



CORPUS ORALES Y LEXICOGRAFÍA

A propósito de Preseaa-Puebla

Niktelol Palacios

EL COLEGIO DE MÉXICO

CORPUS ORALES Y LEXICOGRAFÍA
A PROPÓSITO DE PRESEEA-PUEBLA

ESTUDIOS DE LINGÜÍSTICA
XXXVIII

CENTRO DE ESTUDIOS LINGÜÍSTICOS Y LITERARIOS

CORPUS ORALES Y LEXICOGRAFÍA
A PROPÓSITO DE PRESEEA-PUEBLA

Niktelol Palacios



EL COLEGIO DE MÉXICO

467.97248

P171c

Palacios, Niktelol.

Corpus orales y lexicografía. A propósito de Preseaa-Puebla / Niktelol Palacios. – 1a ed. – Ciudad de México : El Colegio de México, Centro de Estudios Lingüísticos y Literarios, 2022.

174 p. : il., gráf., tablas, mapas ; 22 cm. – (Estudios de lingüística ; 38).

ISBN 978-607-564-388-5

1. Español – Puebla (Estado) – Puebla – Lexicografía.
2. Español – Provincialismos – Puebla (Estado) – Puebla.
3. Español – Español hablado – Puebla (Estado) – Puebla.
4. Sociolingüística – Puebla (Estado) – Puebla.
5. Corpora (Lingüística) – Procesamiento de datos. I. t. II. Ser.

Corpus orales y lexicografía. A propósito de Preseaa-Puebla
Niktelol Palacios

Primera edición, octubre de 2022

D. R. © El Colegio de México, A. C.

Carretera Picacho Ajusco núm. 20

Ampliación Fuentes del Pedregal

Alcaldía Tlalpan

14110, Ciudad de México, México

www.colmex.mx

ISBN 978-607-564-388-5

Impreso y hecho en México

A Carmen Delia y Luz

ÍNDICE

INTRODUCCIÓN	11
1. ANTECEDENTES: CORPUS ORALES DEL ESPAÑOL DE PUEBLA	19
2. EL CORPUS SOCIOLINGÜÍSTICO DE LA CIUDAD DE PUEBLA Y PRESEEA-PUEBLA	27
3. LEMATIZACIÓN Y ANÁLISIS CUANTITATIVO DE PRESEEA-PUEBLA	63
4. RASGOS DE ORALIDAD EN PRESEEA-PUEBLA: REFLEXIONES LEXICOGRAFICAS	93
5. NOTAS LÉXICAS DEL ESPAÑOL DE PUEBLA	139
BIBLIOGRAFÍA	161

INTRODUCCIÓN

El español de Puebla pertenece a la zona dialectal del Altiplano central (Lope Blanch, 1990 [1971]). Quizá por el predominio político, socioeconómico y cultural de la capital del país o por el hecho de que se le percibe como una zona bastante homogénea,¹ representada por los usos lingüísticos de la Ciudad de México, el español de Puebla ha sido poco investigado en la lingüística mexicana. Y es que, como afirma Cabrera (1995: 287) en su estudio sobre urbanización, que bien puede extenderse a la explicación lingüística, la ciudad de Puebla ha desempeñado desde sus orígenes un doble papel en la vida nacional, pues es al mismo tiempo un espacio urbano importante de desarrollo socioeconómico, comercial y cultural en donde se concentra el flujo migratorio del interior del propio estado, de Tlaxcala y de Veracruz, así como de la región sur sureste de México, e incluso de Centroamérica, y un centro de servicios administrativos públicos y de gestión gubernamental para estas poblaciones. Asimismo, por su cercanía con la Ciudad de México, su importancia ha permanecido relativizada como parte de la megalópolis.

En este libro presentamos una propuesta de análisis que combina objetivos, métodos y herramientas de la lingüística computacional, la lexicografía, la lexicología y la sociolingüística. A partir del análisis cuantitativo y cualitativo, automatizado y manual, presentamos el método para lematizar un corpus oral de 108 entrevistas semidirigidas.

¹ En los estudios de dialectología perceptual sobre el español de México se aprecia que, en comparación con la capital del país, los hablantes identifican zonas con mayores diferencias y con identidad propia, es decir, zonas con más “personalidad” (por ejemplo, la península de Yucatán), mientras que el centro de México se considera homogéneo (Morúa y Serrano 2004; Serrano 2019).

Confiamos en que la exposición detallada de las decisiones metodológicas sirva de punto de partida para la preparación de leuarios de otros corpus del Proyecto para el Estudio Sociolingüístico del Español de España y América (Preseea) —y de otros corpus orales regionales— y que, a partir de ese punto de arranque, podamos acercarnos a la variación lexicológica por parcelas referenciales, por el origen etimológico o por la distribución social e identitaria de ciertos ítems.

Fue precisamente el interés lexicográfico el que nos llevó al planteamiento de reunir un corpus de estudio que permitiera dar cuenta de manera integral del español de Puebla, uno que replicara el Corpus del Español Mexicano Contemporáneo (1921-1974) (CEMC),² el cual sirve de base al *Diccionario del español de México* (DEM). Decidimos que nuestro primer paso sería recoger una muestra oral que registrara el habla de la capital, centrada en adultos, para después ampliarla a otros grupos generacionales (principalmente, adolescentes), grupos de migrantes, grupos bilingües y, posteriormente, a la lengua escrita. Ese proyecto también resultaba demasiado ambicioso para nuestros recursos. No obstante, había que empezar a trabajar. Una motivación adicional eran las noticias que Everardo Mendoza nos compartía sobre los avances del *Diccionario del léxico regional de Sinaloa*, cuya legitimidad académica y social se sustenta, precisamente, en subsecuentes estudios dialectales y sociolingüísticos.

En 2011 decidimos arrancar el proyecto siguiendo las directrices del Corpus Sociolingüístico de la Ciudad de México (CSCM) (2011: 22-69) y del Preseea (Moreno Fernández, 2003). La técnica de la entrevista semidirigida tiene una planeación premeditada lingüísticamente para la recolección de datos, y uno de los principales beneficios de la denominada “muestra-tipo” es la posibilidad de recopilar muestras de habla amplias y representativas, socialmente estratificadas. La planeación de la entrevista permite, además, obtener audio de buena calidad. Desde luego, tal planeación en la recolección de los datos su-

² El CEMC, trabajo pionero de estadística lingüística en México, está integrado por muestras (de 2 000 palabras cada una) de 996 textos agrupados en 14 géneros, orales y escritos (para una descripción detallada, véase Lara, Ham y García, 1979b). Estos géneros se estratifican en tres grandes grupos: lengua culta, lengua subcultura y lengua no estándar, y en cada uno de ellos se incluyeron conversaciones o entrevistas grabadas.

pone algunas limitantes, pero consideramos que hemos podido reunir materiales ricos en estilos de habla que van desde interacciones más cercanas a la entrevista hasta la conversación grabada.

De manera paralela, y siempre guiados por nuestros intereses lexicológicos y lexicográficos, fomentamos las investigaciones sobre vocabularios de oficios tradicionales en el estado de Puebla a partir de un método común: la recopilación de vocabulario en su contexto discursivo, para su mejor entendimiento semántico y pragmático (Palacios 2010, 2014, 2020a y 2020b). Partimos de la idea de que sería en estas actividades, así como en los nombres de la comida, la flora y la fauna regionales donde encontraríamos la mayor riqueza denominativa propia (poblanismos de ámbito objetivo).³

El lector puede aguardar, por lo tanto, una guía para la recolección, el preprocesamiento y el procesamiento de datos léxicos a fin de organizarlos en un lecionario, así como un acercamiento panorámico al tratamiento integral del vocabulario del español de Puebla registrado en Preseaa-Puebla. El hilo conductor de este libro son las preguntas ¿qué aportan los corpus orales al conocimiento léxico de una comunidad determinada?, ¿qué tipo de ítems léxicos podemos documentar en los corpus Preseaa?, y ¿qué nos permite conocer este registro sobre la variación léxica (sea que entendamos por ésta la sinonimia referencial,⁴ los *ismos* de

³ Luz Fernández Gordillo (2006: 51-52) explica que, a propósito de la variación diatópica de una lengua histórica, Eugenio Coseriu distingue entre *zona lingüística* y *ámbito objetivo*: “La primera comprende la extensión dialectal en que se utiliza determinado signo lingüístico —por ejemplo, *cajuela*, en la sincronía actual pertenece a la *zona lingüística* del español de México, *patata* a la *zona lingüística* de gran parte del español de España y *mesa* a la *zona lingüística* del español general o común—. El *ámbito objetivo* abarca ‘la región en la que el *objeto* [en contraste con el signo lingüístico que lo denomina] se conoce como elemento del horizonte vital de los hablantes’ (Coseriu 1973: 311) como *vieira*, *acamaya* y *charro*, que vendrían a ser respectivamente un ‘españolismo’ y dos ‘mexicanismos’ de ámbito objetivo”.

⁴ Basados en el modelo de “organon” de Karl Bühler y la triple función del lenguaje humano: *expresión*, *apelación* y *representación* (1961: 41), entendemos por *sinonimia referencial* aquella que en la misma lengua nos permite denominar de forma distinta un mismo referente (objetos y relaciones), por ejemplo, *mejilla* y *cachete* o *llovizna* y *chipichipi* (véase Valadez, 2020). En estos casos no hablamos simplemente de sinonimia, puesto que consideramos que la elección y el uso de una u otra

ámbito objetivo, la variación formal o las modalidades expresivas-estilísticas)?

Debido a que nuestra mirada está indisociablemente ligada a nuestro propio quehacer lexicográfico, se plantea la descripción a partir de su comparación con el español nacional registrado en el leuario del DEM, con lo cual esperamos contribuir a la reflexión sobre los aportes de los corpus regionales orales a la lexicografía integral y contrastiva. Antes de continuar, hacemos eco de lo expresado por Morúa acerca de algunos términos de dialectología y de sociolingüística que usamos de manera amplia a lo largo de este libro:

Si bien los términos mismos *registro*, *variedad*, *dialecto*, *sociolecto*, etcétera, siguen siendo etiquetas laxas con las que se intenta identificar conjuntos de diferencias de uso vinculables a tipos de personas, zonas geográficas o contextos situacionales, y que presuponen la siempre debatible idea de que tales rótulos corresponden a realidades discretas, no parece posible por el momento desligarse de ellas como una manera cómoda de referirse a dichos conjuntos, que en realidad son difíciles de caracterizar y siempre fluctuantes (2011: 492).

El libro se divide en cinco capítulos: en el primero presentamos los antecedentes del estudio del español de Puebla, en particular las propuestas para recoger corpus de oralidad. La revisión parte de la década de 1980 y termina con la reciente publicación del Corpus del Habla de Puebla (Corhpu). Con el fin de brindar un panorama del desarrollo de la lingüística local, hemos dejado de lado trabajos nacionales de gran importancia, como el *Atlas Lingüístico de México* (Lope, 1990-2000) y el Corpus Oral del Español de México (COEM) (en preparación).

En el segundo capítulo describimos a detalle el Corpus Sociolingüístico de la Ciudad de Puebla (CSCP) y su subcorpus Preseea-Pue-

palabra manifiestan diferencias expresivas y apelativas. El lector también puede consultar la interesante discusión que a propósito de la sinonimia referencial hace Luis Fernando Lara (2012b) sobre la definición lexicográfica de los vocabularios jergales.

bla. Si bien se trata de un trabajo que replica el CSCM de Pedro Martín Butragueño y Yolanda Lastra (2011-2015), con un módulo central conformado por las entrevistas semidirigidas según el método de Preseea (Moreno Fernández, 2003) y las pruebas lingüísticas del CSCM, hemos replicado el “Cuestionario sociodemográfico y reticular” del Proyecto de Investigación de la Formación de Dialectos (Fordial) (Villena, Moya, Ávila y Vida, 2003) y un subcorpus de oficios. Pretendemos, pues, compartir de forma minuciosa las decisiones metodológicas con las que trabajamos, así como las adecuaciones que hicimos a las propuestas mencionadas, cuyas publicaciones conviene que el lector consulte para tener información de primera mano.

En el tercer capítulo presentamos las etapas del análisis que nos permitió obtener los 10 715 vocablos⁵ de Preseea-Puebla. Detallamos el proceso seguido —una combinación de análisis automatizados y revisiones manuales (*AntConc* 3.2 [Anthony, 2020], *FreeLing* 4.2 [Padró, 2020] y Palacios y Serralde [en preparación]) de los mismos datos— con dos fines: fomentar el trabajo interdisciplinario (lexicología, lexicografía y lingüística computacional) y exponer con tal claridad las etapas de la investigación que otros equipos interesados en el estudio léxico puedan replicar el método.

En 1994 Raúl Ávila hacía notar que la descripción socioléxica era escasa, en particular si se comparaba con el número de estudios sociolingüísticos sobre variación fonológica. Casi 30 años después la situación ha cambiado poco y los estudios sobre variación léxica siguen siendo escasos. En este contexto, los corpus Preseea, al compartir un mismo método de trabajo, nos pueden arrojar información valiosa tanto para la descripción de una comunidad de habla particular como para los estudios de corte comparativo.

Presentamos la lista de los 100 vocablos con mayor frecuencia absoluta de aparición en nuestro corpus y la comparamos con “El vocabulario fundamental del español de México” (Lara, Ham y

⁵ Entendemos como *vocablo* la palabra que por convención representa el paradigma flexivo o de conjugación; por ejemplo, *amar* (que supone todas las formas conjugadas del verbo) o *gato* y *bueno* (que engloban las formas femenina, masculina, singular, plural, diminutivo, aumentativo, etcétera) (véase Lara, 2006: 138).

García, 1979a). De manera previsible —aunque no por ello menos interesante—, estos 100 vocablos son mayoritariamente de carácter funcional-pragmático (artículos, demostrativos, preposiciones, conjunciones y marcadores discursivos). Además, representan dos terceras partes (un poco más de 66%) de los *tipos*⁶ registrados en nuestro corpus. En busca de un método que permita identificar voces y significados propios, también comparamos la lista completa de los vocablos de Preseea-Puebla con el leuario del *DEM* y encontramos que casi 90% del vocabulario obtenido de las transcripciones de 108 entrevistas está definido en él, es decir, podemos afirmar que el léxico recogido en estas interacciones semidirigidas es parte del español de México en su diversidad y riqueza: vocabulario culto, popular, grosero. Este acercamiento nos parece particularmente importante para el estudio de las variedades regionales, pues evita que marquemos como propias algunas palabras de uso general sólo porque son muy frecuentes en nuestra variedad.

Guiados por el mismo cuestionamiento central —¿qué aportan los corpus Preseea y los corpus orales al conocimiento del vocabulario local?—, en el cuarto capítulo nos acercamos al léxico de Preseea-Puebla que no ha sido definido en el *DEM* (al menos hasta su versión publicada en línea en octubre de 2020). Documentamos en estas voces una enorme riqueza para nuestro estudio: fenómenos propios de la oralidad, como la supresión fónica, los cruces léxicos y los juegos verbales (Briz, 2010; Lara, 1996, y Lope Blanch, 1980a y 1980b); estas unidades permiten una descripción cualitativa más realista de la complejidad discursiva de las entrevistas semidirigidas. También hacemos un acercamiento panorámico a la morfología léxica de esta misma parcela de vocabulario: las palabras no definidas en el *DEM*, pero emparentadas con su leuario.

Por último, en el capítulo 5 nos ocupamos, a manera de ejemplo, de la documentación del significado y los usos de palabras candidatas a ser propias del español de Puebla, es decir, hacemos una propuesta metodológica para la identificación de *ismos* (en nuestro caso, de

⁶ Entendemos por *tipo* cada una de las palabras distintas que aparecen en el corpus; por ejemplo, *mamá*, *madre*, *había*, *hubiera*, *bonita*, *bonitos*, etcétera (véase Lara, 2006).

poblanismos). Dentro del libro, éste es nuestro primer acercamiento al significado, lo cual resulta revelador, pues algunas de las palabras que habíamos indicado que se encuentran ya definidas en el DEM tienen un significado o una estructura argumental distinta en el español de Puebla (como *acusarse* ‘acusar una persona a otra con alguien que puede hacer justicia’, en vez del uso verbal reflexivo ‘acusarse a uno mismo’). Con el fin de tener un acercamiento detallado, hemos elegido únicamente 11 palabras para mostrar la riqueza de la variación semántica. El mismo trabajo deberá repetirse para todo el lecionario, si se quiere obtener un producto de lexicografía integral, aunque, como ya advertimos, Preseea-Puebla sería sólo un módulo del corpus requerido para tal propósito.

Reiteramos nuestro deseo de que el método propuesto sirva de guía y pueda ser replicado por otros equipos que recojan corpus orales y entre cuyos fines se encuentre el estudio lexicológico y lexicográfico. Para nosotros este libro y la publicación de Preseea-Puebla marcan el fin de una primera etapa de investigación que comenzó en 2011. Al mismo tiempo, una vez establecido su lecionario, se abre la posibilidad de acercarnos ya no con una mirada panorámica, sino con detenimiento, a parcelas léxicas (indigenismos, préstamos, locuciones, neologismos), al análisis pormenorizado de la morfología léxica de esta variedad, a la redacción lexicográfica y a las explicaciones de la variación léxica con base en datos reales de lengua oral —con sus ventajas y limitaciones—.

Queremos agradecer a los colaboradores del Corpus Sociolingüístico de la Ciudad de Puebla por permitirnos entrevistarlos, y por el entusiasmo y el compromiso mostrado para que este proyecto se gestara y materializara; a todo el equipo que recogió y transcribió las entrevistas; al Departamento de Sistemas de Información Geográfica de El Colegio de México, en especial a Emelina Nava García; a los dictaminadores anónimos de este trabajo por sus enriquecedoras sugerencias; a los compañeros del DEM, por sus observaciones críticas a las versiones previas de este trabajo, especialmente a Juan Luis Serralde, por su invaluable ayuda; y a Tania Hernández, Erik Franco, Anayeli Hernández Cruz y Carmen Fajardo, por el diálogo permanente, sus lecturas y sus preguntas siempre agudas. Desde luego, cualquier error que persista es sólo responsabilidad nuestra.

1. ANTECEDENTES: CORPUS ORALES DEL ESPAÑOL DE PUEBLA

La necesidad de contar con estudios regionales sobre el español de la zona Puebla-Tlaxcala ha sido una preocupación reiterada de los lingüistas adscritos a la Universidad Autónoma de Puebla desde los años setenta del siglo xx. Investigadores como Adrián Gimete-Welsh, Rosa Montes Miró, James Fidelholtz, José Marcos Ortega y, recientemente, Jorge Adrián Lázaro Hernández han insistido en la necesidad de recoger corpus para la investigación lingüística, trabajo de campo que debería vincularse con la formación de cuadros de investigación y ligarse a los planes y programas de estudio institucionales, y que contribuiría a un mejor entendimiento histórico de la lengua nacional, a la delimitación dialectal del español de México y a la explicación sociolingüística a partir de estudios de lenguas en contacto; además de responder al interés por la descripción formal integral (fónica, morfológica, sintáctica y léxico-semántica) del estado actual del habla, por supuesto, y a la inquietud de elaborar un atlas regional y un diccionario —ya integral, ya contrastivo—.

En esta misma línea de pensamiento surgió la propuesta de levantar el Corpus Sociolingüístico de la Ciudad de Puebla (CSCP), con el subcorpus Preseea-Puebla, que toma como modelo el Corpus Sociolingüístico de la Ciudad de México (CSCM) (2011-2015, coordinado por Pedro Martín Butragueño y Yolanda Lastra), así como las directrices del Proyecto para el Estudio Sociolingüístico del Español de España y América (Preseea) (Moreno Fernández 1993, 1997 y 2003).

EL PERIODO DE 1980 A 2007

En la década de 1980 se comenzó en la Universidad Autónoma de Puebla un proyecto de investigación lingüística en la región Puebla-Tlaxcala bajo la dirección de Adrián Gimete-Welsh, motivado por 1) la escasez de estudios de corte geográfico-lingüístico en la región;¹ 2) la importancia geoestadística de la población indígena y bilingüe; 3) la necesidad de aportar datos lingüísticos a investigaciones históricas, sociológicas y antropológicas, y 4) la pertinencia de la investigación de campo para la formación de estudiantes (1980b: 69-70).

El proyecto Lenguaje, Cultura e Ideología, construido cuidadosamente sobre los conocimientos dialectales y los avances sociolingüísticos, consideraba recoger datos en las siete zonas socioeconómicas del estado de Puebla (Huauchinango, Teziutlán, Ciudad Serdán, Cholula, Puebla, Izúcar de Matamoros y Tehuacán) y en seis localidades de Tlaxcala,² a través de una muestra mínima de ocho informantes por cada localidad, estratificados por sexo (cuatro hombres y cuatro mujeres), edad (cuatro jóvenes y cuatro viejos)³ e instrucción (cuatro cultos y cuatro incultos). Además de la muestra mínima, se pretendía levantar cuestionarios con hablantes de redes específicas para estudiar el léxico patrimonial de cada región; por ejemplo, el de

¹ El Proyecto para la Delimitación de las Zonas Dialectales de México (Lope, 1970b) incluyó entrevistas realizadas en los municipios de Tetela de Ocampo, Acatlán de Osorio y San Martín Texmelucan (Sierra Norte de Puebla, Mixteca poblana y zona centro-oeste del estado, respectivamente). Sin embargo, Gimete-Welsh explica acertadamente que de esta muestra dialectal quedaron fuera áreas importantes para la descripción de la región.

² Las localidades seleccionadas en Puebla fueron Acatlán, Acatzingo, Atlixco, Calpan, Cuetzalan, Chignahuapan, San Pedro Cholula, Huehuetla, Huauchinango, Izúcar de Matamoros, Oriental, Pahuatlán, San Martín Texmelucan, Tehuacán, Tepeaca, Tepeyahualco, Tetela, Teziutlán, Zacapoaxtla y Zacatlán. Las poblaciones de Tlaxcala seleccionadas para el estudio fueron Calpulalpan, Huamantla, Tlaxco, Santa Ana Chiautempan, Tlaxcala y Zacatelco (Gimete-Welsh, 1980a: 83).

³ En cuanto a edad e instrucción formal, se ofrecen en el proyecto las etiquetas que aquí reproducimos, pero no se da un correlato específico sobre el grupo etario ni los años de escolaridad (Gimete-Welsh, 1980a: 74-75).

la caña de azúcar en Izúcar de Matamoros, el del café en Huachinango o el de la manzana en Zacatlán. Todo el material se recogería en grabación magnetofónica y se complementaría con instrumentos de anotación para las observaciones directas de los investigadores.

El modelo a seguir era el *Cuestionario para la delimitación de las zonas dialectales de México* (Lope Blanch, 1970a). El proyecto precisaba los aspectos de capacitación para los investigadores, así como los detalles técnicos para la recolección de datos, la transcripción fonética y el análisis.

En 1985 se publicó el libro *Lenguaje, ideología y clases sociales. Las vecindades en Puebla*, que, de acuerdo con el propio Gimete-Welsh, es un adelanto del proyecto Análisis de los Procesos Ideológicos, el cual era parte del ya descrito Lenguaje, Cultura e Ideología. El libro se sostiene sobre tres pilares teóricos: el materialismo histórico, la teoría del discurso y la teoría semiótica. En un trabajo previo el autor ya había explicado que el objetivo de la investigación sería “el estudio del español popular, la cultura popular y la ideología de las clases sociales que habitan los vecindarios del primer cuadro de la ciudad de Puebla, basándose en el análisis lingüístico y semiótico del discurso de los informantes” (Gimete-Welsh, 1980a: 98).

Carcedo (1983) explica que dentro del mismo proyecto confluían cuatro subproyectos sobre el estudio del habla local: “Lexicografía e ideología del habla popular en Puebla”, “La ideología religiosa de la clase popular poblana vista a través del léxico religioso”, “La penetración ideológica en las clases subalternas a través de los medios de comunicación social” y “La imagen de la mujer en la televisión mexicana y su recepción en las clases populares en la ciudad de Puebla”. Los cuatro proyectos se vinculaban por el método de trabajo, esto es, la obtención de datos mediante entrevistas grabadas, realizadas con un cuestionario que tocaba tres temas principales: economía, política e ideología (Gimete-Welsh y Marroquín, 1985: 28); preguntas diseñadas para “averiguar cómo la ideología orgánica se articula a través de los sistemas culturales y económicos en los sujetos” (Gimete-Welsh, 1980 *apud* Carcedo, 1983: 108).

Pese a la claridad del proyecto y a sus lineamientos teóricos y metodológicos sólidos, no hemos podido encontrar avances o resultados

del mismo, con excepción de una tesis de licenciatura titulada *Léxico religioso e ideología* (Rivera, 1984). Por otra parte, debido a la rotación de los académicos y a la falta de un registro catalográfico institucional que permita el libre acceso a los materiales, no nos ha sido posible establecer en dónde se encuentra el estudio de las vecindades, ni sus condiciones de resguardo.⁴

Posteriormente, entre 2001 y 2006 se hicieron otras propuestas para el estudio del habla de la ciudad asociadas con las líneas y proyectos de los investigadores del Instituto de Ciencias Sociales y Humanidades de la Benemérita Universidad Autónoma de Puebla (BUAP), que versan sobre el habla infantil de Puebla, la descripción del español poblano y la construcción de un corpus del español de Puebla. No obstante, estos corpus no están indizados en el sistema bibliotecario universitario ni se hallan publicados. Lo que hemos podido documentar, a partir de registros curriculares públicos de los propios investigadores, es que se trata de los corpus El Español Hablado de Puebla como Modelo Idiomático del Cambio Lingüístico, proyecto desarrollado en 1999 por Wiaczesław Nowikow en el Instituto de Ciencias Sociales y Humanidades de la BUAP (Nowikow, 2001: 461); Corpus de Habla Infantil Poblana, de 2001, coordinado por Rosa Graciela Montes Miró; Descripción del Español en la Región de Puebla, de 2001 a 2004, y El Español de Puebla: Construcción

⁴ A pesar de haber hablado con algunos profesores y directivos de la institución que conocen el proyecto y aseguran que los materiales están bajo resguardo, ninguno ha podido indicar a cargo de quién o en qué lugar físico se encuentran, ni si están en el Colegio de Antropología Social o en el Centro de Ciencias del Lenguaje. Por tal motivo, no ha sido posible documentar para esta publicación, por ejemplo, el periodo durante el que se recogieron datos, el número de grabaciones acumuladas ni el estado de su cuidado y respaldo físico. No obstante, podemos dar testimonio de que hacia 1995 se continuaban haciendo entrevistas, puesto que, como estudiantes del Colegio de Lingüística y Literatura Hispánica, se nos invitó a hacer una; la única indicación era que el informante viviera en una vecindad del centro de la ciudad de Puebla. Podemos suponer que para entonces las líneas temáticas básicas del cuestionario se habían desvanecido. No hubo instrucciones en torno al equipo de grabación, pero se solicitó que la transcripción fuera fonética usando el alfabeto de la *Revista de Filología Española*. Junto con Rosalba Ramos Huerta, hicimos una entrevista en la casa de un hombre, adulto, de oficio globero, misma que entregamos en formato de cassette de audio y con la transcripción fonética manuscrita.

de un Corpus, de 2005 a 2006. Los responsables de los dos últimos son James L. Fidelholtz, Rosa Graciela Montes Miró, José Marcos Ortega y Guadalupe Blanco. Vale la pena destacar que algunas de estas propuestas se hallan vinculadas con dos de los proyectos más importantes de las últimas décadas para la recogida de datos orales del español: Val.Es.Co (Valencia, Español Coloquial) y Preseea.

Nowikow (2001) detalla el proyecto de un estudio sociolingüístico que toma en cuenta tres variables para la conformación de la muestra de estudio: sexo, edad (distribuida en cuatro grupos, de 16 a más de 65 años) e instrucción educativa (cinco grupos: desde personas sin estudios formales hasta instrucción superior). El cuerpo de datos estaría conformado por entrevistas a 117 informantes, y los grupos sociales tendrían distinto peso estadístico según la demografía local. Para la selección de los entrevistados, se tomaría en cuenta, además, que fueran personas nacidas en Puebla o mexicanos con un mínimo de 10 años de vivir en la ciudad. También se consideraba la posibilidad de incluir a otros hispanohablantes y a extranjeros bilingües no hispanos. En cuanto a la ficha técnica, el tipo de conversación, los criterios de transcripción y las unidades de segmentación del discurso oral, se retomaría la propuesta de Val.Es.Co. (Briz, 1995).

Por otra parte, en las primeras reuniones de equipos de trabajo de Preseea dentro de la Asociación de Lingüística y Filología de América Latina (ALFAL) se propuso recoger el corpus de Puebla. Hasta donde nos ha sido posible documentar, ambos proyectos, Val.Es.Co. y Preseea, permanecieron en fase de planeación.

EL CORPUS DEL HABLA DE PUEBLA (CORHPU)

Para concluir este capítulo, queremos presentar el Corpus del Habla de Puebla (Corhpu), cuyos responsables son Jorge Lázaro, Gerardo Sierra y Guadalupe Huerta. Se trata de un proyecto de lingüística de corpus —quizá sería más conveniente describirlo como parte de la lingüística computacional— desarrollado en el Grupo de Ingeniería Lingüística de la Universidad Nacional Autónoma de México (UNAM) con datos recogidos por estudiantes de la Benemérita

Universidad Autónoma de Puebla en el periodo primavera de 2017 a primavera de 2018.

El proyecto prevé procesar textos de 1990 a 2016 de todo el estado de Puebla para conformar un corpus *multipropósito*, “pues la necesidad que se intenta resolver es contar con un corpus del habla del estado de Puebla. Aun así, lo cierto es que también se incluirán materiales para el mejor conocimiento de la variedad lingüística existente” (Lázaro y Huerta, 2020: 33).

Lingüísticamente —y hasta este momento—, los datos recolectados permiten adscribir el proyecto a los estudios lexicológicos y terminológicos del vocabulario de artes, oficios y aficiones. La información se recogió por un muestreo de cuotas con la variable preestratificatoria *edad* y con cinco informantes por campo de interés. Se abordaron 21 temas y se obtuvieron 120 documentos. El método para la recuperación de información fue la aplicación de un cuestionario.

Los archivos TXT resultantes de la transcripción se cargaron en el Gestor de Corpus (GECO), desarrollado por el Grupo de Ingeniería Lingüística (GIL) de la UNAM.

GECO fue concebido en su totalidad como una aplicación web, enfocada en la construcción colaborativa de corpus. En ese sentido, tiene la finalidad de ser un repositorio central de documentos para una variedad de aplicaciones orientadas al PLN, que pueden ser integradas a GECO por desarrolladores con herramientas de código abierto. Además de esto, tiene un enfoque muy marcado hacia la construcción y publicación de corpus como fin en sí mismo, por lo cual también provee herramientas para dar a conocer los corpus creados a través de un portal web (Sierra, Solórzano y Curiel, 2017: 59).

