

*Seminario del proyecto UAM-Santander*

*Universidad de Estudios Extranjeros de Tokio, 21 mayo, 2018*

# **Origen de <del>, <al> y <el> del español**

## **Distintas frecuencias en el análisis de Pareto**

**Hiroto Ueda (Universidad de Tokio)**

**Antonio Moreno Sandoval (Universidad Autónoma de Madrid)**

# 1. Introducción

Nuestro propósito en esta presentación es doble.

En primer lugar, explicaremos la utilidad del Análisis de Pareto y de la Frecuencia Probabilística.

(Ueda / Moreno Sandoval, 2018)

Para demostrarla intentamos aplicar los métodos estadísticos a la cuestión filológica de la unión y separación de preposición y artículo definido, con especial atención a <del>, <al> y <el>.

Benito Pérez Galdós (1843-1920), Proyecto Gutenberg:

<http://www.gutenberg.org/>

PRESEEA («*Corpus del Proyecto para el estudio sociolingüístico del español de España y de América.*» dir. Francisco Moreno):

<http://preseea.linguas.net/>

CODEA («*Corpus de Documentos Españoles Anteriores a 1800*», dir. Pedro Sánchez-Prieto): <http://corpuscodea.es/>

CORHEN («*Corpus Histórico del Español Norteño*», dir. María Jesús Torrens Álvarez): <http://corhen.es/>

LYNEAL («*Letras y Números en Análisis Lingüísticos*», dir. Hiroto Ueda / Antonio Moreno Sandoval):  
<http://shimoda.llf.uam.es/ueda/lyneal/>

## 2. Hipótesis

### *Hipótesis-1:*

No ha habido la unión de

\*<de el> → <del>, \*<a el> → <al>

(Bybee 2007, Elvira 2009, Ueda 2011, etc.)

en la cronología general de la lengua española.

En su lugar pensamos en los procesos:

DE ILLŌ → <de il(l)o> → <del(o)> → <del>

AD ILLU(M) → <a il(l)o> → <al(o)> → <al>

## *Hipótesis-2:*

La no separación de <del> y <al> no se debe a

\*su frecuencia de uso,

sino a

su rasgo fonético de la vocal final de la preposición

y

su alta utilidad lingüística realizada en forma breve y la amplitud semántica.

### ***Hipótesis-3:***

Entre <del>, <al> y <el> se mantiene una correlación fuerte en su función y forma, lo que ha producido la forma apocopada peculiar común entre estas tres formas unisilábicas.

del(o), al(o); el(o) → del, al; el (desde el principio)

d(e)los, alos; (e)los → delos, alos; los → de los, a los; los

d(e)la, ala; (e)la → dela, ala; la → de la, a la; la

d(e)las, alas; (e)las → delas, alas; las → de las, a las; las

?? [neutro] (?? posterior)

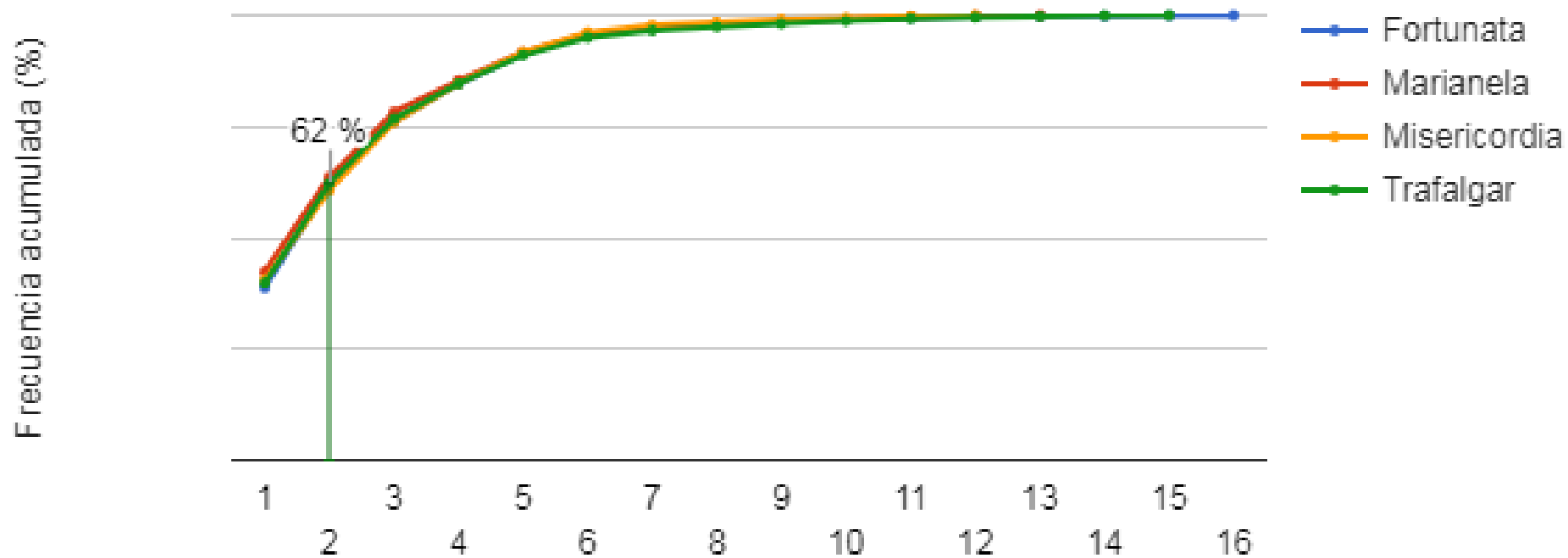
d(e)lo, alo; (e)lo → delo, alo; lo → de lo, a lo; lo

# 3. Unión y separación

## 3. 1. Sincronía

Pareto. FP / 10 mil palabras	Fortunata	Marianela	Misericordia	Trafalgar
IP: Índice de Pareto	0.785	0.833	0.836	0.798
M: Punto de mitad	2	2	2	2
N: Número de datos	16	13	12	15
M / N (%)	13%	15%	17%	13%
A: Acumulación en M.	816	786	836	870
T: Frecuencia total	1 319	1 235	1 379	1 399
A / T (%)	62%	64%	61%	62%

**Tabla 1. Análisis de Pareto. Preposición en cuatro obras de Pérez Galdós**



**Fig. 1. Análisis de Pareto.**

**Preposiciones en cuatro obras de Pérez Galdós**



Rango. FP	Fortunata	Marianela	Misericordia	Trafalgar
1	<b>de : 512 (39 %)</b>	<b>de : 523 (42 %)</b>	<b>de : 559 (41 %)</b>	<b>de : 557 (40 %)</b>
2	<b>a : 304 (23 %)</b>	<b>a : 263 (21 %)</b>	<b>a : 277 (20 %)</b>	<b>a : 313 (22 %)</b>
3	en : 187 (14 %)	en : 180 (15 %)	en : 211 (15 %)	en : 203 (15 %)
4	con : 113 (9 %)	con : 88 (7 %)	con : 121 (9 %)	con : 111 (8 %)
5	por : 88 (7 %)	por : 79 (6 %)	por : 97 (7 %)	por : 89 (6 %)
6	para : 55 (4 %)	para : 50 (4 %)	para : 61 (4 %)	para : 56 (4 %)
7	sin : 20 (2 %)	sin : 16 (1 %)	sin : 22 (2 %)	sin : 22 (2 %)
8	sobre : 11 (1 %)	sobre : 11 (1 %)	hasta : 10 (1 %)	sobre : 11 (1 %)
9	hasta : 9 (1 %)	entre : 7 (1 %)	entre : 8 (1 %)	hasta : 10 (1 %)
10	entre : 7 (1 %)	hasta : 7 (1 %)	desde : 5 (0 %)	desde : 9 (1 %)
11	desde : 5 (0 %)	desde : 6 (0 %)	sobre : 5 (0 %)	entre : 7 (1 %)
12	hacia : 3 (0 %)	hacia : 4 (0 %)	hacia : 3 (0 %)	contra : 4 (0 %)

13	contra : 2 (0 %)	contra : 1 (0 %)	-	ante : 3 (0 %)
14	ante : 1 (0 %)	-	-	hacia : 3 (0 %)
15	bajo : 1 (0 %)	-	-	bajo : 1 (0 %)
16	tras : 1 (0 %)	-	-	-

