



Mutaciones en el acervo léxico del español de la ciudad de México (1997-2022). Un estudio de panel con materiales del corpus PRESEEA

Lexical mutations in Mexico City Spanish (1997-2022): a panel study based on the PRESEEA corpus

Dinorah Beatriz Montiel Pérez
U. Aut. Metropolitana-Iztapalapa

Julio César Serrano Morales
U. Aut. Metropolitana-Iztapalapa

Abstract

In order to determine lexical-semantic changes in Mexico City Spanish, a panel study was conducted with 19 individuals interviewed in the year 2022. These individuals are part of the original PRESEEA - Mexico City corpus, collected between 1997 and 2006. The average time gap between the original and current interviews is approximately 20 years. An analysis of lexicographically-marked vocabulary was conducted. The explored marks included: popular, vulgar, cultured tradition, and loans. Words marked with these characteristics were searched within the interview texts. Nule-hypothesis testing and correspondence analysis were carried out to examine the dependency between the explored lexical categories and sociolinguistic variables such as age, gender, and education level. The results confirm variable phenomena such as age grading, individual change, sociolinguistic leveling, and a shift in attitude towards certain lexical categories.

Keywords: lexical change, lexical-statistics, lexicographic usage labels, correspondence analysis, panel study.

Resumen

Con el fin de determinar cambios léxico-semánticos en español de la Ciudad de México, se realizó un estudio de panel con 19 personas entrevistadas en el año 2022; estas personas forman parte del corpus original PRESEEA - Ciudad de México, levantado entre 1997 y 2006. Entre las entrevistas originales y las actuales existe una distancia temporal promedio de 20 años. Se decidió hacer un análisis del vocabulario que suele ser marcado lexicográficamente en los diccionarios. Las marcas exploradas fueron: popular, grosero, tradición culta y préstamos. Se llevaron a cabo pruebas estadísticas (hipótesis nula y análisis de correspondencia) que permitieran conocer la dependencia entre las categorías léxicas exploradas y las variables sociolingüísticas de edad, sexo y nivel de instrucción. Los resultados confirman fenómenos tales como estratificación por edad, cambio individual, nivelación sociolingüística y un cambio de actitud hacia determinadas categorías léxicas.

Palabras clave: cambio léxico, léxico-estadística, marcas lexicográficas de uso, análisis de correspondencia, estudio de panel.

Citar como: Montiel Pérez, Dinorah Beatriz; Serrano Morales, Julio César (2024). Mutaciones en el acervo léxico del español de la ciudad de México (1997-2022). Un estudio de panel con materiales del corpus PRESEEA. Normas, 14(1), 130-140, doi: 10.7203/Normas.v14i1.28652.

1. Introducción

En las siguientes páginas se abordará el cambio léxico a nivel individual y comunitario en un panel de 19 personas que fueron entrevistadas para el corpus original PRESEEA - Ciudad de México (1997-2006). En particular se analizan vocablos con marcas de uso lexicográficas: *popular*, *grosería*, *préstamo* y *tradición culta*.

Aunque existen algunos antecedentes sobre cambio léxico en español mexicano a partir de

materiales sociolingüísticos (Serrano, 2014; Rivera, 2015; Montiel, 2019), hasta el momento no hay una sola investigación que incorpore un panel de entrevistados con 20 años de distancia temporal. Tras el análisis de 31 entrevistas, se efectuaron pruebas de hipótesis y análisis de correspondencias para determinar si existe correlación o dependencia entre las categorías léxicas exploradas y las variables sociales de edad, sexo y nivel de instrucción.

El artículo se organiza de la siguiente manera: además de esta introducción, se presenta una sección sobre cambio léxico y cómo se ha abordado desde la perspectiva variacionista, lo que permitirá plantear un marco teórico y metodológico sobre el cambio léxico. A continuación, se detalla la metodología, los procedimientos para la selección del léxico y el perfil social de los informantes que formaron el panel. En las tres secciones subsecuentes se revisa la correlación de las categorías léxicas con la edad, el sexo y el nivel de instrucción de los participantes. La última sección corresponde a la discusión de resultados y las conclusiones pertinentes.

2. Cambio léxico y variación: un estudio de panel

El cambio léxico, estudiado desde la sociolingüística, ha tenido poca fecundidad debido al conflicto con la definición laboviana de variable: dos maneras, o más, de decir lo mismo (Labov, 1973). Esto se debe a que el léxico es una categoría abierta, a diferencia de los fonemas o morfemas, cuyo listado finito permite reducir los intercambios entre una variable y otra en determinados contextos.

