisis más exhaustivo en este estudio, ya ha sido documentado en trabajos como los de Enrique-Arias (2012).

Tabla 10. Resultados de la distribución <s> y <ss> en el corpus

	S-	Cumple	-S-	Cumple	-SS-	Cumple
INCOM	262	256 (97,7%)	746	644 (86,3%)	34	20 (58,8%)
CODEA	244	241 (98,7%)	722	688 (95,3%)	21	13 (62%)
CODEMA	250	219 (87,6%)	847	672 (79,3)	55	14 (25,5%)

5. Conclusiones

Como señalábamos al inicio de este estudio, los albores del siglo XVIII trajeron consigo la fundación de la Real Academia Española (1713) y su preocupación por promover un uso correcto de la lengua española. Durante un tiempo, estos sucesos se interpretaron erróneamente como el inicio de un periodo de estabilización lingüística. Sin embargo, en las últimas décadas han salido a la luz trabajos basados en lingüística de corpus que han demostrado que esta centuria presentaba una notable variación en distintos niveles de la lengua (Company Company, 2007; Calderón Campos, 2018; González Sopena, 2023; Blas Arroyo y Puertas Ribés, 2024a; etc.). En nuestro caso, el presente estudio pone el foco en la ortografía de mediados del siglo XVIII, a partir de un corpus compilado gracias a textos procedentes de tres bases de datos textuales: el corpus INCOM (corpus base) y los de CODEA+2022 y CODEMA (corpus de control), con el fin de examinar y comparar las tendencias (orto)gráficas en tres zonas geográficas de la península, así como ahondar, al mismo tiempo, en la formación escrituraria de los escribanos del corpus INCOM.

A la luz de los datos obtenidos en el conjunto del corpus, advertimos que la tendencia general de los escribanos parece mostrar una convergencia con los usos gráficos que la Academia estaba sistematizando en la época. Más que un seguimiento consciente de los preceptos académicos, los resultados sugieren que la institución codificó y dio validez a las soluciones gráficas ya generalizadas en la práctica de la escritura culta. No obstante, al atender a cada subcorpus, se han hallado formas que demuestran vacilaciones ortográficas en un mismo amanuense (por ejemplo, las alternancias registradas en el corpus INCOM: *edad/hedad* en las cartas 17 y 22; o *reciban/recivan* en los textos 40 y 42), así como la convivencia de tradiciones gráficas vinculadas a épocas anteriores (INCOM: *vezino*, *quinze*, etc.), fruto posiblemente de una norma aún en proceso de perfeccionamiento. Se trata, pues, de un periodo de estabilización todavía en marcha, que culminará con la oficialización de las normas ortográficas en el siglo XIX (Martínez Alcalde, 2010: 61).

Ahora bien, esto no niega la existencia de vacilaciones y dudas, comprensibles en un periodo marcado por los cambios constantes en materia ortográfica y en pleno proceso de estandarización, aunque, a nuestro parecer, está lejos de constituir una anarquía ortográfica, como apuntaban autores como Lapesa (1985),

Zamora Vicente (1999), Arnal Purroy (2000), etc. Por tanto, los datos recogidos para este trabajo coinciden con lo expuesto por Ramírez Luengo (2013): más que un caos en el uso de las grafías, podemos advertir ciertas tendencias de uso. Las alternancias documentadas en este corpus, en muchos casos, se advierten en palabras con el mismo lexema e, incluso, en textos del mismo escribano, por lo que nos encontramos ante un periodo de *estandarización en proceso* también en el reino de Valencia, que no había sido estudiado hasta el presente.

En lo que respecta a la formación escrituraria de nuestros escribanos, los resultados del corpus INCOM sugieren que podríamos encontrarnos ante profesionales con un dominio medio-alto de la lengua, y, en gran medida, conocedores de los usos gráficos que imperaban en ese momento. Así lo demuestra la presencia de las grafías cultas latinizantes <ph>, <ch> y <th>, pese a que ya no eran obligatorias; del mismo modo, se advierte una clara preferencia por el empleo de la en posición inicial, rasgo propio de los eruditos. Asimismo, se documenta la <h> en formas del verbo *ser*, característica de la escritura procesal de la época (Almeida Cabrejas, 2014b) y en otras que demuestran ultracorrecciones en caso de duda. De forma similar, el corpus de control de la zona centro (CODEA+2022 e INCOM) predomina la tendencia a seguir las soluciones académicas en la mayoría de los casos. En cambio, en los textos de la zona meridional (CODEMA) se observa que la estandarización gráfica convive con más frecuencia con la realidad fonético-fonológica (como el seseo), de modo que se genera una alternancia mediatizada por la pronunciación que se refleja en el plano escrito.

