

# **Productividad relativa de los esquemas de formación de palabras en el español de México. Actualización 2022**

Relative productivity of word formation schemes in Mexican Spanish. Update 2022

Ramón Zacarías Ponce de León  
Universidad Nacional Autónoma  
de México, Instituto de  
Investigaciones Filológicas,  
Centro de Lingüística Hispánica  
rzacaria@hotmail.com



Recepción: 22 de junio del 2023  
Aceptación: 19 de septiembre del 2023  
doi: 10.22201/enallt.01852647p.2024.78.1074

## Resumen

En este trabajo se presentan estadísticas sobre productividad de la morfología léxica en el español de México. Los datos se obtuvieron de la base neológica Morfolex y los cálculos se llevaron a cabo con el acumulado de neologismos a junio del 2022, el cual asciende a 21 605 palabras. A partir del corpus de neologismos se realizaron conteos de tipos, ocurrencias y *hápx legomena* para cada uno de los esquemas de formación de palabras representados en el corpus. La metodología seguida para estos cálculos es la utilizada en Morfolex desde 2010, presentada detalladamente en un trabajo anterior (Zacarías Ponce de León, 2016). En dicha investigación se mostraron los hallazgos correspondientes a diez mil neologismos. Con la actualización que se da a conocer ahora se reafirma que la base de datos Morfolex es representativa del léxico mexicano y ofrece un panorama de los procesos lexicogénicos más productivos en esta variedad del español. Por último, se hace una comparación entre los datos actuales y los correspondientes a 2016 para constatar la evolución que han experimentado los distintos procesos formativos de léxico.

Palabras clave: productividad; neologismos; Morfolex; *hápx legomena*; morfología léxica

## Abstract

This paper presents statistics on the productivity of lexical morphology in Mexican Spanish. The data were obtained from the Morfolex neologism database, and the calculations were carried out with the accumulated neologisms to June 2022, which amounted to 21,605 words. Counts of types, frequency of occurrence, and *hápx legomena* were calculated from the corpus of neologisms for each word formation scheme contained in it. These calculations followed the methodology used in Morfolex since 2010 and detailed in a previous work (Zacarías Ponce de León, 2016), which reported the findings corresponding to ten thousand neologisms. The update presented here confirms that the Morfolex database is representative of the Mexican lexicon and presents an overview of the most productive word-formation processes in this variety of Spanish. At the end of the work, a comparison is made between the current data and those corresponding to 2016 to explore the evolution of different lexical schemes..

Keywords: Productivity; neologisms; Morfolex; *hápx legomena*; lexical morphology

## 1. Introducción

La morfología léxica del español mexicano ha sido motivo de estudio del proyecto Morfolex desde 2008 (Zacarías Ponce de León, 2013). El principal producto de este proyecto es una base de datos que almacena palabras neológicas e incluye el análisis morfológico de su estructura. Esta base de datos permite agrupar las palabras de acuerdo con el proceso lexicogénico que las formó. Los pormenores de esta base de datos, tanto sus fundamentos como sus procedimientos, pueden consultarse en un trabajo anterior (Zacarías Ponce de León, 2016).<sup>1</sup> El objetivo de este artículo no es presentar la base de datos Morfolex, sino una actualización de la productividad morfológica para un total de 21 605 palabras.

Las palabras que aparecen registradas una sola vez en el corpus son llamadas *hápax legomena*. Este tipo de palabras tienen más probabilidad de haber sido creadas por un proceso productivo que aquellas que aparecen en más ocasiones. Por eso, en Morfolex se utilizan para el cálculo de la productividad relativa. En la siguiente sección se explicará brevemente el tipo de datos que se obtienen del corpus y el cálculo de productividad basado en el conteo de *hápax legomena*.

## 2. Metodología

Esta metodología es homogénea, verificable y replicable. Es homogénea en cuanto a la recolección de neologismos, el análisis morfológico y el etiquetado en la base de datos. Los cálculos de productividad, por su parte, son verificables y simples: el conteo de *hápax legomena* en la base de datos. Finalmente, la metodología es replicable porque puede reproducirse puntualmente en cualquier investigación similar. En publicaciones anteriores de Morfolex esta metodología es descrita con detalle (Zacarías Ponce de León, 2016;

<sup>1</sup> Para consultar las características del proyecto, su metodología y productos, remitimos a los interesados al artículo mencionado.