Escoriza Morera (2017) ha propuesto una forma de análisis de la variación léxica que reduce los contextos (en términos de formalidad) y controla el léxico que se puede utilizar a partir de escritos que los informantes deben completar con las palabras que mejor convengan. Otros estudios de variación en el plano léxico contemplan la agrupación de éste en categorías gramaticales. Rivera (2015) hace un estudio en tiempo real en español de la Ciudad de México donde explora los adjetivos calificativos y obtiene resultados cuantitativos que inciden en cambios sobre el discurso y las formas de describir la realidad entre los dos momentos estudiados. Serrano (2011, 2014), por su parte, agrupó el léxico en categorías que le permitieran distinguir usos innovadores y neologismos; entre sus clasificaciones se encuentran habla juvenil y léxico referido a la tecnología. Con estos antecedentes en mente se planteó realizar un estudio de cambio léxico a través de un estudio de panel que permitiera constatar cambios comunitarios y cambio individual.

Los estudios de panel todavía son una *rara avis* en sociolingüística ya que, para realizarlos, los investigadores se enfrentan a distintos problemas que no existen en los análisis en tiempo aparente. Un estudio de panel consiste en realizar dos pruebas (encuestas, entrevistas, cuestionarios) a la misma muestra, pero con una distancia temporal adecuada para observar cambios entre ambos momentos. La hipótesis detrás es que los hablantes modifican algunos hábitos lingüísticos mientras que otros se mantienen *arraigados* (Buchstaller *et al.*, 2019). De igual forma, aunque el estudio de panel suele enfocarse en el cambio individual, los cambios mostrados en los individuos también pueden estar reflejando mutaciones sociolingüísticas comunitarias. Los estudios de panel, además, permiten la verificación de las observaciones hechas en el *tiempo aparente* (Labov, 1996). Por supuesto, una de las grandes limitaciones de los estudios de panel es que se enfrentan a la baja incidencia de hablantes para el segundo momento, debido a la movilidad social y a la dificultad de generar las mismas condiciones del

primer acercamiento (Cieri *et al.*, 2019).

Para el caso de la presente investigación, se tomó como punto de partida el corpusPRESEEA-Ciudad de México que, entre 1997 y 2006, recogió alrededor de 300 entrevistas (*Corpus Sociolingüístico de la Ciudad de México* [CSCM], Butragueño *et al.*, 2011), por lo que la diferencia temporal entre las primeras entrevistas y las contemporáneas es de 20 años en promedio. A las del CSCM las denominaremos “del año 2000” y a las contemporáneas del “año 2022”, para afectos expositivos.

3. Metodología

A partir de la base de datos de informantes del CSCM se trató de encontrar a la mayor cantidad de hablantes posible para volver a entrevistarlos. En total se localizó a 19 hablantes del corpus original y fueron entrevistados todos durante el año 2022. Además de estos, se realizaron nuevas entrevistas con 12 hablantes jóvenes (también durante 2022) para tener representación de los tres grupos etarios originales. De esta manera el año 2000 está representado por 19 hablantes y el año 2022 por 31 hablantes (v. Tablas 1-2).

2000				2022			
Nivel de instrucción alta				Nivel de instrucción alta			
	Jóvenes	Adultos	Mayores		Jóvenes	Adultos	Mayores
Mujeres	ME-002-31M-97	ME-255-32M-05	ME-213-33M-03	Mujeres	ME-313-31M-22	ME-002-32M-22	ME-255-33M-22
	ME-008-31M-97		ME-242-33M-04		ME-311-31M-22	ME-008-32M-22	ME-213-33M-22
	ME-107-31M-00		ME-264-33M-05		ME-107-32M-22	ME-242-33M-04	ME-264-33M-05
Hombres	ME-197-31H-01	ME-006-32H-97	ME-009-33H-97	Hombres	ME-317-31H-22	ME-197-31H-21	ME-006-33H-22
		ME-254-32H-05	ME-156-33H-01		ME-310-31H-22		ME-254-33H-21
			ME-245-33H-05				ME-009-33H-97
						ME-156-33H-01	ME-245-33H-05
Nivel de instrucción media				Nivel de instrucción media			
Mujeres		ME-005-22M-97		Mujeres	ME-316-21M-22		ME-005-22M-97
		ME-026-22M-98			ME-321-21M-22		ME-026-22M-98
		ME-110-22M-00		Hombres	ME-315-21H-22	ME-106-22H-21	
Hombres	ME-106-21H-00	ME-263-22H-05		ME-318-21H-22	ME-263-22H-22		
Nivel de instrucción baja				Nivel de instrucción baja			
Hombres			ME-262-12H-05	Hombres	ME-312-11M-22		
					ME-319-11M-22		
				Hombres	ME-320-11H-22		ME-262-12H-05
					ME-322-11H-22		

Tablas 1 y 2. Hablantes de los años 2000 y 2022. El código de los informantes es el mismo que se usó en el CSCM. Las entrevistas en sombreado representan a los 12 informantes jóvenes que no formaban parte del CSCM.

