

Informe de los resultados por ámbito temático y contraste entre neología general y especializada

M. Teresa Cabré, Elisabet Llopart y M. Amor Montané

En este informe describimos los resultados obtenidos en el proyecto de acuerdo con tres parámetros. En primer lugar, observamos el índice de interferencia en relación con los préstamos de otras lenguas. A continuación, analizamos la distribución de las unidades léxicas estudiadas de acuerdo con su estructura. Finalmente, valoramos el índice de innovación de los neologismos de acuerdo con la estabilidad en el uso.

1. Índice de interferencia

Área temática	Frecuencia absoluta	Frecuencia relativa
Derecho	13	1,5
Economía	31	4,97
Informática	58	18,1
Medicina	0	0
Medio ambiente	5	0,6
Promedio	21,4	5,034
Mediana¹	13	1,5
General	2092	25,8

Tabla 1. Préstamos

Área temática	Frecuencia absoluta	Frecuencia relativa
Derecho	1	0,1
Economía	16	2,6
Informática	47	14,7
Medicina	0	0
Medio ambiente	1	0,1
Promedio	13	3,5
Mediana	1	0,1
General	63	0,8

Tabla 2. Préstamos híbridos

¹ En el análisis cuantitativo de los resultados por ámbito mostramos, a parte del promedio, la mediana. De acuerdo con Llopis (2012): “La media [el promedio] sufre los efectos de valores extremos, los cuales influyen mucho en su magnitud. La mediana es resistente a esos valores”. Por este motivo, calculamos la mediana, ya que como tenemos valores aislados extremos derivados de las particularidades del léxico de algunos ámbitos, el promedio no refleja los resultados más comunes, sino que en algunos casos se desvía a causa de estos valores extremos.

Área temática	Inglés			Francés			Catalán			Otras lenguas			Total
	Préstamo	Híbrido	Total	Préstamo	Híbrido	Total	Préstamo	Híbrido	Total	Préstamo	Híbrido	Total	
Derecho	1,5	0,1	1,6	-	-	0	-	-	0	-	-	0	1,6
Economía	5	2,6	7,6	-	-	0	-	-	0	-	-	0	7,6
Informática	17,8	14,7	32,5	0,3	-	0,3	-	-	0	-	-	0	32,8
Medicina	-	-	0	-	-	0	-	-	0	-	-	0	0
Medio ambiente	0,6	0,1	0,7	-	-	0	0,1	-	0,1	-	-	0	0,8
Promedio	5	3,5	8,5	0,06	0	0,06	0,02	0	0,02	0	0	0	8,6
Mediana	1,5	0,1	1,6	0	0	0	0	0	0	0	0	0	1,6
General	17,5	0,7	18,2	1,8	0,07	1,9	1,8	0,01	1,8	3,5	0,02	3,5	25,4

Tabla 3. Lenguas de los préstamos y de los préstamos híbridos

Los datos nos permiten observar que el índice de interferencia difiere entre los distintos ámbitos temáticos, y también respecto de los neologismos generales. De hecho, la lengua general es la que presenta un porcentaje de neologismos prestados más elevado, ya que casi una cuarta parte lo son. El índice de interferencia en los ámbitos especializados es siempre menor, siendo el promedio de un 5 %. A pesar de esta situación, también se observa que en la lengua especializada es mayor el porcentaje de préstamos híbridos, es decir, formados por un préstamo o una palabra de la lengua (por ejemplo, *software propietario* u *outputs bancarios*): un 3,5 % de promedio respecto del 0,8 % de la lengua general. El aspecto en el que coinciden los neologismos generales y especializados es en la lengua de mayor interferencia: el inglés. Entre los neologismos generales destaca también la interferencia del catalán, aunque se puede explicar a partir de los textos de vaciado, constituidos esencialmente por prensa de Cataluña.

Si nos centramos exclusivamente en analizar las diferencias entre los préstamos neológicos de los ámbitos especializados, vemos que existen diferencias entre ellos; sobre todo destacan los ámbitos de informática y medicina. Mientras que en informática el porcentaje de préstamos (casi exclusivamente del inglés) es muy elevado (32,8 % en total: 18,1 % de préstamos puros y 14,7 % de préstamos híbridos), el resto de ámbitos presentan un índice de interferencia mucho menor: a informática le sigue el ámbito económico (7,6 % en total), a muchos puntos porcentuales de diferencia; a continuación, la influencia de otras lenguas es escasa en derecho y en medio ambiente; finalmente, en medicina no se ha detectado ningún préstamo neológico. La lengua de mayor interferencia en todos los casos es el inglés.