Lucio Tejeda & Zacarías Ponce de León, 2017). Asimismo, en el sitio morfolex.com puede encontrarse una descripción del proyecto, datos y estadísticas del corpus y publicaciones relacionadas.

### 2.1. *Medición de la productividad*

A fin de contextualizar los datos que se presentan en esta investigación, haremos una breve exposición sobre los fundamentos del cálculo de la productividad mediante el uso de neologismos.

Según lo que señala Vallès Botey (2002), la productividad relativa de los esquemas lexicogénicos de una lengua debe ser medida en el uso real de los hablantes. En tal sentido, hablar de productividad en la morfología léxica debe referirse siempre a palabras de nuevo cuño o, en palabras de esta investigadora, “debe basarse fundamentalmente en neologismos” (2002: 140). Esto significa que un esquema de formación de palabras será productivo si los hablantes lo utilizan constantemente en la creación de nuevo léxico. Vallès Botey afirma que: “ceñirse a los neologismos para medir la productividad actual es coherente desde el punto de vista conceptual [ya que] coincide con la noción básica de productividad más generalizada [...] la habilidad de los hablantes para producir nuevas palabras” (2002: 149).

El otro aspecto importante de la metodología es el uso de palabras que ocurren solamente una ocasión en el corpus; los llamados *hápax legomena*. El método estadístico basado en *hápax* es bastante conocido (Baayen, 1992; Aronoff & Fudeman, 2011). La importancia de estas piezas léxicas es que, evidentemente, no pertenecen a un léxico consolidado y, por lo tanto, su existencia es evidencia de que el esquema lexicogénico con el que se formaron es rendidor y, en función del número de nuevas palabras formadas, puede ser considerado productivo.

### 3. Productividad relativa de los esquemas lexicogénicos

En el Cuadro 1 se presenta la productividad morfológica del español mexicano actualizada a junio de 2022. Se indican los afijos más relevantes en esta variedad de lengua. Se muestran también los esquemas compositivos más productivos, tanto de la composición patrimonial como de la composición culta; así como otros procesos misceláneos. Los conteos se realizaron por Tipos, Ocurrencias y *Hápax legomena*. Esta última columna es la que se utiliza para ordenar los esquemas lexicogénicos según su productividad.

CUADRO 1. Productividad relativa con datos de junio 2022

Proceso morfológico	Ocurrencias		Tipos		<i>Hápax legomena</i>	
	21605		12803		8968	
	Total	%	Total	%	Total	%
Sufijación nombre	<b>6516</b>	<b>30.2</b>	<b>3018</b>	<b>23.6</b>	<b>2023</b>	<b>22.558</b>
-ista	848	3.9	341	2.7	208	2.319
-ero	1094	5.1	356	2.8	207	2.308
-azo	851	3.9	304	2.4	194	2.163
-ear	653	3.0	273	2.1	162	1.806
-ismo	478	2.2	245	1.9	154	1.717
-ense	179	0.8	126	1.0	98	1.093
-izar	256	1.2	127	1.0	91	1.015
-ano	128	0.6	102	0.8	86	0.959
-gate	102	0.5	70	0.5	58	0.647
-ar	189	0.9	82	0.6	57	0.636
-ico	114	0.5	75	0.6	53	0.591
-ería	163	0.8	80	0.6	52	0.580
-esco	76	0.4	58	0.5	50	0.558
Sufijación adjetivo	<b>1136</b>	<b>5.3</b>	<b>679</b>	<b>5.3</b>	<b>491</b>	<b>5.475</b>
-mente	114	0.5	103	0.8	92	1.026
-dad	142	0.7	111	0.9	84	0.937

(continuación)