La mayoría de las entrevistas se llevaron a cabo por *Zoom*, salvo cinco de ellas, donde los participantes optaron por la entrevista presencial¹. Se realizó la transcripción siguiendo las

¹Las entrevistas se hicieron durante la pandemia por Covid 19. Se preguntó a los informantes si preferían la entrevista a distancia a presencial, quienes la escogieron presencial, fue por motivos de trabajo, ya que no

normas de PRESEEA (Butragueño y Lastra, 2011). La lectura completa de cada transcripción permitió seleccionar unidades léxicas que conforman el corpus base de esta investigación. A continuación, se presentan los detalles correspondientes a las marcas lexicográficas que sirvieron de guía para organizar el vocabulario obtenido y los diccionarios usados como referencia.

3.1. Marcas lexicográficas utilizadas

Esta investigación busca entender los procesos de cambio léxico-semántico del vocabulario marcado lexicográficamente como *grosero*, *popular*, *culto*, etc. y su distribución comunitaria en términos de la edad, nivel de instrucción y sexo de los hablantes. La hipótesis central es que estos cambios pueden ser reflejos lingüísticos de mutaciones a nivel social. Para ello nos valimos de algunos diccionarios donde se utilizan marcas lexicográficas de las palabras seleccionadas. Estos fueron:

1) *Diccionario de la lengua española* (DLE). Como es sabido, es un diccionario que pretende abarcar el léxico de los distintos países de habla hispana, y sin embargo indica también algunos usos geográficamente diferenciados. Se decidió utilizar sus marcas por ser el diccionario más general de habla hispana y para poder contrastar con diccionarios especializados en español mexicano.

2) *Diccionario del Español de México* (DEM, 2010). Este diccionario, creado por El Colegio de México, es la obra que mejor refleja la semántica léxica del español utilizado en las entrevistas del CSCM. Es importante señalar que el DEM no es un diccionario de mexicanismos, sino un recuento del español integral tal como se habla y escribe en México.

3) *Diccionario de Mexicanismos* (2023) (DM). Esta obra, publicada por la Academia Mexicana de la Lengua, sí responde a usos muy particulares del español de México. Además, es un diccionario sincrónico en la medida que sólo aparecen lemas que tienen vigencia en la actualidad.

Con estas tres obras se cubre cada uno de los tipos de diccionarios que la lexicografía distingue (Parada Valdés y Academia Mexicana Correspondiente de la Española, 2010): un diccionario general (DLE), uno especial integral (DEM) y uno especial diferencial (DM).

Las marcas de uso consideradas para este trabajo fueron las siguientes:

1. *Popular*: Esta categoría está empatada con la definición del DEM sobre la tradición popular: palabras o locuciones de uso diario que se usan en el ámbito informal. En DM, existen dos marcas al respecto: *popular* y *coloquial*. La diferencia entre ellas estriba que las *populares* se relacionan con un bajo nivel de instrucción; aquí no se utilizará esta diferenciación, pues se considera que estas formas son conocidas por el grueso de la población y responden a la historicidad de la lengua. Ejemplos de estas categorías son: *a todo dar*, *ahí muere*, *alivianar*.
2. *Grosería*: Se prefirió el término *grosería* frente a la marca *vulgar* (propias del DEL y DM, respectivamente), por considerar que ‘vulgar’ es una especificación aún más subjetiva que ‘grosería’. Un término puede o no ser vulgar para un informante, pero una grosería

tenían tiempo fuera de este contexto para realizar la entrevista y se acudió a sus espacios laborales

pertenece a un tipo de palabra que busca ofender o provocar mayor énfasis, por ejemplo: *valer madre, verga, puto*.

3. *Tradición culta*: Estas palabras corresponden a lo definido como ‘tradición culta’ por el DEM. Es léxico que no se marca pero que es utilizado por hablantes con cierto grado de preparación, lo que los distingue de los hablantes legos, como en los vocablos *adocenado, contexto, anacrónico*, entre otras.
4. *Préstamos*: Esta etiqueta se creó para conocer la cantidad de préstamos de lenguas dominantes (el inglés principalmente) que se han incorporado en años recientes. Los préstamos pueden ser palabras ya plenamente integradas al español, como *ok*, o bien palabras que el informante decidió utilizar en el momento de la entrevista en lugar de un término en español, como *open mind* (‘de mente abierta’).