De momento, los datos analizados nos permiten vislumbrar que los escribanos pertenecientes al corpus INCOM eran conocedores de las costumbres escriturarias de la época y trataban de homogeneizar su escritura. No obstante, las vacilaciones encontradas van en consonancia con las dificultades que, posiblemente, encontró la propia Academia a la hora de diseñar una norma que iría cambiando en sus obras, por lo que no podemos pretender que los amanuenses adquiriesen esa doctrina de forma automática (Almeida Cabrejas, 2014b). Sea como fuere, estos testimonios nos permiten reforzar la idea de que no se trata tanto de inconsistencias o incoherencias en los usos gráficos, sino ciertas reiteraciones y regularidades dentro del (sub) corpus (Ramírez Luengo, 2015), aunque, habría que examinar, en un futuro, si esas tendencias consiguen mayor homogeneización en la segunda mitad del XVIII y en las siguientes centurias. Del mismo modo, para comprender de forma completa si la formación de los escribanos es como apuntábamos más arriba, es fundamental también indagar en otros niveles lingüísticos que nos ofrezcan una visión global de sus costumbres escriturarias.

Referencias bibliográficas

- Almeida Cabrejas, Belén (2014a): «Escuchar los textos: el análisis de los textos en el estudio de la fonética y fonología de épocas pasadas», *Linred: Lingüística en la Red*, 11, pp. 1-18.
- Almeida Cabrejas, Belén (2014b): «*Scriptores* con bajo y medio nivel socioeducacional en documentos del siglo XIX del Archivo Municipal de Alcalá de Henares: acercamiento a sus usos gráficos», en Rocío Díaz Moreno y Belén Almeida Cabrejas (eds.), *Estudios sobre la historia de usos gráficos en español*. Axac, pp. 167-210.
- Almeida Cabrejas, Belén, M.^a del Rocío Díaz Moreno y M.^a del Carmen Fernández López (2018): «Los escribanos de 'CODEA'», en M.^a Luisa Arnal Purroy, Rosa M.^a Castañer Martín, José M.^a Enguita Utrilla, Vicente Lagüéns García, M.^a Antonia Martín Zorraquino (coords.), *Actas del X Congreso Internacional de Historia de la Lengua Española*. Diputación Provincial de Zaragoza/Institución 'Fernando el Católico', 2, pp. 1479-1496.
- Álvarez de Miranda, Pedro (2004): «El léxico español desde el siglo XVIII hasta hoy», en Rafael Cano Aguilar (coord.), *Historia de la lengua española*. Ariel, pp. 1037-1064.
- Arnal Purroy, M.^a Luisa (2000): «Cambios grafemáticos en textos notariales aragoneses del siglo XVIII, a la luz de las reformas ortográficas de la RAE», *Archivo de Filología Aragonesa*, 56, pp. 117-136.
- Blas Arroyo, José Luis y Elia Puertas Ribés (2024a): «El arduo despertar de un cambio lingüístico en la distancia comunicativa: paralelismos y divergencias con textos de la inmediatez comunicativa en las relativas oblicuas del primer español moderno», *Philologica canariensis*, 30, pp. 1-34. <https://doi.org/10.20420/Phil.Can.2024.665>
- Blas Arroyo, José Luis y Elia Puertas Ribés (2024b): «INCOM: Un corpus de inmediatez comunicativa para el estudio sociolingüístico del español en su historia», *Cultura, Lenguaje y Representación*, 33, pp. 7-29. <https://doi.org/10.6035/clr.7200>
- Calderón Campos, Miguel (2012): «Variaciones gráficas y fonéticas del español del siglo XVIII en tres corpus hispánicos», en M.^a T. García-Godoy (ed.), *El español del siglo XVIII. Cambios diacrónicos en el primer español moderno*. Peter Lang, pp. 221-254.
- Calderón Campos, Miguel (2018): «Las declaraciones de esencia' del siglo XVIII: un tipo textual para el estudio de la terminología anatómica», *Dynamis: Acta hispanica ad medicinae scientiarumque historiam illustrandam*, 38(2), pp. 427-452.
- Cano Aguilar, Rafael (2018 [2004]): *Historia de la lengua española*. Ariel.
- Company Company, Concepción (2007): «El siglo XVIII en los fondos documentales de México: contribución a la dialectología histórica del español», *Lingüística*, 19, pp. 7-29.
- CODEA+2022 = Grupo de Investigación Textos para la Historia del Español [GITHE]: *CODEA+ 2022 (Corpus de documentos españoles anteriores a 1900)* [en línea], doi: <https://doi.org/10.37536/CODEA.2015> [consultado en: 10/09/2025]