A pesar de la particularidad que parece presentar el ámbito informático en comparación con el resto, la presencia de préstamos del inglés en esta área es un aspecto que ha sido estudiado en varias lenguas románicas (Gouadec 1994, Termcat 2006). De hecho, como la mayoría de innovaciones técnicas se producen en países anglófonos, y por lo tanto la influencia es clara. En derecho se podría destacar la presencia de latinismos, pero en este trabajo no se consideran préstamos. En cuanto a la medicina, sorprende la inexistencia de préstamos, pero es un hecho estudiado (Estopà 1999) que

la terminología médica tiende a la creación de términos con elementos cultos. También hay que tener en cuenta que para este estudio se han descartado los términos que contienen epónimos o las nomenclaturas en latín.

2. Distribución en estructuras

En esta tabla presentamos la distribución de los neologismos terminológicos y generales por estructuras de acuerdo con la frecuencia relativa de aparición. Para seleccionar las estructuras estudiadas nos hemos basado en las formas prototípicas utilizadas para vehicular términos. Por este motivo, en esta tabla no mostramos otras estructuras que han sido analizadas para los neologismos generales, como el adjetivo (A) y el verbo (V).

Área temática	N	NJ	NJJ	NJPN	NJPY	NJPNJ	NJPNPN	NN	NPN	NPNJ	NPNJJ	NPNP	NPNPJ	Otras	Sin cat.
Derecho	14,8	50,1	4	2,3	-	0,3	0,2	-	19,8	3,1	0,1	0,2	0,1	-	4,8
Economía	6,4	45,7	3,2	1,3	-	0,3	-	-	28,7	7,2	-	1,4	0,3	-	5,5
Informática	21,6	57,8	0,9	0,3	-	-	-	0,3	14,4	1,6	-	-	-	-	2,2
Medicina	3,6	78,7	9,3	0,8	1	0,2	-	0,4	4,8	0,6	0,4	-	-	-	-
Medio ambiente	26,6	60	0,8	0,8	0,1	0,3	-	-	10,3	1	-	0,1	-	-	-
Promedio	14,6	58,5	3,6	1,1	0,6	0,3	0,2	0,4	15,6	2,7	0,3	0,6	0,2	0	4,2
Mediana	14,8	57,8	3,2	0,8	0,6	0,3	0,2	0,4	14,4	1,6	0,3	0,2	0,2	0	4,8
General	47	8,6						3,7	2,7					1,6*	

Tabla 4. Estructuras

* Incluye JN (0,9 %), Adv+N, etc. (0,7 %)

La estructura más común en los términos neológicos de los distintos ámbitos de especialidad se corresponde con la forma NJ y es así para cada uno de las cinco áreas estudiadas (derecho, economía, informática, medicina y medio ambiente). La siguiente estructura más común es, dependiendo del área de especialidad, la estructura monoléxica N y la estructura poliléxica NPN. Estos datos son esperables porque en los tres casos se trata de las estructuras prototípicas para vehicular términos. De todos modos, si observamos estos datos con más detalle vemos que el porcentaje de unidades con las diferentes estructuras no es equivalente en los distintos ámbitos del estudio. Si bien hemos visto que la estructura NJ ocupa la primera posición en todos los casos, en la segunda posición encontramos la estructura NPN en derecho (19,8 %), economía (28,7 %) y medicina (4,8 %), pero, en cambio, en el caso de la informática (21,6 %) y el medio ambiente (26,6 %) se encuentran más neologismos terminológicos monoléxicos (N). En relación con la estructura NJ también vemos que, aunque el promedio se sitúa en el 58 % de los casos, encontramos en un extremo el ámbito de

medicina, porque el 78,7 % de los neologismos terminológicos adoptan esta forma. Creemos que este hecho destacable en contraste con el resto de ámbitos puede estar motivado por las recomendaciones de la OMS en relación con las denominaciones del ámbito de la salud. En cuanto al resto de estructuras, podemos comprobar que representan un porcentaje menor respecto de las tres mencionadas.