Entre las principales virtudes de este trabajo de investigación, podemos destacar la presteza con la que sus coordinadores pusieron a disposición pública sus materiales y las facilidades de búsqueda según las variables preestablecidas. No obstante, es posible que ya haya concluido, pues su principal responsable ha dejado de trabajar en la BUAP. El Corhpu se encuentra disponible en la página <<http://www.corpus.unam.mx/geco/portal/index/corphu>>.

COMENTARIOS FINALES

Este breve repaso sobre los proyectos de recolección de datos orales del español de Puebla nos permite observar el ahínco con el que distintos especialistas han emprendido, prácticamente de forma individual, trabajos de investigación con el objetivo de contribuir al conocimiento lingüístico del español de Puebla y la formación de recursos humanos.

No obstante, la pauperización laboral de los docentes universitarios, contratados por tiempo determinado y con excesivas cargas docentes, ha favorecido una alta rotación de personal. Además, la desvinculación entre centros docentes, de investigación y administrativos —como las áreas de bibliotecas— ha dificultado la consolidación de proyectos institucionales que permitan el fortalecimiento de equipos de trabajo, la creación de repositorios y adecuados resguardo y catalogación bibliográficos.

La falta de acceso abierto a los materiales recolectados también ha imposibilitado la continuidad de los proyectos, y, como hemos observado en este recuento, a pesar de que existen muchas similitudes entre ellos y la búsqueda constante de la descripción lingüística empírica sobre datos realistas, cada uno ha vuelto a plantear sus objetivos desde cero; así, los recursos humanos y económicos se han agotado en las fases de planeación.

Por otra parte, resulta interesante para la reconstrucción de la historia lingüística en México, y específicamente de la lingüística en Puebla, observar los acercamientos teórico-metodológicos desde los que se han abordado estas propuestas: el materialismo histórico, la lingüística antropológica, la semiótica discursiva y la lingüística computacional, así como los intereses geolingüísticos, sociolingüísticos y léxico-terminológicos que han dado prioridad a la recolección de tradiciones populares citadinas y rurales. Estos intereses se han consolidado a nivel administrativo, lo que se refleja en las actualizaciones de planes y programas de estudio de la Facultad de Filosofía y Letras, así como en la creación de los centros y programas de investigación —por ejemplo, el Programa de Semiótica y Estudios de la Significación y el posgrado en Ciencias del Lenguaje—, y se ha cumplido con objetivos específicos materializados en tesis y trabajos de grado.

Por último, queremos destacar el interés recurrente —pero todavía inconcluso— por estudiar comunidades de habla y no únicamente demarcaciones político-administrativas. Como hemos podido observar, en los distintos proyectos descritos está explícita o implícita la necesidad de observar de manera conjunta la región Puebla-Tlaxcala (con su complejo entramado social y de contacto lingüístico), así como el papel irradiador de cabeceras municipales de la Sierra Norte del estado de Puebla y de la Mixteca poblana. Añadiríamos la necesidad de estudiar el vínculo del español de la ciudad de Puebla con Xalapa. Deseamos que la publicación de Preseea-Puebla y de este libro contribuyan a fortalecer los cimientos de la investigación lingüística de corte realista y permitan el diálogo con otros proyectos sociolingüísticos y de contacto que se encuentran en marcha.

2. EL CORPUS SOCIOLINGÜÍSTICO DE LA CIUDAD DE PUEBLA Y PRESEEA-PUEBLA

BREVE PANORAMA HISTÓRICO Y DEMOGRÁFICO DEL MUNICIPIO DE PUEBLA

Para entender el mosaico cultural de la ciudad de Puebla conviene repasar algunos datos históricos y demográficos. Su fundación fue un proyecto peculiar de la Corona española para establecer una ciudad de labradores en oposición a los asentamientos de los conquistadores encomenderos, es decir, se trató de un “experimento destinado a lograr que los colonos no dependieran del servicio de los indios” (Hirschberg, 1978: 192). Para la puesta en marcha de esta empresa, confluyeron al menos tres hechos: 1) la preocupación de fray Julián de Garcés, dominico, primer obispo de Tlaxcala, por la falta de autoridad secular que, derivada de la alianza de Cortés con Tlaxcala, suscitó el arribo de “españoles sin honra” que estaban invadiendo los pueblos tlaxcaltecas; 2) el interés franciscano por fundar colonias de labradores españoles, y 3) la necesidad de la Real Audiencia de reducir, mediante el desarrollo de mejores rutas de comunicación entre México y el puerto de Veracruz, el descontento ocasionado por la violenta gestión de Nuño de Guzmán, y su interés estratégico de tener asentamientos en una zona indígena importante (Cholula, Huejotzingo, Tepeaca y Tlaxcala). Sin embargo, las propias disputas entre quienes plantearon la idea de fundar la ciudad (franciscanos, dominicos y la Corona), así como las confusas instrucciones de la Audiencia, llegadas en 1530, ocasionaron, como afirma Julia Hirschberg (1978: 185), que el experimento utópico destinado a combatir la aristocracia terminara por convertirse en la encarnación de ésta.

La fundación oficial de la ciudad data del 16 de abril de 1531, y en la primera Cédula Real que se le otorgó, en 1532, se establece: “es nuestra merced e voluntad que de aquí en adelante se llame o intitule Ciudad de los Ángeles”. El 11 de septiembre de 1862 el nombre cambió a Puebla de Zaragoza (en honor al general Ignacio Zaragoza), por decreto presidencial de Benito Juárez publicado en el *Boletín Oficial* el 25 de septiembre de 1862. Nos parece pertinente añadir que la ciudad de Puebla forma parte del “primer grupo de sitios mexicanos inscritos en la lista del Patrimonio Mundial en el marco de la Convención Mundial de Protección de Sitios con valor Cultural o Natural aprobada por la UNESCO, suscrita en 1972 y ratificada por el Senado en 1984” (Vélez, 2011: 1).¹

Las decisiones estratégicas tanto territoriales como sociodemográficas de la fundación definieron, sin duda, el desarrollo urbano de Puebla y su papel regional. Así puede observarse en su condición actual, pues, como se explica en el *Programa Municipal de Desarrollo Urbano Sustentable de Puebla* (Gobierno del Estado de Puebla y H. Ayuntamiento del Municipio de Puebla, 2016: 35), el valle de Puebla conserva esta función de “territorio bisagra o articulador” dentro del Altiplano central y cumple con la función de servir de contacto entre las regiones central y sur sureste del país: la primera, la más densamente poblada, la de mayor concentración de producción manufacturera y el centro político administrativo del país; la segunda, con recursos naturales sustentables y la mayor diversidad biológica y riqueza cultural (y lingüística, agregamos nosotros)² del país, así como la principal proveedora nacional de energéticos, insumos y productos agropecuarios.

¹ En 1987 se nombró como “bien natural” Sian Ka’an, reserva de la biósfera ubicada en Quintana Roo, y como “bien cultural” el Centro Histórico de la Ciudad de México y Xochimilco, el Centro Histórico de Oaxaca, el Centro Histórico de Puebla y la zona arqueológica de Monte Albán.

² Sobre la diversidad lingüística de la región, el lector puede consultar el magnífico trabajo de Esther Herrera Zendejas “La Babel del sur: el caso de Oaxaca” (2014), donde estudia la diversidad lingüística de ese estado, y el de Sandra Rocío Cruz Gómez (2014), quien presenta un recuento sobre investigaciones sociolingüísticas en lenguas indígenas de Chiapas.

Población

Según el *Censo de Población y Vivienda* (INEGI, 2020),³ el estado de Puebla tiene una población de 6 583 278 habitantes, con una distribución urbana de 79%. El promedio de escolaridad de la población mayor de 15 años es de 9.16⁴ años cursados, es decir, ligeramente por arriba de la secundaria terminada. La población de 3 años o más que habla una lengua indígena representa 9.9% (lo que coloca a la entidad en el octavo lugar nacional, por detrás de Oaxaca, Chiapas, Yucatán, Guerrero, Hidalgo, Quintana Roo y Campeche), y 113 945 habitantes se reconocen como afromexicanos o afrodescendientes (1.7% de la población). En la imagen 1 se muestra el resumen geopoblacional del estado.

Según el *Atlas de los pueblos indígenas* (INPI e Inali, 2020), en Puebla viven 1 094 923 habitantes de pueblos indígenas distribuidos como se presenta en la tabla 1. Las lenguas indígenas más habladas en la entidad son náhuatl, totonaco, mazateco y popoloca.

En la capital, Puebla de Zaragoza, vive un poco más de la cuarta parte de la población del estado (1 692 182, esto es, 25.7%) y cerca de 11% de la población hablante de lengua indígena (117 776 de 1 094 923). Como se observa en la imagen 2, los grupos con más hablantes son nahuas (78 702 hablantes censados), mazatecos (15 262), totonacos (13 609), mixtecos (3 327) y zapotecas (2 184).

³ Nos parece necesario señalar que, cuando se diseñaron y recogieron los datos del corpus, el estado de Puebla tenía una población de 5 779 829 habitantes, con una distribución poblacional urbana de 72%. En la capital, Puebla de Zaragoza, vivía 26% de la población del estado (1 539 819 habitantes), lo que convertía al municipio en el cuarto más poblado del país, por detrás de la delegación (hoy alcaldía) Iztapalapa (con 1 815 786 habitantes), Ecatepec de Morelos (1 656 107) y Tijuana (1 559 683) (INEGI 2010).

⁴ Sirva de referencia saber que el promedio nacional de escolaridad es de 9.64 años de estudio. Puebla ocupa el lugar 26 de las 32 entidades federativas, encabezadas por la Ciudad de México (con un promedio de 11.48 años de escolaridad), Nuevo León (10.76) y Querétaro (10.48). Detrás de Puebla, y para cerrar la lista, se encuentran Guanajuato (9.04), Veracruz (8.75), Michoacán (8.6), Guerrero (8.37), Oaxaca (8.12) y Chiapas (7.78) (INEGI, 2020).

Imagen 1. El estado de Puebla: resumen geopoblacional

Resumen



FUENTE: INEGI, 2020: Información por entidad. Consultado el 25 de septiembre de 2021 en <<http://cuentame.inegi.org.mx/monografias/informacion/pue/default.aspx?tema=me&e=21>>.

- **Capital:** Heroica Puebla de Zaragoza.
- **Municipios:** 217.
- **Extensión:** Representa 1.7 % del territorio nacional.
- **Población:** 6,583,278 habitantes, el 5.2 % del total del país.
- **Distribución de población:**
73 % urbana y 27 % rural; a nivel nacional el dato es de 79 y 21 % respectivamente.
- **Escolaridad:** 9.2 (poco más de tercer año de secundaria); 9.7 el promedio nacional.
- **Hablantes de lengua indígena de 3 años y más:** 10 de cada 100 personas.
A nivel nacional 6 de cada 100 personas hablan lengua indígena.
- **Sector de actividad que más aporta al PIB estatal:** Comercio.
- **Aportación al PIB Nacional en 2019:** 3.3 %.

Tabla 1. Habitantes de pueblos indígenas en el estado de Puebla

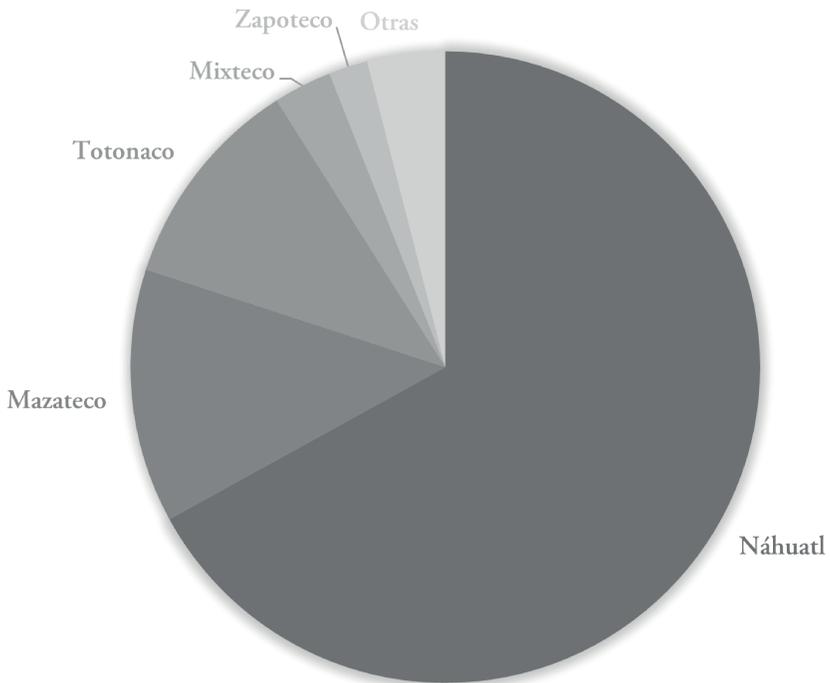
Pueblo indígena	Población
Náhuatl	814 153
Totonaco	152 562
Mazateco	39 853
Popoloca	25 516
Mixteco	19 987
Otomí	14 246
No especificado	10 980
Zapoteco	5 639
Cuicateco	2 135
Maya	1 514
Chinanteco	1 274
Tlapaneco	1 265
Mazahua	976
Mixe	731
Tepehua	680
Tsotsil	554
Ch'ol	440
Tseltal	431
Huasteco	395
Tarasco	313
Popoloca insuficientemente especificado	241
Chatino	171
Huichol	131
Amuzgo	105
Q'anjob'al	96
Chontal insuficientemente especificado	90
Tojolabal	73
Chocholteco	67
Triqui	66
Otras lenguas de América	57
Huave	50
Zoque	33
Mam	32
Chichimeco jonaz	16
Mayo	16

Tabla 1 (Continúa)

Pueblo indígena	Población
Cora	14
Tarahumara	9
K'iche'	8
Tepehuano insuficientemente especificado	2
Akateko	1
Chuj	1
Total	1 094 923

FUENTE: *Atlas de los pueblos indígenas* (INPI e Inali, 2020).

Imagen 2. Lenguas más habladas en la ciudad de Puebla



FUENTE: *Atlas de los pueblos indígenas* (INPI e Inali, 2020).

La presencia de estos 41 grupos indígenas en el estado y de 16⁵ en el municipio de Puebla nos permite vislumbrar la compleja convivencia étnica dentro de la capital poblana. Sin embargo, si bien ésta es la conformación actual de la población hablante de lengua indígena en Puebla, requerimos ahondar en su composición, y en la descripción social y formal de las comunidades, sus lenguas y sus dialectos. Será necesario estudiar el contacto dialectal y lingüístico de estos grupos, sus actitudes ante la lengua y los grados de bilingüismo. Por ejemplo, ¿se trata de comunidades originarias, de migrantes enraizados o de migrantes temporales? ¿Se encuentran aislados o integrados al resto de la población? Más allá de la lengua (o quizá incluso familia lingüística), ¿cuáles son los dialectos que entran en contacto?

Sabemos que durante el siglo xx la migración interna en México estuvo relacionada con procesos de urbanización e industrialización del país; sin embargo, es poco lo que conocemos sobre la conformación de redes comunitarias y de asentamientos de los grupos indígenas en la ciudad de Puebla. Suponemos que no hay hablantes monolingües de lengua indígena que vivan en la capital poblana, pues incluso las comunidades de lengua materna náhuatl, como la de San Miguel Canoa, se han visto obligadas a aprender el español. Como afirma Olivar Espinosa (2020: 66):

La junta auxiliar de San Miguel Canoa es una comunidad que ha estado históricamente vinculada a la ciudad de Puebla, a través de una relación de carácter económico y político que promueve la movilidad de sus habitantes por cuestiones laborales en primer término y, después, para la búsqueda del derecho a la educación y la salud. Lo que ha llevado a que sus habitantes se vean en la necesidad de aprender una segunda lengua, en este caso el español, además del náhuatl, el cual había sido la lengua materna hasta hace poco, generando en la comunidad un estado de bilingüismo permanente en las últimas décadas.

⁵ El *Atlas de los pueblos indígenas* (INPI e Inali, 2020) reporta la presencia de hablantes de las siguientes lenguas dentro del municipio: cuicateco, chatino, chinanteco, ch'ol, huasteco, maya, mazahua, mazateco, mixe, mixteco, náhuatl, otomí, popoloca, q'anjob'al, tarasco, tepehua, tlapaneco, totonaco, tseltal, tsotsil y zapoteco.

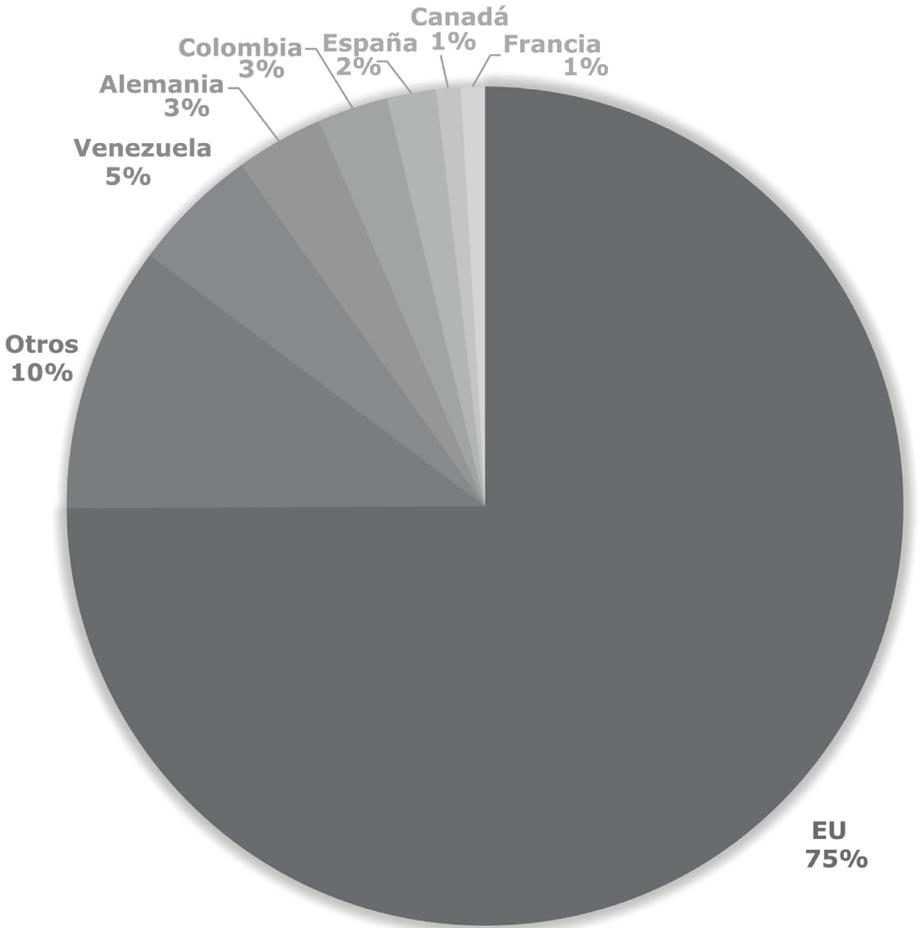
Tenemos noticia de que existe una red de origen mazateco asentada en la Junta Auxiliar de San Aparicio, cuya segunda generación ha nacido en Puebla y se ha integrado escolar y laboralmente a la ciudad, pero no tenemos estudios sociolingüísticos de contacto y variación. Al revisar los datos del *Atlas de los pueblos indígenas*, nos llama la atención el asentamiento en la zona conurbada de huicholes en Atlixco y Cholula, sin presencia en la ciudad de Puebla, y el de población cho'ol en Cuautlancingo. ¿Cuál es su situación migratoria? ¿Cuáles son sus principales problemáticas? ¿Cuáles son sus espacios de inserción?

Durante el periodo 2015-2020, el estado de Puebla ocupó el quinto lugar en emigración y el noveno en inmigración: 152 359 personas salieron para radicar en otras entidades del país y 165 720 llegaron al estado de Puebla. La migración interna proviene de Veracruz (19%), Estado de México (18%), Ciudad de México (16%), Oaxaca (8%) y Tlaxcala (5%). Las causas principales de esta movilidad son “reunirse con la familia” (38.3%), “cambio u oferta de trabajo” (12.4%), “búsqueda de trabajo” (11.4%) y “estudiar” (10.7%). Las entidades hacia donde se desplazaron los habitantes de Puebla fueron exactamente las mismas entidades —lo que nos habla de un vínculo social fuerte entre esta región—, es decir, el Estado de México (15%), la Ciudad de México (13%), Veracruz de la Llave (11%), Tlaxcala y Oaxaca (9 y 6%, respectivamente) (INEGI, 2020).

Flujos migratorios externos

El Instituto Nacional de Estadística y Geografía (INEGI, 2020) reporta que los flujos migratorios externos, esto es, los correspondientes a las personas que cambiaron su residencia entre 2015 y 2020 (30 819 personas nacidas en otros países), provienen mayoritariamente de los Estados Unidos de América (más de 60%, es decir, 19 820), seguidos distantemente por personas originarias de la República Bolivariana de Venezuela (5%). En la imagen 3 puede observarse la distribución de los otros países de origen de los inmigrantes. Nótese que 10% (2 204 inmigrantes) de este movimiento migratorio internacional corresponde a procedencias muy diversas, que enriquecen todavía más este mosaico cultural.

Imagen 3. Migración externa o internacional al estado de Puebla



FUENTE: INEGI, 2020.

Mayoritariamente, los inmigrantes externos o extranjeros se asientan en los municipios de Puebla (7 529 personas), San Andrés Cholula (2 783), Atlixco (1 277) y San Pedro Cholula (1 236).

La principal razón para inmigrar desde los Estados Unidos de América es reunirse con la familia (64.6% ha reportado este motivo).⁶ En segundo lugar están los migrantes de la República Bolivariana de Venezuela (quienes en 30.9% han salido de su país por inseguridad delictiva o violencia). En tercero, de la República Federal de Alemania (de los cuales 35% cambió su residencia por cuestiones laborales). Los inmigrantes colombianos, españoles y canadienses (que en conjunto suman 6%) declaran haber viajado para reunirse con la familia. Por último, queremos destacar la inmigración francesa, que, si bien representa apenas 1% del movimiento migratorio internacional hacia Puebla, reporta que lo hace para estudiar (de 202 personas, 46% inmigró por este motivo).

Transmigración

Puebla es lugar de origen, recepción, repatriación y tránsito migratorio. El paso de transmigrantes centroamericanos y de migrantes del sur de México con destino, principalmente, a los Estados Unidos de América genera un flujo migratorio del que no podemos tener información cuantitativa precisa; sin embargo, gracias a la información recopilada por organizaciones no gubernamentales, defensores de derechos humanos, la Red de Documentación de las Organizaciones Defensoras de Migrantes (Redodem) y, para el caso específico de Pue-

⁶ Es necesario estudiar a detalle los vínculos sobre migración interna y la migración internacional, con especial atención en la manera como la descendencia de los migrantes propicia movilidad social, migración de retorno y contacto lingüístico. Véase, por ejemplo, el interesante trabajo de Vázquez Zúñiga (2020) sobre la migración en el circuito Puebla-Nueva York, con datos de Tulcingo de Valle. La autora analiza los factores que han propiciado la migración unidireccional de la Mixteca poblana hacia los Estados Unidos de América (crisis económica, desempleo, inflación, devaluaciones, desigualdad, marginación falta de oportunidades de desarrollo familiar, etcétera), los cuales fueron la base de la aceleración migratoria de la década de 1980; estudia también la migración circular y el retorno generados tras la política discriminatoria y xenófoba impuesta por Donald Trump.

bla, la Pastoral de Movilidad Humana de la Arquidiócesis de Puebla, se conoce que va en aumento (Sánchez Gavi, 2016; Redodem, 2019).

Entre 2012 y 2013 el Instituto Nacional de Migración (INM) reconoció que el flujo migratorio había aumentado entre 10 y 20% y que, debido a la inseguridad que viven los migrantes en su paso por México, habían buscado nuevas rutas de tránsito, por lo que disminuyó su flujo por el Golfo de México y el Pacífico, y aumentó por el centro del país.

Según datos obtenidos de forma directa por Redodem (de los migrantes que responden sus cuestionarios en alguno de sus 23 albergues, ubicados en 13 estados de la república),⁷ de enero a diciembre de 2019 los transmigrantes en México fueron mayoritariamente hondureños (76%), guatemaltecos (8%) y salvadoreños (7%).

En la ciudad de Puebla hay 6 albergues de la Arquidiócesis destinados a brindar alimento y descanso a los migrantes. De acuerdo con los registros de esta organización, por la capital poblana pasan más de 2 000 migrantes al mes, sobre todo, hondureños (90% de las personas registradas) y salvadoreños. Hasta donde hemos podido investigar, no hay asentamientos de inmigrantes centroamericanos en la ciudad; si bien algunas personas se han establecido en ella, lo han hecho de forma individual y sin regularizar su situación migratoria. Parece pronto para conocer cuál será la influencia de este flujo migratorio en la comunidad y en la (re)construcción de las identidades sociales y lingüísticas de sus diversos participantes.⁸

⁷ Chihuahua, Chihuahua; Guadalupe, Nuevo León; Torreón, Coahuila; San Luis Potosí, San Luis Potosí; Tepic, Nayarit; Irapuato, Guanajuato; Salamanca, Guanajuato; Guadalajara, Jalisco; Celaya, Guanajuato; Querétaro, Querétaro; Huichapan y Atitalaquía, Hidalgo; Ciudad de México; Tierra Blanca y Oluta, Veracruz; Oaxaca e Ixtepec, Oaxaca, y Palenque, Arriaga y Frontera Comalapa, Chiapas.

⁸ Queremos apuntar que, además de la importancia del contacto lingüístico y dialectal entre transmigrantes, hemos podido presenciar (y se ha documentado en reportajes de distintas fuentes periodísticas) que se ha generado, entre quienes piden limosna en esquinas y camiones, una estafa de adultos jóvenes que, siendo mexicanos, “hablan como hondureños”. La trampa se descubre cuando algún agente de migración pide documentación y, para evitar problemas, ellos presentan su documentación oficial como mexicanos y cambian inmediatamente su forma de expresarse. Valdría la pena estudiar este fenómeno en sus distintas caras y, desde el

Otras migraciones

La comunidad libanesa

Las grandes oleadas de inmigrantes libaneses a México comenzaron a finales del siglo XIX, originadas por las reestructuraciones en los sistemas tradicionales de producción del sistema otomano. La reconstrucción de su origen y dispersión resulta difícil debido, entre otras razones, a que:

los conceptos de “árabe” o de “libanés” eran inexistentes para definir una nacionalidad en el siglo XIX. Todos los inmigrantes del Medio Oriente que ingresaban en el país antes de la Primera Guerra Mundial eran considerados como “turcos”, término indebido en la medida a que se refiere a una realidad posterior a la primera conflagración mundial, o “árabes”, concepto prácticamente inexistente en este periodo. Hasta 1920, el término más adecuado debe ser el de “otomano”. Todo el Medio Oriente de hoy, a excepción de Egipto, formaba parte a finales del siglo XIX del Imperio Otomano (Zeraoui, 2006: 17-18).

Sabemos, sin embargo, que la apertura de la política migratoria del gobierno porfirista facilitó la llegada de libaneses al Golfo de México (Tampico, Veracruz y la península de Yucatán) y de palestinos y sirios a Monclova, Saltillo y Monterrey. La migración disminuyó en los años treinta, cuando la política mexicana se volvió más restrictiva (Inclán, 1995).

A su llegada, los libaneses, independientemente de sus profesiones o condiciones sociales previas, se ocuparon en el comercio, primero en la venta de productos que adquirían a comerciantes franceses, pero poco a poco se fueron independizando y comenzaron a importar o producir su propia mercancía. Se convirtieron en promotores de una práctica comercial que se arraigó entre los mexicanos: el comercio

ámbito lingüístico, nos parece que resulta interesante analizar los rasgos que les parecen prominentes a estos hablantes para simular el dialecto hondureño o centroamericano.

por abonos. Para la década de 1940, mediante la apertura de negocios familiares, establecieron algunas de las textileras más importantes de México, y en la actualidad poseen negocios que se encuentran entre los más prósperos de Puebla, ya no sólo en el ramo textil. La diáspora libanesa ha permitido que establezcan redes comerciales prácticamente por todo el mundo.

Ramírez Carrillo (1994: 457-458) afirma que “Un directorio libanés de 1948 menciona a 16 403 personas de ese origen radicadas en la República mexicana; 5 431 de ellas (33%) habitaban en la Ciudad de México, seguidas en orden cuantitativo por 1 550 radicadas en Yucatán (9.5%); 1 447 (7.2%) en Veracruz, y 1 188 (7.2%) en Puebla”. Pese a la falta de registro sistemático, estos datos coinciden cercanamente con los que ofrece Zeraoui (1997: 297-298) cuando explica que la migración árabe en Puebla estaba conformada por 1 270 personas en 1948; el autor agrega que mayoritariamente éstas eran de origen libanés (1 180 personas), seguidas por sirios (78 personas), y que el conjunto apenas estaba representado por palestinos (2 personas). El arraigo libanés en Puebla se facilitó porque en su mayoría formaban parte de la iglesia católica maronita, lo que favoreció su pronta integración a la comunidad poblana.

La asimilación y la aculturación de la comunidad, cuyos miembros estaban arraigados ya con sus familias en México, se perciben prácticamente en cualquier ámbito de la dinámica social, política y económica. Casi todos los miembros de esta comunidad se nacionalizaron y buscaron al mismo tiempo conservar y fomentar sus costumbres y adaptarse al contexto mexicano (Marín y Zeraoui, 2003). La comunidad mexicano-libanesa ha creado distintos espacios de convivencia social que le permiten promover la cultura libanesa y algunas tradiciones árabes, con el fin de preservarlas en las nuevas generaciones y para que, a la llegada de nuevos inmigrantes, éstos puedan sentirse en comunidad y recibir el apoyo de la red.

La apertura del Centro Mexicano Libanés (31 de diciembre de 1972) es un claro ejemplo de convivencia ente las comunidades, pues en él los poblanos pueden conocer, por ejemplo, la gastronomía libanesa. La inclusión y la adaptación de alimentos árabes en la comida mexicana puede notarse, por ejemplo, en la adopción de pan árabe,

jocoque, hummus, dedos de novia, café árabe y algunas prácticas de su preparación y consumo. Además, en Puebla el contacto ha dado lugar a uno de los antojitos tradicionales modernos: los tacos árabes, adaptación de un alimento sirio a la cocina mexicana que se prepara con pan árabe relleno de carne de cerdo condimentada.

Inmigrantes argentinos y chilenos en la década de 1970

No queremos concluir esta mirada panorámica sin al menos mencionar a los inmigrantes chilenos y argentinos llegados a México en la década de 1970 y cuya integración a la vida universitaria de Puebla ha sido relevante, por sus aportes docentes, la fundación de institutos de investigación, la formación de conciencia cívica en los jóvenes universitarios y la participación política en la institución. A todos ellos debemos un estudio propio.