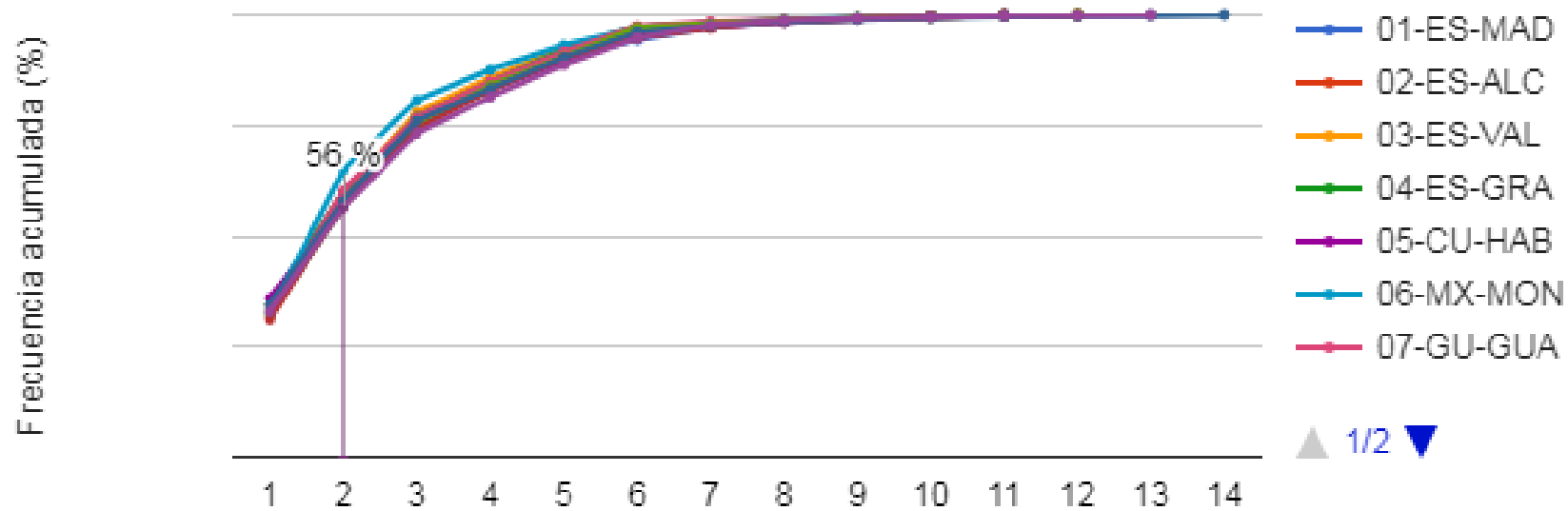
**Tabla 2. Tabla de Pareto.**

**Preposiciones en cuatro obras de Pérez Galdós**

Pareto	ES- MAD	ES- ALC	ES- VAL	ES- GRA	CU- HAB	MX- MON	GU- GUA	CO- MED	PE- LIM	CH- STG	UR- MTV
IP	0.819	0.83	0.832	0.832	0.832	0.852	0.83	0.822	0.825	0.802	0.802
Mitad:	2	2	2	2	2	2	2	2	2	2	2
Número:	12	12	12	12	11	11	12	12	11	14	13
M / N (%)	17%	17%	17%	17%	18%	18%	17%	17%	18%	14%	15%
Acumul.:	461	469	456	453	540	542	581	543	451	521	492
Total:	793	786	768	759	944	835	964	957	783	899	872
A / T (%)	58%	60%	59%	60%	57%	65%	60%	57%	58%	58%	56%

**Tabla 3. Análisis de Pareto.**

**Preposiciones en las ciudades hispanohablantes. PRESEEA**



**Fig. 2. Análisis de Pareto.**

**Preposiciones en las ciudades hispanohablantes. PRESEEA**

R	ES- MAD	ES- ALC	ES- VAL	ES- GRA	CU- HAB	MX- MON	GU- GUA	CO- MED	PE- LIM	CH- STG	UR- MTV
1	de : 266 (34 %)	de : 272 (35 %)	de : 253 (33 %)	de : 261 (34 %)	de : 339 (36 %)	a : 273 (33 %)	de : 307 (32 %)	de : 307 (32 %)	de : 244 (31 %)	de : 310 (34 %)	de : 287 (33 %)
2	a : 195 (25 %)	a : 197 (25 %)	a : 203 (26 %)	a : 192 (25 %)	a : 201 (21 %)	de : 269 (32 %)	a : 274 (28 %)	a : 236 (25 %)	a : 207 (26 %)	a : 211 (23 %)	a : 205 (24 %)
3	en : 133 (17 %)	en : 132 (17 %)	en : 143 (19 %)	en : 130 (17 %)	en : 179 (19 %)	en : 131 (16 %)	en : 159 (16 %)	en : 173 (18 %)	en : 130 (17 %)	en : 162 (18 %)	en : 147 (17 %)
4	por : 68 (9 %)	por : 70 (9 %)	por : 59 (8 %)	por : 60 (8 %)	por : 64 (7 %)	por : 58 (7 %)	con : 78 (8 %)	por : 87 (9 %)	por : 68 (9 %)	con : 65 (7 %)	por : 71 (8 %)
5	con : 54 (7 %)	con : 51 (6 %)	con : 52 (7 %)	con : 56 (7 %)	con : 63 (7 %)	con : 46 (6 %)	por : 65 (7 %)	con : 63 (7 %)	con : 57 (7 %)	por : 64 (7 %)	con : 63 (7 %)
6	para : 32 (4 %)	para : 35 (4 %)	para : 35 (5 %)	para : 38 (5 %)	para : 58 (6 %)	para : 33 (4 %)	para : 56 (6 %)	para : 62 (6 %)	para : 36 (5 %)	para : 51 (6 %)	para : 53 (6 %)
7	entre : 23 (3 %)	hasta : 8 (1 %)	hasta : 6 (1 %)	hasta : 7 (1 %)	entre : 13 (1 %)	entre : 9 (1 %)	hasta : 10 (1 %)	hasta : 9 (1 %)	entre : 17 (2 %)	entre : 13 (1 %)	entre : 25 (3 %)
8	desde : 8 (1 %)	desde : 7 (1 %)	sobre : 6 (1 %)	desde : 5 (1 %)	hasta : 10 (1 %)	hasta : 7 (1 %)	desde : 4 (0 %)	entre : 8 (1 %)	hasta : 13 (2 %)	hasta : 9 (1 %)	hasta : 10 (1 %)
9	hasta : 6 (1 %)	sin : 6 (1 %)	desde : 4 (1 %)	sin : 4 (1 %)	desde : 8 (1 %)	desde : 4 (0 %)	sin : 4 (0 %)	desde : 5 (1 %)	desde : 5 (1 %)	sin : 4 (0 %)	desde : 4 (0 %)

10	sobre : 4 (1 %)	sobre : 4 (1 %)	entre : 3 (0 %)	sobre : 3 (0 %)	sobre : 5 (1 %)	sin : 3 (0 %)	entre : 3 (0 %)	sin : 4 (0 %)	sin : 3 (0 %)	sobre : 4 (0 %)	sin : 3 (0 %)
11	sin : 3 (0 %)	entre : 3 (0 %)	sin : 3 (0 %)	entre : 2 (0 %)	sin : 4 (0 %)	sobre : 2 (0 %)	sobre : 3 (0 %)	sobre : 2 (0 %)	sobre : 3 (0 %)	desde : 3 (0 %)	sobre : 2 (0 %)
12	hacia : 1 (0 %)	hacia : 1 (0 %)	hacia : 1 (0 %)	hacia : 1 (0 %)	-	-	hacia : 1 (0 %)	hacia : 1 (0 %)	-	bajo : 1 (0 %)	contra : 1 (0 %)
13	-	-	-	-	-	-	-	-	-	contra : 1 (0 %)	hacia : 1 (0 %)
14	-	-	-	-	-	-	-	-	-	hacia : 1 (0 %)	-

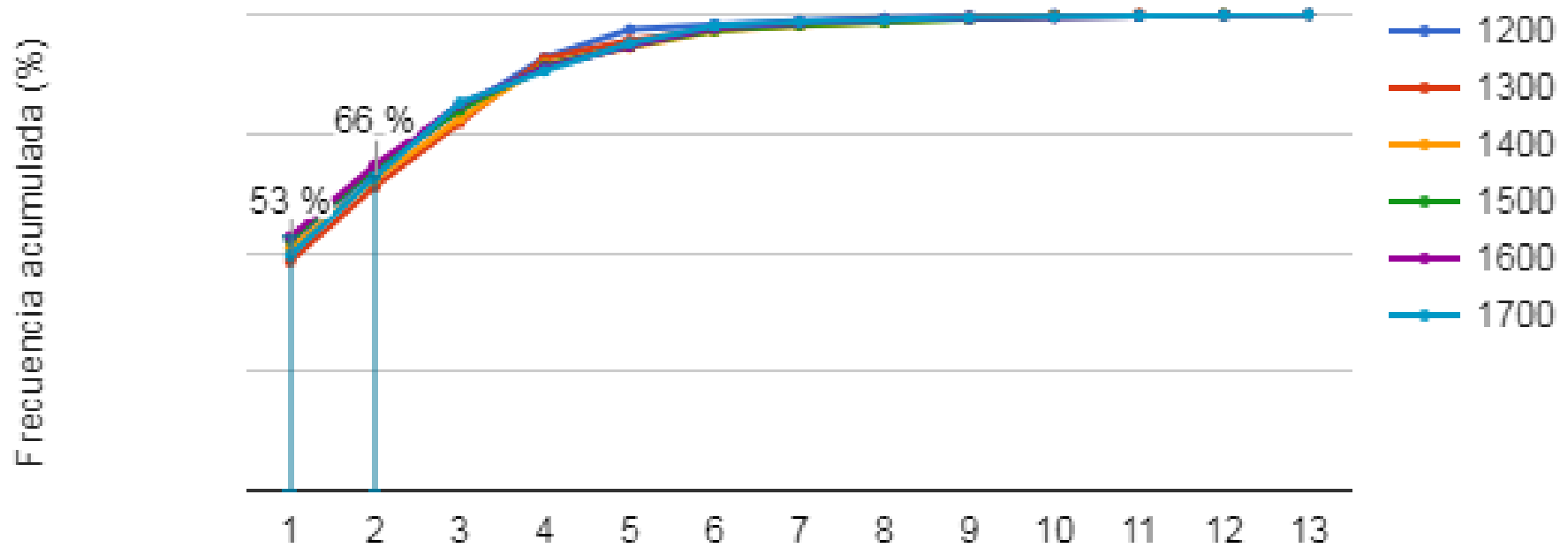
**Tabla 4. Tabla de Pareto. Preposiciones en las ciudades hispanohablantes.**

## **PRESEEA**

### 3. 2. Diacronía

Pareto. FP / 10 mil palabras	1200	1300	1400	1500	1600	1700
Índice de Pareto	0.893	0.879	0.875	0.874	0.877	0.872
M: Punto de mitad	1	2	1	1	1	2
N: Número de datos	11	11	12	13	13	13
M / N (%)	9%	18%	8%	8%	8%	15%
A: Acumulación en M.	781	947	769	831	891	1064
T: Frecuencia total	1496	1484	1515	1577	1672	1615
A / T (%)	52%	64%	51%	53%	53%	66%

**Tabla 5. Análisis de Pareto. Preposición en CODEA crítica, CV y CN**



**Fig. 3. Análisis de Pareto. Preposiciones en CODEA crítica, CV y CN**



R	1200	1300	1400	1500	1600	1700
1	<b>de : 781</b> <b>(52 %)</b>	<b>de : 714</b> <b>(48 %)</b>	<b>de : 769</b> <b>(51 %)</b>	<b>de : 831</b> <b>(53 %)</b>	<b>de : 891</b> <b>(53 %)</b>	<b>de : 798</b> <b>(49 %)</b>
2	en : 223 (15 %)	en : 233 (16 %)	en : 222 (15 %)	en : 238 (15 %)	en : 253 (15 %)	<b>a : 266</b> <b>(16 %)</b>
3	<b>a : 191</b> <b>(13 %)</b>	por : 200 (13 %)	<b>a : 192</b> <b>(13 %)</b>	<b>a : 191</b> <b>(12 %)</b>	<b>a : 211</b> <b>(13 %)</b>	en : 251 (16 %)
4	por : 165 (11 %)	<b>a : 199</b> <b>(13 %)</b>	por : 173 (11 %)	por : 147 (9 %)	por : 130 (8 %)	por : 109 (7 %)
5	con : 88 (6 %)	con : 57 (4 %)	para : 55 (4 %)	para : 65 (4 %)	con : 76 (5 %)	con : 92 (6 %)