Si contrastamos el promedio y la mediana de los distintos neologismos terminológicos por ámbito con los resultados relativos a los neologismos generales de la prensa vemos que los resultados difieren entre sí. En primer lugar, debemos tener en cuenta que en la selección de los candidatos a término se han considerado solamente las estructuras prototípicas que muestran una tendencia mayor a vehicular términos y, por tanto, no se han tenido en cuenta unidades monolíticas con categoría adjetivo o verbo, que en relación con los neologismos generales forman una parte destacada de las unidades nuevas (el 20,7 % del total de neologismos son adjetivos y, en el caso de los verbos, se trata del 6,8 %). En este sentido, estos datos no son comparables porque no tenemos unidades de este tipo en relación con los neologismos terminológicos. Si nos centramos en la estructura N, que sí es comparable, observamos que en el caso de la neología general el 47 % de las unidades nuevas tienen esta estructura frente al 14,6 % de promedio que presentan las unidades terminológicas estudiadas. En relación con el resto de estructuras polilíticas, aproximadamente el 8,6 % de los neologismos generales muestran una estructura NJ, dato que contrasta ampliamente con el promedio del 58,5 % en el conjunto de los ámbitos especializados estudiados. En cuanto a la estructura NPN, el porcentaje de neologismos generales también es menor que en el resto de casos (un 2,7 % frente al 15,6 %). La estructura polilítica en que los datos generales y especializados muestran cierta similitud se corresponde con los casos NN. Además, es el único caso de estructura polilítica en que los datos de los neologismos generales (3,7 %) son superiores al promedio de los neónimos (0,4 %).

3. Índice de innovación: estabilidad en el uso ² (búsqueda en Google Académico)

Las unidades seleccionadas en cada uno de los ámbitos se consideran neológicas de acuerdo con un criterio textual o lexicográfico. En el caso de los neologismos terminológicos de los cinco ámbitos de especialidad tratados en este trabajo (derecho, economía, informática, medicina y medio ambiente), se han seleccionado las unidades que no se encuentran documentadas en un corpus textual determinado, correspondiente al Corpus Técnico del IULA del periodo anterior al estudiado; en el caso de medicina, también se ha utilizado un corpus terminológico, el SNOMED³. En relación con la neología general, como corpus lexicográfico de exclusión se han utilizado dos diccionarios de referencia en lengua castellana, el *Diccionario de la Real Academia Española* (DRAE) y el *Diccionario de uso del español de América y España* de VOX (VOXUSO). Así pues, para obtener más información sobre el carácter neológico de las unidades aplicamos un nuevo filtro correspondiente a Google Académico, que se trata de un buscador que contiene

² Cuando utilizamos el término *neologicidad* hacemos referencia a la estabilidad en el uso.

³ Systematized Nomenclature of Medicine-Clinical Terms (https://www.nlm.nih.gov/research/umls/Snomed/snomed_main.html).

artículos, tesis, libros y resúmenes extraídos de publicaciones académicas, asociaciones profesionales, repositorios en línea, universidades y algunas páginas web. De este modo, según se documentan o no en esta fuente, y el nombre de ocurrencias (o resultados) que presentan, podremos observar si los neologismos obtenidos muestran una estabilidad en el uso más o menos considerable. En este sentido, se trata de estudiar la frecuencia de uso, parámetro relacionado con cualidad de neológico o de neologicidad.

Una vez hemos buscado cada uno de los neologismos en Google Académico⁴ y hemos anotado el número de ocurrencias que presentan en esta fuente, hemos calculado la frecuencia relativa y los hemos agrupado de acuerdo con diversas categorías: si presentan 0 ocurrencias, 1 resultado, entre 2 y 10, entre 11 y 50, 51 y 100, 101 y 1.000, o, finalmente, más de 1.001 ocurrencias.

Área temática	0 oc.	1 oc.	2-10 oc.	11-50 oc.	51-100 oc.	101-1000 oc.	Más de 1001 oc.
Derecho	3,5	1,9	7,4	9,5	6,7	32,4	38,5
Economía	0,5	1,1	10,4	11,7	7,1	40,9	28,4
Informática	0,9	2,2	8,1	9,1	10	29,1	40,6
Medicina	0	6,2	13,3	19,6	13,1	34,4	12,1
Medio ambiente	0	0,6	2,4	8,1	5,4	22,1	62,2
Promedio	0,1	2,4	8,3	11,6	8,5	31,8	36,4
Mediana	0,5	1,9	8,1	9,5	7,1	32,4	38,5
General	5,5	3,3	8,2	11,1	6,2	25,8	40

Tabla 5. Estabilidad en el uso

En general, observamos que las categorías que agrupan un mayor número de neologismos se corresponden con los grupos que indican una frecuencia de uso elevada, de entre 101 y 1.000 resultados, o más. En el otro extremo, las categorías que son menos representativas porque agrupan un número incipiente de unidades son, en primer lugar, las que muestran 0 resultados y, en segundo lugar, las que presentan una ocurrencia en Google Académico. De forma más detallada, vemos que en medicina y medio ambiente no hay ningún neologismo que no se documente en Google Académico y, si nos centramos en las dos últimas categorías, comprobamos que muestran resultados no paralelos porque medicina es el ámbito que presenta un porcentaje menor de neologismos con más de 1.001 ocurrencias (12,1 %), mientras que los neologismos terminológicos de medio ambiente son los que presentan una frecuencia más elevada en esta categoría (62,2 %). Ahora bien, en el caso de medio ambiente este aspecto se ve compensado por el hecho que en la categoría entre 101 y 1.000 resultados es el que presenta un porcentaje mayor que el resto de ámbitos (22,1 %),