Tras el golpe militar a Salvador Allende (11 de septiembre de 1973), simpatizantes de la Unidad Popular (principalmente universitarios, artistas, científicos y literatos) perseguidos por la dictadura de Pinochet (1973-1990) buscaron asilo en México, y los intelectuales mexicanos simpatizantes de las políticas de izquierda los recibieron con agrado.

Prácticamente de manera paralela, la dictadura que gobernó a la república de la Argentina tras el golpe de Estado de 1976 (Proceso de Reorganización Nacional) así como diversas medidas de represión que venían gestándose una década atrás provocaron el exilio masivo de estudiantes y docentes de la comunidad universitaria de Córdoba —Universidad Nacional de Córdoba (UNC) y Universidad Católica de Córdoba (UCC)—. “Según el Registro unificado de víctimas del terrorismo de Estado, entre 1966 y 1983, hubo «7 018 casos de personas víctimas de desaparición forzada y 1 613 víctimas de asesinato». Más de 300 000 argentinos tomaron el camino del exilio sin contar aquellos que se fueron antes del 24 de marzo de 1976” (BUAP, “Los exilios latinoamericanos”).

A las distintas unidades académicas de la Benemérita Universidad Autónoma de Puebla (BUAP) se integraron profesores chilenos y argentinos, y como afirma en entrevista Jaime Estay Reyno:

Más allá de la relativa coincidencia ideológica que puede identificarse entre por una parte los exiliados y, por otra parte, las corrientes y grupos que predominaban en la universidad, ese principio ético de solidaridad se aplicó de manera irrestricta hacia un grupo de personas que huían de la represión de las dictaduras militares del Cono Sur, y con seguridad que ello pasó a formar parte no sólo de la experiencia personal de los trabajadores universitarios de aquel entonces, sino también de los valores que los estudiantes de la época recogieron de su vida universitaria (BUAP, “Los exilios latinoamericanos”).

Territorio

El municipio de Puebla es uno de los 217 que conforman el estado del mismo nombre. Su cabecera municipal, la ciudad de Puebla de Zaragoza, es también la capital del estado y la ciudad principal en la Zona Metropolitana Puebla-Tlaxcala. Tiene una extensión territorial de 556.71 kilómetros cuadrados⁹ y una población de 1 692 181 habitantes (INEGI, 2020), por lo que es el cuarto más poblado del país por detrás de Tijuana (1 922 523 habitantes), Iztapalapa (1 835 486 habitantes) y León (1 721 215). Limita con 14 municipios: 10 poblanos y cuatro del estado de Tlaxcala. La colindancia marca distintas relaciones con los municipios vecinos: hacia el norte y el poniente se forma el corredor industrial, mientras que en los municipios del oriente y el sureste predomina la actividad agropecuaria.

Para el proyecto Preseea-Puebla hemos determinado levantar una muestra de habla delimitada por el criterio de la división ad-

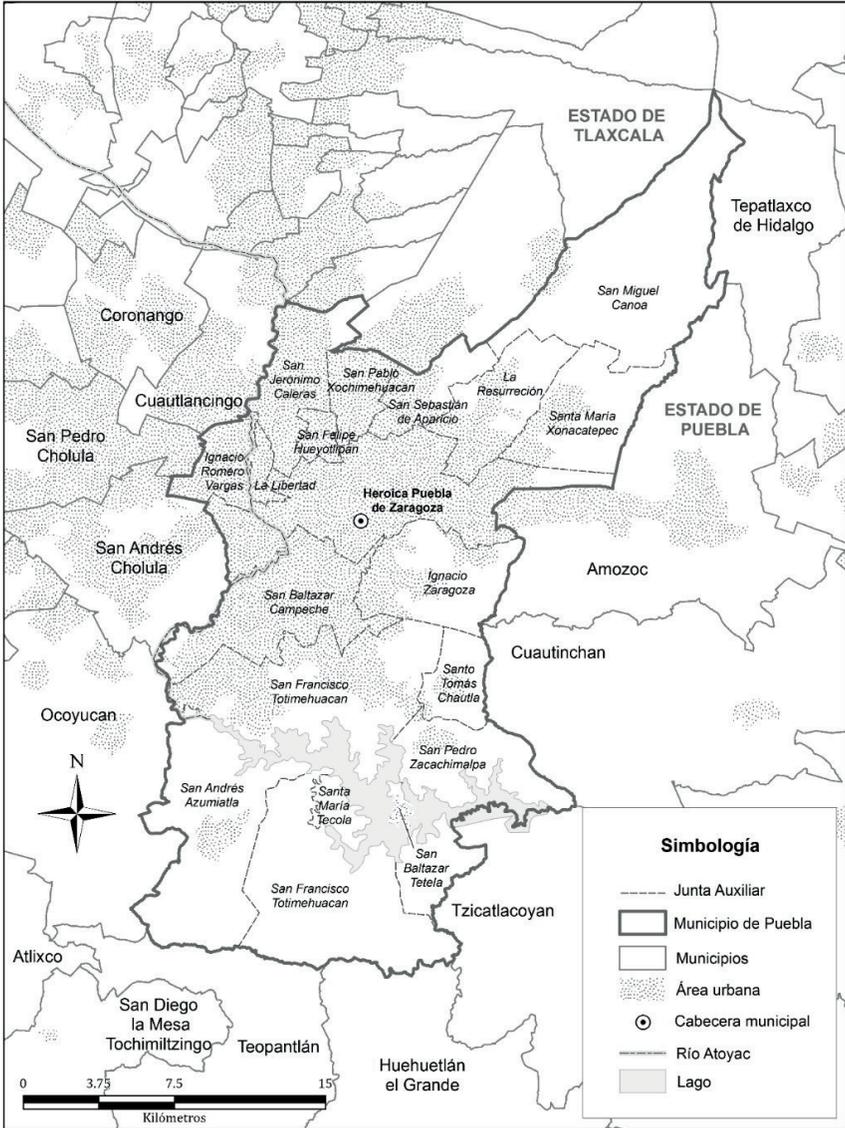
⁹ Para nuestro estudio tomamos en cuenta la información que mejor representa el estado del momento en el que se recogieron las entrevistas, pero sirva de referencia decir que en poco más de 50 años el territorio capitalino ha cambiado en tres ocasiones: de 123.17 kilómetros cuadrados antes del decreto de 1962, a 561.35 kilómetros cuadrados (precisamente con el decreto de 1962), y a 556.71 kilómetros cuadrados, según el dictamen de resolución de límites emitido por el Honorable Congreso del Estado de Puebla el 6 de diciembre de 2013 (Gobierno del Estado de Puebla y H. Ayuntamiento del Municipio de Puebla, 2016: 31).

ministrativa, es decir, por la conformación municipal. Es importante aclarar que la decisión no responde a criterios lingüísticos como el análisis y la delimitación de la comunidad de habla, sino que se basa en las siguientes razones prácticas: la limitación de recursos económicos para sufragar gastos de desplazamiento de los integrantes del proyecto, y las condiciones de inseguridad para visitar algunas zonas de la región. No obstante, como ya hemos descrito al hablar de su población, consideramos que el municipio cumple con los requisitos básicos del Proyecto para el Estudio Sociolingüístico del Español de España y América (Preseaa) (Moreno Fernández, 2003), es decir, se trata de un núcleo urbano principalmente monolingüe de español cuya población ha estado tradicionalmente asentada en el lugar; presenta heterogeneidad sociológica, y es un espacio apropiado para el estudio de la identidad social y lingüística de distintos grupos, ya sea en el conjunto de datos recogidos, ya para el estudio de grupos específicos o de hablantes particulares.

Administrativamente, el municipio se divide en la cabecera municipal o ciudad de Puebla de Zaragoza y 17 juntas auxiliares¹⁰ (Ignacio Romero Vargas, Ignacio Zaragoza, La Libertad, La Resurrección, San Andrés Azumiatla, San Baltazar Campeche, San Baltazar Tetela, San Miguel Canoa, San Felipe Hueyotlipan, San Francisco Totimehuacán, San Jerónimo Caleras, San Pedro Zacachimalpa, San Sebastián de Aparicio, Santa María Xonacatepec, Santo Tomás Chautla, Santa María Tecola y San Pablo Xochimehuacan), como puede verse en la imagen 4.

¹⁰ En el artículo 224 de la Ley Orgánica Municipal (23 de marzo de 2001) se define que: “Las Juntas Auxiliares son órganos desconcentrados de la administración pública municipal y estarán supeditadas al Ayuntamiento del Municipio del que formen parte, sujetos a la coordinación con las dependencias y entidades de la administración pública municipal, en aquellas facultades administrativas que desarrollen dentro de su circunscripción. El vínculo de información e interacción será la Secretaría de Gobernación Municipal o su equivalente en la estructura administrativa” (Gobierno del Estado de Puebla y Orden Jurídico Poblano, 2001: 126).

Imagen 4. Mapa del municipio de Puebla de Zaragoza



FUENTE: elaborado para esta publicación por el Departamento de Sistemas de Información Geográfica de El Colegio de México.

EL CORPUS SOCIOLINGÜÍSTICO
DE LA CIUDAD DE PUEBLA (CSCP)

A principios del año 2011, con un renovado interés por recoger un corpus del habla regional y comenzar a desarrollar distintos proyectos sobre el español de Puebla a los que ello pudiera dar paso, por ejemplo, a la elaboración de un diccionario integral, comenzamos a preparar los recursos para levantar datos regionales bajo dos modelos: el Corpus del Español Mexicano Contemporáneo (CEMC) y los estudios dialectales sobre el habla de Culiacán (Mendoza, 1996, 2003 y 2011). El propósito central de este proyecto fue, desde su comienzo, enfatizar que, si bien el español del Altiplano guarda muchas similitudes debido a la hegemonía cultural de la Ciudad de México, debemos trabajar en el estudio de las hablas locales que lo integran para entender su conformación tanto histórica como actual y confirmar o refutar con datos de uso si la similitud con el español hablado en la Ciudad de México es tan fehaciente como se ha supuesto.

Pronto nos dimos cuenta de que un proyecto de tal envergadura sólo puede desarrollarse en fases subsecuentes y de que posiblemente es tarea para muchos equipos de trabajo. Decidimos comenzar por el estudio de textos orales y adoptamos el método propuesto por Pedro Martín Butragueño y Yolanda Lastra para la conformación del Corpus Sociolingüístico de la Ciudad de México (CEMC, 2011-2015). Nació así el Corpus Sociolingüístico de la Ciudad de Puebla, donde planeábamos recoger distintos módulos de entrevistas, entre ellos los dos que describiremos a continuación: el corpus léxico de oficios y la entrevista de Preseea-Puebla (acompañada por pruebas lingüísticas y un cuestionario sociológico).

*El corpus léxico de oficios,
flora y fauna (2008-2016)*

En el marco de la actualización del Modelo Universitario Minerva (BUAP, 2007) se incluyeron en la currícula obligatoria de la licenciatura en Lingüística y Literatura Hispánica los cursos de Lexicología y Dialectología. Esto nos dio la posibilidad de recopilar, paralelamente

a Preseea-Puebla, vocabulario de parcelas léxicas específicas recogido con un método de trabajo común para el estudio de vocabularios tradicionales. Para ello seguimos la propuesta de Palacios (2010 y 2014), cuya base es el acercamiento pragmático de la lexicología (datos reales y en contexto discursivo como fundamento de la investigación, en vez de respuestas aisladas a cuestionarios) y cuyo interés teórico es conocer la manera en la que confluyen distintos procesos de significación en la construcción de vocabularios especializados que están en la intersección entre la tradición popular y la terminología y que, precisamente por este carácter límite, suelen quedar poco representados en estudios de lengua general y fuera de los estudios terminológicos; dicho de otra forma, se trata de voces muy especializadas como para registrarse en corpus de habla general, y “demasiado populares” para registrarse en documentos terminológicos. Otras veces, para su documentación sólo se cuenta con fuentes secundarias, en especial escritas por antropólogos o etnólogos y recogidas con poca precisión semántica o referencial, pues los objetivos que persiguen dichos trabajos son distintos a los fines léxicos.

Para este corpus se parte del supuesto de que, en las prácticas de oficios tradicionales, además de documentar un vocabulario de alto valor cultural e identitario, pueden observarse rasgos semánticos cognitivos y comunicativos propios de los vocabularios de especialidad (Cabré, 1993; Lara, 2001; Palacios, 2014, 2020a y 2020b). Sirvan para ilustrar los siguientes ejemplos de la conformación del vocabulario de la panadería en su *continuum* entre la lengua común y la especializada. Dentro del léxico común encontraremos palabras como *galleta*, *harina* y *masa*, mientras que en el léxico especializado (que se encuentra en la frontera con el léxico común), voces como *afinar* ‘poner manteca a la mezcla de harina’ o *palote* ‘rodillo grande de madera que sirve para extender la masa’ y, dentro del léxico más especializado, *horno de piso* ‘el que tiene una base refractaria que permite que el pan rústico tenga una cara dura’, y *pan de piso* ‘el que se cuece en dicho horno’. En este vocabulario proliferan (como en todos los vocabularios especializados) denominaciones sintagmáticas que entran en oposición o complementariedad en una clasificación: *pan francés*, *pan español*, *pan de figura*; *horno de columpio*, *horno de pala*, *horno de gaveta*.

También encontramos designaciones que permiten hacer distinciones necesarias para los usuarios especialistas, pero generalmente desconocidas por el hablante no especialista: *tahonero* ‘persona que tiene por oficio hacer el pan’; *panadero* ‘persona que vende el pan, generalmente en la calle’ y *expendedor* ‘persona que vende el pan en un expendio o panadería’.

La propuesta de método combina los estudios tradicionales de la dialectología, especialmente en la selección de los temas (la flora y la fauna locales, la siembra y los oficios característicos de cada comunidad, por ejemplo, la siembra de durazno en Huejotzingo o la elaboración de objetos de piedra en San Salvador el Seco) y la entrevista sociolingüística semidirigida que se organiza mediante redes sociales o se preestratifica por medio de la variable especialidad en el trabajo (maestro, oficial, aprendiz), del sexo (cuando hombres y mujeres tienen responsabilidades diferentes) y de la organización onomasiológica de las terminologías que determinan los temas tratados en la entrevista. Esta organización previa permite ordenar después el vocabulario en áreas temáticas o esferas semánticas.

El objetivo es registrar la mayor cantidad de palabras del vocabulario en el habla del especialista, para lo cual se le entrevista en su lugar de trabajo, en presencia de sus materiales y con la posibilidad de que describa sus herramientas e insumos, narre sus procesos de trabajo y explique las distintas especialidades que éste tiene, puesto que es común que en tales actividades se mantenga una estructura gremial.

Llevar a cabo las entrevistas en el lugar de trabajo tiene otras ventajas: se favorece la producción léxica, porque, al estar en presencia de sus objetos de trabajo, el hablante recuerda un mayor número de palabras asociadas con su especialidad; en caso de estar obviando alguna, el entrevistador (quien no es especialista) puede señalar y preguntar “¿y eso para qué sirve?”, o “¿cómo se usa esto?”. Así se registran más nombres de insumos, herramientas o partes de la maquinaria.

Se pide que por tema se entreviste como mínimo a tres informantes. Para integrar una palabra al corpus, ésta debe aparecer al menos en dos entrevistas, con el fin de evitar incluir hápax en el estudio (Lara, 2006: 152); desde luego, cuando la palabra se registra una sola

vez, pero se puede documentar en fuentes lexicográficas, se privilegia el factor cualitativo sobre el cuantitativo.

Con este método de trabajo se recolectaron más de 120 entrevistas sobre la religiosidad popular, la herbolaria y la medicina tradicional, la siembra, los árboles frutales, las plantas de ornato, las aves, el pulque, la lucha libre, el mariachi, la carnicería, la peletería, los textiles, la panadería, los antojitos —“bocadillos típicos de la cocina mexicana. Muchos de ellos se elaboran con tortilla u otras preparaciones del maíz y se condimentan con alguna salsa picante” (DEM)— y el trabajo con piedra volcánica (como la elaboración de molcajetes, fuentes y distintas figuras de ornato).

Este subcorpus del CSCP es el único que se recogió más allá de los límites del municipio de Puebla, por lo que, además de hablantes de la capital, tenemos datos de la Sierra Norte, la Mixteca, Atlixco, Cholula, Huejotzingo, San Martín Texmelucan, San Salvador el Seco, y de dos poblaciones de Tlaxcala, San Bernardino Contla y San Pablo del Monte.

Preseea-Puebla

Preseea-Puebla está integrado por 108 entrevistas, recogidas en su mayoría en el año 2012, bajo la coordinación de Laura Romero Rangel y Niktelol Palacios, y con la participación de estudiantes del Colegio de Lingüística y Literatura Hispánica de la BUAP. En los años siguientes (hasta los primeros meses de 2017) se levantaron datos con el fin de mejorar la representatividad de algunos sectores y la dispersión geográfica de los colaboradores en un mayor número de zonas de la ciudad. Nuestra colega Stephany Olivar Espinosa contribuyó también con entrevistas en la comunidad bilingüe (español-náhuatl) de la Junta Auxiliar de San Miguel Canoa, material que es parte de su investigación doctoral y que ha aportado una muestra muy interesante a nuestro corpus de conversaciones. Hasta el momento hemos recogido 126 entrevistas y más de 200 horas de conversación grabada; de ellas, hemos elegido 108 para el corpus Preseea y las otras forman parte del subcorpus o corpus de respaldo dentro del CSCP.

LA SELECCIÓN DE LOS INFORMANTES

El método de trabajo de Preseea marca claramente las variables sociales a considerar en la muestra tipo. Por lo tanto, hemos seleccionado a nuestros informantes según tres criterios con valor estratificatorio: la edad, el sexo y la instrucción formal. La muestra está constituida por 18 casillas, y ya que el modelo Preseea deja a cada grupo de trabajo la libertad de considerar el número de entrevistas que se realizarán por cada variable social, nuestro equipo decidió, en una primera fase, entrevistar a 72 informantes (cuatro por cada variable), pero posteriormente se prefirió seguir la decisión de los equipos Preseea-Monterrey y Ciudad de México: recoger seis entrevistas por cada variable. La muestra final quedó conformada como se resume en la tabla 2.

La distribución de los entrevistados queda establecida de la siguiente manera:

1. La variable sexo divide la muestra en hombres y mujeres.
2. En la variable edad hay tres grupos: de 20 a 34 años, de 35 a 54 y de 55 años en adelante.
3. La variable grado de instrucción distingue tres grupos: nivel de instrucción baja, integrado por personas que tienen como máximo la educación primaria (hasta seis años de escolaridad); de instrucción media, que agrupa a quienes cursaron la escuela secundaria, preparatoria o carrera técnica (aproximadamente 12 años de estudios), e instrucción alta, conformado por quienes tienen estudios de licenciatura, técnico superior, especialidad o posgrado (un mínimo de 16 años de estudios formales).

Los informantes se seleccionan según los siguientes criterios:

1. Que hayan nacido en la ciudad de Puebla y jamás hayan vivido fuera del municipio.
2. Que hayan nacido en Puebla y salieran del municipio, pero en el momento de la entrevista tuviesen un mínimo de 20 años de haber regresado.

3. Que hayan nacido fuera del municipio, pero que llegaran antes de los 10 años y tuvieran como mínimo 20 años residiendo en él. Es importante destacar que el lugar de origen no puede ser marcadamente distinto a la capital poblana.

Como ya hemos mencionado, cuando se decidió ampliar el corpus para tener 108 entrevistas, se integró al mismo una pequeña submuestra de entrevistas a hablantes de la Junta Auxiliar de San Miguel Canoa (Olivar, en elaboración).¹¹ Las entrevistas corresponden a las variables H21, H31, H32, M12, M13, M21 y M31, e integramos a Preseea-Puebla tres de ellas, lo que nos permite tener datos sobre población bilingüe náhuatl-español y español-náhuatl que pueden servir como referencia para la comparación con los hablantes monolingües de la ciudad de Puebla o para distintos estudios de contacto lingüístico. En el mapa de la imagen 5 puede notarse con claridad en el noreste la distribución aislada de estos entrevistados en relación con el resto de la muestra.

En la imagen 5 se muestra la distribución geográfica de los entrevistados. Cabe aclarar que se trata de un acercamiento aproximado, puesto que, para salvaguardar el anonimato, no pedimos la dirección de los entrevistados, únicamente preguntamos por su colonia.¹²

LA ENTREVISTA

La entrevista consta de tres partes: la conversación, la aplicación de pruebas lingüísticas y la aplicación de un cuestionario sociológico. Todas las secciones fueron grabadas. Las conversaciones se han pla-

¹¹ Para un adelanto sobre el método de recolección de datos y la estratificación de los 14 hablantes que Olivar entrevista, véase su artículo “La entonación de los enunciados aseverativos del español de contacto de hablantes bilingües de San Miguel Canoa, Puebla: una variedad del español mexicano” (2020).

¹² Una colonia es una localidad sin autonomía jurisdiccional, pero con código postal propio. Es la forma más común de identificar la vivienda en la ciudad de Puebla. En algunos casos, el nombre de la colonia coincide con el de la Junta Auxiliar.

neado para realizarse únicamente entre dos personas, esto es, entre un entrevistado y un entrevistador; sin embargo, en ciertos casos la familia del entrevistado tiene alguna participación, lo que nos permite registrar, dentro de la unidad mayor, pequeñas conversaciones entre varias personas.

Hemos intentado que las grabaciones se efectúen en el ambiente del informante (casa, trabajo, negocio), y únicamente las hemos llevado a cabo en otro sitio (casa del entrevistador o sitio público) cuando el colaborador así lo solicitó. Si la conversación ha sido interrumpida por la llegada de alguna persona, una llamada telefónica, etcétera, continuamos grabando, con el objetivo de obtener diferentes registros de habla del informante. Desde luego, cuando el informante nos solicitó detener la grabación, así lo hicimos, para respetar su privacidad.

Nuestro objetivo fue hacer entrevistas con una duración mínima de 45 minutos, aunque en todos los casos en los que el entrevistado quiso seguir conversando continuamos con la entrevista, por lo cual varias de las interacciones superan las dos horas, y algunas, las tres. En todos los casos asistimos a la entrevista con previa cita habiéndole informado al colaborador que ésta se grabaría en audio —para después ser publicada—, y que constaría de tres partes: una conversación, unas preguntas “que no eran de conocimientos de la escuela” y un cuestionario sociológico que era estrictamente confidencial.

Si bien hemos buscado la realización libre de cada entrevista, bajo el estilo de conversación propio de cada entrevistador y su relación con el entrevistado, en todos los casos hemos propuesto tocar tres temas: la familia, el trabajo y la ciudad. No obstante, ha sido prioridad para nuestro equipo respetar a las personas entrevistadas y sus historias de vida, por lo que, cuando no han querido hablar sobre un tema o quisieron hablar más de uno que de otro, así se ha hecho.

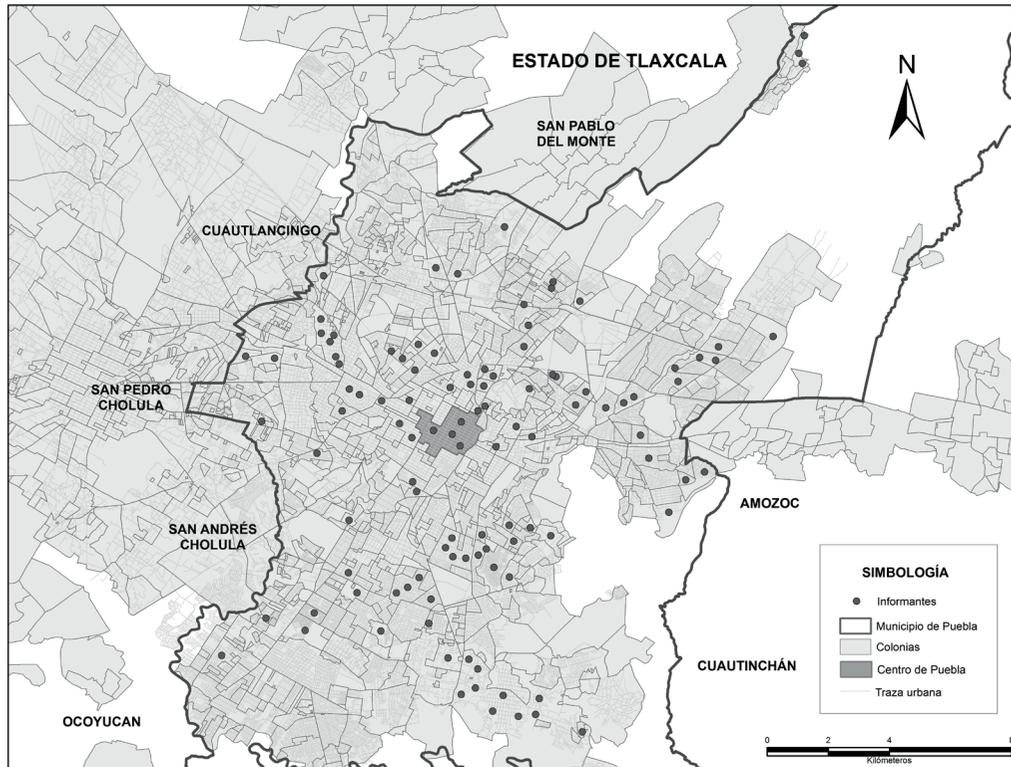
Ahora bien, ya que notamos que los hablantes realizan un cambio de registro según el interlocutor, decidimos aprovechar la diversidad de los integrantes del equipo para hacer la entrevista, es decir, el entrevistador se elegía según se adecuara mejor a las características e intereses del entrevistado.

Tabla 2. Corpus sociolingüístico de la ciudad de Puebla

Edad	1						2						3					
	H			M			H			M			H			M		
Instrucción	1	2	3	1	2	3	1	2	3	1	2	3	1	2	3	1	2	3
Entrevistas	1	7	13	19	25	31	37	43	49	55	61	67	73	79	85	91	97	103
	2	8	14	20	26	32	38	44	50	56	62	68	74	80	86	92	98	104
	3	9	15	21	27	33	39	45	51	57	63	69	75	81	87	93	99	105
	4	10	16	22	28	34	40	46	52	58	64	70	76	82	88	94	100	106
	5	11	17	23	29	35	41	47	53	59	65	71	77	83	89	95	101	107
	12	18	24	30	36	42	48	54	60	66	72	78	84	90	96	102	108	

FUENTE: elaboración propia.

Imagen 5. Distribución de los informantes de Preseea-Puebla



FUENTE: elaborado para esta publicación por el Departamento de Sistemas de Información Geográfica de El Colegio de México.

Los entrevistadores fueron en su mayoría (13 de 15) hablantes de edad 1, instrucción 3, estudiantes de la licenciatura en Lingüística y Literatura Hispánica de la BUAP (siete mujeres y seis hombres): 11 nativohablantes del español del municipio de Puebla que han habitado siempre en él; un hombre cuyo dialecto materno corresponde a la Mixteca poblana, pero que tenía 11 años de radicar en la ciudad cuando hizo las entrevistas, y una mujer que viajaba permanentemente entre la Ciudad de México, el Estado de México y la ciudad de Puebla, y tenía familia en todos esos sitios, aunque radicaba formalmente en Puebla. Las otras dos entrevistadoras son mujeres de edad 2, instrucción 3.

El equipo de entrevistadores y el número de entrevistas realizadas se distribuye según se aprecia en la tabla 3.

Tabla 3. Equipo de entrevistadores de Preseea-Puebla

<i>Entrevistador</i>	<i>Número de entrevistas</i>
Ana Maurel Sánchez de la Barquera Cortez	11
Anayeli Hernández Cruz	24
Angélica Zúñiga Ruiz	10
Carmen Teresa Fajardo Rojas	21
Daniela Saraith Aponte del Moral	1
David Huerta Meza	4
David Jurado Rivera	6
Hugo Carrera Guerrero	7
Karla Roa Suárez	5
Laura Romero Rangel	2
Niktelol Palacios	14
Rogelio Merino Flores	14
Stefany Olivar Espinosa	7
	126

FUENTE: elaboración propia.

LAS PRUEBAS LINGÜÍSTICAS

Después de la entrevista grabada realizamos cuatro pruebas lingüísticas cuyo modelo y orden replican el CSCM (véase 2011: 31-40). A continuación, describiremos las cuatro pruebas recogidas, con el fin de especificar su adecuación para nuestro equipo de trabajo.

Resulta conveniente aclarar que, antes de recoger los datos del corpus, realizamos varios talleres de capacitación para los entrevistadores y levantamos pruebas piloto para asegurarnos de que todos estábamos siguiendo las mismas directrices. Los entrevistadores contamos siempre con un ejemplar impreso de los cuestionarios mientras recogíamos la información.

Lista de palabras

En esta prueba se busca obtener un estilo muy cuidado de pronunciación. El interés se centra en procesos fonéticos como el debilitamiento, el cierre vocálico, los hiatos y diptongos, las consonantes en ataque silábico, las implosivas patrimoniales y las implosivas cultas. Por tratarse de material para el análisis fónico, la calidad de audio de esta prueba fue decisiva para elegir si la grabación sería parte del corpus Preseaa-Puebla o de nuestros archivos de respaldo. Como especifican Pedro Martín Butragueño y Yolanda Lastra, “la inmensa mayoría de las palabras aparecen en el *Atlas Lingüístico de México*, tenido en cuenta para poder ampliar el análisis por medio de la comparación geolingüística, en caso de considerarse interesante” (CSCM, 2011: 35).

Para obtener estas palabras sin que el entrevistador condicionara una respuesta, ideamos el cuestionario 1 (tabla 4), en cuya elaboración usamos distintos recursos, versiones simplificadas de los artículos lexicográficos del DEM, unidades fraseológicas y preguntas de conocimiento general básico. Lo más importante era hacer que las preguntas fueran sencillas para los hablantes. Hicimos varias rondas de pruebas piloto con personas a quienes únicamente les pedimos que respondieran este cuestionario, así fuimos haciendo los ajustes necesarios. Cuando en el cuestionario hay más de una opción para recuperar una palabra, significa que en estas pruebas un núme-

ro considerable de entrevistados no podía responder el ítem con la primera pregunta. Sirva de ejemplo *chiste*: el entrevistador debía leer “Ocurrencia que se dice para hacer reír”, esperaba la respuesta y, si la obtenía, pasaba a “Cómo se les dice a las personas que son originarias de Francia”, pero si no obtenía respuesta a (1), complementaba diciéndole al informante “Como los de Pepito”.