Estas marcas permitieron seleccionar, de las 50 entrevistas, un total de 371 unidades léxicas que produjeron 2392 ocurrencias. Las unidades léxicas del año 2000 fueron 179, con 817 repeticiones (ocurrencias), mientras que en el año 2022 se registraron 296 unidades y 1575 ocurrencias.

	<i>Unidades léxicas</i>				<i>Ocurrencias</i>			
	2000	2022	Total		2000	2022	Total	
Popular	96	123	158	44%	506	773	1279	53%
Grosería	31	58	64	19%	191	541	732	31%
Préstamos	32	63	81	21%	97	160	257	11%
T. culta	20	52	68	15%	23	101	124	5%
	179	296	371	100%	817	1575	2392	100%

Tabla 3. Unidades léxicas y ocurrencias totales

3.2. Pruebas de hipótesis nula y análisis de correspondencia

Con estos datos se efectuaron pruebas de ji-cuadrada para determinar si existe correlación entre los variables sociales *edad, sexo y nivel de instrucción* con las categorías propuestas. La prueba de ji-cuadrada permitió a su vez realizar un *análisis de correspondencia* el cual, de existir dependencia entre variables y categorías, muestra la asociación entre ellas. Específicamente, el análisis de correspondencias

[e]s una técnica descriptiva o exploratoria cuyo objetivo es resumir una gran cantidad de datos en un número reducido de dimensiones, con la menor pérdida de información posible. En esta línea, su objetivo es similar al de los métodos factoriales, salvo que en el caso del análisis de correspondencias el método se aplica sobre variables categóricas u ordinales.

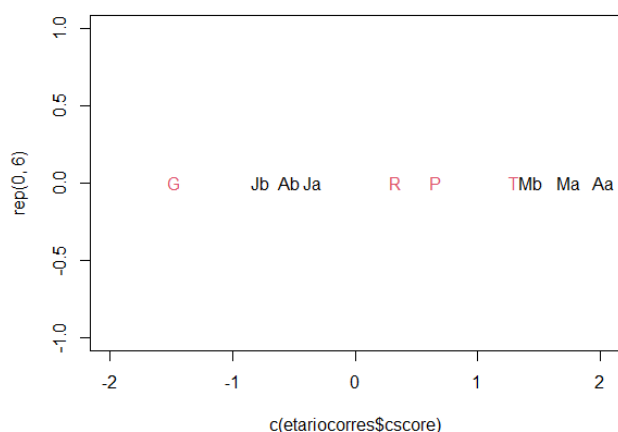
(De la Fuente Fernandez, 2011: 3)

Para poder efectuar el análisis de correspondencia, se tomó en cuenta la cantidad de ocurrencias, dado que al contar las unidades léxicas se traslapan entre algunas categorías, es decir hay léxico popular compartido por toda la muestra y que impide llegar a un 100%. En cambio, las

ocurrencias facilitan esta operación al poderlas contabilizar en su totalidad y evitar traslapes. En la siguiente sección se presentan los resultados de los análisis cuantitativos en el corpus tomando las variables sociales como guía.

4. La variable edad

Los grupos etarios se definieron de acuerdo con lo establecido en el CSCM, a saber: *jóvenes*: de 18 a 34 años, b) *adultos*: de 35 a 54 años y c) *mayores*: de 55 años en adelante. Al realizar la prueba de ji-cuadrada entre edad y categoría léxica, el *p-value* resultante fue de $< 2.2e-16$, por lo que se rechaza que las variables sean independientes. El análisis de correspondencia muestra cómo se asocia esta variable sociolingüística con las categorías léxicas. En la Gráfica 1 se ilustra cómo el léxico marcado como *grosero* se opone por completo al de *tradición culta*. El léxico marcado como *préstamo* y *popular* parecen más cercanos al vocabulario de tradición culta que a las groserías. Los datos etarios muestran que los jóvenes en 2022 son por completo distintos a los adultos del año 2000. Los hablantes jóvenes de ambos momentos parecen seguir las mismas tendencias léxicas, ya que no están tan alejados; como muestra la gráfica, justo entre ambos grupos se encuentran los adultos de 2022, quienes fueron jóvenes en el año 2000. Debe destacarse que la juventud, en ambos periodos, ha favorecido las groserías, dicho vocabulario ha incrementado en la actualidad, y ello se confirma con el hecho de que los jóvenes del año 2000 han aumentado estos porcentajes en la entrevista de 2022. En cuanto a los grupos etarios se pueden observar dos bloques confrontados. Por un lado, se encuentra el bloque de los jóvenes de los periodos junto con los adultos contemporáneos. Por otro lado, el segundo bloque está conformado por los adultos del año 2000 y los mayores de los dos periodos. Los jóvenes del 2000 no emplean el léxico de la misma manera que los adultos de su mismo periodo, en cambio continuaron con un estilo juvenil donde los jóvenes de 2022 también se encuentran representados. El otro bloque nos muestra que los adultos del 2000 continuaron el estilo léxico de los mayores.