CUADRO 1. Productividad relativa con datos de junio 2022

Proceso morfológico	Ocurrencias		Tipos		<i>Hápax legomena</i>	
	21605		12803		8968	
	Total	%	Total	%	Total	%
<i>-ismo</i>	216	1.0	84	0.7	48	0.535
<i>-izar</i>	158	0.7	67	0.5	44	0.491
<i>-ista</i>	116	0.5	50	0.4	31	0.346
<i>-ez</i>	76	0.4	40	0.3	29	0.323
<i>-ear</i>	60	0.3	38	0.3	23	0.256
<i>-ada</i>	18	0.1	16	0.1	12	0.134
<i>-ísimo</i>	10	0.05	10	0.1	10	0.112
Sufijación verbo	<b>2613</b>	<b>12.1</b>	<b>1212</b>	<b>9.5</b>	<b>836</b>	<b>9.322</b>
<i>-ción</i>	469	2.2	224	1.7	154	1.717
<i>-dor</i>	376	1.7	196	1.5	143	1.595
<i>-nte</i>	229	1.1	146	1.1	111	1.238
<i>-o</i>	151	0.7	120	0.9	77	0.859
<i>-miento</i>	189	0.9	99	0.8	65	0.725
<i>-ble</i>	122	0.6	83	0.6	59	0.658
<i>-ada</i>	121	0.6	64	0.5	48	0.535
<i>-e</i>	127	0.6	41	0.3	27	0.301
<i>-dero</i>	60	0.3	36	0.3	26	0.290
<i>-torio</i>	65	0.3	25	0.2	18	0.201
Prefijación nombre	<b>2573</b>	<b>11.9</b>	<b>1651</b>	<b>12.9</b>	<b>1117</b>	<b>12.455</b>
<i>anti-</i>	496	2.3	309	2.4	217	2.420
<i>narco-</i>	452	2.1	225	1.8	154	1.717
<i>na-</i>	113	0.5	92	0.7	75	0.836
<i>ex-</i>	154	0.7	103	0.8	70	0.781
<i>re-</i>	171	0.8	95	0.7	57	0.636
<i>des-</i>	89	0.4	59	0.5	43	0.479

(cont.)

CUADRO 1. Productividad relativa con datos de junio 2022

Proceso morfológico	Ocurrencias		Tipos		Hápax legomena	
	21605		12803		8968	
	Total	%	Total	%	Total	%
<i>pro-</i>	45	0.2	42	0.3	38	0.424
<i>pre-</i>	108	0.5	55	0.4	35	0.390
<i>sub-</i>	113	0.5	55	0.4	33	0.368
Prefijación adjetivo	<b>896</b>	<b>4.147</b>	<b>645</b>	<b>5.038</b>	<b>450</b>	<b>5.018</b>
<i>anti-</i>	187	0.9	128	1.0	86	0.959
<i>in-</i>	76	0.4	48	0.4	32	0.357
<i>pro-</i>	29	0.1	25	0.2	20	0.223
<i>des-</i>	23	0.1	21	0.2	18	0.201
<i>inter-</i>	29	0.1	23	0.2	18	0.201
<i>pre-</i>	41	0.2	26	0.2	18	0.201
<i>pos-</i>	36	0.2	22	0.2	16	0.178
<i>ex-</i>	28	0.1	21	0.2	15	0.167
<i>extra-</i>	25	0.1	18	0.1	14	0.156
Prefijación verbo	<b>524</b>	<b>2.425</b>	<b>349</b>	<b>2.726</b>	<b>251</b>	<b>2.799</b>
<i>re-</i>	216	1.0	123	1.0	84	0.937
<i>des-</i>	122	0.6	81	0.6	59	0.658
<i>pre-</i>	32	0.1	20	0.2	11	0.123
<i>sobre-</i>	14	0.1	13	0.1	11	0.123
<i>co-</i>	14	0.1	11	0.1	9	0.100
<i>in-</i>	7	0.03	6	0.05	5	0.056
<i>con-</i>	2	0.01	2	0.02	2	0.022
Parasíntesis	<b>308</b>	<b>1.4</b>	<b>151</b>	<b>1.2</b>	<b>107</b>	<b>1.193</b>
<i>en-</i> <i>-ar</i>	176	0.8	70	0.5	46	0.513
<i>a-</i> <i>-ar</i>	72	0.3	37	0.3	31	0.346
<i>des-</i> <i>-ar</i>	30	0.1	21	0.2	14	0.156

(cont.)