La dinámica de esta prueba fue muy ágil y, para la mayoría de los hablantes, divertida. Dio pie a una interacción en la que el entrevistado actuaba como si estuviera en un juego de adivinanzas; incluso tenemos registro de interacciones del tipo “A ver, échame otra”, “Ésa estuvo fácil”, “¡Ah, chihuahua! ¡Cuál será?”. Los hablantes supusieron que en esta sección se estaba poniendo a prueba su memoria y agilidad mental. No obstante, habrá que esperar a que se realicen estudios fónicos para determinar si la velocidad y la entonación de la respuesta espontánea afectan en algo la pronunciación cuidada que se buscaba.

El cuestionario sirvió como descanso entre la entrevista y las siguientes pruebas, las cuales resultaron más cansadas, en especial la de “Seguridad” y el “Cuestionario sociológico”.

Prueba de seguridad lingüística

Hemos replicado íntegramente la prueba de Pedro Martín Butragueño y Yolanda Lastra (CSCM, 2011: 36). Le leemos al informante pares de palabras y le preguntamos cuál es la que emplea; en una segunda ronda de lectura de los pares le preguntamos si ha escuchado “a otros” decir una forma distinta a la que él emplea y, por último, cuál es la que él considera más correcta. El cuestionario incluye un apartado de seguridad léxica (donde se pregunta, por ejemplo, por pares como *gelatina* / *jaletina* y *búho* / *tecolote*), y uno de seguridad gramatical (*ayer vinimos* / *ayer venimos* o *la sartén* / *el sartén*).

Lectura de texto

La tercera prueba es la lectura de un texto preparado para el CSCM, que puede consultarse íntegro en el corpus coordinado por Pedro Martín Butragueño y Yolanda Lastra (CSCM, 2011: 37). Cabe aclarar

que nosotros no hemos usado esta prueba como lectura de comprensión, sino para obtener un registro cuidadoso de pronunciación. Para ello, les hemos solicitado a los informantes hacer una o más lecturas del texto en silencio y que nos avisaran cuando estuviesen listos para hacer una lectura en voz alta. Les hemos informado con antelación sobre esta prueba, con el fin de que, en caso de necesitarlos, tuvieran cerca sus anteojos de lectura.

Tabla 4. Cuestionario 1. Lista de palabras para la obtención de datos de fonética segmental

<i>Contexto</i>	<i>Preguntas</i>	<i>Fuente</i>	<i>Respuesta</i>
Deblitamiento vocálico	1. Ocurrencia que se dice para reír. 2. Como los de Pepito.	Equipo*	chiste
	Cómo se les dice a las personas que son originarias de Francia.	DEM	franceses
	Arma de fuego de cañón corto que se sujeta con una sola mano (seña)	DEM	pistola
	1. Cuando alguien se arrepiente por decir o hacer algo, se dice que se da... (seña) de pecho. 2. Cuando hay una pelea cuerpo a cuerpo, ¿qué le propina un hombre a otro?	Equipo	golpes
	Cuerpo civil encargado de vigilar el orden público, garantizar la seguridad de los ciudadanos y perseguir a los delincuentes.	Equipo	policía
	Qué se dan los novios en los labios como muestra de cariño.	Equipo	besos
	Parte que constituye el cerebro, especialmente la parte comestible de algunos animales. Se hace con ellos tacos de...	Equipo	sesos
	Personas que están encarceladas.	DEM	presos
	Cuando una persona hace dieta es porque quiere bajar algunos...	Equipo	kilos
	1. Cómo se le llama cuando hace esto (suspirar). 2. Cuál es la acción de respirar profundamente por algún sentimiento intenso, especialmente por amor.	Equipo	suspirar

* El equipo que elaboró esta prueba e hizo el pilotaje antes de comenzar la recolección formal del corpus estuvo integrado por Ana Maurel Sánchez de la Barquera Cortez, Anayeli Hernández Cruz, Angélica Zúñiga Ruiz, Hugo Carrera Guerrero, Laura Romero Rangel, Niktelol Palacios y Rogelio Merino Flores.

FUENTE: elaboración propia.

Prueba de actitudes y creencias

Esta prueba integra preguntas personales con las de opinión general. Los temas que toca son México como país, la ciudad de Puebla, el informante y la norma lingüística. El apartado contiene las mismas 17 interrogantes recogidas en el cuestionario del CSCM (2011: 38-40), con una mínima adecuación en algunas de ellas para incluir las características sociodemográficas de la ciudad de Puebla:

- 4.¹³ ¿Quiénes hablan mejor, y en qué orden: los yucatecos, los costeños, los norteños, los poblanos o los del D. F.?¹⁴ ¿Por qué?
11. ¿Cuándo se hablaba mejor, en la Puebla de hace 20 o 30 años o en la de ahora?¹⁵ ¿Por qué?
12. ¿Cree que se han perdido las formas tradicionales de hablar de los barrios?¹⁶ ¿Por qué?
13. ¿Quiénes hablan mejor, los que han llegado a Puebla del interior del estado o los de Puebla capital?¹⁷
14. ¿Quiénes hablan mejor y en qué orden: los del barrio de San Antonio, los del centro, los de San Manuel, los de La Vista o los de La Margarita?¹⁸

¹³ Para mayor referencia, presentamos el número original de la pregunta en el cuestionario.

¹⁴ En el CSCM se pregunta “¿Quiénes hablan mejor y en qué orden: los yucatecos, los veracruzanos, los norteños, los del D. F.?”.

¹⁵ La pregunta original dice: “¿Cuándo se hablaba mejor, en el D. F. de hace 20 o 30 años o en el de ahora?”.

¹⁶ En esta pregunta hemos cambiado colonia por barrio, porque estos asentamientos propios del Centro Histórico de la ciudad son los más tradicionales del tejido social poblanos. Su trazo data de los siglos fundacionales y en ellos se conserva una fuerte cohesión social, la vida en torno a las fiestas parroquiales y la migración indígena, que se mezcla con las familias de mayor arraigo.

¹⁷ En el CSCM se pregunta “¿Quiénes hablan mejor, los venidos a México de provincia o los de México de siempre?”.

¹⁸ En el CSCM se pregunta: “¿Quiénes hablan mejor y en qué orden: los de Tepito, los de Xochimilco, los de Coyoacán, los de Satélite?”. Hemos modificado estas opciones por el barrio de San Antonio (barrio tradicional cuya fama de peligrosidad data desde que era zona de tolerancia de la ciudad, para convertirse después en el espacio de dominio de la banda delictiva más conocida de la capital pobлана, y

Estas pequeñas modificaciones y la réplica de la prueba han permitido comenzar a explorar las diferencias en las actitudes y creencias lingüísticas entre los entrevistados. Palacios y Pesqueira (2018) analizaron las respuestas a cinco preguntas de este cuestionario: 1) ¿quiénes hablan mejor, los hombres o las mujeres?, 2) ¿quiénes hablan mejor, los jóvenes, los padres o los abuelos?, 3) ¿los de qué barrio hablan mejor? (el informante debe elegir entre un listado de barrios de su ciudad), 4) ¿qué es eso de “hablar mejor”?, y 5) ¿qué es hablar bien?

Compararon las respuestas de 24 informantes: 12 del CSCM y 12 del CSCP (seis mujeres y seis hombres de cada ciudad, con diferentes niveles de instrucción y de distintas edades). Retomaremos aquí algunos resultados que nos parecen interesantes sobre las preguntas (4) y (5). Las autoras encontraron que para los hablantes de la Ciudad de México “hablar mejor” está relacionado principalmente con seguir las normas gramaticales, mientras que para los poblanos el peso del adjetivo “mejor” está anclado a las normas sociales, en un estándar. En la Ciudad de México se valora el deber como el cumplimiento de las reglas internas de la lengua y en Puebla el deber del hablante como usuario en su trato frente a otro. Palacios (2015) ya había comentado

que también es emblemático por el narcomenudeo y las cuestiones sociales asociadas a la migración marginal, el Centro (es decir, el primer cuadro de la ciudad, en donde viven pocas personas: una zona que mezcla los espacios de trabajadores de los comercios diversos, estudiantes universitarios, dueños de casonas viejas y dueños de espléndidos departamentos acordes con un moderno programa de gentrificación), la colonia San Manuel (próxima a la Ciudad Universitaria, es ejemplo del crecimiento demográfico planeado que posibilita la vida familiar y social ordenada y tranquila, modelo de urbanización de los años setenta e icono de la clase media poblana), La Vista (fraccionamiento cerrado de casas lujosas en la zona de reciente crecimiento comercial, que opera como núcleo urbano emergente, por lo que ha dado origen a la formación de corredores económicos con grandes centros comerciales y zonas de comercio exclusivo, restaurantes, hoteles y nuevas vialidades) y La Margarita (unidad habitacional construida por el Instituto del Fondo Nacional de la Vivienda para los Trabajadores, que modificó el crecimiento horizontal hacia el vertical en la planeación urbana de la ciudad; se trata de un espacio de viviendas de interés social promovido por la Confederación de Trabajadores de México, fundado en 1979; un núcleo urbano periférico, integrado por más de 5 300 casas, y una de las zonas más densamente pobladas de la ciudad, cuyas necesidades de seguridad, agua, alumbrado, transporte, etcétera, son difíciles de satisfacer).

que el peso de la norma social en la muestra de Puebla era “producto de la tradición y la valoración simbólica de la comunidad”.

Aunque no era parte de los objetivos de ese trabajo, Palacios y Pesqueira mencionan que vale la pena resaltar que la mayoría de los informantes poblanos considera su variante muy prestigiosa, y señala Puebla como el lugar donde se habla mejor el español. En cambio, los capitalinos tienen respuestas más generalizadas. Cinco mencionaron que la Ciudad de México es donde mejor se habla, pero también respondieron en “el centro de México”, “provincia”, “Guadalajara”, “Xalapa”, “Guanajuato”.

Si bien se trata de un acercamiento exploratorio y no del estudio de toda la prueba de actitudes y creencias, podemos observar diferencias interesantes y percibir que, desde el diseño, el cuestionario del CSCM buscó, por ejemplo, oponer regiones de habla perceptualmente muy bien diferenciadas; en el CSCP agregamos la oposición de hablas entre la Ciudad de México y la capital poblana, con el fin de saber si había un extrañamiento por parte de los entrevistados o si surgían respuestas como “hablamos igual”, y en ningún caso ocurrió esto.

A partir de los resultados de Palacios y Pesqueira también podemos identificar que, en concordancia con las creencias sobre la pregunta “¿Qué es hablar bien?”, para los informantes de Puebla el trato con el otro y el deber como usuarios de la lengua son lo más importante a fin de hablar bien. Tomar como modelo el habla de quienes tienen autoridad por su conocimiento no parece relevante, como sí lo es para los informantes de la Ciudad de México. Esto parece coincidir con los datos sobre qué es “hablar mejor” y cómo los capitalinos le dan un poco más de peso al nivel de instrucción, mientras que para los poblanos de la muestra esto no tuvo importancia.

El cuestionario sociológico

Finalmente, para concluir la entrevista, hemos recogido un cuestionario sociológico integrado por 72 preguntas. Éste se tomó del Proyecto de Investigación sobre la Formación de Dialectos (Fordial) (Villena, Moya, Ávila y Vida, 2003) y tiene valor postestratificatorio para la investigación. También hemos hecho algunas adecuaciones

tomando en cuenta la información sociodemográfica de la ciudad de Puebla (INEGI, 2010), en particular para plantear las categorías de ingresos económicos.

Fordial “es un estudio de la formación de variedades lingüísticas en los contactos producidos en los núcleos urbanos a consecuencia de los movimientos poblacionales ligados a la inmigración procedente de las áreas rurales” (Villena *et al.*, 2003: 607). En este cuestionario se indaga la influencia de:

las redes sociales de los informantes consideradas como el mejor procedimiento para observar y analizar la integración en la estructura urbana o, contrariamente, la lealtad a la localidad de origen; (2) las actitudes individuales sobre ciertos indicadores de ruralidad y de urbanismo; (3) el contacto con los medios de comunicación (Villena *et al.*, 2003: 619).

Las respuestas al cuestionario permitirán hacer estudios sobre modos de vida (Lastra y Martín, 2000). Esta información nos parece particularmente relevante para el estudio postestratificatorio de hablas de líderes lingüísticos y de subgrupos sociales.

COMENTARIOS FINALES

Preseaa-Puebla, compuesto por 108 entrevistas recogidas entre 2012 y 2017 en Puebla de Zaragoza, capital del estado —donde residen 1 692 181 habitantes que hablan, además de español, 16 lenguas indígenas, y registran flujos migratorios de origen, recepción, repatriación y tránsito—, es un corpus de acceso abierto que brinda la posibilidad de estudiar el español de Puebla, el de ciertas comunidades y el de individuos con características específicas. El material recogido permite analizar la manera en la que se (re)construye la identidad social y lingüística de la comunidad y de los migrantes, ya hacia el interior, ya en comparación con otras, y posibilita la descripción de la variación, el contacto y el cambio lingüístico (fónico, morfosintáctico, léxico, sintáctico y pragmático e identitario). Los estudios pueden

tomar en cuenta las variables sociales preestratificadoras (sexo, edad e instrucción educativa formal) o postestratificadoras (provenientes del cuestionario sociológico). Preseea-Puebla, pues, brinda información valiosa para el estudio lingüístico de base empírica.

En lo que resta del libro nos centraremos en la descripción de aspectos léxicos del español de Puebla, pero quedan abiertas la posibilidad y la invitación para continuar con diversos estudios de los datos recogidos.

La estratificación de la muestra nos permitió dar cuenta del complejo entramado social de nuestra comunidad de habla, aunque las entrevistas permiten reconstruir la transformación sociodemográfica de las décadas de 1970 y 1980, caracterizada por la migración campo-ciudad.

A nivel léxico nos planteamos dos hipótesis: las entrevistas semi-dirigidas de Preseea-Puebla nos permiten recoger vocabulario general del español de México en sus diversas manifestaciones estilísticas (culto, técnica, coloquial, grosera, peyorativa, etcétera), es decir, no por tratarse de material oral encontraremos únicamente léxico coloquial o léxico regional. La segunda es que, por tratarse de registro de habla, podremos documentar el uso de ítems y de fenómenos propios de la oralidad que difícilmente aparecen en otro tipo de corpus (juegos de palabras, cruces léxicos, abundantes procesos de derivación y el registro de fenómenos propios de las hablas de solidaridad, poblanismos, neologismos).

3. LEMATIZACIÓN Y ANÁLISIS CUANTITATIVO DE PRESEEA-PUEBLA

La ingeniería lingüística, la lingüística de (o con) corpus y los estudios de lexicología y lexicografía han venido desarrollándose paralelamente desde hace varias décadas. Entendemos la ingeniería lingüística como “la aplicación de las técnicas informáticas al desarrollo de aplicaciones que incluyen componentes relacionados con el lenguaje y el habla” (Llisterri, 2003: 10). En cuanto a los estudios de lexicología y lexicografía, replicamos el método de base pragmática, es decir, estudiamos “la complejidad de la significación verbal manifiesta en los usos reales del español” (Lara, 2017: 6).

Gracias a la ingeniería lingüística existe un gran número de aplicaciones que permiten procesar textos con diversos fines. Específicamente para la lexicografía computacional y el análisis cuantitativo del lenguaje natural podemos obtener listas de palabras distintas (*tipos*), lematizarlas (agruparlas en *vocablos*), indicar su frecuencia de aparición, hacer distintos cálculos estadísticos (por ejemplo, de coaparición o de dispersión), así como recuperar un determinado número de palabras previas y posteriores a la analizada, lo que nos da la posibilidad de estudiarlas en su contexto (*concordancia*).

En el caso de la lingüística mexicana, la relación, más que paralela, es convergente, puesto que la lingüística computacional surgió con el desarrollo del Proyecto Diccionario del Español de México (Lara, Ham y García, 1979b; Sierra 2011).¹ Luis Fernando

¹ Como señala Gerardo Sierra (2011: 475), en la bibliografía especializada se ha difundido, erróneamente, que el trabajo de principios de la década de 1980 de la editorial Collins y la Universidad de Birmingham para construir un corpus

Lara, director del DEM, relata en el “Prólogo” del diccionario que “en 1973 era todavía raro utilizar la computadora electrónica en la investigación lingüística” (DEM, 2010); sin embargo, el equipo interdisciplinario del diccionario desarrolló el Analizador gramatical del DEM (García, 1979) y conformó el Corpus del Español Mexicano Contemporáneo (CEMC). Este libro procura darle continuidad a esa línea de trabajo. Por lo tanto, en este capítulo presentaremos una descripción del léxico recogido en las 108 entrevistas semidirigidas de Preseea-Puebla, con base en herramientas y procesos del análisis computacional.

Aspiramos, por una parte, a continuar con el diálogo entre la aplicación de la ingeniería lingüística y el trabajo lexicológico, y, por otra, a presentar detalladamente el método para el análisis de los textos, desde el preprocesamiento de las transcripciones hasta la documentación de vocablos en el lecionario del *Diccionario del español de México* (versión en línea consultada en octubre de 2020).

Hemos diseñado una propuesta *ad hoc* que permite el *preprocesamiento* o preparación de los textos para la extracción automatizada (Palacios y Serralde, en preparación), la cual consiste en eliminar los encabezados, los caracteres no Unicode, las distintas marcas y etiquetas que forman parte del etiquetado obligatorio de Preseea (marcas de turno, comentarios del transcriptor, ruidos, registro de pronunciación, etcétera) y los caracteres no alfabéticos. Posteriormente usamos herramientas de análisis lingüístico computacional disponibles de manera gratuita en la web: *AntConc* 3,² y *FreeLing*,³ para la identifi-

informatizado de lengua oral y escrita con una muestra representativa del inglés británico fue el precursor de la lexicografía computacional y de la lexicografía basada en corpus. El equipo del Diccionario del Español de México (DEM) es el pionero en la construcción de un corpus de lengua con fines lexicográficos y en generar una línea de investigación de lingüística computacional, puesto que su trabajo comenzó en la década de los años setenta.

² *AntConc* es un *software* libre desarrollado por Laurence Anthony en la Universidad de Waseda, Japón. Permite realizar análisis lingüístico de corpus (uno o varios textos en formato TXT), y se encuentra disponible para su descarga en <<https://www.laurenceanthony.net/software/antconc/>>.

³ Como explica su desarrollador, Padró (2011: 13), “FreeLing es una librería de código abierto para el procesamiento multilingüe, que proporciona una amplia

cación de ocurrencias, tipos, *n-grams* y la lematización. A la serie de procesos automáticos siguió la de procesos manuales para la corrección de distintos tipos de errores.

Las 108 transcripciones de Preseea-Puebla nos arrojaron un texto plano de arranque de 1 481 753 palabras gráficas (ocurrencias o *tokens*) y 38 907 tipos.⁴ Tras el análisis mixto y la corrección de errores —humanos y automatizados— obtuvimos un corpus de estudio integrado por 1 303 362 palabras, 36 023 tipos y 10 715 vocablos. Explicaremos a continuación de forma detallada cómo llegamos a estos resultados. Por ahora, sólo queremos destacar que, comparativamente con Preseea-Ciudad de México y Preseea-Monterrey, hallamos que estos corpus arrojan datos similares: superan los 30 000 tipos y 1 200 000 palabras gráficas.

ANÁLISIS AUTOMATIZADO Y MANUAL DE PRESEEA-PUEBLA

El análisis de los datos buscó un equilibrio entre la automatización en el procesamiento de los datos léxicos a partir del estudio de los textos escritos (conjuntos de transcripciones) y el análisis lingüístico manual. A continuación detallaremos las fases del estudio con el deseo de que pueda ser replicado por otros equipos de trabajo que analicen corpus orales y corpus regionales con objetivos léxicos.

Una vez concluidas las revisiones de las 108 transcripciones de Preseea-Puebla, el material se guardó sin formato (esto es, con la extensión TXT), para facilitar su (pre)procesamiento automatizado.

Como se detalla en Palacios y Serralde (en preparación), a partir del documento “Reglas de transcripción”, Juan Luis Serralde diseñó

gama de funcionalidades de análisis para varios idiomas”. Entre las lenguas que es posible analizar con *FreeLing* se encuentran inglés, español, portugués, italiano, francés, alemán, ruso, catalán, gallego, croata y esloveno.

⁴ Sirva de referencia decir que un análisis equivalente en *AntConc* para los corpus Preseea-Monterrey y Preseea-México, ambos integrados por 108 entrevistas, arroja 1 197 331 *tokens* y 31 049 tipos para el primero, y 1 282 417 *tokens* y 32 764 tipos para el segundo. Agradezco a Pedro Martín Butragueño por proporcionarme la información sobre el cscm.

ad hoc el preprocesamiento del material para la obtención de los datos léxicos buscados. En primer lugar se eliminaron las etiquetas y el encabezado de los archivos; después, las marcas de inicio de turno (número y marca de hablante, por ejemplo “E:”, “I:”, etcétera), y, posteriormente, los caracteres que no fueran Unicode, para que Python pudiera hacer su procesamiento. También se suprimieron las marcas que indican algún ruido, las anotaciones del transcriptor y las de pronunciación y, finalmente, los caracteres no alfabéticos que pudieran haber quedado (erratas de etiquetado, errores al teclear o caracteres no visibles, como los tabuladores). En el último paso nos percatamos de que hacer una limpieza detallada requeriría multiplicar las órdenes para considerar excepciones motivadas por el error humano, y decidimos apegarnos a un proceso automatizado tan elegante como fuera posible, esto es, con el menor número de indicaciones, sin considerar excepciones ni casos particulares (generalmente, provenientes de erratas o de variaciones en la interpretación de las reglas de transcripción). Esto permite que el sistema pueda aplicarse a otros corpus con el mismo etiquetado. Cuando presentemos la revisión manual, daremos cuenta del tipo de erratas que pudimos identificar y que se conservaron debido a esta decisión.

Si bien cabría la pregunta de si sería mejor generar desde el origen una transcripción de tipo ortográfico para los análisis léxicos, debemos recordar que el proyecto Preseea considera una serie de etiquetas mínimas de transcripción a fin de que los materiales sean comparables, etiquetas que responden a la marcación de fenómenos lingüísticos de distinta índole. Y si bien solicita la entrega de 18 transcripciones sin marcas, nosotros estamos trabajando con el corpus completo (108 transcripciones).

A partir de la programación en Python 3.8.2,⁵ se obtuvo un primer acercamiento cuantitativo al corpus: palabras gráficas (*ocurrencias*), palabras distintas (*tipos*), riqueza léxica⁶ y *n-grams* (es decir, secuen-

⁵ Python es un lenguaje de programación multiparadigma, es decir, permite distintos tipos de programación. Su autor principal y responsable es Guido van Rossum. Se encuentra disponible en <<https://www.python.org/about/>>.

⁶ La *riqueza léxica* es el cálculo de cuántas palabras diferentes se utilizan en un texto, lo cual ayuda a determinar su complejidad para el procesamiento de in-

cias de n palabras) de todo el corpus. Demos un ejemplo del texto plano analizado:

- (1) I: mi tesis consistía en un plan en un plan en un manual de de publicidad para para pequeñas y medianas empresas pues era un manual básico de de publicidad gráfica para las llamadas pymes consistía en en seguir unos pasos como como una diferenciación de mercado como la elección de los medios por dónde publicitarse como consejos de qué de de cómo sustentar sus mensajes ante su público

[PUEB-H13-015]⁷

Este texto tiene 68 palabras gráficas (ocurrencias o *tokens*), 44 de las cuales son palabras distintas (tipos). Los tipos que se repiten son *de* (9 veces), *en* (5), *como* (4), *un* (4), *para* (3), *consistía* (2), *plan* (2) y *publicidad* (2). También se identifican 40 vocablos o formas que representan el paradigma flexivo o de conjugación: las palabras invariables como *en*, *de*, *para*, *pues* e *y* tienen un vocablo idéntico (*en*, *de*, *para*, *pues* e *y*), mientras que las flexiones y conjugaciones se agrupan en español, generalmente bajo el vocablo masculino singular y la forma de infinitivo; sirvan de ejemplo los tipos *empresas*, cuyo vocablo es *empresa*, y *consistía*, cuyo vocablo es *consistir*.

Antes del preprocesamiento del etiquetado de las transcripciones para trabajar con un texto plano, teníamos 1 481 753 palabras gráficas (*tokens*) y 38 907 tipos. Una vez hecha la primera limpieza, trabajamos con un corpus (conjunto de transcripciones en formato sólo texto) de 1 303 362 palabras y 36 023 tipos, con un cálculo de riqueza léxica de 0.11. En la tabla 5 se muestran los

formación y el nivel de vocabulario del que dispone su autor. Para una explicación detallada de los tipos de estadística Type-Token en Python, véase Delgado, 2019.

⁷ Recuérdese que la etiqueta de la entrevista que presentamos entre corchete al final de la cita se forma de la siguiente manera: el primer bloque de información indica las variables estratificadoras de Pressea (sexo, edad y nivel educativo) tras el guion, y con tres dígitos se indica el número de la entrevista dentro de Pressea-Puebla. De tal manera que debe entenderse que este ejemplo está tomado de la entrevista 15, que se hizo a un hombre de edad 1 (entre 20 y 34 años) y de instrucción alta.

n-grams o unidades simples y en combinatoria (hasta 4 palabras) más frecuentes.

Los resultados se organizan en cada columna por orden descendente según su frecuencia absoluta de aparición. La combinatoria es en su mayoría discursiva y no léxica, es decir, responde a la creación del discurso libre, y sólo en pocos casos, a locuciones y marcadores discursivos —*a lo mejor, cómo se llama, haz de cuenta (que), o sea, y todo eso, y ya*—. En las palabras simples (*1-grams*), como podría preverse, encontramos principalmente elementos funcionales (artículos, conjunciones y preposiciones). No obstante, queremos destacar la elevada frecuencia de aparición de *no*⁸ y *sí*, misma que puede explicarse por la importancia de estas palabras para la interacción dialógica.

Los *2-grams* (o combinaciones más frecuentes de dos palabras) arrojan unidades discursivas de interés para el análisis de la intensificación y la atenuación discursiva.⁹ Por otra parte, se destaca, mas no sorprende, el elevado número de apariciones del nombre de la ciudad donde se recogen los datos: *la_ciudad_de_Puebla*. Consideramos que esto es resultado de las preguntas que guían la conversación (infancia, vida cotidiana, las costumbres y tradiciones de la ciudad) e indicio de la conciencia con la que los entrevistados hablaban de su ciudad en oposición a otras, como es esperable en este tipo de corpus, lo que tomamos como una buena señal en nuestra búsqueda de léxico identitario.

La lematización automatizada del corpus se hizo con el programa *FreeLing*. De este análisis se obtuvieron 17 206 candidatos de estudio. A partir de esta lista de vocablos, comenzamos una revisión manual que nos permitió descartar erratas (a) y otras unidades que decidimos

⁸ Véase el estudio de Orozco (2014) sobre el valor de *¿no?* en las entrevistas sociolingüísticas. Desde luego, sería necesario replicar el estudio para conocer las características específicas de su uso en Preseea-Puebla, pero suponemos que seguirá un camino similar.

⁹ Los estudios sobre atenuación en Preseea-Puebla han arrojado datos sumamente interesantes sobre la manera de construir el discurso en esta comunidad, y han permitido apoyar la idea de que se trata de un fenómeno sociopragmático altamente vinculado con la comunidad de habla. Véanse Palacios (2017, 2019 y en edición), y Cestero (2020).

Tabla 5. Los 10 *n*-grams más comunes (1-grams a 4-grams) en Preseaa-Puebla. Análisis de Python 3.8.2

1-grams	2-grams	3-grams	4-grams
que: 48 976	o sea: 5 523	no no no: 1 112	no no no no: 280
no: 47 816	no no: 4 219	yo creo que: 1 030	haz de cuenta que: 243
y: 45 828	en la: 4 092	sí sí sí: 881	ya no ya no: 160
de: 41 280	de la: 3 987	no o sea: 808	sí sí sí sí: 152
a: 34 306	lo que: 3 825	y este y: 774	la ciudad de Puebla: 147
la: 31 004	sí sí: 3 494	a lo mejor: 754	este cómo se llama: 143
sí: 26 114	a la: 3 387	o sea no: 531	en la en la: 128
en: 20 818	ya no: 2 915	a mí me: 517	eso es lo que: 121
me: 19 220	y ya: 2 767	es lo que: 506	a mí no me: 121
el: 18 882	en el: 2 743	y todo eso: 446	yo me acuerdo que: 120

FUENTE: elaboración propia.

eliminar para esta fase del estudio. Observamos dos tipos de problemas ocasionados desde la transcripción:

- Aquellos que en el texto original corresponden a un error en el etiquetado; por ejemplo, falta de una marca de cierre, lo que ocasionó dos tipos de problemas en el análisis: a) que durante el preprocesamiento se juntaran palabras (*ahbueno*, *menosy* o *tar-dedonde*); en estos casos, se reconocen con claridad las palabras que se unieron gráficamente (*ah_bueno*, *menos_y* o *tarde_donde*), y b) “falsas palabras”, que corresponden a la descripción de la pronunciación dentro del texto (por ejemplo, *cul*, que aparece en la siguiente cadena: “cool <~cul>”), o emisiones incompletas que no tuvieron una correcta marcación, como *ante*, que viene de un texto donde el entrevistado hace una pausa y repite después la palabra completa: “ante- / antecesor, digo”.
- Las erratas en la escritura, cuya aparición, pese a las múltiples revisiones, es inevitable (*acamapanados* por *acampanados* o *popo-pe* por *popote*).

Para las erratas de tipo (aii), fue necesario revisar cada concordancia a fin de evitar la eliminación, por ignorancia, de una palabra sólo porque el analista la desconociera. Cabe destacar que el programa relaciona mediante un guion bajo las palabras que, intuye, han de conformar una misma unidad, como los ejemplos que presentamos a continuación (4b):

- Nombres propios (Agatha_Christie, Action_Man ‘juguete producido por las compañías Hasbro y Palitoy’, Africam_Safari, Colegio_de_Bachilleres, Estadio_Cuauhtémoc, Gustavo_Moscoso, Zacapoxtla).
- Siglas (Afore ‘Administradora de Fondos para el Retiro’, BMW ‘Bayerische Motoren Werke AG’, INAOE ‘Instituto Nacional de Astrofísica, Óptica y Electrónica’, IMSS ‘Instituto Mexicano del Seguro Social’).
- Indicaciones de nombres propios dentro del texto. Recuérdese que se transcriben usando una o dos letras siguiendo el orden del abecedario A, B, C... AA, BB.

Si bien por el momento hemos eliminado la onomástica de nuestra investigación, se trata de unidades léxicas importantes para el conocimiento de la comunidad (¿qué se individúa mediante el nombre propio y cómo esta individualización da cuenta de los valores colectivos e identitarios?). Lejos de descartar el estudio de estas unidades por su valor altamente referencial, creemos que sería más pertinente preguntarnos por las huellas lingüísticas (por ejemplo, los topónimos) y culturales (por ejemplo, los nombres propios de personas a partir de ciertas figuras públicas, de ideologías nacionalistas o extranjerizantes, de íconos culturales) que nos permiten conocer o reconstruir.