Gráfica 1. Análisis de correspondencia entre cuatro categorías léxicas (en color rosa) y la variable *edad* (en negro). (Para leer la presente gráfica y las siguientes: P= categoría léxica *popular*, G= *grosería*, R= *préstamos*, T= *tradición culta*; J= *jóvenes*, A=*adultos*, M= *may-*

ores; las letras *b* y *a* que acompañan a las siglas de edad refieren al momento de la entrevista: a= año 2000, b= año 2022).

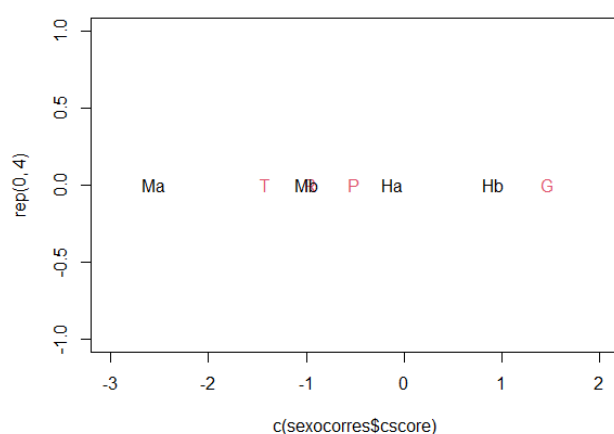
Los hablantes mayores de ambos momentos también tienen un comportamiento léxico similar, aunque el léxico de tradición culta está más asociado a los mayores del presente. Las palabras populares y los préstamos también son más cercanos al habla de los mayores del año 2022 que a otros grupos etarios.

5. La variable sexo

En el caso de las diferencias entre hombres y mujeres también se observan cambios interesantes, aunque ha de resaltarse el hecho de que para el año 2000 hay 10 mujeres y 9 hombres, mientras que en la muestra de 2022 hay 16 mujeres y 15 hombres. El *p-value* resultado de la ji-cuadrada fue de $<2.2e16$, por lo que se acepta que hay dependencia entre el sexo y las categorías léxicas.

La distribución de la variable sexo, de acuerdo con el análisis de correspondencia (v. Gráfica 2), coloca a las mujeres del año 2000 en completa oposición a los hombres del año 2022, mientras que las mujeres actuales se aproximan al léxico de los hombres del pasado (año 2000). La distribución de las categorías léxicas vuelve a mostrar que las palabras marcadas como groserías constituyen la categoría más alejada de las otras tres.

En general, la Gráfica 2 muestra que el léxico de las mujeres en 2000 cambió mucho respecto al levantamiento de 2022. Las mujeres en el presente se asocian mejor con el léxico de tradición culta y los préstamos; mientras que en el pasado apenas favorecieron el léxico de tradición culta. Los cambios entre los hombres también son interesantes. En el año 2000 eran más cercanos a la tradición culta, al léxico popular y a los préstamos, mientras que en la actualidad favorecen las groserías.



Gráfica 2. Análisis de correspondencia entre cuatro categorías léxicas y la variable *sexo*. En color negro, M corresponde a *mujeres* y H a *hombres*.

