

Visualización cartográfica por el método de la vecindad más próxima
y cuantificación de difusión léxica de cambios ortográficos
– Interpretaciones geográficas e históricas
de las grafías dobles iniciales <ss> del español antiguo

Hiroto Ueda (Universidad de Tokio)
Leyre Martín Aizpuru (Universidad de Sevilla)

§1. Los aspectos lingüísticos del español en su historia inicial son posibles de observar en los documentos notariales de siglos IX a XIII, emitidos en el norte de la península ibérica (Fig. 1):



Fig. 1

Los corpus lingüísticos históricos CORHEN («Corpus Histórico del Español Norteño», María Jesús Torrens Álvarez, CSIC) y CODEA «Corpus de Documentos Españoles Anteriores a 1800», Pedro Sánchez-Prieto Borja y GITHE) recogen los documentos concernientes y demuestran el lugar y la fecha de la emisión. Nosotros nos encargamos de construir el sistema de búsqueda, tratamientos estadísticos, visualización gráfica y cartográfica (LYNEAL «Letras y Números en Análisis Lingüísticos»: Fig. 2):

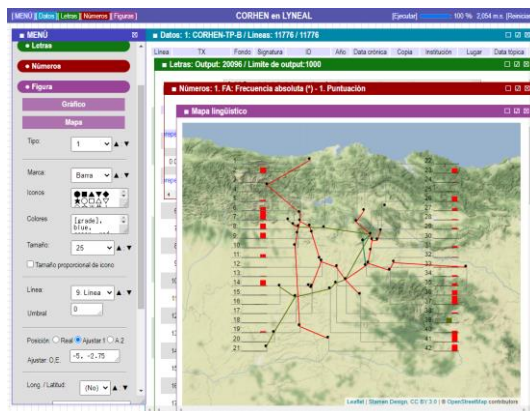


Fig. 2

§2. En esta ocasión, tratamos las grafías dobles iniciales de palabra #ss, que se observan en el español antiguo (Fig. 3):



Fig. 3
ssellada con su ssello (CODEA-8, Valladolid, 1283)

En los estudios anteriores, se han presentado varias observaciones al respecto: la distinción semántica entre #ss y #s: #sson (verbo *ser*) : #son (sustantivo, ‘sonido’) que hay en *Dança general de la muerte*, Sevilla, 1520 (Morreale, 1996: 158, 173); la relación con la consonante interior + <ss>, por ejemplo, *Perssia, penssar* (Sánchez-Prieto Borja, 1998: 138; Martín Aizpuru, 2019: 353-354); los caracteres de las letras cursivas (Sánchez-Prieto Borja, 2004: 442), los usos preferentes de determinados escribanos (Martín Aizpuru, 2019: 350-351). En adición, en esta ocasión proponemos interpretar la historia y geografía de la doble grafía inicial #ss, en contraste con la doble grafía interior -ss-.

El mapa de la Fig. 4 muestra la distribución de frecuencias relativas de #ss / (#ss + #s) y el mapa de la Fig. 5, la distribución de frecuencias relativas de -ss- / (-ss- + -s-):

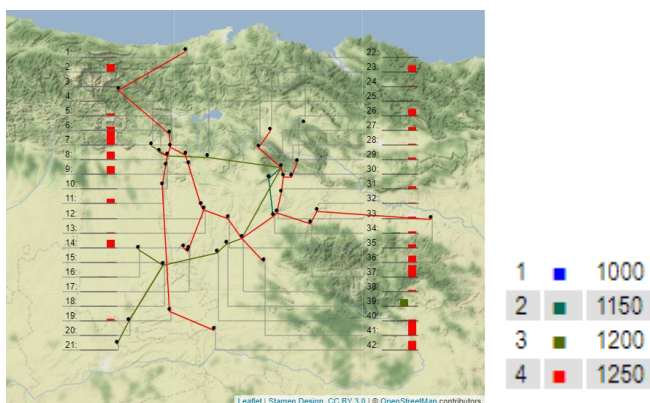


Fig. 4. #ss / (#ss + #s)

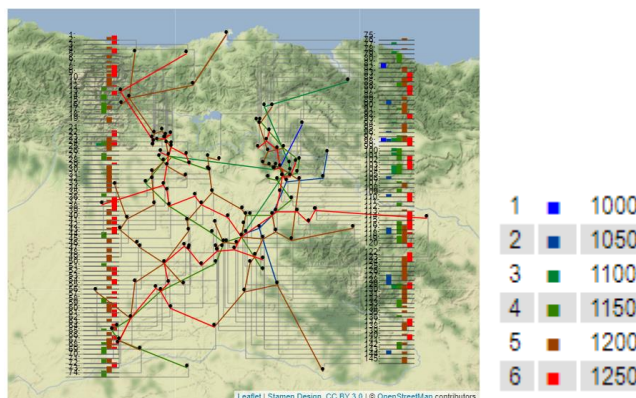


Fig. 5. -ss- / (-ss- + -s-)