La siguiente fase de la revisión manual nos permitió evaluar “errores sistemáticos” o limitaciones en la lematización del propio programa. La presencia de clíticos impide la detección del lema infinitivo (*abrazándolos*, por ejemplo, se lematiza de forma idéntica: *abrazándolos* → *abrazándolos*, *besándonos* → *besándonos*, *muévale* → *muévale*, *pásamelo* → *pásamelo*, etcétera).

En algunos casos, los prefijos y los elementos cultos que toman la primera posición de la palabra se analizan como compuestos, y el programa marca una separación de elementos lematizando la segunda en masculino singular, por ejemplo: *biomagnética* → *bio_magnético* y *exconvento* → *ex_convento*. En estos casos sólo hemos hecho la corrección del análisis eliminando el guion y sumando sus ocurrencias al vocablo correspondiente.

Cabe destacar que el programa no hace lo mismo con otros elementos cuyo lema registra idéntico al tipo: *hipertensión*, que lematiza como *hipertensión*, *megabyte* → *megabyte*, *microempresas* → *microempresa*, *mililitros* → *mililitro*, *rehacer* → *rehacer* o *subsecuentes* → *subsecuente*.

Como podemos ver en los siguientes ejemplos, los sufijos apreciativos se analizan de distintas formas; algunas muestran reinterpretaciones morfológicas muy interesantes: *aparatazos* se lematiza como *aparata*, *bebecita* → *beber_cita*, *maquinita* → *maqui_nit*, *perengani-to* → *perengano*, *pesitos* → *peso*, *lejitos* → *lejío*, *limpiecitas* → *limpia*, *mochillos* → *mocho*, *morrillo* → *morrillo*, *negrita* → *negrito*, *noblezón* → *nobleza*, *pantaloncitos* → *pan_talo*, *pegalón* → *pe_galón*, *recamarotas* → *recamar_ota*, *repicita* → *repi*, *silloncito* → *sil_lo*.

En su mayoría los diminutivos, aumentativos y despectivos se lematizan de forma idéntica al tipo, como en *maquinitas* → *maquinitas*, *jovencillos* → *jovencillos*, *jovencito* → *jovencito*, *maicito* → *maicito*, *poblanitos* → *poblanitos* o *poquitín* → *poquitín*.

Como es esperable, los participios son particularmente difíciles de estudiar de manera automatizada. La identificación de su función adjetiva o sustantiva no es correctamente predicha y se marca como verbo; esto ocurre especialmente en contextos de frontera, como en las oraciones con verbo predicativo. Por ejemplo, en (2) y (3) el programa lematiza *abusada* bajo el vocablo *abusar*, y *afectada* como *afectar* (a pesar de estar antecedido por un artículo):

- (2) I: me dijo sí eres **abusada** y ya me empezó a jalar y yo me iba con ellos a los juzgados y mira léete esto y mira léete esto

[PUEB-M13-035]

- (3) I: sí ese día ya yo fui la **afectada** del 15 de septiembre porque como mi esposo se puso a tomar con él

[PUEB-M12-027]

Ocurre lo mismo incluso en los participios irregulares que se especializan como adjetivos o sustantivos (letra): *manuscrita* → *manuscibir*, y con otras palabras ya lexicalizadas, como *abogadas* → *abogar*.

Hay algunos sobreañálisis de pseudocompuestos, que ocurren cuando el programa no puede predecir la palabra como unidad, pero reconoce al menos una de sus partes, como si fuese una voz independiente: *gastalón* → *gas_talón*, *mirreyes* → *mir_rey*, *mitotero* → *mitote_ro*, *pitayas* → *pi_taya*, *pitufos* → *pi_tufo* o *rastas* → *ras_tas*.

En algunas formas conjugadas, el programa no puede predecir el infinitivo y lematiza de forma idéntica al tipo: *cierro* → *cierro*, *cobijo* → *cobijo*, *portas* → *porta*, *referí* → *referí*, *saluda* → *saluda* o *supe* → *supe*.

Finalmente, queremos comentar los casos en que voces de tradición popular y mexicanismos tampoco se lematizan o se lematizan de forma errónea: *melena* → *meleno*, *mesereaban* → *mesereaban*, *mesereando* → *mesereando*, *nahuales* → *nahuales* y *relajientos* → *relajientos*.

El estudio sistemático y detallado del análisis automatizado podría redundar en propuestas de mejoras para estos lematizadores. Como hemos visto, resultan de gran interés para el desarrollo teórico-metodológico sobre la delimitación y el reconocimiento de la palabra en las lenguas naturales: presencia de clíticos, pseudomorfemas cultos, (pseudo)derivación apreciativa y multifuncionalidad, principalmente de participios regulares e irregulares.

La revisión manual de los resultados nos permitió eliminar 6 656 errores o sobreinterpretación del programa (esto es, 38.69% de los datos obtenidos automáticamente). Pasamos así de 17 206 vocablos lematizados por *FreeLing* a 10 715 candidatos de estudio. En la tabla 6 ejemplificamos cómo hemos reorganizado el material, tras haber eliminado las erratas, errores del programa y duplicaciones identificadas hasta este momento del proceso de análisis.¹⁰

¹⁰ En esta reorganización sumamos las ocurrencias a los vocablos correspondientes, por lo que se apreciarán diferencias entre las tablas 5 y 7.

Tabla 6. Corrección de vocablos y eliminación de duplicados

<i>Tipo</i>	<i>Vocablo FreeLing</i>	<i>Vocablo corregido</i>	<i>Vocablo registrado (sin duplicar)</i>
abogadas	abogar	abogado	
abogadazo	abogado	abogado	abogado
abrázala	abrázala	abrazar	
abrazándolos	abrazándolos	abrazar	abrazar
abrazándose	abrazándose	abrazar	
abrazos	abrazo		abrazo
ábrele	ábrele	abrir	
ábrelo	ábrelo	abrir	abrir
abu	abu	abuelo	
abue	abue	abuelo	
abuelo	abuelo	abuelo	abuelo
agüelita	agüelita	abuelo	

FUENTE: elaboración propia.

Es importante recordar que estamos analizando únicamente la forma de estas palabras, es decir, no nos hemos detenido en el estudio semántico, por ejemplo, de polisemia ni de homonimia. En cuanto a la función gramatical de estas palabras, hace falta un análisis detallado de las ocurrencias. Por ahora, y gracias a que en un primer momento decidimos conservar el acento diacrítico entre adjetivos y pronombres demostrativos, hemos podido ubicar estas categorías como vocablos independientes. Sin embargo, no ocurre lo mismo con artículos definidos y pronombres, que se mantendrán bajo un mismo vocablo.

LOS 100 VOCABLOS MÁS FRECUENTES DE PRESEEA-PUEBLA

Los 100 vocablos más frecuentes de Preseea-Puebla corresponden a 75.17% del total del corpus analizado. En conjunto, suman 979 796 ocurrencias de 1 303 362 que lo conforman. Atribuimos este resultado a que se trata de textos orales. En la tabla 7 presenta-

mos el listado de voces con su frecuencia absoluta de aparición en Preseea-Puebla y hemos decidido comparar el resultado con los 100 vocablos más frecuentes del *vocabulario fundamental del español*, que resultaron del análisis del CEMC.

Las palabras de contenido léxico más frecuentes dentro de Preseea-Puebla son mayoritariamente verbos (*haber, ser, ir, decir, estar, tener, saber, hacer, dar, gustar, poder, llegar, creer, querer, llevar, salir*), algunos de ellos gramaticalizados o verbos de soporte, y otros de uso pragmático atenuante (como *creer* y *saber*), es decir, su alta frecuencia no necesariamente está vinculada a su contenido léxico sino a sus funciones gramaticales y pragmáticas.

Los sustantivos más comunes son denominaciones temporales cuya función podría vincularse con la narrativa de la entrevista (*año, día, tiempo*), con las relaciones de parentesco (*gente, hijo, madre y padre*) —posiblemente porque uno de los módulos de conversación es, precisamente, “la familia”— y con hiperónimos o palabras de amplio alcance semántico (*cosa, gente, casa, vez*). En este listado aparecen otras palabras (*mucho, bueno, nada, bien, poco, mejor*) polifuncionales, por lo que se requiere de un estudio detallado de sus contextos de uso para asignarles categoría gramatical y función.

En la tabla 7 puede observarse que 76 de las 100 palabras coinciden en ambas listas. Si bien la coincidencia es muy alta, nos centraremos en explorar algunas de las diferencias:

- 1) En Preseea-Puebla hemos lematizado bajo el vocablo *el* los tipos *el* y *los*; bajo *la*, los tipos *la* y *las*, y, bajo *le*, los tipos *le* y *les*.
- 2) Debido a la naturaleza del corpus estudiado, hemos decidido enlistar, como vocablos, unidades que tienen un alto valor dialógico y expresivo, como *eh, ah, ajá, mm, mjú*.
- 3) Destaca la presencia de deícticos; consideramos que esto se debe a que el corpus se conforma por entrevistas cara a cara; por ejemplo, en esta lista se registran como muy frecuentes los locativos *acá, allá*.
- 4) Habíamos adelantado que el uso de *no* es muy alto en nuestro corpus (ocupa la segunda posición) por su valor pragmático dentro de las entrevistas sociolingüísticas (véase Orozco, 2014).

VOCABLOS DE PRESSEA-PUEBLA

Una vez que obtuvimos la lista de 10 715 vocablos de Preseea-Puebla y siempre guiados por el interés de saber cuánto de este vocabulario es parte de la riqueza léxica del español de México, y si hemos podido registrar léxico regional (entendido como el que pertenece al español de Puebla) en nuestro corpus oral, decidimos hacer una comparación entre este listado y el lemario del *DEM*.¹¹

Esperamos que otros equipos Preseea —o de otro tipo de corpus regionales— comiencen a lematizar sus datos y que, a partir de una metodología común, podamos hacer estudios sobre variación léxica, ya sea a partir de la comparación entre nuestros lemarios completos o en ciertas parcelas léxicas.¹² La obtención del lemario es, pues, un paso preparatorio para futuros estudios. En esta aproximación, nos hemos centrado en la comparación formal, pero será indispensable en una segunda etapa de estudio hacer un análisis de significados.

¹¹ Antes de presentar los resultados, es conveniente hacer un breve recordatorio. El Diccionario del Español de México es un proyecto de lexicografía integral dirigido por Luis Fernando Lara en El Colegio de México que ha dado origen a la publicación de cinco diccionarios impresos (cada uno incluye el lemario del anterior con base en criterios estadísticos): *Diccionario fundamental del español de México (DFEM)* (1982), *Diccionario básico del español de México (DBEM)* (1986), *Diccionario del español usual en México (DEUM)* (1996) y *(DEUM2)* (2009) y *Diccionario del español de México (DEM)* (2010). En 2019 se publicó en línea la nueva edición, *DEM2*, que además cuenta con versiones para Android e IOS. A partir de este momento, el diccionario en línea contendrá actualizaciones cuya velocidad necesariamente aventajará a la obra impresa, aunque no la sustituirá. Cabe aclarar que esta misma velocidad de renovación propiciará que, al salir la versión impresa de nuestro libro, algunos resultados no coincidan con la versión disponible en línea del *DEM*.

¹² Véanse, por ejemplo, los estudios sobre indigenismos presentados en una primera versión como ponencias en la Asociación de Lingüística y Filología de América Latina (*ALFAL*) 2021 y posteriormente elaborados como artículos por Pedro Martín Butragueño y Yolanda Lastra en coautoría sobre Preseea-Ciudad de México, y por Erik Franco y Niktelol Palacios sobre Preseea-Puebla (“Primera aproximación a la variación léxica en el corpus sociolingüístico de la Ciudad de México: el caso de los indigenismos” y “Presencia y uso de indigenismos en Preseea-Puebla”, respectivamente). A partir de una metodología común, se hizo un primer acercamiento a la variación léxica y la vitalidad de estas voces en dos comunidades del Altiplano.

Tabla 7. Los 100 vocablos más frecuentes en el CEMC y en Preseea-Puebla

F_TOT	CEMC	F_TOT	Preseea-Puebla	F_TOT	CEMC	F_TOT	Preseea-Puebla	F_TOT	CEMC	F_TOT	Preseea-Puebla	F_TOT	CEMC	F_TOT	Preseea-Puebla
71 514	la	50 547	que	10 543	decir	10 725	le	4 052	mucho	5 170	eh	2 597	llegar	2 756	hijo
66 920	que	49 452	no	10 495	tener	10 085	por	4 037	aquí	4 960	hacer	2 571	él	2 717	él
63 497	de	48 641	y	10 427	me	9 769	con	4 029	ése	4 695	sí	2 508	poner	2 698	crear
59 476	y	43 150	de	9 816	ya	9 575	como	3 678	querer	4 584	ah	2 503	entre	2 671	tu
51 466	el	40 403	haber	9 646	más	9 447	porque	3 507	también	4 419	otro	2 392	ni	2 632	querer
49 302	en	40 220	la	9 287	todo	9 316	todo	3 460	saber	4 375	ver	2 379	nuestro	2 510	luego
44 820	a	36 938	ser(v)	9 235	pues	9 062	te	3 393	te	4 106	cuando	2 321	llevar	2 476	aquí
36 116	un	36 032	a	9 019	o	8 902	decir	3 389	nos	4 071	también	2 316	cosa	2 418	gente
36 016	ser(v)	31 944	el	9 009	sí	8 546	qué	3 335	bueno	4 009	al	2 245	parte	2 412	tú
33 704	se	27 333	sí	8 846	hacer	8 500	estar	3 326	año	3 968	mm	2 198	sólo	2 302	casa
31 475	no	21 657	un	8 523	le	8 455	ajá	3 325	deber	3 907	cómo	2 195	pasar	2 251	llevar
30 999	los	21 277	en	8 348	pero	7 979	más	3 275	vez	3 611	pasar	2 147	ahora	2 153	siempre

20 747	las	19 492	me	7 294	poder (v)	7 878	tener	3 215	dos	3 602	del	2 120	dejar	2 086	ay
19 883	por	19 235	pues	7 021	yo	7 466	ahí	3 188	sin	3 558	dar	2 105	ella	2 056	ella
18 781	con	18 964	ya	6 515	otro	7 375	entonces	3 180	entonces	3 533	gustar	2 087	alguno	2 039	ni
18 690	del	16 248	yo	6 161	si	7 060	eso	3 032	hasta	3 528	nada	2 080	tiempo	2 018	poco
17 823	su	15 569	este	5 749	mi	6 837	para	3 020	éste	3 412	bien	2 070	después	2 007	ahora
17 521	ir	14 791	se	5 626	dar	6 447	mucho	2 890	venir	3 385	poder	2 062	hombre	1 982	mejor
17 450	haber	13 503	lo	5 084	porque	6 241	muy	2 809	ahí	3 371	año	2 023	les	1 949	día
15 545	para	13 448	pero	5 060	ese	6 162	así	2 791	día	3 304	cosa	2 008	poco	1 914	allá
13 521	lo	12 703	o	4 915	cuando	5 797	su	2 772	mismo	2 943	padre	1 987	salir	1 876	mí
11 894	este	12 143	mi	4 913	muy	5 739	bueno	2 720	usted	2 875	madre	1 978	donde	1 769	tiempo
11 752	estar	11 786	mjú	4 846	uno	5 640	ese	2 710	nada	2 874	hasta	1 966	luego	1 767	ellos
11 214	al	11 210	uno	4 491	qué	5 314	saber	2 692	sobre	2 801	vez	1 920	primero	1 745	acá
11 040	como	10 818	ir	4 159	así	5 195	nos	2 610	bien	2 796	llegar	1 915	desde	1 719	salir
													1 101 322	979 796	

FUENTE: elaboración propia.

VOCABLOS DOCUMENTADOS AUTOMÁTICAMENTE

Como ya hemos explicado, nuestro punto de partida para describir la composición del léxico obtenido en Preseea-Puebla es compararlo con el leuario del *DEM*. Presentamos en esta sección el listado que resultó del cotejo de ambos leuarios. De 10 715 vocablos de Preseea-Puebla, 87.84% (9 413 vocablos) coincide, al menos en la forma, con entradas ya publicadas en el *DEM* (para saber si coinciden también en significado, se requiere analizar cada contexto de uso y cada acepción de los artículos lexicográficos del *DEM*).

Al observar los resultados, nos damos cuenta de que nuestro corpus refleja la diversidad léxica del habla del registro semiformal: igualmente se encuentran palabras no marcadas (*abuela, matemáticas, noticia, sombrero, tiempo, amar, coser, morir, querer, divertido, grande, huraño*) que de tradición culta (*australopiteco, biosicosíntesis, cinéfilo, heideggeriano, homoparental, neoconservadurismo, oleaginosa*), variaciones de arraigo popular (*abajar* ‘bajar’, *ajuntarse* ‘vivir en pareja sin casarse’), coloquialismos (*caguama* ‘botella de cerveza de aproximadamente 1 litro’, *cachar* ‘descubrir a alguien haciendo algo prohibido’, *machetear* ‘estudiar con mucha dedicación’) o voces groseras u ofensivas (*carajo, gato* ‘sirviente’, *naco, pendejo, puto*).

Aunque en este acercamiento no es nuestro objetivo hacer una clasificación en áreas temáticas, sí podemos asegurar que en Preseea-Puebla se registra léxico de materias diversas; sirva de ejemplo mencionar palabras de distintas ciencias (*alzheimer, célula, simbiosis, fluoxetina*), técnicas (*deshebradora, fotocopiadora, software, sonómetro*), oficios (*albañil, artesano, baterista, bizcochero, cadenero* ‘persona encargada de cuidar la entrada de un bar, un antro, etcétera’, *franelero* ‘persona que ayuda a los automovilistas a estacionarse, cuida y lava coches estacionados en la calle a cambio de una propina’, *gasero, mariachi, memelera, soldador*), de gastronomía (*chile en nogada, mazapán, mole, strudel, tamal*), deportes (*corredor de fondo, entrenador, fútbol, gol, taekwondo*), música (*danzonero, narcocorrido, jazz, percusión, piano, sinfonola, twist, reggaetón*) y religión (*biblia, budista, católico, dios, epifanía, rezar, talmud, querubín*).

VOCABLOS DOCUMENTADOS MANUALMENTE

Ahora bien, para dar continuidad al método que proponemos en este libro, nos dimos a la tarea de hacer una revisión manual de los 1 154 vocablos que no aparecieron en el análisis automatizado. Esto quiere decir que buscamos una a una las palabras de la lista en la página electrónica del diccionario. Desde luego, tras esta primera fase de análisis, no podemos decir que se trata de voces propias del español de Puebla, y ni siquiera podemos asegurar que no estén definidas en el DEM (explicaremos esto más adelante). Presentamos por ello una revisión detallada buscando palabra por palabra en el diccionario, aprovechando recursos como el listado de vocablos que “guardan cierta similitud con el que se busca” y las búsquedas avanzadas de su página de internet.

Queremos insistir en el valor que damos a la complementariedad de métodos de investigación: automatizados y revisión manual. En esta nueva revisión hemos encontrado algunas diferencias que responden a limitaciones del sistema de búsqueda; a dificultades del lematizador para identificar ciertas formas lingüísticas (como los verbos pronominales); a la variación ortográfica, que da indicios, por una parte, de los procesos de adaptación de los préstamos y, por otra, de las actitudes ante su ortografía (norma culta), y a la identificación de palabras complejas (compuestos sintagmáticos, colocaciones y locuciones). Todo ello enriquece el diálogo y aporta ideas para mejorar tanto el proceso de consulta automatizado como la transcripción del corpus.

Como es esperable, la comparación automatizada requiere que la coincidencia en la cadena comparada sea plena. Debido a que eliminamos signos de puntuación, las 18 interjecciones de la tabla 8 no pudieron recuperarse automáticamente, pero sí están definidas en el diccionario.

Diferencias en la formalización del vocablo

FreeLing no lematiza verbos pronominales, por lo que, al comparar nuestra lista con las entradas de los artículos lexicográficos del DEM, no se identificaron los 52 verbos de la tabla 9. Al revisar las

Tabla 8. Interjecciones de Preseea-Puebla cuya entrada en el DEM tiene signo de admiración

	Preseea-Puebla	DEM
1.	ah	¡ah!
2.	ajá	¡ajá!
3.	ay	¡ay!
4.	caray	¡caray!
5.	chin	¡chin!
6.	ea	¡ea!
7.	eh	¡eh!
8.	hola	¡hola!
9.	jijo	¡jijo!
10.	oh	¡oh!
11.	ojalá	¡ojalá!
12.	órale	¡órale!
13.	pun	¡pun!
14.	quihubo	¡quihubo!
15.	quihúbole	¡quihúbole!
16.	újule	¡újule!
17.	uta	¡uta!
18.	uy	¡uy!

FUENTE: elaboración propia.

concordancias (en *AntConc*), corroboramos que su lematización debe ser pronominal, lo que nos permitió hacer las correcciones pertinentes en nuestro lemario.

Otro grupo discordante de palabras está motivado por las decisiones ortográficas de nuestro equipo de transcripción. Esto ocurre en palabras (casi todas son préstamos) cuya ortografía admite variación, es decir, que tienen más de una forma correcta de escribirse en español de México, como lo señala el propio DEM. En la tabla 10 indicamos la manera en la que se registró la palabra en nuestras transcripciones y la forma en que se registra la entrada en el DEM; podemos observar, además, que en el diccionario se da cuenta de la variante que hemos elegido.

Tabla 9. Verbos de Presea-Puebla que están en el *DEM* como entradas pronominales

	Presea-Puebla	<i>DEM</i>		Presea-Puebla	<i>DEM</i>
1.	abstener	abstenerse	27.	enrolar	enrolarse
2.	accidentar	accidentarse	28.	ensañar	ensañarse
3.	adueñar	adueñarse	29.	envarar	envararse
4.	aguar	aguararse	30.	escabullir	escabullirse
5.	antojar	antojarse	31.	esforzar	esforzarse
6.	apellidar	apellidarse	32.	esmerar	esmerarse
7.	apropiar	apropiarse	33.	estrellar	estrellarse
8.	arrepentir	arrepentirse	34.	explayar	explayarse
9.	arrodillar	arrodillarse	35.	friquear	friquearse
10.	atrever	atreverse	36.	inconformar	inconformarse
11.	ausentar	ausentarse	37.	ingeniar	ingeniarse
12.	chispar	chisparse	38.	maravillar	maravillarse
13.	congraciar	congraciarse	39.	mofar	mofarse
14.	desbarrancar	desbarrancarse	40.	osificar	osificarse
15.	desentender	desentenderse	41.	pandear	pandearse
16.	destrampar	destramparse	42.	percatar	percatarse
17.	desvivir	desvivirse	43.	persignar	persignarse
18.	embolsar	embolsarse	44.	petatear	petatearse
19.	encabronar	encabronarse	45.	quejar	quejarse
20.	encaprichar	encapricharse	46.	rebelar	rebelarse
21.	encariñar	encariñarse	47.	regodear	regodearse
22.	enclave	enclavarse	48.	resignar	resignarse
23.	endeudar	endeudarse	49.	retractar	retractarse
24.	engentar	engentarse	50.	santiguar	santiguarse
25.	enmariguanar	enmariguanarse	51.	tambalear	tambalearse
26.	enorgullecer	enorgullecerse	52.	vanagloriar	vanagloriarse

FUENTE: elaboración propia.

Tabla 10. Variación ortográfica entre Preseea-Puebla registrada en el lema del DEM

	Preseea-Puebla	DEM
1.	bufet	buffet (También bufet)
2.	casete	cassette (También casete)
3.	confeti	confetti (También confeti)
4.	cuzca	cusca (También cuzca)
5.	desharrapado	desarrapado (También desharrapado)
6.	espagueti	spaghetti (También se escribe como espagueti)
7.	estatus	status (También se escribe estatus)
8.	folclórico	folklórico (También folclórico)
9.	folklor	folklore (También folklor, folclore o folclor. Se pronuncia folclór)
10.	full	ful (También full)
11.	garaje	garage (También garaje)
12.	guion	guión (También guion)
13.	ícono	icono (También ícono)
14.	ok	okey (Se abrevia ok)
15.	período	periodo (También período)
16.	restaurant	restaurante (También se pronuncia y se escribe restorán y en ocasiones se escribe restaurant y se pronuncia así)
17.	restorán	
18.	stencil	esténcil (También se escribe stencil)
19.	videocasetera	videocassettera (También videocasetera o videocasetera)

FUENTE: elaboración propia.

Esta descripción nos permite mostrar que, sin importar cuánto se expliciten los criterios de transcripción y cuántas revisiones se hagan de los materiales, siempre habrá diversidad en su aplicación. Cabe preguntarse ¿a qué se debe dicha variación? El caso de los préstamos refleja, por una parte, el grado de su adopción y adaptación, y, por otro, da información valiosa sobre posturas extranjerizantes o castellanizantes entre los transcriptoros.

Para fines prácticos, tras la identificación de estas voces hemos buscado unificar la ortografía privilegiando el uso más común entre nuestros transcriptoros.

Además de las variaciones registradas en la tabla 10, donde vimos que la decisión del equipo aparece en el *DEM*, en la tabla 11 presentamos 26 decisiones ortográficas distintas entre Preseesa-Puebla y la entrada del *DEM*, sin que en el diccionario haya registro sobre esa variación. Se trata de casos diversos. Puede observarse que, en algunos, los transcriptoros hemos preferido en Preseesa-Puebla el préstamo no adaptado, mientras que el *DEM* opta por la forma adaptada, y en otros, ocurre justo a la inversa. Además de los distintos fenómenos de variación hallados agregamos, cuando lo consideramos pertinente, información del *Diccionario de la lengua española (DLE)* o de otras fuentes.

Esta revisión nos ha permitido reflexionar sobre el proceso de transcripción y en algunos casos hemos decidido volver para hacer correcciones. Cambiamos *bisquet* por *bísquet*, *diésel* por *diesel*, *fólder* por *folder*, *grafiti* por *grafitti*, *hándicap* por *handicap*, *kínder* por *kinder*, *marihuano* por *mariguano*,¹³ *memorándum* por *memorandum*, *referéndum* por *referendum*, *sándwich* por *sandwich*, *sicología* por *psicología* y *ultimátum* por *ultimatum*.

Particularmente interesantes nos parecen los casos de *tarjea* por *atarjea* y *tecojote* por *tejojote*. Ambas son formas de uso común en la tradición popular y vale la pena preguntarse si deberían registrarse

¹³ Cabe decir que el *DEM* registra la variación *marihuana* y *mariguana*, incluso proporciona una entrada para cada ortografía, pero no hay correspondencia para el agente que la consume, el cual sólo se presenta bajo la entrada *mariguano*. No obstante, la ortografía *marihuano* queda como cabo suelto bajo la entrada “moto² s y adj (Popular) Marihuano”.

como variantes léxicas (dos vocablos distintos) o fónicas (un solo vocablo, uno de ellos con marca de pronunciación). Hemos optado por decisiones diferentes, analicemos las razones.

Veamos primero *atarjea* y *tarjea*. Tenemos únicamente dos apariciones en el corpus, una para cada forma; la primera se registra en la entrevista de un hombre joven de instrucción alta, la segunda, en la de un hombre adulto de instrucción baja.

Corominas indica que, si bien su etimología es incierta, se trata de una voz de origen árabe cuya primera documentación proviene de 1527 (*Ordenanzas de Sevilla*). Sabemos que la *a* átona de inicio de palabra lleva a inestabilidad y pérdida cuando se reinterpreta como parte del artículo que la antecede, es decir, *la atajea* se simplificó y reinterpreto por *la tarjea* (Leal, 1993: 171). Esta pérdida de *a* inicial se registra muy tempranamente en castellano, de tal manera que la documentación en el Corpus Diacrónico del Español (Corde) nos permite atestiguar 10 contextos de uso en seis documentos distintos: tres del siglo xvii, dos del siglo xviii y uno del siglo xx. Consideramos pertinente para el español de Puebla registrar ambas formas como usuales.

En el caso de *tecojote* y *tejocote* registramos seis usos: cinco de la metátesis del nahuatlismo dichas por dos hablantes, un hombre y una mujer mayores, de instrucción media, y la forma *tejocote* sólo apareció una vez, dicha por una mujer adulta de instrucción alta. Pese a ser común en el habla, debido a que no hallamos documentación en fuentes escritas, decidimos registrarla en el corpus con la ortografía *tejocote* con la etiqueta de pronunciación <~tecojote>.

El resto de las palabras de la tabla 11 permanece sin modificación, pero una vez que encontramos que están definidas en el DEM con el significado del corpus, aunque la entrada tenga una ortografía distinta, las contamos como coincidencia.

Algunos de los vocablos de Preseea-Puebla tienen forma plural, de diminutivo o son pluriléxicos; todos estos casos, como puede verse en la tabla 12, no se identificaron automáticamente, puesto que no coinciden con las entradas del DEM, sino que se encuentran registrados como acepciones bajo una entrada simple, como se indica en la columna correspondiente.

Tabla 11. Distintos fenómenos de variación ortográfica entre Preseca-Puebla y la entrada del DEM

	Preseca-Puebla	DEM	DLE	Otras fuentes
1.	bísquet	bísquet	bísquet	
2.	boiler	bóiler		
3.	cártel	cartel	cartel ² (También cártel)	
4.	cuetón	cohetón		
5.	chelo	cello	chelo	
6.	diésel	diesel	diésel	
7.	eslogan	slogan	eslogan	
8.	estrés	stress	estrés	
9.	exprés	express	exprés	
10.	fólder	folder	fólder	
11.	grafiti	grafitti	grafiti	
12.	hándicap	hándicap	hándicap	
13.	huazontle	huazontle (también guauzontle o huazontli)		Conabio:* huazontle
14.	kínder	kínder	kínder	
15.	marihuana	mariguano		
16.	memorándum	memorándum	memorándum	
17.	quichua	quechua	quichua	
18.	récord	record	récord	
19.	referéndum	referéndum	referéndum	
20.	sándwich	sándwich	sándwich	
21.	sicología	psicología	sicología	
22.	tarjea	atarjea		
23.	tecojote	tejocote		
24.	tráiler	tráiler	tráiler	
25.	ultimátum	ultimátum	ultimátum	
26.	zodiaco	zodíaco	zodiaco	

* sv quelite, consultado en diciembre 2020, en <<https://www.biodiversidad.gob.mx/diversidad/alimentos/quelites>>.

FUENTE: elaboración propia.

Tabla 12. Vocablos de Preseea-Puebla registrados como acepciones en el DEM

	Preseea-Puebla	DEM
1.	a diario	sv diario (A) diario
2.	a dónde	sv dónde A <i>dónde</i>
3.	a prisa	sv prisa A (toda) prisa, de prisa o con (mucho) prisa o (muchas) prisas
4.	jándale!	sv andar ¡Ándale!
5.	anterior a	sv anterior Anterior a
6.	carnitas	sv carne Carnitas
7.	celos	sv celo Celos
8.	de veras	sv veras De veras
9.	en boga	sv boga En boga
10.	en torno	sv torno En torno a
11.	jamás de los jamases	sv jamás Jamás de los jamases
12.	por favor	sv favor Por favor
13.	o sea	sv ser Es decir, o sea, esto es

FUENTE: elaboración propia.