6	contra : 16 (1 %)	para : 35 (2 %)	con : 52 (3 %)	con : 53 (3 %)	para : 64 (4 %)	para : 60 (4 %)
7	para : 11 (1 %)	contra : 16 (1 %)	contra : 17 (1 %)	ante : 17 (1 %)	ante : 18 (1 %)	ante : 13 (1 %)
8	ante : 9 (1 %)	ante : 12 (1 %)	ante : 11 (1 %)	sobre : 10 (1 %)	sin : 11 (1 %)	sin : 8 (0 %)
9	sin : 7 (0 %)	sin : 9 (1 %)	sin : 11 (1 %)	sin : 9 (1 %)	contra : 4 (0 %)	sobre : 7 (0 %)
10	sobre : 4 (0 %)	sobre : 6 (0 %)	sobre : 9 (1 %)	contra : 8 (1 %)	desde : 4 (0 %)	desde : 3 (0 %)
11	entre : 1 (0 %)	entre : 3 (0 %)	entre : 3 (0 %)	hasta : 4 (0 %)	hasta : 4 (0 %)	entre : 3 (0 %)
12	-	-	desde : 1	entre : 3	sobre : 4	hasta : 3

		(0 %)	(0 %)	(0 %)	(0 %)
13 -	-	-	desde : 1	entre : 2	contra : 2
			(0 %)	(0 %)	(0 %)

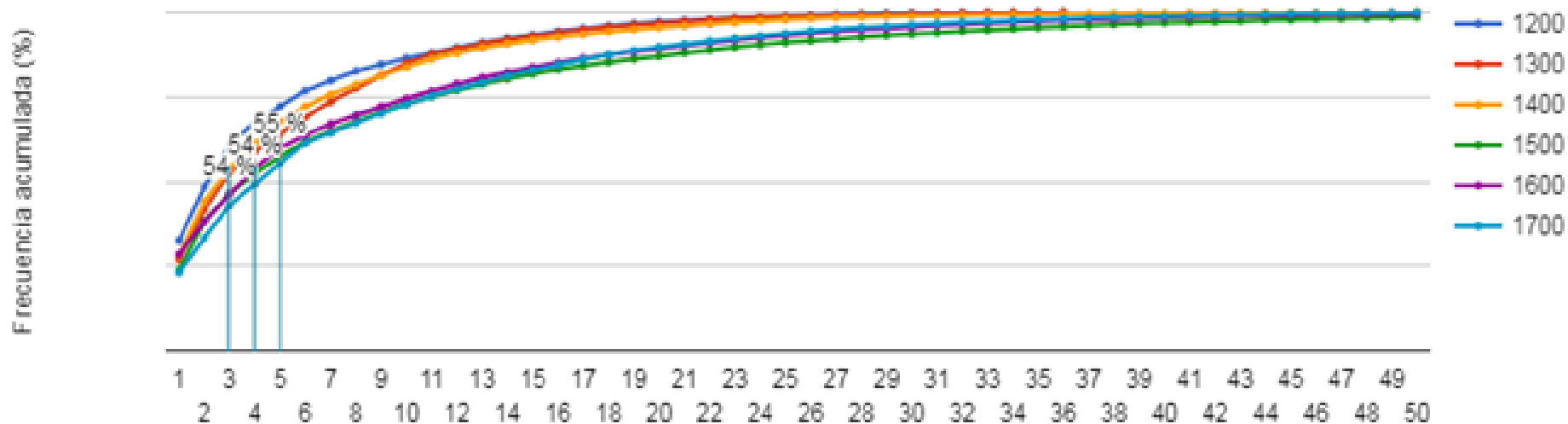
**Tabla 6. Tabla de Pareto. Preposición en CODEA crítica, CV y CN**

FP / 100 mil palabras	1200	1300	1400	1500	1600	1700
Formas unidas	2 788	2 500	3 077	3 375	2 555	2 284
Formas separadas	877	753	766	1 016	1 838	1 725

**Tabla 7. Número de palabras en CODEA crítica, CV y CN**

Pareto. FP / 100 mil palabras	1200	1300	1400	1500	1600	1700
Índice de Pareto	0.968	0.963	0.959	0.922	0.937	0.937
M: Punto de mitad	3	3	3	4	4	5
N: Número de datos	36	36	50	76	60	52
M / N (%)	8%	8%	6%	5%	7%	10%
A: Acumulación en M.	1875	1447	1781	1975	1977	1680
T: Frecuencia total	3089	2743	3300	3775	3673	3036
A / T (%)	61%	53%	54%	52%	54%	55%

**Tabla 8. Análisis de Pareto. Unión y separación de prep. + art. en CODEA paleográfica, CV y CN**



**Fig. 4. Análisis de Pareto.**

**Unión y separación de prep. + art. en CODEA paleográfica, CV y CN**

<b>del: 1003 (32 %)</b>	Unión: prep. frecuente
<b>antel: 1 (0 %)</b>	Unión: prep. poco frecuente
<b>por la: 144 (5 %)</b>	Separación: prep. frecuente
<b>contra el: 1 (0 %)</b>	Separación: prep. poco frecuente

**Tabla 9. Tabla de Pareto. Unión y separación de prep. + art. en CODEA paleográfica, CV y CN**

R	1200	1300	1400	1500	1600	1700
1	<b>del: 1003 (32 %)</b>	<b>del: 737 (27 %)</b>	<b>del: 927 (28 %)</b>	<b>del: 895 (24 %)</b>	<b>del: 1046 (28 %)</b>	<b>del: 701 (23 %)</b>
2	<b>dela: 487 (16 %)</b>	<b>dela: 402 (15 %)</b>	<b>dela: 519 (16 %)</b>	<b>dela: 547 (14 %)</b>	<b>de la: 350 (10 %)</b>	<b>del: 701 (23 %)</b>
3	<b>al: 385</b>	<b>al: 308</b>	<b>al: 335</b>	<b>al: 305 (8 %)</b>	<b>dela: 301</b>	<b>al: 293</b>

R	1200	1300	1400	1500	1600	1700
	(12 %)	(11 %)	(10 %)		(8 %)	(10 %)
4	delos: 207 (7 %)	delos: 166 (6 %)	delos: 238 (7 %)	delos: 228 (6 %)	al: 280 (8 %)	de la: 190 (6 %)
5	alos: 146 (5 %)	alos: 144 (5 %)	enla: 208 (6 %)	delas: 190 (5 %)	en la: 221 (6 %)	en el: 188 (6 %)
6	por la: 144 (5 %)	por la: 136 (5 %)	alos: 150 (5 %)	enla: 167 (4 %)	delos: 143 (4 %)	en la: 183 (6 %)
7	de la: 93 (3 %)	ala: 122 (4 %)	por la: 119 (4 %)	de la: 112 (3 %)	en el: 117 (3 %)	con el: 91 (3 %)
8	delas: 86 (3 %)	por el: 109 (4 %)	delas: 100 (3 %)	ala: 108 (3 %)	de los: 94 (3 %)	ala: 87 (3 %)
9	en el: 60	enla: 106	en la: 84	alos: 107	delas: 94	delos: 85

R	1200	1300	1400	1500	1600	1700
	(2 %)	<b>(4 %)</b>	(3 %)	<b>(3 %)</b>	<b>(3 %)</b>	<b>(3 %)</b>
10	en la: 57 (2 %)	<b>delas: 97</b> <b>(4 %)</b>	por el: 83 (3 %)	<b>enel: 91 (2 %)</b>	de las: 86 (2 %)	de los: 84 (3 %)
11	de los: 49 (2 %)	en la: 77 (3 %)	<b>ala: 79 (2 %)</b>	en la: 78 (2 %)	por la: 86 (2 %)	<b>enel: 73</b> <b>(2 %)</b>
12	<b>ala: 45 (1 %)</b>	por los: 41 (1 %)	de la: 57 (2 %)	por la: 76 (2 %)	<b>ala: 77 (2 %)</b>	con la: 65 (2 %)
13	por el: 44 (1 %)	en los: 38 (1 %)	<b>delo: 52</b> <b>(2 %)</b>	<b>delo: 72 (2 %)</b>	<b>enel: 67</b> <b>(2 %)</b>	a la: 59 (2 %)
14	<b>delo: 42</b> <b>(1 %)</b>	<b>delo: 36</b> <b>(1 %)</b>	<b>enlos: 40</b> <b>(1 %)</b>	por el: 57 (2 %)	a la: 56 (2 %)	por el: 59 (2 %)
15	en los: 32	en las: 30	de los: 33	<b>enlas: 53</b>	por el: 55	<b>ael: 49</b>



R	1200	1300	1400	1500	1600	1700
	(1 %)	(1 %)	(1 %)	(1 %)	(1 %)	(2 %)
16	<b>enel: 32</b> (1 %)	de la: 27 (1 %)	<b>enlas: 31</b> (1 %)	en el: 50 (1 %)	<b>delo: 50</b> (1 %)	<b>enla: 49</b> (2 %)
17	a los: 27 (1 %)	<b>alo: 25</b> (1 %)	<b>enel: 29</b> (1 %)	a la: 42 (1 %)	<b>enla: 48</b> (1 %)	de las: 46 (2 %)
18	en las: 23 (1 %)	<b>enlas: 22</b> (1 %)	<b>alás: 23</b> (1 %)	de los: 38 (1 %)	con la: 37 (1 %)	<b>deel: 46</b> (2 %)
19	por los: 21 (1 %)	por las: 17 (1 %)	en el: 23 (1 %)	por los: 36 (1 %)	de el: 35 (1 %)	<b>delas: 34</b> (1 %)
20	<b>enla: 18</b> (1 %)	<b>alás: 14</b> (1 %)	por los: 20 (1 %)	<b>alás: 33 (1 %)</b>	a los: 34 (1 %)	<b>alós: 30</b> (1 %)
21	a el: 12	a los: 13	a los: 17	<b>alo: 33 (1 %)</b>	<b>alós: 33</b>	de el: 27