Al comparar los dos mapas, nos damos cuenta de que la distribución de la #ss inicial es amplia y esporádica (Fig. 4), mientras que la de -ss- interior es amplia también,

pero en contraste, más frecuente que la anterior (Fig. 5). De las líneas que combinan los puntos hablaremos más adelante (§3).

Trask (2000, *s.v. four-part analogy*) explica la analogía de cuatro partes y toma como ejemplo el dialecto inglés del este de Estados Unidos: *dive* ‘bucear’ (presente) - *dove* (pasado) en lugar de *dived* (pasado, forma normal) bajo la influencia de *drive* ‘conducir’ (presente) - *drove* (pasado) y presenta la ecuación (1), la que sustituimos por la (2):

$$(1) A : B :: C : ? \quad \text{drive} : \text{drove} :: \text{dive} : ?, \quad ? = \text{dove}$$

$$(2) X_i : X_j :: Y_i : ? \quad \text{drive} : \text{drove} :: \text{dive} : ?, \quad ? = Y_j = \text{dove}$$

En la ecuación (2), X y Y representan verbos *drive* y *dive*, e *i* y *j*, la forma de presente y la de pasado, respectivamente. De manera igual que esta relación analógica, al aplicar X y Y a la posición interior e inicial e *i* y *j* a la forma fonológica y a la ortográfica, respectivamente, podemos establecer la siguiente relación paralela:

$$-/s/- : -ss- :: \#/s/ : ?, \quad ? = \#ss.$$

En la misma relación, lo normal es que la pareja izquierda, *-/s/- : -ss-*, que es la causa, sea suficientemente más preponderante que la pareja derecha, *\#s : \#ss* (efecto), lo que acabamos de confirmar en Fig. 4 y 5.

§3. En los dos mapas anteriores (Fig. 4, 5), hemos combinado los puntos reactivos por la línea de la vecindad más próxima. El programa de la visualización cartográfica instalado en el sistema LYNEAL combina cada punto relevante con el otro que se encuentra más cerca geográficamente. Para ello, hemos utilizado el método de la vecindad más próxima del análisis de conglomerado (*cluster*). Como los fenómenos se expanden en la cercanía, la línea de vecindad más próxima es realista. La red que combina los puntos reactivos suele ser complicada (Fig. 6), mientras que la línea de vecindad más próxima es sencilla (Fig. 4). En las líneas de irradiación cuyo centro es el punto medio de longitud y latitud de todos los puntos reactivos, es difícil de encontrar el sentido lingüístico del mismo centro. El método del casco convexo (*convex hull*) que envuelve los puntos más exteriores puede incluir dentro los puntos no reactivos, por lo que generaliza excesivamente el área (véase la Fig. 8, donde los puntos 4, 10, 30, 39, 40 están incluidos en el área del año 1250 indebidamente).

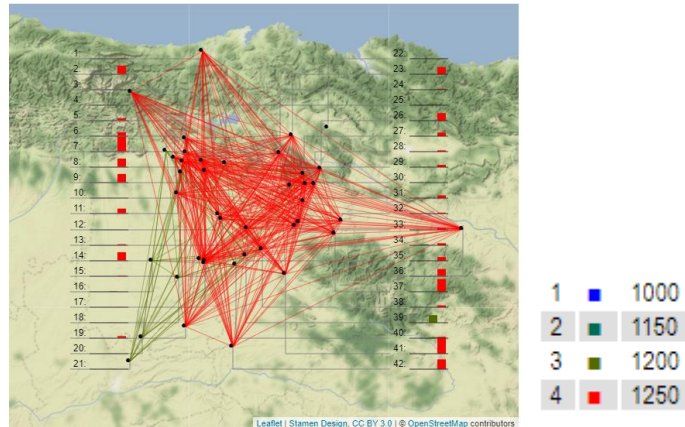


Fig. 6. #ss / (#ss + #s)