Al revisar la lista de vocablos, nos dimos cuenta de que algunas formas identificadas como monoléxicas son pluriléxicas, y, como tales, se encuentran definidas en el DEM (tabla 13).

Finalmente, en la tabla 14 registramos variaciones de género y número entre el lema identificado por *FreeLing* y la entrada del DEM.

COMENTARIOS FINALES

Hemos propuesto un método mixto (automatizado y manual) para la lematización de los corpus Preseea, que bien puede replicarse en cualquier otro corpus oral o regional una vez que se tiene un texto plano de análisis. La combinación de procesos nos ha permitido hacer correcciones en cada una de las etapas del procesamiento y del análisis cuantitativo de los datos que resumimos en la imagen 6.

Tabla 13. Vocablos pluriléxicos en Preseesa-Puebla documentados gracias al DEM

	Preseesa-Puebla	DEM
1.	best	best seller
2.	cabelludo	sv cuero <i>Cuero cabelludo</i>
3.	cake	hot cake
4.	chantilly	sv crema <i>Crema Chantilly</i>
5.	full	sv ful <i>A ful, al ful o hasta el ful</i>
6.	generis	sui generis
7.	gó	a go gó
8.	grosso	grosso modo
9.	hot	hot cake hot dog
10.	inalámbrico	sv teléfono <i>Teléfono inalámbrico</i>
11.	manuscrita	sv letra <i>Letra manuscrita</i>
12.	placero	sv taco <i>Taco placero</i>
13.	priori	sv a priori
14.	rocanrol	sv rock <i>Rock and roll</i>
15.	run	sv home <i>Home run</i>
16.	seller	best seller
17.	striptease	strip-tease
18.	topic	trending topic
19.	trending	trending topic
20.	vivendi	sv modus <i>Modus vivendi</i>

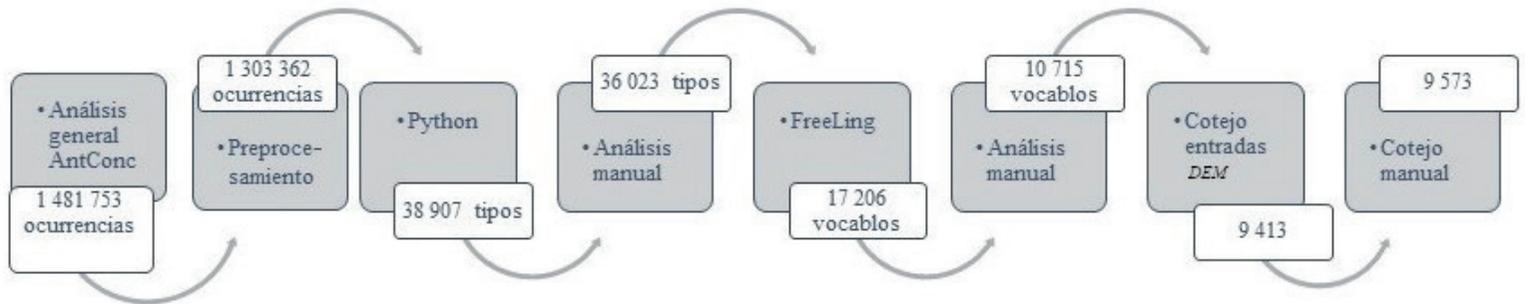
FUENTE: elaboración propia.

Tabla 14. Variación de género y número en el vocablo de Preseea-Puebla y la entrada en el DEM

	Preseea-Puebla	DEM		Preseea-Puebla	DEM
1.	abarrote	abarrotés	22.	marchanta	marchante
2.	alhajas	alhaja	23.	materna	materno
3.	altibajo	altibajos	24.	memelas	memela
4.	arra	arras	25.	misceláneo	miscelánea
5.	autoparte	autopartes	26.	montacarga	montacargas
6.	bastonero	bastonera	27.	morra	morro
7.	chaval	chavalo	28.	nagua	naguas
8.	clavados	clavado	29.	plomado	plomada
9.	comilla	comillas	30.	quedado	quedada
10.	concubina	concubino	31.	refaccionario	refaccionaria
11.	coqueto	coqueta	32.	regalía	regalías
12.	cortadora	cortador	33.	rejega	rejego
13.	cosquilla	cosquillas	34.	rendidora	rendidor
14.	crío	cría	35.	shorts	short
15.	declaratorio	declaratoria	36.	vacación	vacaciones
16.	drinks	drink	37.	vival	vivales
17.	fundidor	fundidora			
18.	gárgara	gárgaras			
19.	lépera	lépero			
20.	ligada	ligado			
21.	manchamanteles	manchamantel (también manchamanteles)			

FUENTE: elaboración propia.

Imagen 6. Proceso de análisis cuantitativo



FUENTE: elaboración propia.

Los resultados que produjo el estudio automatizado nos permiten reflexionar —teórica y metodológicamente— sobre la construcción de la unidad palabra en español y la construcción del vocablo-entrada en lexicografía.

Tradicionalmente, se dice que en el español de México se prefiere conservar la ortografía de la lengua que provee el préstamo léxico, sin embargo, hemos visto distintos ejemplos de variación ortográfica de la misma palabra, incluso de variación entre la ortografía del préstamo y la de formas derivadas. La variación ortográfica documentada puede ser objeto de un estudio pormenorizado sobre fijación, adaptación de préstamos y actitudes y creencias lingüísticas ante la norma.¹⁴

En este primer acercamiento panorámico a los datos hemos podido corroborar nuestra primera hipótesis, es decir, las entrevistas semidirigidas de Preseea-Puebla nos permiten recoger vocabulario general del español de México en sus diversas manifestaciones estilísticas (culto, técnica, coloquial, grosera, peyorativa, etcétera). Dicho de otra forma, no por tratarse de material oral encontramos únicamente léxico coloquial o léxico regional. Tras el análisis, podemos atestiguar que el vocabulario recogido en Preseea-Puebla es común al español de México, puesto que cerca de 90% forma parte del leuario del diccionario integral del español de México (*DEM*).

No podemos afirmar todavía qué proporción del porcentaje restante, 10%, es parte del español regional o si únicamente no se encuentra registrado en el *DEM* porque se trata de innovaciones léxicas (voces de uso común posterior a la sincronía práctica del *CEMC*), porque todo corpus tiene límites de registro o porque, aunque usual, no responde a sus criterios de macroestructura. En futuros estudios será importante hacer una descripción detallada sobre el nivel de lengua que puede asignarse a cada palabra del leuario de Preseea-Puebla, mismo que podría estudiarse en comparación con las marcas de uso (nivel de lengua) que ofrece el propio *DEM* o con las variables estratifi-

¹⁴ Considérese que todos los transcripores son hablantes cultos y de áreas de especialidad que generan una permanente reflexión sobre la lengua: estudiantes del último semestre —o egresados— de la licenciatura de lingüística y literatura hispánica y estudiantes de posgrado en lingüística y educación.

catorias del corpus. También puede estudiarse desde el *modo* y el *tenor* (Halliday, 1978) de cada entrevista, puesto que encontramos interacciones tanto de solidaridad como jerárquicas.

Visto de forma integral, Preseea-Puebla recoge una muestra de vocabulario que abarca distintos temas (oficios, juegos infantiles, alimentos, ciencia, tecnología), y la entrevista semidirigida favoreció que reuniéramos tanto vocabulario tradicional como voces de reciente acuñación. Queremos reiterar la importancia de continuar con estudios que abarquen, además de la forma, el análisis de cada uno de los contextos de uso.

Por último, podemos apuntar que el método que proponemos en este estudio se centra en la obtención de unidades monoléxicas. El reconocimiento de colocaciones y locuciones requeriría una planeación distinta en el procesamiento de los datos, aunque la base que proponemos sigue las mismas fases: el estudio automatizado de co-ocurrencia seguido de la lectura cuidadosa de las concordancias para distinguir tanto la fijación como la idiomatización. Sólo para ejemplificar, diremos que en la revisión de concordancia hemos identificado locuciones como *valer un cacahuete* ‘no importar nada’ —construye donde quieras, güey/ a mí me vale un cacahuete (PUEB_H12_010)—, *echar los perros* ‘pretender a alguien’ —y todavía el médico legista me echa los perros (M22_062)—, *a las vivas* ‘atento, pendiente de algo’ —sí/ tiene uno que estar ahí a las vivas (M11_023)—, *estar dado al cuas* ‘estar alguien o algo en muy mal estado’ —entonces era un billar/ medio cantina/ y era una vecindad [...] o sea la casa estaba fracturada en tres dueños/ y mm/ toda dada al cuas (M33_104)— y *de buenas a primeras* ‘de forma inesperada’ —en agosto y de buenas a primeras/ sin decir agua va quitaron les quitaron la plaza sin dar una sola explicación (H31-073)—.¹⁵

¹⁵ Sobre el estudio específico de colocaciones y locuciones en Preseea-Puebla, recomendamos los trabajos de Carmen Fajardo Rojas (2016) y Xóchitl Tavera-Cervantes (2021).

4. RASGOS DE ORALIDAD EN PRESEEA-PUEBLA: REFLEXIONES LEXICOGRÁFICAS

El siguiente paso en nuestra propuesta es acercarnos, al menos panorámicamente, al análisis cualitativo del leuario de Preseea-Puebla. Los objetivos de este acercamiento son 1) continuar describiendo la composición léxica de este corpus oral, 2) reflexionar sobre el posible tratamiento lexicográfico de este subconjunto léxico con miras a la lexicografía integral y diferencial, y, por lo tanto, 3) proponer un método que nos permita justificar cuándo un vocablo efectivamente es característico del español de Puebla.

En el acercamiento cuantitativo (hecho a partir de la comparación con el leuario del *DEM*), observamos que su núcleo corresponde, al menos formalmente, al español de México. Vimos además que el leuario de Preseea-Puebla registra tanto voces cultas como populares y groseras, y que en él se registra vocabulario de distintas áreas temáticas. Continuaremos con el mismo hilo conductor y, en esta aproximación, estudiaremos el vocabulario que no pudimos documentar en el *Diccionario del español de México*:¹ recuérdese que del cotejo de Preseea-Puebla con el leuario del *DEM* obtuvimos 1 142 vocablos que no se encuentran en el diccionario.

Organizamos el análisis en tres grandes secciones que nos permiten enfatizar la riqueza de los corpus orales para el estudio del habla y su riqueza expresiva. En la primera hemos clasificado las unidades que consideramos que son parte de la creatividad y la riqueza

¹ Recuérdese que la versión en línea del *DEM* está en permanente crecimiento y revisión. Los datos que aquí presentamos corresponden al cotejo con la versión publicada en línea en octubre de 2020.

discursiva oral (Briz, 2000; Lara, 1996; San Vicente, 2017, y Zimmermann, 1996).

En la segunda ejemplificamos vocablos que, si bien no están en el DEM, tienen relación formal (derivados, compuestos o parasintéticos) con un vocablo que sí aparece definido en este diccionario —por ejemplo, en el leuario no encontramos *autocontrol*, *berrinchudo*, *deschongar(se)* ni *picahielos*, pero sí *control*, *berrinche*, *chongo*, *picar* y *hielo*, además, como parte de la macroestructura del diccionario, encontramos tablas de infijos, una de prefijos y otra de sufijos, donde se registra *auto-*, *des-*, etcétera—.

En la tercera sección damos ejemplos de palabras de arraigo en el español de México que hasta el momento no se encuentran en el diccionario, pero seguramente, cuando aparezcan en su corpus, se integrarán a su leuario (como *ayocote* ‘frijol casi del tamaño de una haba’, *charpe* ‘horquilla con un resorte que sirve para lanzar objetos como pequeñas piedras’ o *querubín* ‘dentro de la tradición cristiana, ángel que, ubicado en el segundo coro celestial, se dedica a alabar a Dios’), así como neologismos o voces de uso común posterior a la sincronía práctica del DEM, que posiblemente se registren en el CEMC2 (*boxer* ‘prenda interior que cubre de la cintura hasta el principio de los muslos’, *engargolar* ‘unir hojas de papel introduciendo un arillo o espiral a través de los agujeros que se han hecho para ese fin en algunos de sus bordes’ o *queer* ‘identidad sexual o de género que problematiza las posiciones identitarias fijas o estáticas y, al mismo tiempo, las disidencias sexuales y de género’).

EXPRESIONES DE ORALIDAD

Las unidades léxicas que estudiaremos en este primer apartado han sido seleccionadas a partir de lo que intuitivamente consideramos característico de la riqueza discursiva oral. Algunas de estas voces son ocurrencias momentáneas, otras tienen mayor arraigo en la comunidad, incluso algunas se reproducen en la lengua escrita cuando se quiere caracterizar la oralidad. Dado que las unidades aquí estudiadas tienen rasgos disímiles, tenemos distintas propuestas para su tratamiento lexicográfico. Por ejemplo, nos parece que *bici* (de bicicleta), *quimio* (de quimioterapia) y *chacha* (de muchacha ‘sirvienta’) podrían convertirse en entradas del

Tabla 15. Creaciones léxicas con *-ción*

<i>Preseea-Puebla</i>	
1	acoplación
2	caducación
3	nombración
4	probación

FUENTE: elaboración propia.

diccionario como variantes léxicas, como ya ocurre en el *DEM* con las entradas *pa* y *para*, *pa* y *papá*, *ma* y *mamá* o *nomás* y *nada más*. *Chacha* requeriría una marca que dé cuenta de su uso despectivo o peyorativo. En cambio, *buzo caperuzo* (‘alerta, atento a algo’), *estudihambre* o *lengüeriquear*, si bien gozan de cierto arraigo en el habla, requieren de un contexto conversacional específico para que su producción sea adecuada a la situación comunicativa y, sobre todo, para ser comprendidas, por lo que no sugeriríamos su introducción al leuario del diccionario de lengua. No obstante, insistimos en que el registro de su uso espontáneo es de gran valía para distintos estudios, en especial para los de oralidad.

Una vez delineado este marco general, comenzaremos a dar ejemplos de distintos fenómenos. En la tabla 15 presentamos sustantivos deverbales formados con el sufijo *-ción*, que indican ‘el resultado de la acción expresada por el verbo’. Se trata de innovaciones léxicas que, si bien se generaron a partir de los esquemas morfológicos del español, responden a creaciones posiblemente efímeras,² espontáneas —unas veces voluntarias y festivas y otras a lapsus—, que ejemplificamos en (4) y (5). Cabe destacar la variación léxica que ellas representan, pues existen otras voces más difundidas de las cuales podemos decir, al menos aparentemente, que tienen el mismo significado (*acoplamiento*, *caducidad*, *nombramiento* y *prueba*):

² Este fenómeno de creación léxica dio lugar en años recientes a una moda lingüística que tuvo amplia difusión en memes con la imagen de un gato. Sirvan de ejemplo los siguientes textos, cuya autoría no nos ha sido posible documentar: “Hice la enamoración, entonces me hicieron la engañación y ahora estoy en la lloración”, “Preparándome para ir a la estudiación”, “Me apetece la morición” o “Me voy a hacer la automorición”. Como puede verse en estos ejemplos, además de con los recursos morfológicos, se juega con combinaciones sintácticas anómalas en español (por ejemplo, la combinación con el verbo *hacer*), con cambios léxicos (como *automorición* por *suicidio*) y con la ortografía *-sión*, *-ción*.

- (4) E: oye / ¿y para bailar? con / puro chaparrito ¿cómo le hacías?
 I: ¡ay no! hubieras visto esas / para las vueltas ¿no? (risa) pues fíjate que el que me gustaba estaba más o menos a mi vuelo / pues ya más o menos como que le agarrábamos pero siempre y cuando ahí siempre te dicen / pues / “déjame bailarla” ¿no? o o “pásala” / sí había chaparritos / pero a veces no me daban vueltas sino que el más lo que era con los chaparritos eran los pasos y ya / ahí haz de que cuenta de que es una **acoplación** de que tú te das la vuelta solita y te volteas para allá y así
- [PUEB-M12-026]

- (5) I: hay carne ya / que tiene por caducar / eh tienen en este / fecha de **caducación** / mm lácteos lo que antes nada más / el de enlatados lo que se usaban era los chiles
- [PUEB-M32-101]

En la tabla 16 mostramos voces que provienen de acortamientos (apócopos, aféresis y desgastes), de cruces léxicos y de voces en el linde morfofonológico. Todas se caracterizan por ser parte de interacciones informales, y su relación con la forma no acortada resulta transparente. Como puede verse, algunos de ellos son muy comunes y de uso general en español (sirvan de ejemplo *bici*, *depa*, *profe*, *quimio*, *refri* y *uni*). En ninguno de estos casos el cambio en la forma conlleva un cambio en el significado.

Nos parece que únicamente *chacha* y *mijo* presentan un comportamiento semántico distinto, puesto que *muchacha* no significa ‘mujer joven’ sino ‘sirvienta’, y su uso es peyorativo,³ y *mijo* puede analizarse como un desgaste fónico en el que el posesivo,

³ El DEM registra *muchacha* como forma lexicalizada femenina, y define “s f 1 Sirvienta: llamar a la *muchacha*, conseguir *muchacha*”, y, como segunda acepción, “2 *Muchacha de entrada por salida* Sirvienta que no habita en la casa para la que trabaja, sino que va y viene periódicamente”. Mientras tanto, el significado ‘joven’ se define, como es esperable, bajo la entrada *muchacho* “s Persona que no ha llegado todavía a la edad adulta, particularmente la que se encuentra en la adolescencia: los *muchachos del pueblo*, una *muchacha de pelo corto*”.

de naturaleza átona, que termina con la misma vocal con la que inicia la segunda palabra, se fusiona en la pronunciación, esto es, pasa de *mi hijo* a *mijo* (el sustantivo conserva sus propiedades, por lo que permite la formación del femenino, del diminutivo, etcétera), unas veces conservando el significado ‘descendiente de alguien’ (por ejemplo en “*Mijo* nació en Zacapoaxtla”), otras en calidad de forma de tratamiento afectiva y solidaria, como vocativo (“Me decía la directora: ‘¿Qué vas a estudiar, *mijita*? Estudia comercio’ ” o “Qué onda, *mijo*, ¡jalas?”).

Puede observarse que en esta misma lista hemos incluido *boni* y *bueni*, que podrían considerarse como formas truncadas de *bonito* y *buenito*, respectivamente; sin embargo, consideramos que la “i” final se emplea como un pseudomorfema de diminutivo usual en el habla dirigida a niños o aññada (como *amigui*, *chiqui*, *mami* y *papi*).

Enseguida presentamos algunos juegos verbales⁴ que se lematizaron como vocablos en el análisis automatizado, puesto que formalmente se corresponden con palabras monoléxicas. Ninguno de estos ejemplos corresponde a verdaderos vocablos o palabras de la lengua (Franco, en preparación; Frenk, 1953; Lara, 2012a, y Lope, 1980a y 1980b), pero dan prueba de la creatividad de los hablantes, así como de los recursos semánticos y formales con que cuentan la lengua y los hablantes para su formación. Dentro de la tabla 17, los ejemplos 14 (*buzo caperuzo*), 15 (*chacho chicho*) y 20 (*taka taka*) fueron identificados en el análisis de concordancias, pues en la lista aparecieron como vocablos independientes (*buzo*, *caperuzo*, *chacho*, *chicho* y *taka*).

⁴ Para la definición de *juego verbal* seguimos a Franco (en preparación): “El *juego verbal* se caracteriza por la manipulación intencionada de los distintos componentes que estructuran un sistema lingüístico mediante el manejo de una o más técnicas discursivas, dando como resultado una palabra, una frase o un texto breve que trae consigo un efecto humorístico, ingenioso, irónico, irreverente, eufemístico, etcétera; valgan como ejemplo casos como: ¿*Qué hongo?* como variante del saludo ¿*Qué onda?*; *Voy echar una firma* para comunicar que se va a ir a orinar; *empacar* con el sentido de ‘comer’; *Si tienes lavadora, ahí lavemos* para despedirse de alguna persona o *Ya bailó Bertha* como eufemismo de *Ya valió verga*”.

Tabla 16. Acortamiento

	Acortamiento	Forma sin acortar	Acortamiento	Forma sin acortar		
Apócope	1.	bara	barato	2.	beis	beisbol
	3.	bici	bicicleta	4.	bra	brassier
	5.	bro	brother	6.	cel	celular
	7.	compu	computadora	8.	conta	contador
	9.	depa	departamento	10.	depre	depresión
	11.	derma	dermatología	12.	dírec	director
	13.	doc	doctor	14.	electro	electrocardiograma
	15.	fut	futbol	16.	gober	gobernador
	17.	greco	(lucha) grecolatina	18.	inge	ingeniero
	19.	lic	licenciado	20.	logo	logotipo
	21.	micro	microbús	22.	obvi	obvio
23.	pele	película	24.	peni	penitenciaria	
25.	peque	pequeño	26.	presi	presidente	
27.	profe	profesor	28.	quimio	quimioterapia	
29.	refri	refrigerador	30.	uni	universidad	
31.	voli	volibol				
Aféresis	32.	bus	autobús	33.	chacha	muchacha
	34.	ñora	señora	35.	jámonos!	vámonos
Acortamiento	36.	maso	más o menos	37.	mijo*	mi hijo
	38.	porfa	por favor	39.	pueque	puede que
Pseudodiminitivo	40.	bueni	bueno	41.	boni	bonito

* En el DEM sí aparece un artículo lexicográfico para *mijo*, pero las acepciones corresponden respectivamente a la planta herbácea (*Panicum miliaceum*) y a su grano.

FUENTE: elaboración propia.

Tabla 17. Juegos verbales

<i>Fenómeno de estudio</i>	<i>Ejemplo</i>	<i>Proceso</i>
Cruce léxico	1. bizcabuela	bizca + abuela
	2. buapachoso	BUAP + guapachoso
	3. buaperro	BUAP + perro
	4. combeber	convivir + beber
	5. chochentas	chochear + ochentas
	6. estudi hambre	estudiante + hambre
	7. indiosado	indio + endiosado
	8. indiosincrasia	indio + idiosincrasia
	9. ingeniatura	ingeniería + licenciatura
	10. panzar	panza + pasar (pasar de panzazo)
Derivación	11. loquero	loco + -ero
	12. mamitis	mamá + -itis
Simplificación	13. sanseacabó	
Rima	14. buzo caperuzo	
	15. chacho chicho	
Juegos fónicos	16. chesco	
	17. fucho	
	18. gruexo [gruekso]	
Eufemismo	19. ingue	
Duplicación	20. taka taka	
Acrónimo	21. pipope	

FUENTE: elaboración propia.

Entendemos por *cruce léxico* “la fusión o superposición de dos palabras en un solo lexema” (Rodríguez, 1989: 357). En el caso de los juegos verbales, el cruce permite mayoritariamente la formación de unidades de significado composicional.⁵

Revisemos los ejemplos de la tabla 17 comenzando por *bizcabuela* (quizá debería escribirse *biscabuela*, aunque hemos optado por la primera ortografía para dar cuenta con mayor transparencia de las partes que lo integran): se trata de un cruce entre las palabras *bizca* y *abuela* (6) cuyo soporte es la vocal compartida al final y al principio de ambas palabras, que permite nombrar de forma lúdica y cariñosa a la bisabuela. El uso podría estar motivado por un personaje de la serie de televisión *El Chavo del Ocho*: Doña Nieves Frías viuda de Limón Aguado (Doña Nieves), la bisabuela de la Chilindrina, a quien este personaje infantil nombra siempre como “mi bizcabuela”. Ambos personajes eran representados por la misma actriz, María Antonieta de las Nieves.

- (6) I: no me gusta / es que mi abuelita le digo “ay es que si contigo no cambió / no va a cambiar con los demás”
 E: sí por una parte es bonito que mantenga la / la relación y que
 I: mjú
 E: pues el niño tenga / a su mamá / a su [abuela a su]
 I: [a su bisabuela / a su **bizcabuela**]
 E: (risa)
 I: también tiene
 E: ¡ah sí?
 I: sí también vive mi abuelita

[PUEB-M11-021]

⁵ Autores como González Avilés (2018) y Makri (2012), basados en la propuesta de Cusin-Berche 1999, usan el término *compocación* —formado precisamente por el proceso que nombra (de *composition* ‘composición’ y *truncation* ‘acortamiento’)— para distinguirlo de *blend*. En ambos procesos, la unión de lexemas produce una pérdida fónica, pero en el segundo, el enlace se da sobre un elemento fónico común. Por ejemplo, la unión de *estudiante* con *hambre* se hace sobre la vocal <a>, por lo que se clasificaría como *blend*, mientras que *dieciantos*, formado sobre la base de los números 16 a 19 y la palabra *tantos*, que no tiene un sonido que sirva de gozne, se analizaría como *compocación*. En este estudio emplearemos *cruce* para englobar ambos fenómenos.

Los dos juegos siguientes, *buapachoso* y *buaperro*, nombran de manera festiva a los estudiantes de la universidad pública más grande del estado de Puebla: se trata de cruces entre la sigla BUAP, y las palabras *guapachoso*, que el DEM define como “adj (*Popular*) Cuyo ritmo o baile es alegre y sensual”, y *perro* ‘persona que muestra abiertamente interés amoroso o sexual por otra’. Las siglas de la universidad fueron UAP hasta 1987, y BUAP desde ese año, cuando el Congreso del Estado la declaró institución Benemérita y la denominación festiva se actualizó con el cambio, es decir, pasó de decirse que alguien era (g)uapachoso a *buapachoso*.

Combeber es otro juego que no responde a la creación momentánea, sino que goza de cierta tradición en el habla. En él se cruzan las palabras *convivir* y *beber* y se destaca la idea de que, en una reunión, “el acercamiento y la camaradería entre los miembros de un grupo” (DEM) se acompañará con la ingesta de bebidas alcohólicas.

(7) I: que este / también nos reunimos es- bueno nosotros normalmente lo que hacemos es este / nos reunimos para comer / ah no es cierto [(risa)]

E: [(risa)]

I: okey para dar el grito⁶ [(risa)]

E: [(risa)]

I: para **combeber** [ah no es cierto]

E: [(risa)]

I: ¡no! ¡no! eso no con- no sí nada más este nos ponemos a a echar relajo ¿no? pero sanamente / bueno en mi familia casi no se acostumbra a tomar / entonces este / pues nos ponemos a / a estar ahí tranquilos este cenando todo / y ya cuando es lo del grito pues ya gritamos ¿no? [(risa)]

E: [(risa)]

[PUEB-M12-028]

⁶ *Dar el grito* se refiere a la celebración del “grito de la Independencia”, hecho que simboliza el inicio de la guerra de Independencia de México y que se conmemora el 15 de septiembre con una fiesta con comida tradicional mexicana, especialmente el pozole y el tequila.

Otros juegos con cierta tradición, es decir, que no son ocurrencias momentáneas, son *chochentas*: cruce que permite señalar que quien fue joven en la década de 1980 está *chocho*, “Que no le funciona la cabeza por ser de edad avanzada, que está decrepito” (DEM). La broma se basa en que se trata de personas adultas, pero no viejas; si no fuese así, la expresión tendría un carácter ofensivo en vez de lúdico. *Estudihambre*, cruce entre *estudiante* y *hambre*, enfatiza que durante la etapa escolar de la vida se tiene poco dinero y es común pasar hambre, y *panzar*, juego que reduce a unidad monoléxica la locución *pasar de panzazo*, definida en el DEM como “Aprobar un examen o una prueba con la calificación mínima”.

Tenemos tres juegos más que consideramos con menos tradición, dos de ellos ofensivos: *indiosado* e *indiosincrasia*. Se basan en la similitud fónica entre *indio* y el inicio de las palabras *endiosado* e *idiosincrasia*, con las cuales sólo hay un segmento de diferencia. Al colocar al inicio la palabra *indio*, se añade el significado “Persona ignorante, incivilizada” (DEM). *Ingeniatura* es, de este grupo, posiblemente el juego más espontáneo, y responde al cruce entre ingeniería y licenciatura. Tradicionalmente se ha distinguido entre estos dos tipos de estudios universitarios; sin embargo, es posible recibir un título, por ejemplo, de licenciado en ingeniería en sistemas computacionales.

Otro recurso para la formación de juegos es agregar sufijos o pseudosufijos. En el corpus documentamos dos formaciones híbridas. En la primera, el sufijo *-ero*, que permite formar sustantivos que significan ‘profesión u oficio’ (en este caso *loquero*), se agrega a la base, *loco*; esto es, se trataría del oficio de ‘la persona que se dedica a atender a los locos’. Se emplea para nombrar lúdicamente a psicólogos o psiquiatras. Por su parte, *mamitis*, formado con el sufijo culto *-itis* ‘inflamación’, permite señalar que una persona tiene un apego excesivo por su madre.

Otro recurso del juego es la rima. Por ejemplo, a (*ponerse*) *buzo* ‘estar listo para algo, ponerse atento’ se agrega *caperuzo*,⁷ para formar

⁷ Consideramos que en español de México la palabra *caperuzo* sólo se usa en este juego y nunca como palabra independiente. Nuestras búsquedas con restricción a México en distintos corpus (CEMC, Corde, CREA y Corpes XXI) coinciden

una rima consonante. En *chacho chicho* se registra primero un acortamiento de *muchacho* a *chacho*, que rima con *chicho* “que es muy bueno en lo que hace o a lo que se dedica; que es el mejor o más destacado” (DEM). En este caso, además de la rima, la segunda palabra sí aporta significado, y hemos podido documentar una forma ampliada del mismo (no podemos asegurar si el que presentamos fue reducción del mayor o viceversa): *el muchacho chicho de la película gacha*; incluso hay una canción del grupo mexicano de rock El Tri cuyo título es “Muchacho chicho”, que repite el juego completo.

Hemos agrupado tres juegos con cambio fónico: las palatalizaciones *chesco* por *refresco*, *fucho* por *futbol*, y *gruexo* [gruekso] por *grueso* “Difícil, complicado o peligroso”.

Restan tres ejemplos más: *ingue*, que es un acortamiento eufemístico de *chingue* (*su madre*). Esta pronunciación se generalizó en los años noventa debido a un conductor de televisión, y posiblemente estuvo motivada por una prohibición sobre lo que podía decirse en aquel momento en ese medio. Hemos clasificado como juego *taka taka*, duplicación de una voz inexistente en español que sirve como gentilicio peyorativo de ‘japonés’. Desconocemos el origen del mismo, pero podría deberse a que esta combinación de sonidos y sílabas resulta prominente a los mexicanos al escuchar hablar japonés.