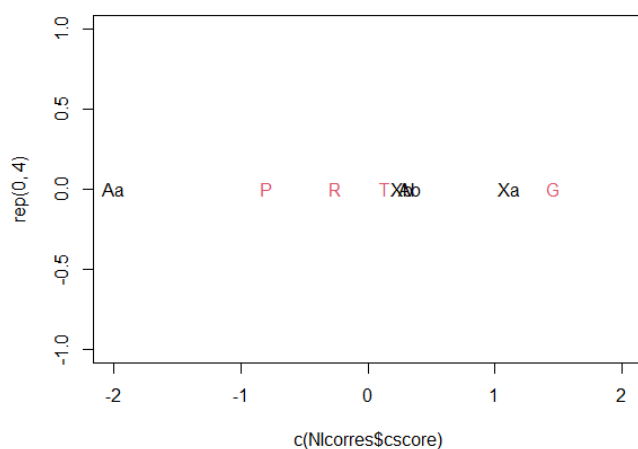
La Gráfica 2 permite apreciar que el léxico de las mujeres de la actualidad se ha aproximado al de los hombres del pasado. Aunque la brecha sociolingüística entre hombre y mujeres sigue

siendo amplia, parece que las mujeres tienden a aproximarse al léxico masculino con el paso de los años. Este resultado es compatible con lo que ha reportado Rivera (2015) sobre un acercamiento léxico-semántico entre hombres y mujeres en el campo de los vocablos adjetivos.

6. La variable nivel de instrucción

Esta variable se tuvo que reconfigurar, debido a que sólo se pudo entrevistar en el año 2022 a un hablante del nivel de instrucción baja. De acuerdo con el CSCM, el nivel de instrucción alta (NIA) incluía a personas desde nivel de licenciatura (trunco o concluido) hasta el nivel de posgrado. El nivel de instrucción media (NIM) contemplaba a hablantes con secundaria terminada (12 años de escolaridad) hasta aquellos que hubieran ingresado o terminado el bachillerato. Finalmente, el nivel de instrucción baja (NIB) contemplaba a personas sin ningún tipo de instrucción hasta los que habían concluido los estudios básicos (primaria, equivalente a 6 años de escolaridad).

En el año 2000 hay 13 personas del NIA, cinco del NIM y sólo una del NIB. Ante esta disparidad numérica, se optó por englobar a los informantes con instrucción media y baja en un solo grupo (NIMB), en ambos momentos; de esta manera, la distribución por niveles de instrucción en el año 2022 fue la siguiente: 17 informantes del NIA y 12 del NIMB.



Gráfica 3. Análisis de correspondencia entre cuatro categorías léxicas y la variable *nivel de instrucción*. En color negro, A corresponde al nivel de instrucción alto y X a los otros niveles de instrucción.

Al efectuar la ji-cuadrada para conocer si existía relación entre los niveles de instrucción de los dos momentos y las categorías léxicas se obtuvo un *p-value* de $< 2.2e-16$, por lo que se acepta que existe dependencia entre esta variable y las cuatro categorías léxicas del estudio.

El análisis de correspondencias mostró que la mayor oposición se encuentra entre los niveles de instrucción NIA y NIMB del año 2000. Los niveles de instrucción en la actualidad no presentan diferencias tan acusadas, es decir, parece que la brecha sociolingüística entre ellos (léxico-semántica, específicamente) parece disminuir notablemente. Esto ya había sido observado

por Serrano (2014) respecto al léxico sustantivo y una serie de variables morfológicas en español de la Ciudad de México.

Las categorías léxicas se distribuyen de manera semejante respecto a las variables anteriores: las groserías se encuentran siempre muy alejadas de las tres restantes. El léxico popular, los préstamos y las palabras de tradición culta son cercanas a los dos niveles de instrucción del presente.

Las groserías están más cercanas al NIMB en el pasado que en el presente y el NIA incrementó considerablemente su acercamiento a dicha categoría con el paso del tiempo, lo que sugiere, nuevamente, un acercamiento léxico en la actualidad entre los hablantes de niveles de instrucción diferenciados, al menos respecto al uso que hacen del léxico interdicto.

7. Discusión y conclusiones

El análisis estadístico del vocabulario marcado lexicográficamente en español de la Ciudad de México, a partir de datos de entrevistas sociolingüísticas con un panel de 19 personas, permitió observar cómo ciertas unidades léxicas mutan de manera significativa respecto a su correlación con rasgos sociales, ya que presentan movimientos importantes en el tiempo que pueden sugerir un cambio de actitud ante unidades léxicas populares, groseras, ante los préstamos y palabras de la tradición culta.

Se pudieron identificar también fenómenos interesantes como la estratificación por edad. Por ejemplo, los jóvenes no modificaron sus preferencias léxicas, por el contrario, las incrementaron en un lapso de 20 años. Por su parte, en el grupo de los adultos se identificó una retención importante de su habla de juventud, especialmente de los préstamos. Aunque se observa una ruptura generacional, tanto los jóvenes del año 2000 como los adultos de 2022 retuvieron su distribución del léxico marcado lexicográficamente, mientras que los adultos de 2000 se asimilan más a las distribuciones de los hablantes mayores. Se podría concluir entonces que el habla de la juventud define el perfil léxico de los individuos, al menos hasta su etapa adulta. Por supuesto, habrá que hacer más investigaciones, con otro tipo de vocabulario, para confirmar estos hallazgos.