CUADRO 1. Productividad relativa con datos de junio 2022

Proceso morfológico	Ocurrencias		Tipos		<i>Hápax legomena</i>	
	21605		12803		8968	
	Total	%	Total	%	Total	%
<i>des-izar</i>	20	0.1	14	0.1	8	0.089
<i>en-ecer</i>	6	0.03	6	0.05	6	0.067
<i>en-izar</i>	1	0.005	1	0.01	1	0.011
<i>a-izar</i>	3	0.01	2	0.02	1	0.011
Composición culta	<b>2503</b>	<b>11.6</b>	<b>2015</b>	<b>15.7</b>	<b>1322</b>	<b>14.741</b>
<i>mega</i>	289	1.3	200	1.6	131	1.461
<i>super</i>	206	1.0	177	1.4	118	1.316
<i>multi</i>	160	0.7	122	1.0	65	0.725
<i>mini</i>	88	0.4	77	0.6	63	0.702
<i>micro</i>	129	0.6	101	0.8	58	0.647
<i>bio</i>	111	0.5	91	0.7	57	0.636
<i>ciber</i>	181	0.8	119	0.9	55	0.613
<i>semi</i>	84	0.4	77	0.6	51	0.569
<i>logo</i>	81	0.4	60	0.5	40	0.446
<i>metro</i>	63	0.3	45	0.4	34	0.379
Composición patrimonial	<b>2510</b>	<b>11.6</b>	<b>1559</b>	<b>12.18</b>	<b>1187</b>	<b>13.236</b>
N+N	830	3.8	632	4.9	489	5.453
V+N	947	4.4	472	3.7	366	4.081
N+Adj	183	0.8	129	1.0	102	1.137
Adj+Adj	148	0.7	116	0.9	90	1.004
N+de+N	63	0.3	46	0.4	36	0.401
Adj+N	77	0.4	56	0.4	34	0.379
Adj+i+Adj	81	0.4	31	0.2	17	0.190
Prep+N	64	0.3	21	0.2	14	0.156



(cont.)

CUADRO 1. Productividad relativa con datos de junio 2022

Proceso morfológico	Ocurrencias		Tipos		Hápax legomena	
	21605		12803		8968	
	Total	%	Total	%	Total	%
V+Adv	17	0.1	14	0.1	13	0.145
N++Adj	21	0.1	16	0.1	12	0.134
Traslapamiento*	954	4.4	781	6.1	611	6.813
Acortamiento	174	0.8	116	0.9	84	0.937
Acortamiento compuesto	747	3.5	486	3.8	363	4.048
Acronimia y siglación	143	0.7	133	1.0	120	1.338

\* También conocidos como cruces o *blends*. Se refiere a palabras del tipo burrocracia, cenicebrijo, tuitertura, suaperro, etcétera.

El cuadro anterior representa la productividad relativa de los procesos de formación léxica más relevantes en el español de México. Se desglosan los procesos afijales distinguiendo entre sufijos, prefijos y circunfijos. Cada una de estas agrupaciones se desglosa a su vez por afijo. Solo se presentan los afijos más productivos. De la misma manera, se desglosan los diferentes tipos de composición y también otros esquemas formativos menos relevantes.

### 3.1. Estadísticas por esquema lexicogenésico

Se presenta ahora los resúmenes de los esquemas lexicogenésicos más importantes, sufijación, prefijación y composición, con el objetivo de ofrecer una visión más general. En cuanto a la sufijación (Cuadro 2), la más productiva, y de manera considerable, es la de nombre; le sigue la de verbo y por último la de adjetivo. El panorama cambia con la prefijación (Cuadro 3). De nuevo, es la de sustantivo la más productiva, pero la de adjetivo se coloca en segundo

lugar, dejando la prefijación de verbo como menos productiva. En general, observamos que la sufijación total (37.3%) es mucho más productiva que la prefijación total (20.3%). Respecto a la composición patrimonial (13.2%) y a la composición culta (14.7%), como era de esperarse, se colocan en un nivel inferior al de la prefijación y al de la sufijación (Cuadro 4). Destacan en este caso, los dos tipos de composición que se utilizan más en español, el N+N y el de V+N, con niveles semejantes de productividad.