El último ejemplo que presentaremos es el acrónimo *pipope*, etiqueta relevante e identitaria dentro de este corpus. El acrónimo significa ‘pinche poblano pendejo’, y se usa como gentilicio ofensivo cuando lo dice alguien externo a la comunidad; no obstante, entre los poblanos se convirtió en un juego con la reinterpretación “pieza poblana perfecta”, y nos ha sorprendido que algunos de nuestros estudiantes universitarios (adultos jóvenes de entre 18 y 25 años) ya

con esta intuición, pues sólo encontramos una concordancia de uso: “Los no militarizados, los no asalariados de las profesiones escalofriantes, también debemos comportarnos con valor, decisión, voluntad y astucia a fin de sobrevivir con donaire. O sea: buzo caperuzo” (*El batallador*, Gilberto Chávez, 1986). Llama la atención por ello que el *Diccionario de americanismos* presente la entrada caperuzo, -a con un significado cercano al de la rima: “Persona muy capaz que destaca en un ámbito determinado”, marcada como popular y de uso en Chile y Bolivia.

sólo conocen el segundo significado y ubican la palabra como voz identitaria.⁸

También hemos considerado como rasgos de oralidad el uso de algunos prefijos gradativos: *hiper-*, *micro-*, *mega-*, *re-* (con sus variantes *rete-* y *reteque-* y *super-*). Estos mismos aparecen en otros derivados del corpus (como *hipotensión*, *hiperactivo*, *microficha*), pero aquí nos ocuparemos únicamente de los que se usan para intensificar confiriendo mayor expresividad al discurso. Los intensificadores pueden aparecer juntos (como *supermega* o *superultramega* o combinarse con aumentativos y diminutivos, como *mega+árbol+ote*). Permiten esta gradación sustantivos, adjetivos, adverbios y verbos: *hipertrabajador*, *megaarbolote*, *megaborracho*, *megachiquito*, *mega-greñero*, *megaconsentido*, *megaendeudado*, *microgato*, *reagresivo*, *rebién*, *rebueno*, *recaro*, *recomer*, *rechillón*, *rechistoso*, *rechupar*, *redifícil*, *reencantar*, *refeo*, *reindispensable*, *relargo*, *reloco*, *remandar*, *rerrico*, *retepelear*, *retequeapurarse*, *retequenojar*, *retierno*, *superaguado*, *superbién*, *superbonito*, *superbueno*, *super-camioneta*, *supercansado*, *supercerca*, *supercinéfilo* (destaca en este ejemplo la unión al culto), *supercurso*, *superdescombinado*, *superdiferente*, *superdifícil*, *superdócil*, *superexigente*, *superguapo*, *superimportante*, *superinteligente*, *superinteresante*, *superlleno*, *superpeligroso*, *supermegaborracho*, *superplural*, *superpobre*, *superrápido*, *superseguro*, *super temprano*, *superultramegaayudar*.

(8) I: es que vas a Veracruz y de veras te tratan **rebién** / son a todo dar
[PUEB-M23-068]

(9) I: me paro / me **retequeapuro** / dejo hecho mi quehacer y ya me voy a trabajar / vengo en la tarde / vengo / hago mi comida y ya
[PUEB-M21-059]

(10) I: para que te pueda hipnotizar alguien / necesitas relajarte a / un grado en el que llegues a un estado mental // y para mí eso es

⁸ Se trata de un proceso de revaloración social que no es poco común. Sirva a manera de ejemplo complementario lo que ha ocurrido con *chilango*, gentilicio despectivo para los habitantes de la Ciudad de México que dio origen a campañas como “Ser chilango es chido”, a una revista con este nombre y a que esta denominación, ofensiva para algunos adultos, sea, para los habitantes más jóvenes de esta ciudad, una forma no marcada, es más, sea la etiqueta para autodenominarse.

superdifícil porque yo estoy pues / no no / no me puedo relajar así como que pon tu mente en blanco y ojos de huevo cocido / nada de esas cosas

[PUEB-M33-103]

- (11) I: me daba mucho mucho coraje que en tercero de prepa ellas sacaban / por ejemplo yo sacaba ocho DD sacaba ocho ¡y ellas sacaban nueve! pero era porque / las dos se sentaban atrás y las dos se **superultramegaayudaban**

[PUEB-M13-033]

También encontramos dobles verbales con y sin los prefijos *a-* y *en-* (*em-*) cuyo significado es el mismo, pero no su registro de habla. Por ejemplo, *abajar* y *bajar*, *ajuntar(se)* y *juntar(se)* ‘vivir junta una pareja sin casarse’, *emborrar* y *borrar*, *emprestado* y *prestado* y *empicar* y *en picada* ‘en caída libre’. Un doblete por sonorización, *segundar* y *secundar*, en el que posiblemente el cambio haga más transparente para los hablantes el significado ‘hacer segunda’, y, por último, documentamos un doblete del adjetivo *polviento* y *polvoriento*. Ejemplificamos en (12) y (13).

- (12) I: un mes / dos meses antes les había dicho que yo ya no me iba **ajuntar** que porque me iba muy mal con las chicas // pues mi última pareja pues me trató / me vio la cara ¿no?

[PUEB-M12-025]

- (13) I: y nos agarró la patrulla / y este y ora que nos pide que nos metiéramos a un callejón y que nos **abajáramos** los pantalones y ya mi cuate le dijo “no / ¡la manga qué!” dice “mejor llévame” “no, que te los bajas” “la manga” / mi cuate pues era más flaquito que yo pero era medio valentón / “no cabrón” dice y me dice / me dice a mí “tú tranquilo / tú ponte atrás de mí” me dice / porque como él sabía yo que venía de de del pueblo / este pues como que él me cuidaba

[PUEB-H22-048]

VOCES NO DEFINIDAS EN EL *DEM* PERO EMPARENTADAS MORFOLÓGICAMENTE CON SU LEMARIO

En Preseea-Puebla registramos palabras que, si bien no están definidas en el *DEM*, se emparentan morfológicamente con algunas de sus entradas, lo cual nos puede dar información sobre la tradición lexicográfica y sobre la productividad discursiva de ciertos morfemas. Nos centraremos en el segundo aspecto. Se trata de voces que responden a los procesos regulares de formación de palabras del español, y si bien su significado puede ser fácilmente reconstruido composicionalmente, su documentación nos permite describir los procesos lexicogenéticos del español de México (Zacarías, 2017 y 2018). En este sentido, los datos pueden ser o no relevantes para la lexicografía, según los criterios de inclusión de entradas de cada obra, pero, sin duda, son de interés para el estudio de la variación.

Para la identificación de redes o familias de palabras hemos hecho una consulta manual palabra por palabra en el buscador de la versión electrónica del *DEM*, el cual contiene una herramienta de gran valor para el estudio morfológico, puesto que nos avisa, por ejemplo, “Los siguientes vocablos guardan cierta similitud con el que se busca” (imagen 7). Cuando la lista ofrecida automáticamente no resultó útil para nuestro estudio, partimos de nuestro conocimiento lingüístico para introducir palabras emparentadas que creíamos que estarían definidas en el *DEM*.

Una vez obtenida la lista de parentesco comenzamos a organizar las palabras de Preseea-Puebla según el proceso de formación de palabras correspondiente. Para ello, retomamos la información de morfología léxica que aparece en la sección “Apoyos del *DEM*” en la versión electrónica del diccionario. Compartimos la idea de que los morfemas no tienen significado propio, sino que lo construyen al entrar en combinatoria;⁹ así debe entenderse la glosa que se presenta tras cada morfema en este apartado.

Comenzaremos por enlistar los procesos de derivación. Primero presentaremos una sección de prefijos, después de sufijos y, finalmente, de parasíntesis. Cada apartado se organiza alfabéticamente, por lo que se encuentran intercalados prefijos y sufijos “plenos”, numerales y cuantifica-

⁹ Para una revisión crítica sobre cuál es la unidad mínima con significado, véase Aronoff (1985) (en particular §2.1, pp. 7-17), y Lara (2004).

dores, elementos griegos y latinos que toman el lugar de prefijos o sufijos, y acortamientos modernos que toman el lugar de prefijos y entran en la composición de palabras.

La descripción de los morfemas que entran en combinatoria parte de su relación formal con el vocablo que aparece como entrada en el DEM, el cual se presenta en la tercera columna de las tablas 18 a 22. En ellas indicamos la categoría gramatical de cada palabra según los contextos de uso de nuestro corpus. En ocasiones se trata de un proceso directo de derivación entre la palabra documentada en Preseea-Puebla y la entrada del DEM, como en *antibala* → *anti-* + *bala*, o en *bicentenario* → *bi-* + *centenario*. En otros hay más de un proceso implicado o una relación indirecta en la formación de palabras; en ese caso hemos decidido presentar sólo uno de ellos. Privilegiamos el apartado de prefijación; por ejemplo, *antialacránico* → *anti-* + *alacrán* + *-ico* se registra en la sección de prefijos, bajo *anti-*. Estamos conscientes de que se trata de una decisión arbitraria, pero consideramos que establecer, al menos, un grado de relación formal con el leuario del diccionario y documentar la palabra en uso discursivo es suficiente para comprender su riqueza morfológica y semántica.

Antes de presentar la tabla que incluye los sufijos y los elementos cultos que toman el lugar del sufijo y entran en la composición de palabras, queremos dar cuenta de manera independiente de los derivados con sufijos apreciativos que consideramos que ya están lexicalizados, por lo que nos parece que deberían formar parte del leuario del diccionario.

Como hemos argumentado previamente (Palacios, 2020c: 443), entendemos por *lexicalización* tanto el resultado como el conjunto de procesos morfológicos y sintácticos que permiten que una unidad léxica o una sintáctica se convierta en un nuevo signo lingüístico, es decir, un nuevo elemento del leuicon cuyo significado propio hace necesario incluirlo en el diccionario de lengua. Agregaríamos que proponemos su inclusión como entrada independiente: *alucín* (*alucinación*),¹⁰ *arillo* (*aro*), *bovedilla* (*bóveda*), *corralón* (*corral*), *chalina* (*chal*), *chicharrín* (*chicharrón*), *filmina* (*filme*), *jetón* (*jeta*), *lambiscón* (*lamer*), *mandilón* (*mandil*), *patineta* (*patín*), *regañón* (*regañar*) y *remendón* (*remendar*).

¹⁰ Presentamos entre paréntesis la palabra incluida en el DEM que guarda relación con la de Preseea-Puebla.

Tabla 18. Palabras de Preseea-Puebla relacionadas con alguna entrada del DEM: prefijación

	Preseea-Puebla	Entrada en el DEM	Categoría
anti- 'oposición o contrariedad'	1. antialacránico	alacrán	adj
	2. antiasalto	asalto	adj
	3. antibala	bala	adj
	4. anticolonial	colonial	adj
	5. antidemocracia	democracia	sust
	6. antimoda	moda	sust y adj
	7. antipartidista	partidista	sust y adj
	8. antipedagógico	pedagógico	adj
	9. antipolítico	político	adj
	10. antisequestro	secuestro	adj
	11. antitaurino	taurino	adj
auto- 'a sí mismo, de sí mismo'	12. autocompasión	compasión	sust
	13. autocontrol	control	sust
	14. autopromoción	promoción	sust
bi- 'dos'	15. bicentenario	centenario	sust
bio- 1 'que se relaciona con la vida'	16. biomagnético	magnético	adj
	17. biopsicósíntesis	síntesis	sust
2 'que se relaciona con la biología'	18. biotecnología	tecnología	sust
co- 'reunión o participación de alguien o algo en lo expresado por la base'	19. coarticulado	articulado	adj
	20. codependiente	dependiente	adj
des- 'negación, privación o inversión del significado de la forma base'	21. desagradecidamente	agradecido	adv
	22. desaburrirse	aburrirse	vbo
	23. desatorar	atorar	vbo
	24. descalcificar	calcificar	vbo
	25. descasar	casar	vbo
	26. desclasamiento	clase	sust
	27. descongelado	congelado	adj
	28. descontrol	control	sust
	29. desempastado	empastado	adj
	30. desempastar	empastar	vbo
	31. desestresar	stress	vbo
	32. desestructurado	estructurado	adj
	33. desestructurar	estructurar	vbo
	34. desevolucionar	evolucionar	vbo
	35. desinteresar	interesar	vbo
	36. desproporcionar	proporcionar	vbo
	37. destender	tender	vbo
	38. desvalorado	valorar, sobrevalorado	adj

Tabla 18 (Continúa)

		Preseea-Puebla	Entrada en el DEM	Categoría
dis- 'dificultad o anomalía'	39.	discapacidad	capacidad	sust
	40.	disfuncionalidad	funcionalidad	sust
eco- acortamiento de 'ecológico'	41.	ecoturístico	turístico	adj
electro-* 'eléctrico o de la electricidad'	42.	electroóptico	óptico	adj
ex- 'abandono de un vínculo, cargo o posición'	43.	exalumno	alumno	sust
	44.	excompañero	compañero	sust
	45.	exconvento	convento	sust
	46.	exmujer	mujer	sust
	47.	exnovio	novio	sust
	48.	expareja	pareja	sust
extra- 'situación externa a otra cosa o más allá de ella'	49.	extracurricular	curriculum	adj
	50.	extranormal	normal	adj
fono-, fonía- 'que se relaciona con el sonido o la voz humana'	51.	fonomímica	mímica	sust
hemo- 'relacionado con la sangre'	52.	hemodinámico	dinámico	adj
	53.	hemodinámicamente	dinámico	adv
hiper- 'muy elevado sobre su estado normal'	54.	hiperactivo	activo	adj
hipo- 'menos de, por debajo de lo normal'	55.	hipotiroidismo	tiroides, tiroideo	sust
im- 'negación de lo significado por la base'	56.	impopular	popular	adj
inter- 'Entre uno o varios, en medio de'	57.	interactivo	activo	adj
	58.	interactuar	actuar	vbo
	59.	interdenominacional	denominación	adj
	60.	interreligioso	religioso	adj
micro- 'muy pequeño, de dimensiones menores a las normales'	61.	microelectrónico	electrónico	sust
	62.	microempresa	empresa	sust
	63.	microfilm	film	sust
minus- 'que es menos que lo normal o lo común'	64.	minusválido	válido	adj
multi- 'que tiene muchos elementos o se repite varias veces'	65.	multidisciplinario	disciplinario	adj
	66.	multifactorial	factorial	adj
	67.	multinivel	nivel	adj
	68.	multiuso	uso	adj
	69.	multivitamínico	vitamínico	adj
narco- acortamiento de 'narcotráfico'	70.	narcocorrido	corrido	sust

Tabla 18 (Finaliza)

		Preseea-Puebla	Entrada en el DEM	Categoría
neo- 'que se constituye como una novedad a partir de otra cosa de la misma especie'	71.	neoconservadurismo	conservadurismo	sust
	72.	neoliberalismo	liberalismo	sust
para- 'parecido o paralelo con otra cosa, o estado o carácter semejante a otra'	73.	paranormal	normal	adj
peri- 'en torno a, alrededor de'	74.	peridural	epidural	adj
poli- 'que tiene muchos o varios elementos'	75.	polideportivo	deportivo	sust
pre- 'anterior a otro o antes de otro'	76.	precabildo	cabildo	sust
	77.	precongelado	congelado	adj
	78.	prenacional	nacional	adj
psico- 'de la actividad mental, la conducta y los procesos psíquicos'	79.	psicométrico	métrico	adj
	80.	psicomotricidad	motriz	sust
radio- 'que se relaciona con las radiaciones'	81.	radiotécnico	técnico	sust y adj
	82.	recalcular	calcular	vbo
	83.	recomer	comer	vbo
	84.	redensificar	densificar	vbo
	85.	redirigir	dirigir	vbo
	86.	reingreso	ingreso	sust
	87.	reinscripción	inscripción	sust
	88.	remarcar	marcar	vbo
	89.	repoblado	poblado	adj
	90.	reprogramar	programar	vbo
	91.	reubicar	ubicar	vbo
	92.	rezumar	zumbar	vbo
retro-* 'hacia atrás'	93.	retrovisor	visor	adj
semi- 'medio, casi'	94.	semiescolarizado	escolaridad	adj
	95.	semihúmedo	húmedo	adj
	96.	semimilitarizado	desmilitarizado	adj
	97.	semipadrote	padrote	adj
	98.	semipeatonal	peatón	adj
	99.	semiurbano	urbano	adj
	sub- 'que ocurre por debajo de cierta superficie o de cierto grado; subordinado a otro'	100.	subcampeón	campeón
101.		subcorpus	corpus	sust
102.		subdirección	dirección	sust
103.		subespecialidad	especialidad	sust
ultra- 'que se ubica más allá de o del otro lado de algo'	104.	ultracostumbrismo	costumbrismo	sust
	105.	ultracostumbrista	costumbrista	adj

FUENTE: elaboración propia.

Imagen 7. Resultado de búsqueda en el DEM: “Los siguientes vocablos guardan cierta similitud con el que se busca”



Buscar



BÚSQUEDA AVANZADA

* CODEPENDIENTE *

no se ha incluido entre las entradas del diccionario. Sin embargo, "codependiente" aparece en los siguientes artículos:

farmacodependiente

adj y s m y f Que es adicto a los fármacos o que padece farmacodependencia: “Decían que Marilyn era una rubia *farmacodependiente*”, “Hay círculos de apoyo para *farmacodependientes*”

🐦 [Twitter #DEMColmex](#)

codependiente

Los siguientes vocablos guardan cierta similitud con el que se busca:

condescendiente
contendiente
dependiente
independiente

Lotería de palabras

divulgación

s f

1. Acto de divulgar algo: la *divulgación de una teoría*, “Su música no tuvo mucha *divulgación*”, “Trataron de contener la *divulgación de las ideas revolucionarias*”
2. De *divulgación* Tratándose de obras, conferencias, etc. que tiene como finalidad dar a conocer un tema determinado, abordándolo de manera que pueda ser comprendido o apreciado por el público en general: *libro de divulgación científica*, “Escribe artículos de *divulgación* sobre la cultura japonesa”

[Leer artículo completo](#)

FUENTE: DEM, <<https://dem.colmex.mx>>.

En relación con los sufijos diminutivos, queremos presentar dos ejemplos, uno de pérdida y otro de reinterpretación (falso diminutivo): *chambra* por *chambrita*, cuando la forma diminutiva nombra específicamente la prenda de abrigo para bebé que cubre el torso y tiene las mangas largas (Torres, 1970: 246); y *cema* por *cemita*, comida tradicional poblana.

En las tablas 19 y 20 enlistamos alfabéticamente los sufijos documentados. Desde luego, no se trata de una descripción exhaustiva de las posibilidades derivativas del español, ni siquiera de Preseea-Puebla, estudio que sería por demás interesante desde un acercamiento de variación morfológica y para replicar o dar seguimiento a la línea de trabajo que brillantemente desarrolla Ramón Zacarías, esto es, “un acercamiento al léxico del español y a las estructuras morfológicas que los hablantes tienen disponibles para ejercer su creatividad lingüística. También [...] de retratar el habla que los mexicanos ponemos en práctica todos los días” (2018: 9). Por ahora, nos centraremos en estudiar las palabras de nuestro corpus que no aparecen definidas en el *DEM* pero que son parte de la misma familia léxica de otras que sí están en el diccionario.

Enlistamos en la tabla 20 los adverbios terminados en *-mente* cuyos adjetivos base se definen en el *DEM*.

En la tabla 21 agrupamos las formaciones circunflejas de verbos y adjetivos documentadas en Preseea-Puebla. Encontramos verbos parasintéticos formados con *a-* *-ar*, *des-* *-ar* y *en-* (*em-*) *-ar*, y también adjetivos deverbales formados con el sufijo *-ado*. En estos últimos casos, y, a falta del verbo en el leuario del *DEM*, indicamos ambos procesos en la primera columna, es decir, enlistamos como formaciones verbales de las que se desprende el adjetivo.

Para recapitular lo mostrado en las tablas, podemos afirmar que el análisis de nuestros datos refuerza la misma tendencia que hemos venido señalando: Preseea-Puebla permite estudiar voces y recursos léxicos cultos y populares; además, podemos documentar la formación de palabras de especialidad de distintos ámbitos laborales. Como sucede en general con el español, los procesos de sufijación son más comunes. Hemos documentado 105 formaciones con prefijos, 496 con sufijos (516 si contamos los adverbios en *-mente* como un tipo de derivación sufijal) y 35 parasintéticas. Los derivados analizados

son mayoritariamente nombres (227 adjetivos y 273 sustantivos). Los sufijos más productivos son *-ado* (registrado en 88 palabras), *-ero* (en 56), *-ear* (en 42) y *-ar* (en 38). Consideramos que esto coincide con el comportamiento gramatical y semántico del corpus: por una parte, se registra un gran número de adjetivos deverbales y de verbos (algunos, posiblemente neológicos), y, por otra, una de las esferas más significativas en esta lista de palabras de Preseea-Puebla es el registro de oficios. En la tabla 22 daremos cuenta de los compuestos registrados.

En todos los casos revisados, los dos lexemas que intervienen en su formación se encuentran definidos en el *DEM*. No obstante, en algunos de ellos no parece que el significado se desprenda únicamente de los que tienen las unidades que los integran. Por ejemplo, *asaltacunas* ‘persona que tiene la costumbre de sostener relaciones amorosas con otra mucho más joven que ella’ (14) o *milhojas* ‘pan elaborado por varias capas de hojaldre’:

(14) I: entonces haz de cuenta que de ella y yo decíamos “tú te quedas con Q” / porque así se llamaba Q y “tú te quedas con R” porque casi se llevan por uno o dos años los dos / pero cuando entraron tenían como ocho o nueve años

E: ¡ay! ¡estaban chiquitos!

I: sí pero se veían bien bonitos [(risa)]

E: ¡pinches **asaltacunas**!

I: sí nosotras pues ya teníamos como catorce

[PUEB-M12-026]

(15) I: sí de vez en cuando voy allá / me gustan mucho los picones y me gustan mucho las / este ¿cómo se llaman? / **milhojas** / el pan de **milhojas** me gusta muchísimo creo que / en ningún lado lo hacen como ahí / y hay otras panaderías y entonces íbamos allá y mamá compraba el pan para / aparentemente toda la semana

[PUEB-M23-070]

Algunas de las palabras que presentamos son en realidad derivados de un compuesto, pero, dado que el compuesto no se registra en el lecionario del *DEM*, decidimos describirlos aquí, y proponemos que ambos, el derivado y el compuesto, están lexicalizados. Ejemplos de este

tipo son *motocarga* ‘vehículo de tres ruedas, con motor y batea para transportar cargas ligeras’ y *motocarguista* ‘persona que conduce una motocarga, especialmente quien lo hace por oficio’; *videojugador* ‘persona que juega videojuegos’ y *videojuego* (también incluido en tabla 22) “juego electrónico que se visualiza en una pantalla” (DLE).

Para terminar este apartado, presentamos algunos préstamos en proceso de adaptación que se relacionan con una entrada ya existente en el DEM: *breaker* ‘interruptor de luz; switch’ (en el DEM se define *break* y *no break*); *serpentario*, cuya etimología, según indica el DLE, es “Del Ingl. *serpentarium*, y este der. del lat. *serpens*, *-entis* ‘serpiente’, según el modelo del ingl. *aquarium* ‘acuario’”; *telemarketing* ‘sistema de mercadotecnia en el que el vendedor usa el teléfono u otro medio de comunicación electrónico para contactar con clientes potenciales’ (en el DEM, *marketing*); *tester* ‘persona que se dedica a probar productos para después participar en su campaña de venta’ (en el DEM, *test*), y *webcam* ‘cámara digital que se conecta a un dispositivo electrónico, generalmente una computadora, para capturar y transmitir imágenes a través de la web’ (en el DEM, *web*).

De manera global podemos decir que este apartado nos ha permitido observar el enriquecimiento léxico a partir de los recursos morfológicos generales del español. Sin duda, hace falta un estudio más detallado de la morfología léxica de Preseea-Puebla; sin embargo, creemos haber establecido una base descriptiva para futuras comparaciones con otros corpus orales que nos permite conocer si la productividad de cada recurso y morfema específico es similar en otras hablas locales.

En el aspecto lexicográfico, consideramos que la riqueza de las 656 palabras organizadas según los procesos de derivación, composición y parasíntesis en relación con una entrada del DEM brindará al lector información sobre la riqueza morfológica y pragmática del español de México. Coincidimos plenamente con Zacarías (2017: 123) en que “un análisis profundo de la morfología léxica, con especial atención en los patrones productivos, revela importantes aspectos de la identidad lingüística de una sociedad”. Por ello, consideramos que estas palabras —más los compuestos y préstamos en proceso de adaptación morfológica— deben estudiarse cuidadosamente para conocer si son candidatas a integrarse en el lecionario del diccionario.

Tabla 19. Palabras de Preseea-Puebla relacionadas con alguna entrada del DEM: sufixación

	Preseea-Puebla	DEM	Categoría
-a, -e, -o 'resultado de una acción'	1. carraspeo	carraspear	sust
	2. desfase	desfasar	sust
	3. embarco	embarcar	sust
	4. jaloneo	jalonear	sust
	5. monitoreo	monitorear	sust
	6. remojo	remojar	sust
	7. tontera	tontería	sust
-ada '1 resultado de una acción'	8. arrastrada	arrastrar	sust y adj
	9. barrancada	barranca	sust
	10. cateada	catear	sust
	11. chiripada	chiripa	sust
	12. mesereada	mesero	sust
	13. mochada	mochar(se)	sust
	14. nalgada	nalga	sust
	15. novelada	novelar	adj
2 'que se hace con el instrumento nombrado por la base'	16. camionada	camión	sust
	17. parrillada	parrilla	sust
3 'que es propio de un sujeto determinado'	18. chamacada	chamaco	sust
	19. chistosada	chistoso	sust
	20. fresada	fresa	sust
	21. huevonada	huevón	sust
	22. putada	puta	sust
	23. salvajada	salvaje	sust
-ado '1 la característica que resulta de la acción indicada por la base'	24. abreviado	abreviar	adj
	25. acuartelado	acuartelar	adj
	26. adentrado	adentrarse	adj
	27. albergado	albergar	adj
	28. alrevesado	al revés	adj
	29. avecindado	avecindarse	adj
	30. balaceado	balacear	adj
	31. botado	botar	adj
	32. calibrado	calibrar	adj
	33. cicatrizado	cicatrizar	adj
	34. clausurado	clausurar	adj
	35. conceptualizado	conceptualizar	adj
	36. convocado	convocar	adj
	37. chamuscado	chamuscar	adj
	38. choteado	chotear	adj

Tabla 19 (Continúa)

	Preseea-Puebla	DEM	Categoría
	39. debrayado	debrayar	adj
	40. depositado	depositar	adj
	41. desabrochado	desabrochar	adj
	42. descalabrado	descalabrar	adj
	43. desconectado	desconectar	adj
	44. desesperanzado	desesperanza	adj
	45. desligado	desligar	adj
	46. desvirtuado	desvirtuar	adj
	47. diagnosticado	diagnosticar	adj
	48. distanciado	distanciar	adj
	49. electrocutado	electrocutar	adj
	50. encabronado	encabronarse	adj
	51. encarrilado	encarrilar	adj
	52. enfrascado	enfrascarse	adj
	53. envidiado	envidiar	adj
	54. escampado	escampar	adj
	55. escolarizado	escuela, escolaridad	adj
	56. escriturado	escritura	adj
	57. estafado	estafar	adj
	58. estresado	stress	adj
	59. estropeado	estropear	adj
	60. etiquetado	etiquetar	adj
	61. fincando	fincar	adj
	62. flasheado	flash	adj
	63. fracturado	fracturar	adj
	64. grafitado	grafitti	adj
	65. hospedado	hospedarse	adj
	66. ideado	idear	adj
	67. incrementado	incrementar	adj
	68. influenciado	influenciar	adj
	69. inhalado	inhalar	adj
	70. insensibilizado	sensibilizar, insensible	adj
	71. lustrado	lustrar	adj
	72. madreado	madrear	adj
	73. mafiado	mafia	adj
	74. martajado	martajar	adj
	75. masticado	masticar	adj

-ado

i 'la característica que resulta de la acción indicada por la base'

Tabla 19 (Continúa)

	Preseea-Puebla	DEM	Categoría
	76. matado	matarse	adj y sust
	77. mermado	mermar	adj
	78. militarizado	militar, desmilitarizar	adj
	79. mochado	mochar(se)	adj
	80. negado	negar	adj
	81. orinado	orinar	adj
	82. pactado	pactar	adj
	83. paniqueado	pánico	adj
	84. parchado	parchar	adj
	85. penado	penar	adj y sust
	86. penalizado	penalizar	adj
	87. personalizado	personalizar	adj
	88. prendado	prenda	adj
	89. referenciado	referencia	adj
	90. regenerado	regenerar	adj
	91. repujado	repujar	sust
	92. rescatado	rescatar	adj
	93. resonado	resonar	adj
	94. revocado	revocar	adj
	95. revolcado	revolcar	adj
	96. rondado	rondar	adj
	97. saqueado	saquear	adj
	98. sepultado	sepultar	adj
	99. situado	situar	adj
	100. sonsacado	sonsacar	adj
	101. sopleteado	soplete	adj
	102. suicidado	suicidar	adj
	103. tachado	tachar	adj
	104. transitado	transitar	adj
	105. trepado	tregar	adj
	106. vacilado	vacilar	adj
	107. visitado	visitar	adj
	108. visualizado	visualizar	adj
	109. zarandeado	zarandear	adj
	110. boleado	bola	adj
	111. chapado	chapa	adj

-ado

¹ la característica que resulta de la acción indicada por la base'

² la característica que resulta de la acción indicada por la base

Tabla 19 (Continúa)

		Preseea-Puebla	DEM	Categoría	
-aje 1. 'el resultado de la acción hecha por el agente indicado en la base'	112.	ambulanteaje	ambulante	sust	
	113.	boletaje	boleto	sust	
2. 'el conjunto de elementos que conforman un todo'	114.	gramaje	gramo	sust	
	115.	clientelar	cliente	adj	
-al, -ar 'que algo es característico o abundante en relación con lo indicado por la base'	116.	chaval	chavalo / chavo	sust	
	117.	demonial	demonio	adj	
	118.	grupal	grupo	adj	
	119.	organizacional	organización	sust	
	120.	peatonal	peatón	adj	
	121.	rodal	rueda	sust	
	122.	oratoriano	oratorio	sust	
-ano 'que se relaciona con aquello expresado por la base; que algo o alguien tiene las cualidades o las costumbres correspondientes a ella'	123.	angustiante	angustiar	adj	
	124.	cagante	cagar	adj	
	125.	castrante	castrar	adj	
	126.	desgastante	desgastar	adj	
	127.	deteriorante	deteriorar	adj	
	128.	esclavizante	esclavizar	adj	
	129.	estresante	stress	adj	
	130.	fermentante	fermentar	adj	
	131.	formante	formar	sust	
	-ante 'el agente que realiza la acción de la raíz verbal'	132.	frustrante	frustrar	adj
		133.	impactante	impacto	adj
		134.	interpretante	interpretar	sust
		135.	intimidante	intimidar	adj
		136.	itinerante	itinerario	adj
		137.	postulante	postular	sust
		138.	relajante	relajar	adj
		139.	suavizante	suavizar	sust
	-anza 'el acto o el resultado de lo nombrado por el verbo base'	140.	tratante	tratar, trata	sust
		141.	traumante	traumar	adj
142.		mudanza	mudar	sust	
143.		acicarar	acicate	vbo	
-ar Morfema flexivo que pospuesto a la raíz forma verbos	144.	agendar	agenda	vbo	
	145.	alocar	alocado	vbo	
	146.	auditar	auditoría	vbo	
	147.	avorazar	avorazado	vbo	