R	1200	1300	1400	1500	1600	1700
	(0 %)	(0 %)	(1 %)		(1 %)	(1 %)
22	con la: 11 (0 %)	<b>enlos: 13</b> (0 %)	<b>alo: 17 (1 %)</b>	<b>porla: 32</b> (1 %)	<b>alas: 29</b> (1 %)	por lo: 27 (1 %)
23	<b>enlas: 10</b> (0 %)	<b>enel: 12</b> (0 %)	en las: 15 (0 %)	<b>enlos: 31</b> (1 %)	con el: 29 (1 %)	<b>porel: 25</b> (1 %)
24	con el: 9 (0 %)	por lo: 8 (0 %)	por las: 15 (0 %)	de las: 29 (1 %)	de lo: 26 (1 %)	a los: 22 (1 %)
25	de las: 9 (0 %)	en el: 7 (0 %)	de las: 14 (0 %)	a los: 24 (1 %)	por los: 26 (1 %)	<b>delo: 22</b> (1 %)
26	por las: 9 (0 %)	<b>enlo: 6</b> (0 %)	a las: 11 (0 %)	con el: 20 (1 %)	en las: 24 (1 %)	por la: 22 (1 %)
27	<b>enlos: 8</b>	ante los: 5	en los: 9	<b>conel: 20</b>	para el: 19	<b>alas: 19</b>

R	1200	1300	1400	1500	1600	1700
	<b>(0 %)</b>	(0 %)	(0 %)	<b>(1 %)</b>	(1 %)	<b>(1 %)</b>
28	<b>alas: 5 (0 %)</b>	de los: 5 (0 %)	a la: 6 (0 %)	en los: 20 (1 %)	para la: 17 (0 %)	de lo: 13 (0 %)
29	a la: 4 (0 %)	a la: 4 (0 %)	de lo: 6 (0 %)	<b>porel: 19 (1 %)</b>	ante el: 16 (0 %)	a las: 12 (0 %)
30	<b>alo: 3 (0 %)</b>	con los: 4 (0 %)	con las: 3 (0 %)	de lo: 16 (0 %)	con las: 15 (0 %)	en las: 12 (0 %)
31	<b>ael: 2 (0 %)</b>	de las: 4 (0 %)	con los: 3 (0 %)	<b>enlo: 15 (0 %)</b>	<b>deel: 15 (0 %)</b>	<b>conel: 11 (0 %)</b>
32	por lo: 2 (0 %)	entre los: 4 (0 %)	<b>conel: 3 (0 %)</b>	por las: 15 (0 %)	a las: 14 (0 %)	en los: 11 (0 %)
33	<b>antel: 1</b>	a las: 1	<b>conla: 3</b>	<b>porlos: 15</b>	a el: 13 (0 %)	para el: 9

R	1200	1300	1400	1500	1600	1700
	(0 %)	(0 %)	(0 %)	(0 %)		(0 %)
34	con los: 1 (0 %)	con el: 1 (0 %)	<b>conlos: 3</b> (0 %)	<b>deel: 14 (0 %)</b>	en los: 13 (0 %)	con las: 8 (0 %)
35	contra el: 1 (0 %)	en lo: 1 (0 %)	<b>enlo: 3 (0 %)</b>	para la: 14 (0 %)	<b>alo: 9 (0 %)</b>	<b>conla: 8</b> (0 %)
36	sobre los: 1 (0 %)	sobre la: 1 (0 %)	por lo: 3 (0 %)	<b>conlos: 13</b> (0 %)	<b>conla: 9</b> (0 %)	ante el: 7 (0 %)
37	-	-	sobre lo: 3 (0 %)	para el: 13 (0 %)	<b>enlas: 9</b> (0 %)	para la: 7 (0 %)
38	-	-	con el: 2 (0 %)	en las: 12 (0 %)	con los: 7 (0 %)	<b>porlo: 7</b> (0 %)
39	-	-	contra lo: 2	ante los: 8	<b>enlos: 7</b>	a el: 5 (0 %)

R	1200	1300	1400	1500	1600	1700
			(0 %)	(0 %)	(0 %)	
40 -	-		<b>porla: 2</b> (0 %)	<b>antel: 8 (0 %)</b>	a lo: 6 (0 %)	<b>alo: 5 (0 %)</b>
41 -	-		<b>ael: 1 (0 %)</b>	con los: 7 (0 %)	por las: 6 (0 %)	con los: 5 (0 %)
42 -	-		<b>antel: 1 (0 %)</b>	por lo: 7 (0 %)	por lo: 6 (0 %)	por las: 5 (0 %)
43 -	-		con la: 1 (0 %)	<b>porlas: 7</b> (0 %)	<b>ael: 5 (0 %)</b>	a lo: 4 (0 %)
44 -	-		<b>conlas: 1</b> (0 %)	sobre lo: 7 (0 %)	<b>conel: 5</b> (0 %)	con lo: 4 (0 %)
45 -	-		<b>contra la: 1</b>	a las: 6 (0 %)	desde el: 5	por los: 4

R	1200	1300	1400	1500	1600	1700
			(0 %)		(0 %)	(0 %)
46 -	-	-	contra los: 1 (0 %)	a lo: 6 (0 %)	para las: 5 (0 %)	<b>porla: 4</b> (0 %)
47 -	-	-	en lo: 1 (0 %)	ante el: 6 (0 %)	<b>conlos: 3</b> (0 %)	desde el: 3 (0 %)
48 -	-	-	fasta el: 1 (0 %)	con la: 6 (0 %)	en lo: 3 (0 %)	en lo: 2 (0 %)
49 -	-	-	<b>porlas: 1</b> (0 %)	para lo: 6 (0 %)	<b>enlo: 3 (0 %)</b>	<b>enlos: 2</b> (0 %)
50 -	-	-	sobre la: 1 (0 %)	<b>antelos: 5</b> (0 %)	<b>porel: 3</b> (0 %)	para lo: 2 (0 %)
51 -	-	-	-	<b>conla: 5 (0 %)</b>	con lo: 2	entre los: 1

R	1200	1300	1400	1500	1600	1700
					(0 %)	(0 %)
52 -	-	-	-	contra el: 5 (0 %)	<b>conlas: 2</b> <b>(0 %)</b>	sobre el: 1 (0 %)
53 -	-	-	-	sobre la: 5 (0 %)	para lo: 2 (0 %)	-
54 -	-	-	-	con las: 4 (0 %)	para los: 2 (0 %)	-
55 -	-	-	-	contra lo: 4 (0 %)	<b>porlos: 2</b> <b>(0 %)</b>	-
56 -	-	-	-	en lo: 4 (0 %)	sobre la: 2 (0 %)	-
57 -	-	-	-	para los: 4	<b>anteel: 1</b>	-

R	1200	1300	1400	1500	1600	1700
				(0 %)	<b>(0 %)</b>	
58	-	-	-	de el: 3 (0 %)	contra el: 1 (0 %)	-
59	-	-	-	<b>parala: 3</b> <b>(0 %)</b>	contra la: 1 (0 %)	-
60	-	-	-	contra la: 2 (0 %)	sobre el: 1 (0 %)	-
61	-	-	-	sobre los: 2 (0 %)	-	-
62	-	-	-	<b>ael: 1 (0 %)</b>	-	-
63	-	-	-	<b>conlas: 1</b> <b>(0 %)</b>	-	-



R	1200	1300	1400	1500	1600	1700
64	-	-	-	contra los: 1 (0 %)	-	-
65	-	-	-	<b>contra el: 1</b> <b>(0 %)</b>	-	-
66	-	-	-	desde el: 1 (0 %)	-	-
67	-	-	-	entre el: 1 (0 %)	-	-
68	-	-	-	entre los: 1 (0 %)	-	-
69	-	-	-	hasta el: 1 (0 %)	-	-

R	1200	1300	1400	1500	1600	1700
70	-	-	-	para las: 1 (0 %)	-	-
71	-	-	-	<b>paralo: 1</b> (0 %)	-	-
72	-	-	-	<b>paralos: 1</b> (0 %)	-	-
73	-	-	-	<b>porlo: 1</b> (0 %)	-	-
74	-	-	-	sobre el: 1 (0 %)	-	-
75	-	-	-	sobre las: 1 (0 %)	-	-
76	-	-	-	<b>sobrelo: 1</b>	-	-

R	1200	1300	1400	1500	1600	1700
				(0 %)		

Tipo	1200	1300	1400	1500	1600	1700	SUMA
Unión	16	15	24	34	24	20	133
- Unión. prep. frec.	15	15	23	29	22	20	124
- Unión. prep. poco. f.	1	0	1	5	2	0	9
Separación	20	21	26	42	36	32	177
- Separación. prep. frec.	18	18	22	32	25	25	140
- Separación. prep. poco. f.	2	3	4	10	11	7	37
TOTAL	36	36	50	76	60	52	310

**Tabla 10. Casos tipo de Unión y separación de prep. + art. en CODEA paleográfica, CV y CN**

Frecuencia	Unión	Separación	Total
Prep. frecuente	124	140	264
Prep. poco frecuente	9	37	46
Total	133	177	310

**Tabla 11. Cruce Frecuencia y Unión - Separación en CODEA paleográfica, CV y CN**

## 4. Discusión

4. 1. Origen de <del>, <al> y <el>

Español moderno: el porque/ del <el> en:

*el, los, la, las, lo; del, al*

Lati/n: (a) Mene/ndez Pidal :: (b) Alvar / Pottier

(a) **ILLE**, ILLOS, ILLA, ILLAS, ILLUD; DE ILLO, AD ILLUM

(b) **ILLUM**, ILLOS, ILLA, ILLAS, ILLUD; DE ILLO, AD ILLUM

Español medieval en orígenes (siglo X-XI):