Respecto al contraste entre hombres y mujeres, se encontró que los hombres son los líderes del cambio léxico en las categorías seleccionadas, pues las mujeres del año 2022 presentan un patrón léxico similar al de los hombres en el pasado (año 2000). Con todo, hoy en día la diferencia entre el léxico de hombres y mujeres sigue siendo amplia, especialmente en el léxico grosero, que fue el que más aumentó con el tiempo.

Respecto al nivel de instrucción se aprecia un cierre de la brecha sociolingüística en 20 años. En el año 2000 la diferencia en el uso del léxico de *tradición culta* entre niveles de instrucción alto y medio/bajo era muy amplia, y sorprende que en la actualidad esta diferenciación haya casi desaparecido, según el análisis de correspondencias. Esta cercanía entre niveles implicó un movimiento de ambos grupos; específicamente el nivel de instrucción alto es el que más se acercó hacia el NIMB y no a la inversa. Este acercamiento en el tiempo real entre hablantes de distinta instrucción escolar ya se había documentado en las investigaciones de Serrano (2014) y Rivera (2015), y estos datos corroborarían aquellas observaciones.

La marca de uso que más cambió su frecuencia fue el léxico de las groserías, ya que tuvo un

incremento en las tres variables sociales revisadas. Ello podría estar indicando un cambio de actitud sociolingüística (cada vez más favorable) ante este tipo de vocabulario (V. Tabla 4).

Categoría	2000	2022
Popular	Jóvenes, hombres, NIA	Mayores, mujeres, común
Grosería	Jóvenes, hombres, NIMB	Jóvenes, hombres, común
Préstamo	Jóvenes, hombres, NIA	Adultos, mujeres, común
Tradición culta	Mayores, hombres, NIA	Mayores, mujeres, común

Tabla 4. Evolución del perfil sociolingüístico de los hablantes que favorecen significativamente el vocabulario marcado lexicográficamente en español de la Ciudad de México.

La Tabla 4 permite destacar lo siguiente: 1) los hombres lideraban los cuatro tipos de vocabulario en el año 2000, pero para 2022, las mujeres dominan tres de los cuatro campos: popular, préstamos y léxico de tradición culta; 2) el perfil del hablante que dice groserías sigue siendo similar, aunque ahora no se correlaciona con ningún nivel de instrucción; 3) el grupo de mayor edad siempre estará vinculado al léxico de tradición culta; 4) los préstamos parecen caracterizar a una generación, ya que los mismos informantes que los favorecieron en el pasado lo siguen haciendo en la actualidad; finalmente, 5) el nivel de instrucción dejó de ser significativo, por lo que tendríamos una comunidad de habla nivelada sociolingüísticamente, al menos respecto a esta categoría. Habrá que investigar si el uso de redes sociales de internet está teniendo un efecto en estos procesos de nivelación sociolingüística.

Esta investigación sociolingüística del léxico a través de marcas de uso nos permite entender cómo ciertas variables sociales se correlacionan significativamente con tipos específicos de vocabulario, lo que puede ser un reflejo de las transformaciones sociales que ha tenido la comunidad en 20 años. Los datos permiten constatar los criterios de los hablantes ante el léxico marcado y tabuizado como las groserías, algo que hemos documentado en Montiel (2019) y Montiel y Serrano (2023).

Otra contribución del estudio radica en la confrontación de los cambios individuales y los cambios comunitarios. Si bien en el panel se estudia la mutación del léxico individual, dichos cambios se pueden constatar a nivel grupal, dado que el informante avanza en los grupos etarios y transforma sus usos léxicos de acuerdo con su sexo y nivel de instrucción. De acuerdo con algunas observaciones sociolingüísticas a nivel fonético (Brook *et al.*, 2018; Nahkola y Saanilahti, 2004) los cambios más importantes en el lenguaje se dan entre los 18 y 25 años. En este ejercicio encontramos mutaciones comunitarias y cambios individuales que se producen a lo largo de toda la vida de los panelistas, lo que indica una diferencia respecto a los cambios en el nivel gramatical, que suelen constreñirse a partir de la edad adulta. Aunque se documentaron procesos estables (la estratificación por edad o la persistencia de los préstamos en la generación de jóvenes del 2000), en las variables sexo y nivel de instrucción se insinúan cambios sociolingüísticos en favor del léxico explorado.