CUADRO 2. Productividad de la sufijación

	Ocurrencias		Tipos		<i>Hápax legomena</i>	
	Total	%	Total	%	Total	%
Sufijación nombre	6516	30.2	3018	23.6	2023	22.558
Sufijación adjetivo	1136	5.3	679	5.3	491	5.475
Sufijación verbo	2613	12.1	1212	9.5	836	9.322
<b>Totales</b>	<b>10 265</b>	<b>47.5</b>	<b>4 909</b>	<b>38.3</b>	<b>3 350</b>	<b>37.3</b>

CUADRO 3. Productividad de la prefijación

	Ocurrencias		Tipos		<i>Hápax legomena</i>	
	Total	%	Total	%	Total	%
Prefijación nombre	2573	11.9	1651	12.9	1117	12.455
Prefijación adjetivo	896	4.147	645	5.038	450	5.018
Prefijación verbo	524	2.425	349	2.726	251	2.799
<b>Totales</b>	<b>3 993</b>	<b>18.5</b>	<b>2 645</b>	<b>20.6</b>	<b>1 818</b>	<b>20.3</b>

CUADRO 4. Productividad de la composición patrimonial

	Ocurrencias		Tipos		<i>Hápx legomena</i>	
	Total	%	Total	%	Total	%
N+N	830	3.8	632	4.9	489	5.453
V+N	947	4.4	472	3.7	366	4.081
N+Adj	183	0.8	129	1.0	102	1.137
Adj+Adj	148	0.7	116	0.9	90	1.004
N+de+N	63	0.3	46	0.4	36	0.401
Composición	<b>2510</b>	<b>11.6</b>	<b>1559</b>	<b>12.18</b>	<b>1187</b>	<b>13.24</b>

En el Cuadro 5 se comparan todos los procesos de formación de palabras registrados en Morfolex; se incluye la composición culta, los cruces léxicos, la acronimia, los acortamientos y las siglas. Se presentan primero los casos de afijación, posteriormente los de composición, luego los de cruce léxico y, por último, el resto de los procesos.

CUADRO 5. Productividad global por esquema lexicogenésico

	Ocurrencias		Tipos		<i>Hápx legomena</i>	
	Total	%	Total	%	Total	%
Sufijación	10265	47.5	4909	38.3	3350	37.3
Prefijación	3993	18.5	2645	20.6	1818	20.3
Circunfijación	308	1.4	151	1.2	107	1.2
Composición patrimonial	2510	11.6	1559	12.18	1187	13.2
Composición culta	2503	11.6	2015	15.7	1322	14.8
Traslapamiento	954	4.4	781	6.1	611	6.8
Acortamiento	174	0.8	116	0.9	84	0.94
Acortamiento compuesto	747	3.5	486	3.8	363	4.04
Acrónimos y siglas	143	0.7	133	1.0	120	1.34

El panorama de la morfología léxica puede apreciarse mejor de manera gráfica. En la Figura 1 se incluyen todos los procesos. Para mayor claridad, solo se grafica la columna correspondiente a la productividad calculada mediante las palabras *hápax legomena*.

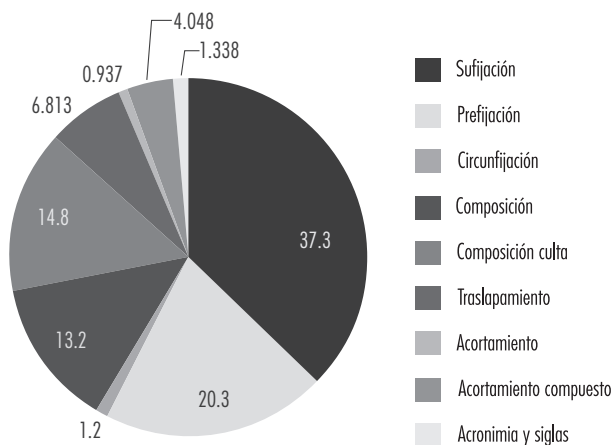


FIGURA 1. Productividad global por esquema lexicogenésico

#### 4. Comparación de productividad 2016–2022

Ahora se presenta la comparación entre las estadísticas del 2022 y las estadísticas previas, que corresponden al 2016 (Cuadro 6). Se comparan los procesos generales a nivel de *hápax legomena*, es decir, el grado de productividad relativa.