- CODEMA = Grupo de investigación ARINTA, corpus CODEMA (Corpus Diacrónico de Documentación Malagueña). Disponible en: <https://corpora.uah.es/charta/index.php?action=corpus> [consultado en: 20/09/2025].
- Díaz Moreno, M.^a del Rocío (2014): «Una aproximación al uso de abreviaturas en documentos de Bilbao del siglo XVIII», Belén Almeida Cabrejas y M.^a del Rocío Díaz Moreno (eds.), *Estudios sobre historia de los usos gráficos en español*. Axac, pp. 143-153.
- Frago, Juan Antonio (2010): *El español de América en la Independencia*. Taurus.
- Freixas Alás, Margarita (2016): «Orígenes de la Ortografía de la RAE: primeras aportaciones de Bartolomé Alcázar (1715) y de Adrián Conik (h. 1716)», *Revista de Filología Española (RFE)*, 96(1), pp. 113-148.
- García-Godoy, M.^a Teresa (ed.) (2012a): «Introducción», en M.^a Teresa García-Godoy (ed.), *El español del siglo XVIII. Cambios diacrónicos en el primer español moderno*. Peter Lang, pp. 9-18.
- García-Godoy, M.^a Teresa (ed.) (2012b): *El español del siglo XVIII. Cambios diacrónicos en el primer español moderno*. Peter Lang.
- Gómez Asencio, José Jesús (2008): «El trabajo de la Real Academia Española en el siglo XVIII (y después)», *Península. Revista de Estudios Ibéricos*, 5, pp. 31-53.
- Gómez Seibane, Sara y José Luís Ramírez Luengo (2007): *El castellano de Bilbao en el siglo XVIII: materiales para su estudio*. Universidad de Deusto.
- González Sopeña, Inmaculada (2023): «Confusión de sibilantes y neutralización -r/-l en los inventarios de bienes andaluces del siglo XVIII», *Études Romanes de Brno*, 44, pp. 141-162.
- INCOM = Grupo de investigación *Sociolingüística* (UJI), corpus INCOM (=Corpus de textos de inmediatez comunicativa. Disponible en <https://auth.sketchengine.eu/> [última consulta 24 de agosto de 2025].
- Lapesa Melgar, Rafael (1985): *Historia de la lengua española*. Gredos.
- Martínez Alcalde, M.^a José (1990): «Las ortografías de Mayans y Bordazar: el fondo de una polémica», *Revista de Filología Española*, 70, pp. 143-159.
- Martínez Alcalde, M.^a José (2007): «Codificación y norma del español en el siglo XVIII: su enseñanza desde una perspectiva historiográfica», *Dieciocho: Hispanic enlightenment*, 30(1), pp. 119-130.
- Martínez Alcalde, M.^a José (2010): *La fijación ortográfica del español: norma y argumento historiográfico*. Peter Lang, Colección Fondo Hispánico de Lingüística y Filología.
- Mondéjar, José (2007): «De la antigüedad y de la naturaleza de las hablas andaluzas», en Juan Antonio Moya Corral y Marcin Ryszard Sosinski (coords.), *Las hablas andaluzas y la enseñanza de lengua: actas de las XII Jornadas sobre la Enseñanza de Lengua Española*. Universidad de Granada, pp. 13-24.
- NLTLE = Real Academia Española. *Nuevo tesoro lexicográfico de la lengua española* [en línea]. <http://buscon.rae.es/nltle/SrvltGUILoginNltle>.
- Puertas Ribés, Elia (2026): *Textos para la historia del español XVII Archivo Municipal de Castellón de la Plana: Selección, edición y análisis lingüístico de cartas administrativas (Siglo XVIII)*. Editorial Universidad de Alcalá.
- Ramírez Luengo, José Luís (2006): «Materiales para la historia de la lengua española en Centroamérica: algunos documentos dieciochescos (1703-1758)», *Ámbitos: revista de estudios de ciencias sociales y humanidades*, 16, pp. 119-135.
- Ramírez Luengo, José Luís (2012): «Notas sobre las tendencias gráficas del español colombiano en la época de las Independencias (1830)», en José Luís Ramírez Luengo (coord.), *Por sendas ignoradas. Estudios sobre el español del siglo XIX*. Axac, pp. 167-182.
- Ramírez Luengo, José Luís (2013): *Una descripción del español de mediados del siglo XVIII. Edición y estudio de las cartas de M. Martierena del Barrando (1757-63)*. Axac.
- Ramírez Luengo, José Luís (2015): «La muerte de una letra: empleo y decadencia de la <ç> en la escritura del español (siglos XVI-XVIII)», *Diálogo de la Lengua*, 7, pp. 22-35.
- Ramírez Luengo, José Luís (2020): «Una propuesta metodológica para el estudio histórico de la ortografía: el ejemplo de la <tt> en el español moderno (1550-1800)», *Philologica canariensis*, 26, pp. 35-55.
- Real Academia Española (1726 [1963]): «Discurso Proemial de la Orthographía de la Lengua Castellana», En *Diccionario de Autoridades*, vol. I (LXI-LXXXIV). Edición facsímil. Gredos.
- Real Academia Española (1741): *Orthografía de la lengua castellana*. Imprenta de la Real Academia Española.
- Rosenblat, Ángel (1951): «Las ideas ortográficas de Bello», *Obras completas de Andrés Bello. Estudios gramaticales*. Ministerio de Educación, 5, pp. 9-138.
- Sánchez-Prieto Borja, Pedro (2011): *La edición de textos españoles medievales y clásicos. Criterios de presentación gráfica*. Cilengua.
- Zamora Vicente, Alonso (1999): *La Real Academia Española*. Espasa.