*elo, elos, ela, elas; delo, alo*

*Glosas: G.E. Glosas Emilianenses (siglo X) / G.s. Glosas Silenses (siglo XI)*

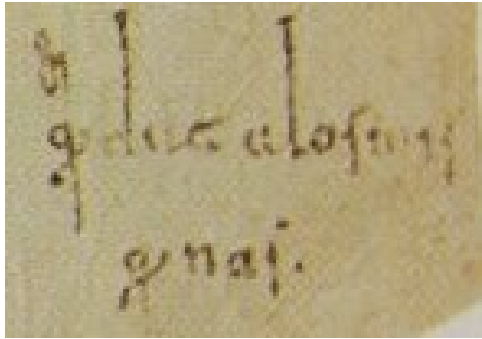


Fig. 5. *G. E.* f. 69r. 48: *q<ui> dat alosmis|q<ui>nos*



Fig. 6. *G. E.* f. 71v. 85: *d<e>la probatione*

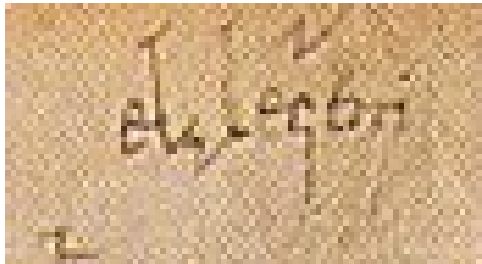


Fig. 7. *G. E.* f. 71v. 86: *ela legem*

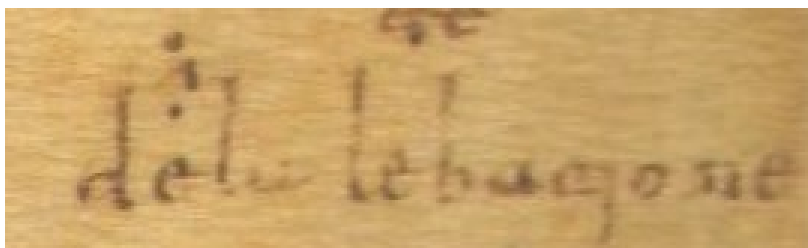


Fig. 8. G. S. f. 311r. 21: dele lebatione

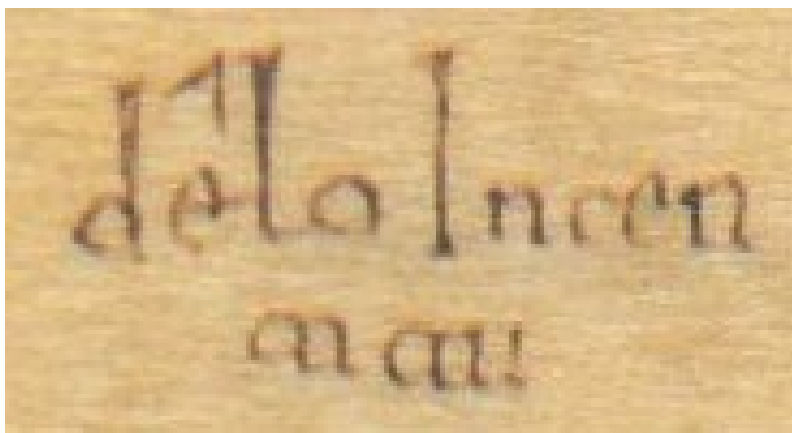


Fig. 9. G. S. f. 312v. 38: dele incen|turu

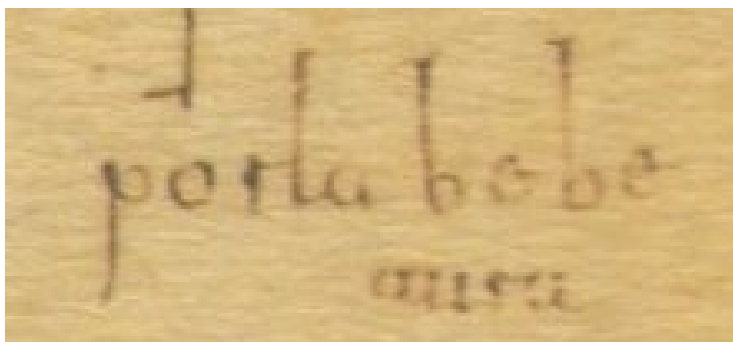


Fig. 10. G. S. f. 313r. 55: porla bebe|tura



FA	1200	1300	1400	1500	1600	1700
<b><i>de el</i></b>	<b>2</b>	<b>1</b>	<b>0</b>	<b>14</b>	<b>53</b>	<b>23</b>
<i>de la</i>	89	38	89	242	420	117
<i>del</i>	800	732	1187	1752	1195	389
<i>dela</i>	402	412	680	1088	364	181
Total	1 293	1 183	1 956	3 096	2 032	710

**Tabla 12. Frecuencia absoluta de <de el>, <de la>, <del>, <dela>  
en CODEA**

FA	900	1000	1100	1200
<i>de la</i>	0	0	17	60
<i>del</i>	1	1	19	151
<i>dela</i>	1	0	8	91
Total	2	1	44	302

**Tabla 13. Frecuencia absoluta de <de el>, <de la>, <del>, <dela> en CORHEN**

## (1) Retorromance

---

Prep	il	igl	la	l'	ils	las
a	al	agl	alla	all'	als	allas
cun	cul	cugl	culla	cull'	culs	cullas
da 'desde'	dal	dagl	dalla	dall'	dals	dallas
de	dil	digl	della	dell'	dils	dellas
en	el	egl	ella	ell'	els	ellas
per 'para'	pil	pigl	pella	pell'	pils	pellas
sin 'sobre'	sil	sigl	silla	sill'	sils	sillas
sper 'al lado de'	spel	spegl	spella	spell'	spels	spellas
sur 'encima de'	sul	sugle	sulla	sull'	suls	sullas

---

## (2) Italiano

---

Prep.	il	lo	l'	la	i	gli	le
a	al	allo	all'	alla	ai	agli	alle
con	[col]	[collo]	coll'	colla	[coi]	[cogli]	colle
da 'desde'	dal	dallo	dall'	dalla	dai	dagli	dalle
di 'de'	del	dello	dell'	della	dei	degli	delle
in	nel	nello	nell'	nella	nei	negli	nelle
per 'para'	pel	(per lo)	(per l')	(per la)	pei	(per gli)	(per le)
su 'sobre'	sul	sullo	sull'	sulla	sui	sugli	sulle

---

[col] ~ con il; [collo] ~ con lo; [coi] con i; [cogli] con gli.

### (3) Português

---

Prep.	o	a	os	as		
	a	ao	à	aos	às	
	de	do	da	dos	das	
	em	no	na	nos	nas	
	por	por	pelo	pela	pelos	pelas

---

### (4) Francés

---

Prep.	le	les	
	à	au	aux
	de	du	des

---

## (5) Español

Prep. el

a al

de del

Orden descendente, envolvente:

retorromance > italiano > portugués > francés > español

## 4. 2. Frecuencia de uso

Las primeras dos preposiciones más frecuentes son <de> y <a>. Los documentos notariales presentan más variaciones de las primeras preposiciones, donde compiten otras dos: <en> y <por>.

La frecuencia de uso no puede ser causa del cambio lingüístico, sino más bien su efecto.

La causa más plausible del cambio lingüístico de unidades lingüísticas sería lógicamente su característica lingüística, en nuestro caso, de su forma, significado y función de las unidades en cuestión, <del> y <al>.

Nos fijamos en la frecuencia de la forma separada <de la>, que es bastante alta, desde el principio y a lo largo de siglos.

Si la causa principal de la forma <del> fuera su frecuencia de uso, la de su pareja femenina <dela> también debería serlo.

La diferencia fundamental entre <del> y <dela> radica en que <del> era una unidad monoléxica desde el principio mientras que <dela> competía siempre con <de la> por razones expuestas en la sección anterior.



Nos referimos a la relación envolvente que hay entre las lenguas románicas que se realiza en manera que las unidades de la lengua de izquierda del orden descendente, retorromance > italiano > portugués > francés > español, envuelven las de las lenguas de derecha.

La razón por la que ocurre este fenómeno no puede ser de carácter numérico, que presentaría distintos grados de unión y separación en la manifestación de formas, sino de carácter estructural, tal como presentan los paradigmas donde observamos un estado perfecto de integración.

## 5. Conclusión

### *Hipótesis-1:*

No ha habido la unión de

\*<de el> → <del>, \*<a el> → <al>

(Bybee 2007, Elvira 2009, Ueda 2011, etc.)

en la cronología general de la lengua española.

En su lugar pensamos en los procesos:

DE ILLŌ → <de il(l)o> → <del(o)> → <del>

AD ILLU(M) → <a il(l)o> → <al(o)> → <al>

## *Hipótesis-2:*

La no separación de <del> y <al> no se debe a

\*su frecuencia de uso,

sino a

su rasgo fonético de la vocal final de la preposición

y

su alta utilidad lingüística realizada en forma breve y la amplitud semántica.