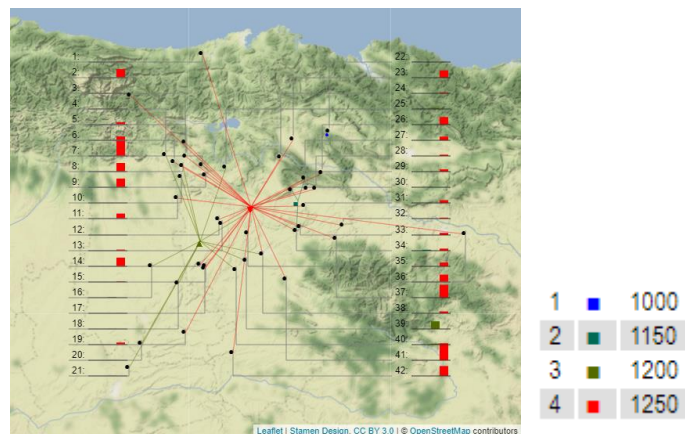


Fig. 7. #ss / (#ss + #s)

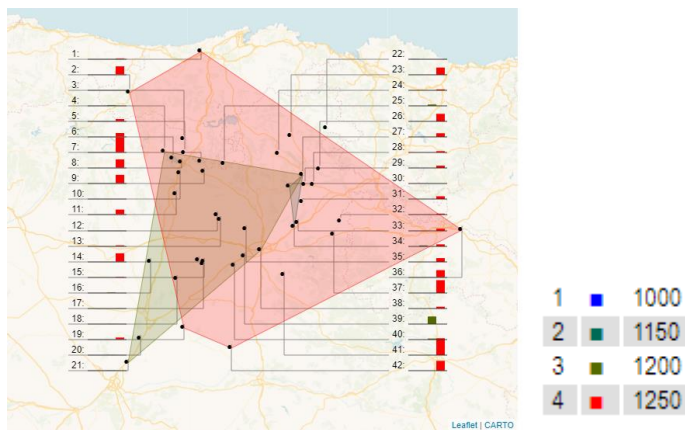


Fig. 8. #ss / (#ss + #s)

§4. En los siglos XII y XIII, la doble -ss- en la posición intervocálica representaba la /s/ sorda (*passar* /pa'sar/) en oposición a -s- en la misma posición que representaba /z/ sonora (*casa* /'kaza/); en el siglo XIV ambos confluyeron en la misma grafía y fonema -s- /s/ (*pasar* /pa'sar/, *casa* /'kasa/). A continuación, trataremos la doble <ss> desaparecida tanto en la posición interior como inicial.

La escuela lingüística del siglo XIX, llamada de los neogramáticos (*Junggrammatiker*), consideraba que en principio un cambio fonológico se aplica a todas las palabras sin excepción (Osthoff and Brugmann, 1878, cf. Lehmann 1967: 204). Por

otra parte, en el siglo XX se ha propuesto la idea de ‘difusión léxica’, donde el progreso del cambio fonológico varía según las palabras (Wang, 1969).

Nosotros, en vez de admitir la dicotomía excluyente de ambas posiciones, proponemos una interpretación gradual. En primer lugar, representamos el estado del cambio fonológico $X \rightarrow Y$ sin excepción en la Fig. 9 (f: frecuencia; t: tiempo). En realidad, suele haber una extensión de X después del cambio (p: palabras que se mantienen sin cambio en el tiempo posterior) y unos precursores de Y antes del cambio (a: palabras que se presentan con cambio fonológico antes del cambio mayoritario, en el tiempo anterior):



Fig. 9. $X \rightarrow Y$ / Fig. 10. a: excepción anterior / p: excepción posterior

Por lo tanto, definimos el grado de difusión léxica (D) de la siguiente manera:

$$D = (a + p) / (X + Y)$$

El grado D se aproxima al mínimo ($\rightarrow 0$) cuando $a = p \rightarrow 0$ (Fig 11); y se aproxima al máximo ($\rightarrow 1$) cuando $a + p \rightarrow X + Y$, es decir, X y Y coexisten sin cambio alguno de $X \rightarrow Y$ (Fig 12):



Fig. 11. Aproximación al mínimo: $D = 0$ / Fig. 12. Aproximación al máximo: $D = 1$

Al aplicar el grado de difusión léxica (D) al cambio ortográfico $-ss- \rightarrow -s-$ en posición interior de palabra del corpus CODEA, obtenemos la Fig. 13.