⁴ En el caso de los neologismos generales se ha utilizado una muestra representativa y aleatoria de un 10 % del total de neologismos sobre el total de 8.099 unidades, por la imposibilidad de buscarlos todos manualmente.

que acostumbran a mostrar valores similares entre ambas categorías. Por ejemplo, en el caso de las unidades del ámbito del derecho, que muestran un porcentaje parecido de unidades documentadas con entre 101 y 1.000 ocurrencias (32,4 %) y con más de 1.001 resultados (38,5 %).

Una vez hemos comparado los resultados de los distintos ámbitos de especialidad, vamos a comparar también el promedio y la mediana que presentan estas unidades con los neologismos generales. En esta línea, podemos afirmar que los resultados son similares. De todos modos, en el caso de los neologismos generales se observa una tendencia a obtener resultados más elevados en las categorías inferiores (de 0 y 1 ocurrencias), ya que suman un 8,9 % de las unidades, mientras que la mediana de los neologismos terminológicos es del 2,4 %. En este sentido, cabe tener en cuenta que el filtro utilizado (Google Académico) no contiene textos generales procedentes de la prensa, de ensayos, etc., sino, como su nombre indica, centrados en el ámbito académico. Así pues, este hecho podría explicar esta diferencia.

A continuación presentamos los resultados anteriores en forma de gráfico para observar la curva de estabilidad que presentan los neologismos de los diferentes ámbitos y, de este modo, visualizar las similitudes y diferencias.

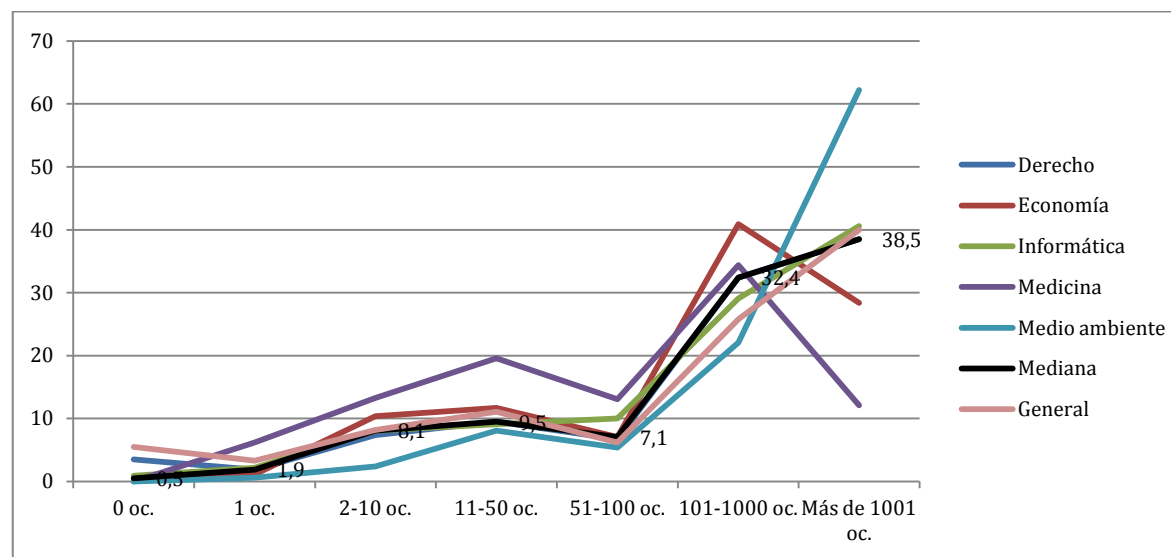


Gráfico 1. Curva de estabilidad

Para empezar, podemos ver que en todos los ámbitos la curva va escalando posiciones desde la categoría de 0 y 1 ocurrencias hasta la categoría de 51 a 100 ocurrencias. En este punto, se observa una disminución en todos los casos y, a continuación, la curva vuelve a subir. Finalmente, en la última categoría, se observan algunas diferencias relevantes. En este punto, tanto en el caso de medicina como de economía el porcentaje de unidades documentadas con más de 1.001 disminuye, mientras que en el resto de ámbitos sube, especialmente en el caso de las unidades de medio ambiente.