### ***Hipótesis-3:***

Entre <del>, <al> y <el> se mantiene una correlación fuerte en su función y forma, lo que ha producido la forma apocopada peculiar común entre estas tres formas unisilábicas.

del(o), al(o); el(o) → del, al; el (desde el principio)

d(e)los, alos; (e)los → delos, alos; los → de los, a los; los

d(e)la, ala; (e)la → dela, ala; la → de la, a la; la

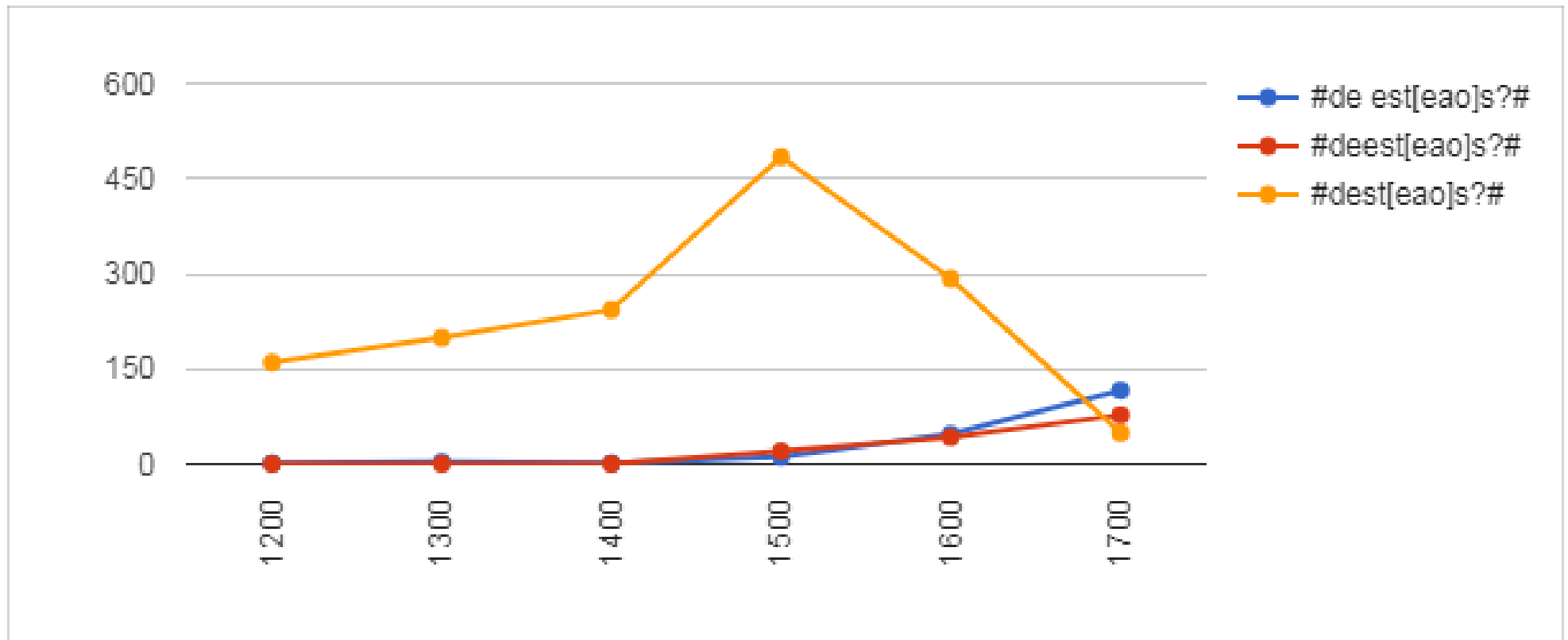
d(e)las, alas; (e)las → delas, alas; las → de las, a las; las

?? [neutro] (?? posterior)

d(e)lo, alo; (e)lo → delo, alo; lo → de lo, a lo; lo

FA	1200	1300	1400	1500	1600	1700
#de est[eao]s?#	1	3	1	11	47	116
#deest[eao]s?#	0	0	0	20	42	77
#dest[eao]s?#	160	199	243	484	292	49

**Tabla 15. Frecuencia absoluta  
de <de este>, <deeste>, <deste>  
en CODEA (CV, CN)**



**Fig. 11. Frecuencia absoluta  
de <de este>, <deeste>, <deste>  
en CODEA (CV, CN)**

Podemos pensar que las formas unidas y las separadas de <de> + demostrativo y las mismas de artículo definido, <delos, dela, delas, delo> y <delos, dela, delas, delo>, son variantes gráficas, sin diferencia morfológica.

Sin embargo, las formas unidas constantes <del> y <al> no pueden ser variantes morfológicas de \*<de el> y \*<a el>, puesto que las últimas no existen ni han existido en la historia. <del> y <al> han sido y son formas únicas y especiales, lo mismo que <el>.