CUADRO 6. Comparativo de productividad 2016–2022

	<i>Hápax legomena</i> 2016 (4618)		<i>Hápax legomena</i> 2022 (8968)	
	Total	%	Total	%
Sufijación	1738	37.6	3350	37.3
Prefijación	1198	25.9	1818	20.3
Circunfijación	67	1.45	107	1.2

(continuación)

CUADRO 6. Comparativo de productividad 2016–2022

	<i>Hápax legomena</i> 2016 (4618)		<i>Hápax legomena</i> 2022 (8968)	
	Total	%	Total	%
Composición	1047	22.7	2509	28.0
Traslapamiento	263	5.7	611	6.8
Acortamiento	50	1.1	84	0.94
Acortamiento compuesto	167	3.6	363	4.04
Acónimos y siglas	88	2.0	120	1.34

Al analizar el cuadro anterior se comprueba que los principales procesos de formación de palabras, la afijación y la composición, se mantienen como los más productivos. Sin embargo, se nota un cambio relevante: la composición en general, sumando la patrimonial y la culta, se ha robustecido, al pasar de un grado de productividad de 22.7 a 28. Esta alza en la productividad de la composición trae como consecuencia una disminución en la productividad relativa en el resto de los procesos. Esto es particularmente cierto para la prefijación, que pasó de ser el segundo proceso en orden de productividad a ser el tercero, cediendo su lugar a la composición. Una posible explicación al aumento de productividad de la composición es el contexto de la pandemia por COVID-19, ya que este acontecimiento generó gran cantidad de neologismos (Zacarías Ponce de León, 2022), muchos de ellos formados con los sustantivos *COVID* (*coviflu*, *covicumpleaños*) y *corona* (*coronabicho*, *coronabobo*). En cuanto al proceso más productivo de formación de palabras, la sufijación se mantiene claramente en el primer lugar, con una pequeña disminución en la estadística reciente.

## 5. Observaciones finales

Las estadísticas presentadas en esta nota nos proporcionan un panorama de la morfología léxica del español de México a través del estudio de la productividad relativa de los esquemas de formación de palabras. El conteo de palabras *hápax legomena* ha sido un parámetro importante en la configuración del corpus Morfolex, ya que mediante este es viable establecer un grado de productividad relativa. Uno de los objetivos de este proyecto ha sido la difusión de los cálculos de productividad relativa de nuestra variedad de lengua. La metodología del proyecto para el cálculo de la productividad (Zacarías Ponce de León, 2016), a partir de una base de datos neológica, es la única hasta el momento que ofrece la productividad de todos los esquemas léxicos de manera relativa. Aunado a esto, el acopio constante de datos permite comparar dos momentos diferentes en el corpus para observar los cambios en productividad. Esto es justamente lo que se ha presentado en este trabajo, la comparación entre los datos de 2016 con los de 2022. Consideramos que, como resultado de dicha comparación, es posible afirmar que la base de datos Morfolex ofrece un retrato adecuado del comportamiento de la morfología léxica del español mexicano.

## 6. Referencias

- Aronoff, Mark, & Fudeman, Kirsten (2011). *What is morphology?* Malden: Wiley-Blackwell.
- Baayen, R. Harald (1992). Quantitative aspects of morphological productivity. En Geert Booij & Jaap van Marle (Eds.), *Yearbook of morphology 1991* (pp. 109–149). Dordrecht: Kluwer Academic.
- Lucio Tejeda, Sofía, & Zacarías Ponce de León, Ramón (2017). Actualización del corpus Morfolex a partir de la 23ª edición del Diccionario de la Lengua Española. *Estudios de Lingüística Aplicada*, 66, 109–134.
- Vallès Botey, Teresa (2002). La productividad morfológica en un modelo dinámico, basado en el uso y en los usuarios. En M. Teresa Cabré, Judit Freixa & Elisabet Solé (Eds.), *Lèxic i neologia* (pp. 139–157). Bar-

celona: Observatori de Neologia, Institut Universitari de Lingüística Aplicada, Universitat Pompeu Fabra.

Zacarías Ponce de León, Ramón (2013). El buzón buscapalabras. Procesos de formación de neologismos. *Anuario de Letras Hispánicas. Glosas Hispánicas* (Vol. 2, pp. 81–89). Ciudad de México: Facultad de Filosofía y Letras, Universidad Nacional Autónoma de México.

Zacarías Ponce de León, Ramón (2016). Morfología léxica en el español actual de México: neología y productividad. *Estudios de Lingüística Aplicada*, 64, 11–31.

Zacarías Ponce de León, Ramón (2022). Neologismos relacionados con la pandemia de COVID-19 en México. *Lingüística Mexicana. Nueva Época*, 4(1), 7–29.