Las limitaciones del trabajo son algo obvias: la muestra consta de 19 hablantes en panel, complementada con 12 entrevistas de hablantes jóvenes (31 entrevistas totales). También muestra las limitaciones cuantitativas de los materiales de lengua oral, que normalmente ofrecen menor riqueza léxica que la lengua escrita en estilos más formales. Sin embargo, creemos

que la implementación del análisis de ji-cuadrada y de correspondencias permitió llegar a certezas cuantitativas que alientan a realizar más investigaciones en este cruce entre sociolingüística y lexicografía. Puede explorarse, por ejemplo, si los cambios en estas categorías funcionan igual en otros dialectos del español o si son cambios que se han producido sólo en la Ciudad de México, por supuesto, atendiendo a las motivaciones de las transformaciones documentadas. También puede explorarse las mismas categorías léxicas en otro tipo de materiales (por ejemplo, en redes sociales de internet).

En general, este ejercicio permitió reconocer el valor de las marcas lexicográficas de uso para rastrear cambios léxico-semánticos en corpus sociolingüísticos y la utilidad de recursos estadísticos como la ji-cuadrada y el análisis de correspondencias para obtener conclusiones empíricamente robustas en lexicografía.

8. Referencias bibliográficas

- Academia Mexicana de la Lengua. (2010). *Diccionario de Mexicanismos*. Siglo XXI.
- Brook, M., Jankowski, B., Konnelly, L., & Tagliamonte, S. (2018). 'I don't come off as timid anymore': Real-time change in early adulthood against the backdrop of the community. *Journal of Sociolinguistics*, 22(4), 351-374. <https://doi.org/10.1111/josl.12310>
- Buchstaller, I., & Wagner, S. E. (Eds.). (2019). *Panel studies of variation and change*. Routledge.
- El Colegio de México. (2010). *Diccionario del español de México*. Retrieved from <https://dem.colmex.mx/Inicio>
- Cieri, C., & Yaeger-Dror, M. (2019). Alternative sources of panel study data. In I. Buchstaller & S. E. Wagner (Eds.), *Panel studies of variation and change* (pp. 53-72). Routledge.
- De la Fuente Fernández, S. (2011). *Análisis de correspondencias simples y múltiples*. Universidad Autónoma de Madrid.
- Escoriza Morera, L. (2017). Semántica léxica y sociolingüística variacionista: Las marcas sociolingüísticas en la descripción semántica del léxico. *Rilce. Revista de Filología Hispánica*, 33(3), 1297-1319.
- Labov, W. (1996). *Principios del cambio lingüístico. Vol. 1: Factores internos*. (P. Martín Butragueño, Trad.). Gredos.
- Martín Butragueño, P., & Lastra, Y. (Eds.). (2011). *Corpus sociolingüístico de la ciudad de México: Materiales de PRESEEA-México*. El Colegio de México.
- Montiel, D. (2019). *Cambio semántico en español. Disfemismos socioeconómicos en una red social de clase media* [Master's thesis]. Universidad Nacional Autónoma de México.
- Montiel, D., & Serrano, J. (2023). Disponibilidad léxica y exclusión social. Datos sociolingüísticos de la Ciudad de México. *Lingüística Mexicana*, 5(1), 65-95.
- Nahkola, K., & Saanilahti, M. (2004). Mapping language changes in real time: A panel study on Finnish. *Language Variation and Change*, 16(02). <https://doi.org/10.1017/S0954394504162017>

Labov, W. (1973). *Sociolinguistic Patterns*. University of Pennsylvania Press.

Real Academia Española. (2001). *Diccionario de la lengua española*. Espasa Calpe.

Rivera Vidal, A. (2015). *Sociolingüística de los adjetivos en español mexicano* [Master's thesis]. Universidad Nacional Autónoma de México.

Serrano, J. (2011). Retracción e innovación léxica en español de la ciudad de México: 1970-2000. In P. Martín Butragueño (Ed.), *Realismo en lingüística. Primer coloquio de cambio y variación lingüística* (pp. 189-213). El Colegio de México.

Serrano Morales, J. C. (2014). *Procesos sociolingüísticos en español de la Ciudad de México. Estudio en tiempo real* [Doctoral dissertation]. El Colegio de México.

9. Anexo de abreviaturas

CSCM= Corpus Sociolingüístico de la Ciudad de México

NIA= Nivel de instrucción alta

NIM= Nivel de instrucción media

NIB= Nivel de instrucción baja

NIMB= Nivel de instrucción media y baja