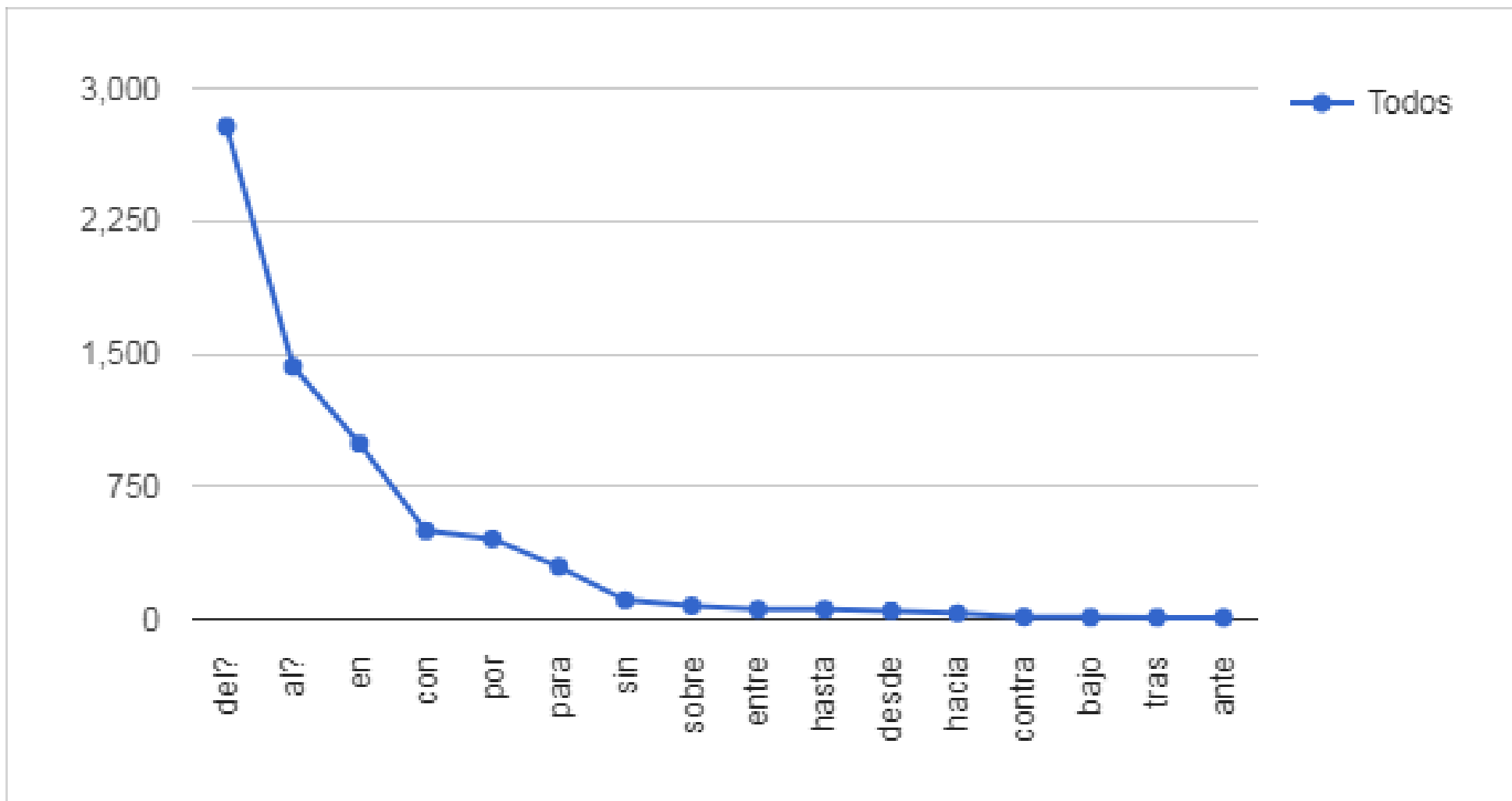
# Apéndice-1. Análisis de Pareto

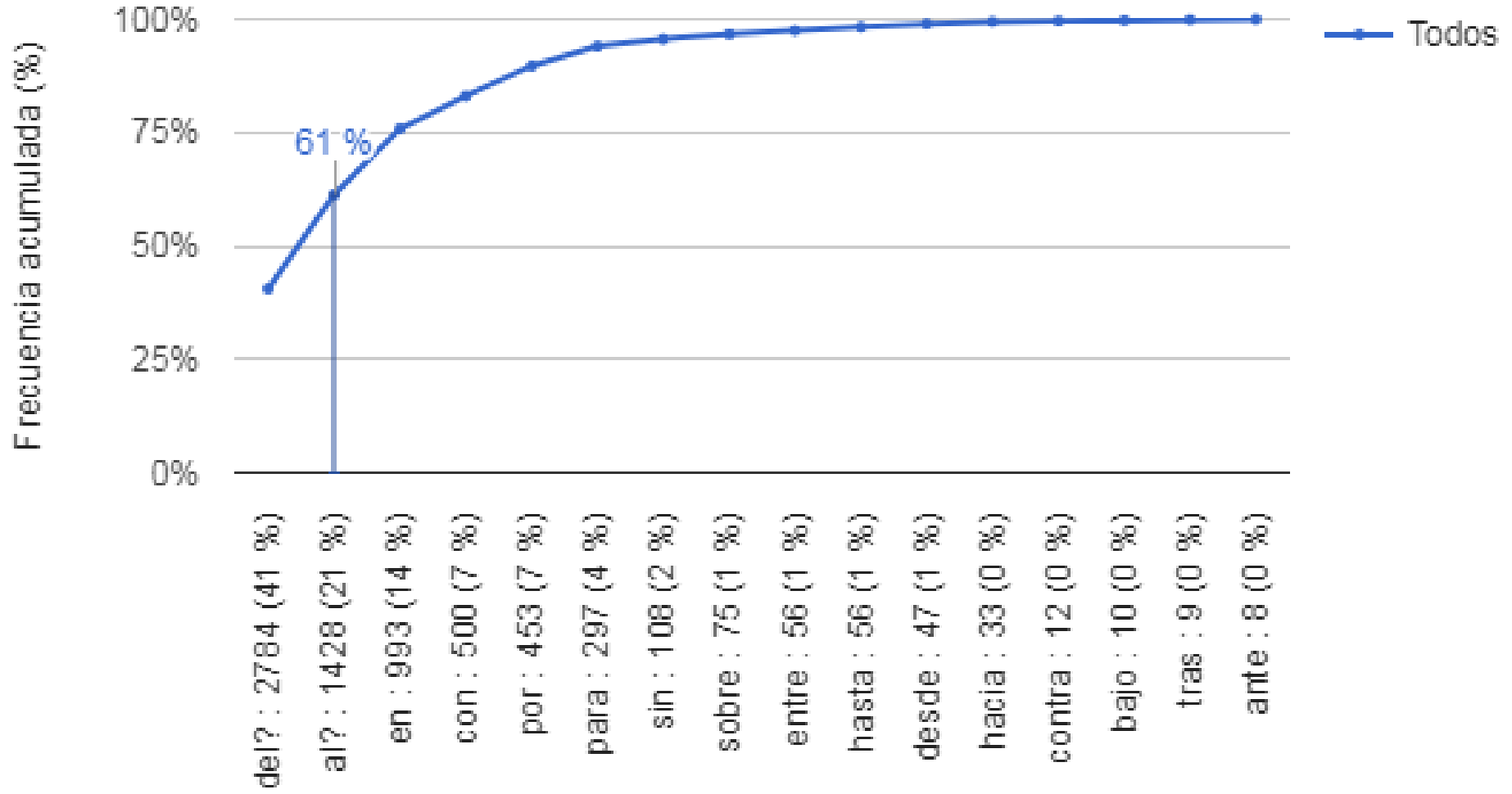
*Marianela* (1878), Benito Pérez Galdós

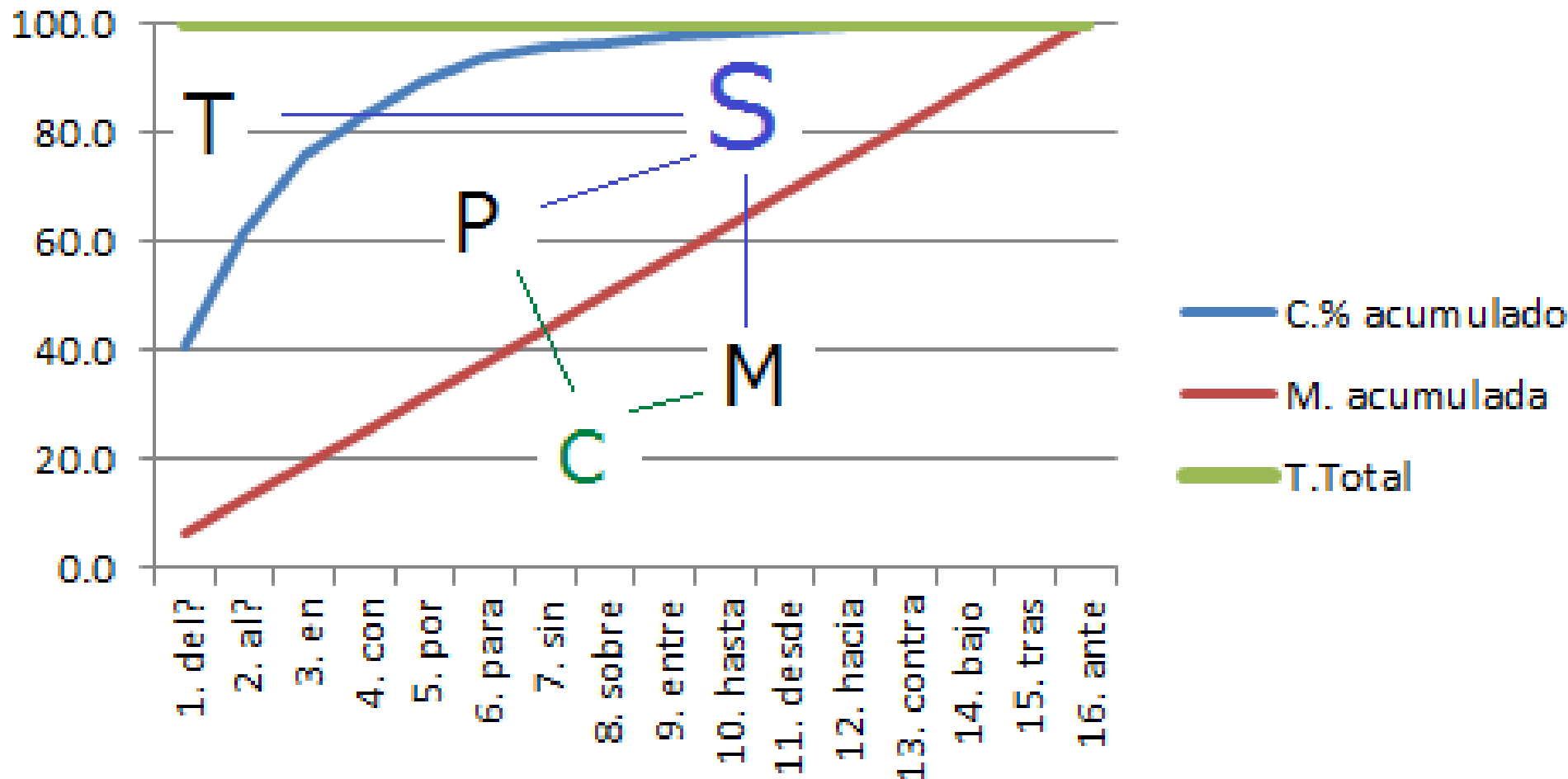
FA	Todos
de	2 784
a	1 428
en	993
con	500
por	453
para	297
sin	108
sobre	75
entre	56

hasta	56
desde	47
hacia	33
contra	12
bajo	10
tras	9
ante	8

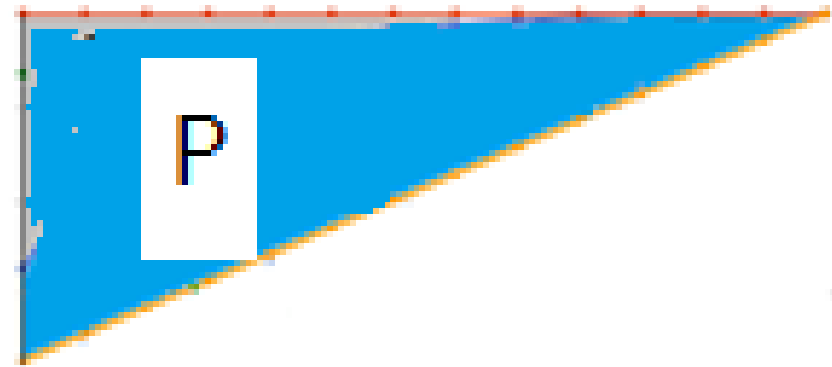
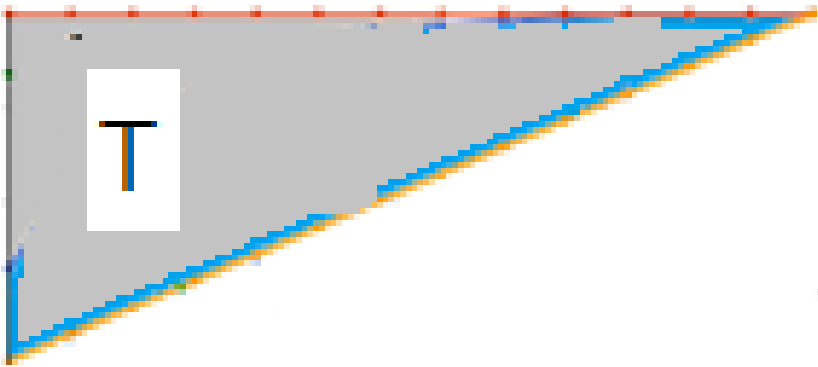
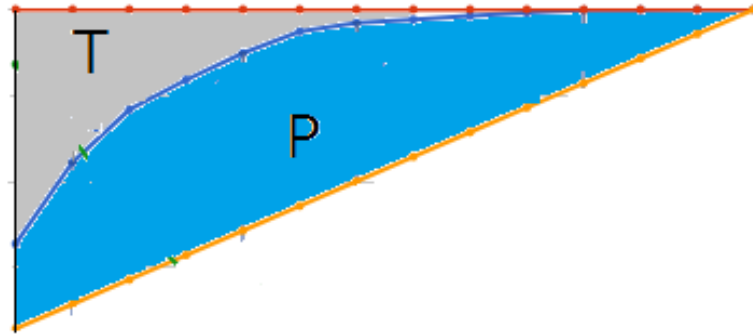
---







$$IP = P / (T + P)$$



Mínimo:  $IP = 0$  ( $P = 0$ )

/

Máximo:  $IP = 1$  ( $T = 0$ )

$$S = T + P + M$$

$$C = P + M$$

$$T = S - P - M = S - C$$

$$P = C - M$$

Por consiguiente, el Índice de Pareto (IP) es:

$$IP = P / (T + P) = (C - M) / (S - C + C - M) = (C - M) / (S - M)$$

Prep.	Frec.	%	C.% acumulado	Media	M. acumulada	T.Total
1. del?	2 784	40.5	40.5	6.3	6.3	100.0
2. al?	1 428	20.8	61.3	6.3	12.5	100.0
3. en	993	14.5	75.8	6.3	18.8	100.0
4. con	500	7.3	83.1	6.3	25.0	100.0
5. por	453	6.6	89.6	6.3	31.3	100.0
6. para	297	4.3	94.0	6.3	37.5	100.0
7. sin	108	1.6	95.5	6.3	43.8	100.0
8. sobre	75	1.1	96.6	6.3	50.0	100.0
9. entre	56	0.8	97.5	6.3	56.3	100.0
10. hasta	56	0.8	98.3	6.3	62.5	100.0



Prep.	Frec.	%	C.% acumulado	Media	M. acumulada	T.Total
11. desde	47	0.7	99.0	6.3	68.8	100.0
12. hacia	33	0.5	99.4	6.3	75.0	100.0
13. contra	12	0.2	99.6	6.3	81.3	100.0
14. bajo	10	0.1	99.8	6.3	87.5	100.0
15. tras	9	0.1	99.9	6.3	93.8	100.0
16. ante	8	0.1	100.0	6.3	100.0	100.0
Suma	6 869	100.0	1429.8	100.0	850.0	1 600.0

$$P: s(C) - s(M) \quad 1429.8 - 850 = 580$$

$$T + P: s(T) - s(M) \quad 1600 - 850 = 750$$

$$IP: P / (T + P) \quad 580 / 750 = 0.773$$

## Apéndice-2. Frecuencia probabilística

Frecuencia Absoluta (FA), CODEA:

FA	1200	1300	1400	1500	1600	1700
<b>de el</b>	<b>2</b>	<b>1</b>	<b>0</b>	<b>14</b>	<b>53</b>	<b>23</b>
de la	89	38	89	242	420	117
del	800	732	1187	1752	1195	389
dela	402	412	680	1088	364	181
Total	1 293	1 183	1 956	3 096	2 032	710