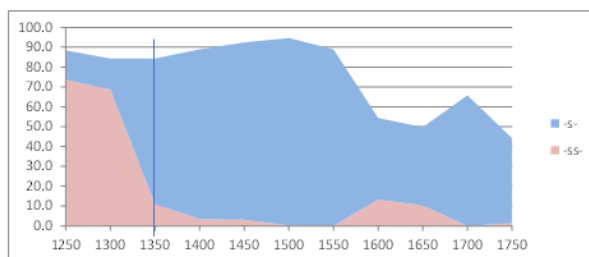


Fig. 13. $-ss- \rightarrow -s-$, $D = .087$

Sin embargo, el grado de difusión léxica (D) obtenida con las frecuencias

absolutas no es comparable con otro que presente frecuencias diferentes. Por ello, utilizamos las frecuencias relativas en lugar de las absolutas, donde la suma de las frecuencias relativas es siempre 1, de modo que la totalidad de suma de frecuencias relativas de todos los años es igual al número de años (N). De esta manera la ecuación del grado de difusión de frecuencias relativas se vuelve sencilla:

$$D' = (a + p) / N$$

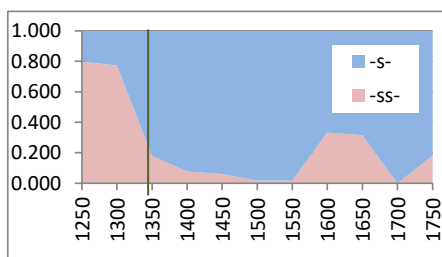


Fig. 14. -ss- → -s-, D' = 0.146

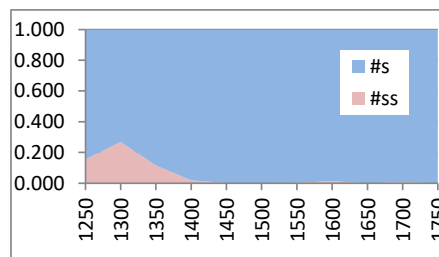


Fig. 15. #ss → #s, D' = 0.053

En la Fig. 15, #s es preponderante a través de los siglos, de modo que calculamos D por #ss / (#ss + #s) al considerar que hubo el cambio #ss → #s en 1250, con el residuo posterior (p) de #ss.

En las Figuras 14 y 15, -ss- y #ss coinciden parcialmente desde 1250 hasta 1500. Sin embargo, la frecuencia relativa de -ss- es preponderante en 1250 y 1300 (Fig. 14), mientras que la de #ss nunca ha llegado a la supremacía relativa (Fig. 15). Por esta razón, -ss- debe ser causa de #ss, y no a la inversa.

§5. Finalmente, hemos averiguado la presencia o ausencia de -ss- y #ss (0 : 1~) en 624 documentos de los años 1050-1250 de CODEA y preparado la siguiente tabla:

Documentos	#ss: 0	#ss: 1~	Suma	Valor esperado	#ss: 0	#ss: 1~	Suma
-ss- : 0	123	10	133	-ss- : 0	93.1	39.9	133.0
-ss-: 1~	314	177	491	-ss-: 1~	343.9	147.1	491.0
Suma	437	187	624	Suma	437.0	187.0	624.0

Tabla 1. Relación entre -ss- y #ss en los documentos de 1050 a 1250 de CODEA

En la tabla observamos las frecuencias significativamente altas de -ss-:0 + #s:0; y -ss-:1~ + #ss:1~ (123; 177: $\chi^2 = 40.6$, $p < 0.01$). Por consiguiente, podemos afirmar que si -ss- está ausente en un documento, #ss tiende a ausentarse; y si -ss- está presente, también #ss tiende a estarlo en el mismo documento.

Referencia citada:

Lehmann, W. P. (1967). *A Reader in Nineteenth Century Historical Indo-European*

- Linguistics*. Bloomington: Indiana University Press.
- Martín Aizpuru, L. (2019), *La «norma lingüística» de la cancillería real castellana (1230-1312): Fernando III, Alfonso X, Sancho IV y Fernando IV*, Tesis doctoral, Universidad de Salamanca.
- Morreale, M. (1996), «La Dança general de la muerte (II)», *Revista de Literatura Medieval*, VIII, pp. 111-177.
- Sánchez-Prieto Borja, P. (1998), *Cómo editar los textos medievales. Criterios para su presentación gráfica*. Madrid. Arco / Libros.
- _____. (2004), «La normalización del castellano escrito en el siglo XIII. Los caracteres de la lengua: Grafías y fonemas» en Rafael Cano (coord.) *Historia de la Lengua Española*, pp. 423-448. Barcelona. Ariel.
- Trask, R. L. (2000), *The Dictionary of Historical and Comparative Linguistics*, Edinburgh University Press.
- Wang, W. S-Y. (1969), «Competing Changes as a Cause of Residue», *Language*, vol. 45, pp.9-25.

saikinrinho-sp.docx 2020/8/31