## Frecuencia Relativa (FR):

FR. (%)	1200	1300	1400	1500	1600	1700
<b>de el</b>	<b>0.2</b>	<b>0.1</b>	<b>0</b>	<b>0.5</b>	<b>2.6</b>	<b>3.2</b>
<i>de la</i>	6.9	3.2	4.6	7.8	20.7	16.5
<i>del</i>	61.9	61.9	60.7	56.6	58.8	54.8
<i>dela</i>	31.1	34.8	34.8	35.1	17.9	25.5

Atributo	1200	1300	1400	1500	1600	1700
Palabra	73 316	90 987	119 607	185 010	106 694	49 150

$$FN (<de el> \text{ en } 1200) = 2 / 73316 * 100\ 000 = 2.72$$

FN.PI.:100 mil palabras	1200	1300	1400	1500	1600	1700
<i>de el</i>	2.7	1.1	0	7.6	49.7	46.8
<i>de la</i>	121.4	41.8	74.4	130.8	393.6	238
<i>del</i>	1091.2	804.5	992.4	947	1120	791.5
<i>dela</i>	548.3	452.8	568.5	588.1	341.2	368.3

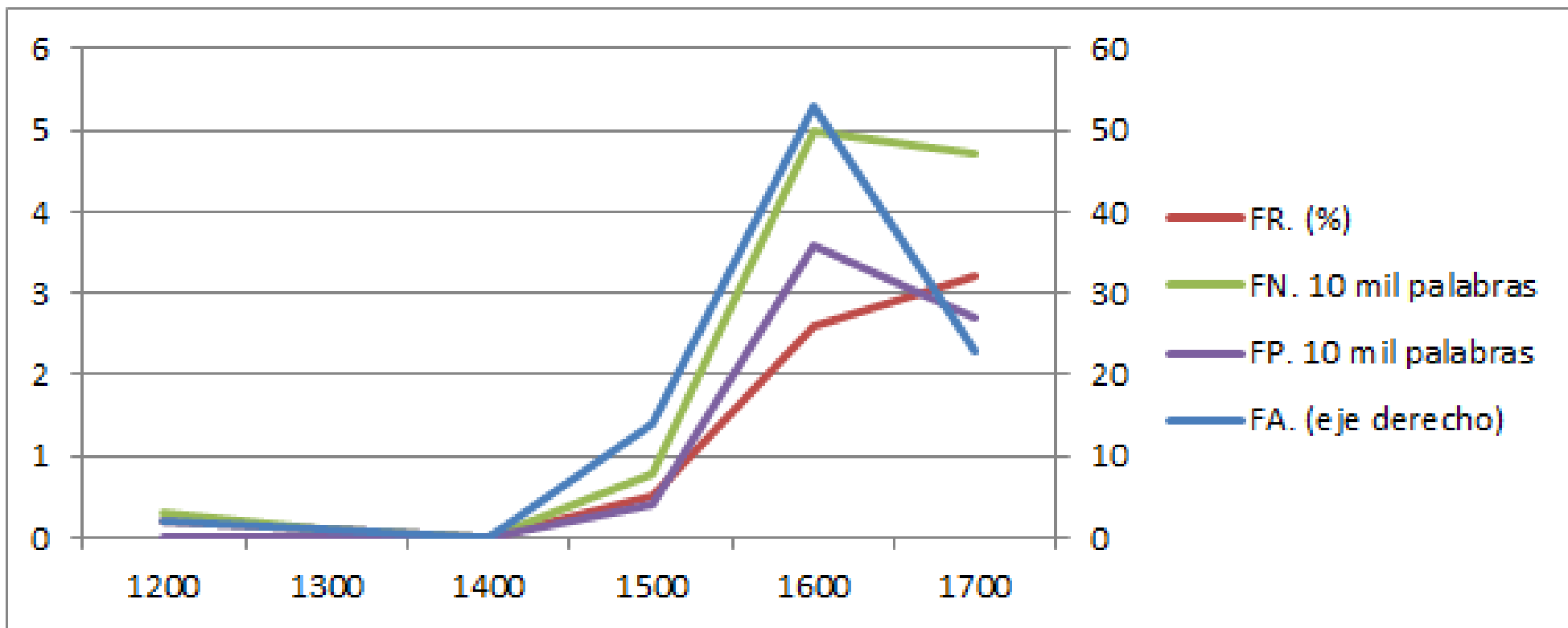
# Frecuencia probabilística

Significatividad ( $s$ , función  $S$ ) que da el  $x$  éxitos en  $n$  ensayos de un evento dotado de la probabilidad binomial expectativa ( $e$ ):

$$s = \text{BinS}(x, n, e) = \text{BINOMDIST}(x-1, n, e, 1)$$

$$e = \text{BinE}(x, n, s)$$

FP. 10 mil palabras	1200	1300	1400	1500	1600	1700
<i>de el</i>	0.2	0	0	3.7	35.5	27.1
<i>de la</i>	94	27.7	57.6	112.2	350.5	190.5
<i>del</i>	1004	737.3	927	895.5	1046.5	701.8
<i>dela</i>	487.1	402.8	519.2	547.6	301.1	308.2



## Programa-1: BinS

Function BinS(x, n, e)

'Significatividad s (x: ocurrencia, n: ensayos, e: probabilidad expectativa)

If x = 0 Then BinS = 0: Exit Function

BinS = Application.BinomDist(x - 1, n, e, 1)

End Function

## Programa-2: BinE

Function BinE(x, n, s)

'Probabilidad expectativa e (x: ocurrencia, n: ensayo, s: significatividad)

Dim i, k, r, mn, mx, sc: If x = 0 Then BinE = 0: Exit Function

r = 10 ^ 6: mx = r 'precisión: máximo de búsqueda

Do While k < 1000

```

i = (mx + mn) / 2 'punto medio entre máximo y mínimo
BinE = i / r 'candidato de la probabilidad expectativa
sc = BinS(x, n, BinE) 'significatividad propia del candidato
If sc < s - 1 / r Then 'Si sc no llega a s-1/r ...
    mx = i 'bajar el máximo de la búsqueda al punto medio
ElseIf sc > s + 1 / r Then 'Si sc sobrepasa a s-1/r...
    mn = i 'subir el mínimo al punto medio
Else 'Si sc entra en el ámbito de  $s \pm 1/r$  ...
    Exit Do 'salir del bucle

End If

k = k + 1

Loop

End Function

```



## Referencias

- Almela, Ramón / Pascual Cantos, Pascual / Sánchez, Aquilino / Sarmiento, Ramón / Almela, Moisés. 2005. *Frecuencias del español. Diccionario y estudios léxicos y morfológicos*. Madrid: Editorial Universitas.
- Alvar, Manuel / Pottier, Bernard. 1983. *Morfología histórica del español*. Madrid: Gredos.
- Ávila Muñoz, Antonio Manuel. 1999. *Léxico de frecuencia del español hablado en la ciudad de Málaga*. Málaga: Universidad de Málaga.
- Bybee, Joan. 2007. *Frequency of Use and the Organization of Language*. Oxford University Press.
- Davies, Mark. 2006. *A frequency dictionary of Spanish. Core vocabulary for learners*. New York: Routledge.
- Elvira, Javier..2009. *Evolución lingüística y cambio sintáctico*. Bern: Peter Lang.

- García Hoz, Víctor. 1953. *Vocabulario usual, vocabulario común y vocabulario fundamental*. Madrid: C.S.I.C.
- García Hoz, Víctor. 1976. *El vocabulario general de orientación científica y sus estratos*. Madrid: C.S.I.C.
- Hopper, Paul J. / Closs Traugott, Elizabeth. 2003. *Grammaticalization*. (2nd ed.) Cambridge University Press.
- Juilland, Alphonse & Chang-Rodríguez, Eugenio. 1964. *Frequency Dictionary of Spanish words*. The Hague: Mouton.
- Justicia, Fernando. 1995. *El desarrollo del vocabulario. Diccionario de frecuencias*. Granada: Universidad de Granada.
- Kataoka, Kozaburo. 1982. *Romansugo hikaku bunpou*. (Gramática comparativa de lenguas románicas). Tokio. Asahishuppan, 170-172.
- Kenny, Anthony. 1982. *The Computation of Style*. Oxford. Oxfordshire.

Kobayashi, Satoshi. *Manuale di grammatica italiana*. (en japonés). Tokio. Hakusuisha.

Koch, Richard. 1997. *The 80/20 Principle*. London: Nicholas Brealey.

Menéndez Pidal, Ramón. 1968. *Manual de gramática histórica española*. Madrid: Espasa-Calpe.

\_\_\_\_\_. 1980. *Orígenes del español. Estado lingüístico de la península ibérica hasta el siglo XI*. Madrid: Espasa-Calpe.

Saussure, Ferdinand de. 1971. *Curso de lingüística general*. Buenos Aires: Editorial Losada.

Ueda, Hiroto. 1987 / 2007. *Frecuencia y dispersión del vocabulario español*, Instituto de Estudios Lingüísticos, Universidad de Estudios Extranjeros de Tokio.

<https://lecture.ecc.u-tokyo.ac.jp/~cueda/kenkyu/goi/frec-disp/frec-disp-0.pdf>

- \_\_\_\_\_. 2011. *Supeingo bunpou handobukku*, en japonés. Tokio: Kenkyusha.
- \_\_\_\_\_. 2017. "Formación histórica de <del> y <al>" (en japonés), *Boletín de la Asociación Japonesa de Hispanistas*, 24, pp.2-3.
- \_\_\_\_\_. 2018. *Análisis de datos cuantitativos para estudios lingüísticos*.  
<https://lecture.ecc.u-tokyo.ac.jp/~cueda/gengó4-numeros/doc/numeros-es.pdf>
- \_\_\_\_\_/ Moreno Sandoval, Antonio. 2018. "Unión y separación de preposición y artículo definido del español. Observaciones con la frecuencia probabilística en el análisis de Pareto", comunicación oral en *X Congreso Internacional de Lingüística de Corpus*, Cáceres, España, 9-11 mayo, 2018.
- Zipf, George K. 1936. *The Psycho-biolology of Language*. New York: George Roughledge & Sons.