Tabla 19 (Continúa)

	Preseea-Puebla	DEM	Categoría
	148. barbar	barba	vbo
	149. bardar	barda	vbo
	150. bronquearse	bronca	vbo
	151. colapsar	colapso	vbo
	152. contactar	contacto	vbo
	153. curricular	currículo, curriculum	vbo
	154. derrumbar	derrumbe	vbo
	155. deshielar	deshielo	vbo
	156. desplomar	desplome	vbo
	157. enseriar	en serio	vbo
	158. esposar	esposas	vbo
	159. estresar	stress	vbo
	160. gestar	gestación	vbo
	161. gradar	grado	vbo
	162. grifar	grifo	vbo
	163. impactar	impacto	vbo
	164. infundar	infundado	vbo
	165. liderar	líder	vbo
	166. lindar	linde	vbo
	167. lloviznar	llovizna	vbo
	168. machuelar	machuelo	vbo
	169. mafiar	mafia	vbo
	170. maliciar	malicioso	vbo
	171. mecatear	mecate	vbo
	172. mellar	mella	vbo
	173. misionar	misionero	vbo
	174. pautar	pauta	vbo
	175. potenciar	potencia	vbo
	176. prejuiciar	prejuicio	vbo
	177. promocionar	promoción	vbo
	178. publicitar	publicitario / publicista	vbo
	179. sincerar	sincero	vbo
	180. zanjar	zanja	vbo
-ar Morfema flexivo que pospuesto a la raíz forma verbos	181. herbolario	hierba, herbario	sust
-azo 1 'que algo es resultado de una acción intensa o violenta'	182. almohadazo	almohada	sust
	183. batazo	bat	sust
	184. calcetazo	calceta	sust
	185. cinchazo	cincho	sust

Tabla 19 (Continúa)

	Preseea-Puebla	DEM	Categoría	
-azo 1 'que algo es resultado de una acción intensa o violenta'	186.	cinturonazo	cinturón	sust
	187.	cojinazo	cojín	sust
	188.	fierrazo	fierro	sust
	189.	golpetazo	golpe	sust
	190.	pleitazo	pleito	sust
	191.	trapazo	trapo	sust
	192.	zoquetazo	zoquetiza	sust
2 'que resulta de una acción rápida o intempestiva'	193.	chicotazo	chicote	sust
	194.	escopetazo	escopeta	sust
	195.	golazo	gol	sust
	196.	ruiderazo	ruido	sust
-ble 'que algo o alguien es digno o merece la realización de una acción'	197.	abastecible	abastecer	adj
	198.	distribuable	distribuir	adj
	199.	pasable	pasar	adj
	200.	rescatable	rescatar	adj
	201.	retroactible	retroactivo	adj
-cida 'que mata o elimina algo'	202.	sarricida	sarro	adj y sust
-ción 'el resultado de la acción expresada por la base'	203.	globalización	global	sust
	204.	igualación	igualar	sust
	205.	impermeabilización	impermeabilizar	sust
	206.	intimidación	intimidar	sust
	207.	masificación	masa	sust
	208.	monopolización	monopolio	sust
	209.	redensificación	densificación	sust
	210.	sanación	sanar	sust
	-ico 'que tiene la cualidad indicada por la base'	211.	claustrofóbico	claustrofobia
-(i)dad 'abstracción de una cualidad'	212.	equitatividad	equitativo	sust
	213.	manualidad	manual	sust
	214.	visceralidad	visceral	sust
-dera 'nombres de utensilios o instrumentos que sirven para realizar determinada acción'	215.	caminadera	caminar	sust
-dero, -dera 1 'el resultado repetido o continuado de lo expresado por la base'	216.	cambiadero	cambiar	sust
	217.	compradera	compra	sust
	218.	enredadero	enredar	sust
	219.	estudiadera	estudiar	sust
	220.	lavadera	lavar	sust
	221.	llenadera	llenar	sust
	222.	lloradera	llorar	sust

Tabla 19 (Continúa)

	Preseea-Puebla	DEM	Categoría
-dero, -dera 1 'el resultado repetido o continuado de lo expresado por la base'	223. peleadero	pelear	sust
	224. sacadera	sacar	sust
-dor 'mecanismo, dispositivo o aparato que realiza una acción'	225. casadero	casar	adj
	226. alisador	liso	sust
	227. barredor	barrer, barrendero	sust
	228. deshebrador	deshebrar	sust
	229. enchinador	chino	sust
	230. humidificador	humo	sust
	231. navegador	navegar	sust
	232. pulidor	pulidora	sust
	233. respirador	respirar	sust
	234. triturador	triturar	sust
-dor, -dora 'que alguien o algo tiene la cualidad o costumbre producida por una acción'	235. vaporizador	vaporizar	sust
	236. vibrador	vibrar	sust
	237. abortador	abortar	adj
	238. calificador	calificar	sust
	239. capacitador	capacitar	sust
	240. cateador	catear	sust
	241. concentrador	concentrar	sust
	242. consentidor	consentir	adj
	243. chacoteador	chacotear	adj
	244. degustador	degustación	sust
	245. desarrollador	desarrollar	sust
	246. enriquecedor	enriquecer	adj
247. escalador	escalar	sust	
248. facilitador	facilitar	sust	
249. identificador	identificar	sust	
250. mamador	mamar	sust	
251. notificador	notificación	sust	
252. rehabilitador	rehabilitación	sust	
253. solapador	solapar	sust	
254. tratador	tratar	sust	
255. violador	violar	sust	
256. voceador	vocear	sust	
-dora 1 'lugar o agente que realiza la acción expresada por la base'	257. alquiladora	alquilar	sust
	258. armadora	armar	sust
	259. certificadora	certificar	sust

Tabla 19 (Continúa)

		Preseea-Puebla	DEM	Categoría
2 'máquina, mecanismo o dispositivo que realiza una acción'	260.	copiadora	copia	sust
	261.	deshebradora	deshebrar	sust
	262.	punteadora	puntear	sust
	263.	teñidora	teñir	sust
-dura 'conjunto de los elementos expresados por la base'	264.	jabonadura	jabón	sust
-ear* Morfema flexivo que pospuesto a la raíz forma verbos. Es muy productivo en neologismos	265.	alcahuetear	alcahuate	vbo
	266.	atarear	atareado	vbo
	267.	bolsear	bolsa	vbo
	268.	cachetear	cachete	vbo
	269.	cachondear	cachondo	vbo
	270.	campanear	campana	vbo
	271.	campechar	campechano	vbo
	272.	cartear	carta	vbo
	273.	cascarear	cascarita	vbo
	274.	clarear	claro	vbo
	275.	chantajear	chantaje	vbo
	276.	chelear	chela	vbo
	277.	choquear	shock	vbo
	278.	dobletear	doblete	vbo
	279.	esmaltear	esmalte	vbo
	280.	esnorquelear	snorkel / esnorquel	vbo
	281.	flashear	flash	vbo
	282.	grafitear	grafitti	vbo
	283.	gritinear	gritar, gritón	vbo
	284.	hojalatear	hojalata	vbo
	285.	jetear	sv jeta Echar(se) una jeta	vbo
	286.	mensajear	mensajero, mensaje	vbo
	287.	meserear	mesero	vbo
	288.	mochilear	mochila	vbo
	289.	monear	mona	vbo
	290.	nalguear	nalga	vbo
	291.	navajear	navaja	vbo
	292.	necear	neocio	vbo
	293.	orejear	oreja	vbo
	294.	padrotear	padrote	vbo
295.	palotear	palo	vbo	
296.	pastear	pasto	vbo	

Tabla 19 (Continúa)

	Preseea-Puebla	DEM	Categoría	
-ear* Morfema flexivo que pospuesto a la raíz forma verbos. Es muy productivo en neologismos	297.	perrear	perro	vbo
	298.	pilotear	piloto	vbo
	299.	putear	puto / puta	vbo
	300.	rocanrolear	rock and roll	vbo
	301.	terapear	terapia	vbo
	302.	testear	test	adj
	303.	tristear	triste, tristeza	vbo
	304.	truquear	truco	vbo
	305.	vagabundear	vagabundo	vbo
	306.	ventanear	ventana	vbo
-era 1 'recipiente que contiene o en que se coloca algo'	307.	alcanforera	alcanfor	sust
	308.	coquera	coca, cocaína	sust
	309.	harinera	harina	sust
	310.	molera	mole	adj
	311.	pecera	pez	sust
2 'ocupación u oficio de una persona'	312.	boletera	boleto	sust
	313.	danzonera	danzón	sust
	314.	memelera	memela	sust
	315.	violetera	violeta	sust
	316.	conera	cono	sust
3 'accesorio o prenda de vestir o una parte de ella'	317.	estuchera	estuche	sust
	318.	resortera	resorte	sust
-ería 1 'estado de la característica de la base'	319.	holgazanería	holgazán	sust
	320.	monería	mono ²	sust
	321.	mamonería	mamón	sust
2 'actividad que se realiza como oficio'	322.	putería	puta	sust
3 'lugar donde se da un servicio o se vende lo indicado por la base'	323.	barbería	barba	sust
	324.	calcetería	calceta	sust
	325.	cemitería	cemita	sust
	326.	marisquería	marisco	sust
	327.	ostionería	ostión	sust
	328.	pollería	pollo	sust
	329.	pozolería	pozole	sust
	330.	galletería	galleta	sust
	331.	hojalatería	hojalata	sust
	332.	leñería	leña	sust
	333.	restaurantería	restaurante	sust

Tabla 19 (Continúa)

	Preseea-Puebla	DEM	Categoría
3 'lugar donde se da un servicio o se vende lo indicado por la base'	334. sedería	seda	sust
	335. talachería	talacha	sust
	336. tortería	torta	sust
4 'abundancia o el conjunto de lo expresado por la base'	337. milperío	milpa	sust
	338. chingonería	chingar	sust
	339. sangrerío	sangre	sust
-ero 1 'ocupación u oficio de una persona'	340. aluminiero	aluminio	sust
	341. bizcochero	bizcocho	sust
	342. boyero	buey	sust
	343. cadenero	cadena	sust
	344. carguero	carga, cargador	adj
	345. cetrero	cetrería	sust
	346. charolero	charola	sust
	347. chirrionero	chirrión	sust
	348. churrero	churro	sust, adj
	349. empanadero	empanada	sust
	350. fayuquero	fayuca (también falluca)	sust
	351. franelero	franela	sust
	352. gatillero	gatillo	sust
	353. grafitero	grafitti	sust
	354. grillero	grillo	sust, adj
	355. grupero	grupo	adj
	356. hojalatero	hojalata	sust
	357. joyero	joya	sust
	358. ladrillero	ladrillo	sust
	359. maraquero	maraca	sust
	360. marmolero	mármol	sust
	361. nevero	nieve	sust
	362. papero	papa	sust
	363. pollero	pollo	sust
	364. repostero	repostería	sust
	365. rocanrolero	rock and roll	sust
	366. sonidero	sonido	sust
	367. talachero	talacha	sust
	368. tarjetero	tarjeta	sust
	369. textilero	textil	sust
	370. torilero	toril	sust
	371. tubero	tubo	sust
372. viñero	viña	sust	

Tabla 19 (Continúa)

	Preseea-Puebla	DEM	Categoría	
2 'nombres de objetos, utensilios o muebles para poner lo expresado por la base'	373.	conero	cono	sust
	374.	cortinero	cortina	sust
	375.	cucharero	cuchara	sust
	376.	cunero	cuna	sust
3 'lugar donde abunda lo indicado por la base'	377.	avispero	avispa	sust
	378.	botanero	botana	sust
	379.	baciniquero	bacínica	sust
	380.	empedrado	empedrado	sust
	381.	marranero	marrano	sust
	382.	putero	puta	sust
4 que una condición es propia de la cualidad manifestada por la base	383.	cacahuatero	cacahuate	sust
	384.	carero	caro	adj
	385.	carroñero	carroña	sust y adj
	386.	colero	cola	sust
	387.	guajolotero	guajolote	adj
	388.	llanero	llano	adj
	389.	mañanero	mañana	adj
	390.	mochilero	mochila	sust
	391.	negrero	negro	adj
	392.	pistero	pistear	sust
	393.	playero	playa	adj
	394.	rollero	rollo	adj
	395.	marranero	marrano	sust
-esco que algo o alguien tiene la cualidad propia de la base	396.	porresco	porro	adj
-ez resultado abstracto de una característica o una cualidad	397.	adultez	adulto	sust
	398.	borreguez	borrego	sust
	399.	codez	codo	sust
	400.	inmediatez	inmediato	sust
	401.	jodidez	jodido	sust
	402.	peladez	pelado	sust
-grafo, -grafía 'propio de la escritura'	403.	mastografía	mama	sust
-ía* 'dignidad, jurisdicción, oficio o lugar donde se ejerce lo expresado por la base'	404.	contraloría	contralor	sust
-(t)ico* 'que tiene la cualidad o característica expresada por la base'	405.	diabético	diabetes	sust
	406.	ideático	idea	adj
	407.	mediático	medio	adj

Tabla 19 (Continúa)

	Preseea-Puebla	DEM	Categoría	
-'(t)ico* 'que tiene la cualidad o característica expresada por la base'	408.	púbico	pubis	adj
	409.	simbiótico	simbiosis	adj
	410.	vandálico	vándalo	adj
-ido, -ida 'que alguien o algo tiene la cualidad o se encuentra en el estado de lo significado por la base'	411.	advertido	advertir	adj
	412.	demolido	demoler	adj
	413.	derruido	derruir	adj
	414.	despedido	despedir	adj
	415.	fortalecido	fortalecer	adj
	416.	percibido	percibir	adj
	417.	permanecido	permanecer	adj
	418.	recluido	recluir	adj
	419.	regido	regir	adj
	420.	residido	residir	adj
	421.	sacudido	sacudir	adj
-iento 'que posee la cualidad, sensación o condición expresada por la base'	422.	relajiento	relajo	adj
-ficar* hacer o convertir algo en lo expresado por la base	423.	lotificar	lote	vbo
-il 'que algo es propio o característico de lo significado por la base'	424.	gangsteril	gángster	adj
-ismo 1 'el nombre de una corriente del pensamiento en cualquier disciplina'	425.	chamanismo	chamán	sust
	426.	estilismo	estilo	sust
	427.	priismo	priista	sust
2 'la actitud o el comportamiento de alguien'	428.	cuatachismo	cuate	sust
	429.	pistolerismo	pistolero	sust
	430.	protagonismo	protagonista	sust
	431.	raterismo	ratero	sust
3 'la exaltación de las costumbres de otras culturas o los usos y préstamos de palabras de una lengua en contraste con otra'	432.	chilanguismos	chilango	sust
-ista 1 'ocupación u oficio de una persona'	433.	anatomista	anatomía	sust
	434.	bajista	bajo	sust
	435.	biblista	Biblia	sust
	436.	camarista	camarero	sust
	437.	consumista	consumismo	sust
	438.	intensivista	intensivo	sust
	439.	jazzista	jazz	adj
	440.	montañista	montaña	sust
	441.	ortopedista	ortopedia	sust
	442.	paracaidista	paracaídas	sust
	443.	referencista	referencia	sust

Tabla 19 (Continúa)

	Preseea-Puebla	DEM	Categoría	
-ista 1 'ocupación u oficio de una persona'	444.	relajista	relajo	sust
	445.	zootecnista	zootecnia	sust
2 'partidario o seguidor de un movimiento, estilo o ideología'	446.	maoísta	mao	sust
	447.	cognitivo	conocimiento	adj
-ivo, -iva 1 'que se relaciona o es propio de lo significado por la base'	448.	normativa	normativo	sust
	449.	confrontativa	confrontar	adj
2 'condición indicada por la base'	450.	llevativo	llevar	adj
	451.	optativo	optar	adj
	452.	participativo	participar	adj
	453.	supresivo	suprimir	adj
	454.	batiza	bat	sust
-iza* 'gran cantidad de golpes fuertes que alguien da o recibe'	455.	madriza	madrear	sust
	456.	putiza	puto / puta	sust
	457.	tranquiza	trancazo	sust
	458.	corretiza	corretear	sust
'resultado de la acción expresada por la base'	459.	regañiza	regañar	sust
	460.	alcoholizar	alcohol	vbo
-izar 'acción derivada de la base'	461.	estupidizar	estupidez, estúpido	vbo
	462.	gelatinizar	gelatina	vbo
	463.	globalizar (globalizado)	global	vbo
	464.	microfonizar	micrófono	vbo
	465.	peatonizar	peatón	vbo
	466.	potencializar	potencial	vbo
	467.	sociabilizar	social, sociable	vbo
	468.	enamoradoizo	enamorado / enamorar	adj
-izo 'que alguien frecuentemente experimenta lo expresado por la base'	469.	dietología	dieta	sust
-logía 'ciencia o estudio de una disciplina'	470.	metrología	metro	sust
	471.	computólogo	computadora	sust
-logo 'estudioso de una disciplina'	472.	comunicólogo	comunicador	sust
	473.	nutriólogo	nutrición	sust
	474.	odontólogo	odontología	sust
	475.	patólogo	patología	sust
	476.	radiólogo	radiología	sust

Tabla 19 (Finaliza)

	Preseea-Puebla	DEM	Categoría	
-miento 'el resultado de la acción o el acontecimiento significado por la base'	477.	chocamiento	choque	sust
	478.	encabronamiento	encabronarse	sust
	479.	entubamiento	entubar	sust
	480.	perseguiamiento	perseguir	sust
	481.	desdibujamiento	desdibujar	sust
-nomía* 'estudio o conjunto de leyes de una disciplina'	482.	archivonomía	archivo	sust
-oso 'que algo o alguien tiene en abundancia lo indicado por la base'	483.	azufroso	azufre	adj
	484.	bilioso	bilis	adj
	485.	castroso	castrar	adj
	486.	chocoso	chocante	adj
	487.	cizañoso	cizaña	adj
	488.	compañeroso	compañero	adj
	489.	encimoso	encimar	adj
	490.	espirituosos	espiritual	adj
	491.	granoso	grano	adj
	492.	guerroso	guerra	adj
	493.	hippioso	hippie	adj
	494.	hostigoso	hostigar	adj
	495.	piojoso	piojo	adj
	496.	rastoso	rasta	adj
497.	sexoso	sexo	adj	
498.	yerboso	yerba	adj	
-teca 'lugar donde se guarda algo'	499.	planoteca	plano	sust
-udo* 'que algo o alguien tiene o hace en abundancia lo indicado por la base'	500.	berrinchudo	berrinche	adj y sust
	501.	broncudo	bronca	adj
	502.	cabelludo	cabello	adj
	503.	chipotudo	chipote	adj
	504.	mocudo	mocosos, moco	adj y sust
	505.	suertudo	suerte	adj

* Unidades cuyo significado no se toma del DEM.

FUENTE: elaboración propia.

Tabla 20. Palabras de Preseea-Puebla relacionadas con alguna entrada del DEM: adverbios terminados en -mente

	Preseea-Puebla	DEM
1.	aparecidamente	aparecido
2.	apuradamente	apurado
3.	calmadamente	calmado
4.	clasistamente	clasista
5.	coincidentemente	coincidente
6.	coloquialmente	coloquial
7.	diferentemente	diferente
8.	gastronómicamente	gastronomía
9.	generacionalmente	generacional
10.	impresionantemente	impresionante
11.	incongruentemente	incongruente
12.	inteligentemente	inteligente
13.	laboralmente	laboral
14.	mayoritariamente	mayoritario
15.	médicamente	médico
16.	módicamente	módico
17.	obscenamente	obsceno
18.	pragmáticamente	pragmático
19.	psicológicamente	psicológico
20.	visualmente	visual

FUENTE: elaboración propia.

Tabla 21. Palabras de Preseea-Puebla relacionadas con alguna entrada del DEM: parasíntesis

		Preseea-Puebla	DEM	Categoría
a- -ar	1.	achipotar	chipote	vbo
	2.	apalear	palear	vbo
	3.	apanicar	pánico	vbo
	4.	aperrar	perro	vbo
a- -ar → -ado	5.	agringado	gringo	adj
	6.	apantallado	pantalla	adj
	7.	apoblanado	poblano	adj
des- -ar	8.	deschavetar	chaveta	vbo
	9.	deschongarse	chongo	vbo
	10.	descontextuar	contexto	vbo
	11.	desmallar	malla	vbo
	12.	desnucar	nuca	vbo
	13.	emberrinchar	berrinche	vbo
	14.	embolar	bola	vbo
	15.	emperezar	pereza	vbo
16.	emputar	puto	vbo	
en-, em- -ar	17.	encamar	cama	vbo
	18.	encañonar	cañón	vbo
	19.	encartonar	cartón	vbo
	20.	enchinar	chino	vbo
	21.	enfiestar	fiesta	vbo
	22.	engruesar	grueso	vbo
	23.	enrielar	riel	vbo
	24.	ensalivar	saliva	vbo
	25.	ensobretar	sobre	vbo
	26.	entambar	tambo	vbo
	27.	entiesar	tieso	vbo
	28.	entronar	trono	vbo
29.	enviciar	vicio	vbo	
en- -ar → -ado	30.	enlelado	lelo	adj
	31.	enlodado	lodo	adj
	32.	ensebado	sebo	adj
	33.	entaconado	tacón	adj
	34.	entamalado	tamal	adj
	35.	entarimado	tarima	adj y sust

FUENTE: elaboración propia.

Tabla 22. Palabras de Preseea-Puebla relacionadas con alguna entrada del DEM: composición

	Preseea-Puebla	DEM	Categoría
1.	asaltacunas	asaltar, cuna	sust
2.	audiolibro	audio, libro	sust
3.	cortabolsas	cortar, bolsa	sust
4.	cortagallos	cortar, gallo	sust
5.	cubrebocas	cubrir, boca	sust
6.	fisicomatemático	física, matemático	sust
7.	jugoterapia	jugo, terapia	sust
8.	magnetoterapia	magneto, terapia	sust
9.	malhablado	mal, hablar	adj
10.	malhecho	mal, hacer	adj
11.	malpasar	mal, pasar	vbo
12.	malviajado	mal, viajar	adj
13.	milhojas	mil, hoja	sust
14.	montacarguista	montar, carga	sust
15.	motosierra	motor, sierra	sust
16.	mototaxista	moto, taxista	sust
17.	pelapapas	pelar, papa	sust
18.	picahielos	picar, hielo	sust
19.	portabultos	portar, bulto	sust
20.	sacagolpes	sacar, golpe	sust
21.	videojuego	video, juego	sust
22.	videojugador	video, jugador	sust

FUENTE: elaboración propia.

VOCABULARIO DE PRESEEA-PUEBLA SIN DOCUMENTAR
EN EL LEMARIO DEL DEM

Para concluir la descripción del leuario de Preseea-Puebla, nos falta dar cuenta de 312 palabras que no hallamos ni directa ni indirectamente en el DEM y que son parte fundamental de la riqueza léxica que el corpus ofrece. En la lista siguiente encontramos candidatos a neologismos y también candidatos a ser palabras del español de Puebla, especialmente aquellas que nombran la comida local.

Listado de palabras de Preseea-Puebla

aerobics	campante	crush
after	canilla	cuatrimestre
albanene	canola	cuenda
alzheimer	cantabar	cuentahabiente
amarillista	carnaza	cúter
aparcar	carriola	chacharear
arpilla	celibato	chainazo
asturiano	cháchara	champurrado
atipujado	cianótico	chanfle
atipujar	cienciología	changarro
autismo	cintillo	charpe
ayocote	cizalla	chemo
background	clam	chichicaxtle
bisnes o business	clamato	chichón
boli	clonazepam	chilate
bonsai	cocker	chileatole
botarga	coda	chilposo
boxer	combi	chilpozotle
braguero	congolio /	chimisclán
bullying	congoleum	chinamite
búngalo	conurbado	chipileño
bye	cool	chitón
calesa	cover	chiviar
camorra	croqueta	chocomilk

choquilla	encarrar	guarrez
chorear	endorfina	guaxmole
choro	engargolar	gym
chorrillo	epifanía	hapkido
chucho	eslinga	happy
chunche	espray	hawaiano
chupavelas	estrobo	heavy
chupirul	factoraje	herbolaria
chusma	fan	hi
chuza	fashion	hipster
daddy	fax	homoparental
dark	feeling	house
darqueto / darketo	fisioterapia	huapilla
dead	fiu	huehue o huéhuere
decálogo	fluoxetina	inbox
del_tingo_al_tango	fondant	jacuzzi
demo	forever	jagüete
derrapón	forja	jamaiquino
desgarriate	fragor	jito
displayar	free	jobito
developer	friend	karaoke
dianética	fufis	karma
difuminar	galantina	lambada
discapacitado	garapiñado	larín
dóberman	garigoleado	led
dojo	garlopa	lencho
dopado	gaznate	light
école	gilipollas	linguililinguili
económico-	gioser	logo
administrativo	gratificante	lumbociatalgia
ecualizar	gratificar	macuil
embabucar	greco	maí
embustero	guácala	mainstream
empatía	guácatelas	man
empático	guáchatelas	marquetería
encarrado	guamazo	máster

matachín	pitufu	schnauzer
mazacote	play	serigrafía
mazapán	please	seudonombre
men	posdesfile	shopping
michelada	privanza	shuffle
mirrey	punching_bag	sin_tapujos
molonquear	punchis_punchis	six
molote	queer	sleeping
money	querubín	soterrado
motate	quiz	spa
muégano	rambla	speech
nachos	rapel	speechless
networking	rave	stand by
nini	recalcitrar	stop
novopan	reggaeton /	strudel
octaedro	reguetón	suaje
ocumo /	reguetonero	sumerio
ocume	relax	supertazón
over	repellar	table_dance
pachtli /	reposit	tachún-tachún
pastle	resistol	taekwondista
pailería	respingar	taekwondo
pailero	retinopatía	tagger
palanqueta	right	talavera
palma camedor	rinopatía	talmud
pamba	rocker	tartán
pancho	roquero	tazo
pandroso	rosticería	teibolera
papanata	rostizado	termo
papanicolau /	rostizar	tesiteca
papanicolaou	rugby	tibetano
party	sacalepuntas	tijuanense
paspartú	saga	timbiriche
paste	salinista	tinga
piercing	santillanista	tip
pinscher	sauna	tirito

tlalitos	tupper	wannabe
tlayoyo	tutú	wc
tlayuda	twist	world_music
top	underground	xonacate
top secret	unicel	yolixpa
toto	unisex	zambar(se)
touch	vendepatria	zancón
tozudo	viene-viene	zapear
triatlón	vírgula	
tugurio	vocho	

Este listado final, aunque cuantitativamente reducido (un poco menos de 3% del leuario de Preseea-Puebla), es rico en fenómenos léxicos de estudio, y muchas de las palabras que lo integran son candidatas para su inclusión en el leuario de un diccionario. Para esto último se requiere hacer antes un estudio lexicológico y su documentación exhaustiva.

Habremos de destacar, por una parte, que encontramos algunos préstamos y palabras de acuñación reciente, por ejemplo, *bullying* ‘abuso y maltrato persistente de una persona más fuerte o más poderosa a otra’, *fax* ‘aparato por el cual se emite y se recibe por la línea telefónica un documento escrito’, *michelada* ‘bebida que se prepara con jugo de limón y cerveza en un vaso escarchado con limón y sal, puede llevar algún tipo de salsa o chile en polvo’, *networking* ‘redes sociales para vincularse profesionalmente con colegas y clientes potenciales’, *nini* ‘joven que ni estudia ni trabaja’, *piercing* ‘arete, arillo u otro objeto de joyería que se introduce en una perforación para adornar el cuerpo de una persona’.

Por otra parte, hemos registrado palabras que gozan de gran arraigo en el español de México y son parte importante de la cultura mexicana. Sirvan de ejemplo *champurrado* ‘atole de agua, masa de maíz, chocolate y piloncillo, que puede llevar pinole y canela’, *charpe* ‘pieza, generalmente de madera, en forma de Y, de cuyos extremos se amarra una tira elástica con que se lanzan proyectiles pequeños’, *chichón* ‘abultamiento de la piel que se produce por un golpe en la cabeza’, *chimisclán* ‘bizcocho sin azúcar, con forma de rombo, similar al cocol, aunque sin barniz de huevo ni ajonjolí espolvoreado’, *choquilla* ‘olor desagradable que pueden tener algunos alimentos como el huevo si no se comen inmediatamente

después de guisados y el de las cosas cuando se lavan y secan mal', *garapiñado* 'tratándose de frutos secos como cacahuates y nueces, que se carameliza y adquiere una consistencia crujiente', *herbolaria* 'disciplina que, basada en conocimientos tradicionales, usa productos de origen vegetal para la prevención y curación de enfermedades', *mazacote* 'masa o mezcla espesa y pegajosa', *paste* 'empanada de harina de trigo, horneada, rellena de diversos alimentos', *sacalepuntas* 'persona engreída y buscapleitos', *rostizado* 'tratándose de carne, que se cuece directamente al calor de una hoguera ensartada a una espada giratoria' y *timbiriche* 'juego de estrategia en el que los oponentes deben trazar cuadros sobre una misma hoja donde se han marcado previamente con puntos una cuadrícula; en turnos consecutivos los jugadores unen horizontal o verticalmente dos puntos y quien logra cerrar el cuadrado gana el punto y quien acumula más puntos gana el juego'.

COMENTARIOS FINALES

A partir de este primer acercamiento cualitativo al léxico de Preseea-Puebla podemos confirmar la segunda de nuestras hipótesis: por tratarse de un corpus oral, hemos podido documentar contextos de uso de unidades que difícilmente aparecen en otro tipo de corpus, por ejemplo, juegos de palabras, creaciones morfológicas momentáneas, cruces léxicos, abundantes derivados, dobletes y variación morfológica. Nos parece importante resaltar que la importancia del registro de estas unidades es su presencia misma en contextos de uso real, sin importar el número de veces que se registre la misma unidad.

Para concluir este primer acercamiento, clasificaremos algunas de las palabras tratadas en grupos que esbozan futuras líneas de estudio lexicológico, terminológico y lexicográfico:

- 1) Nombres de bebidas, alimentos y platillos tradicionales de la gastronomía mexicana, varios de los cuales corresponden a indigenismos: *ayocote*, *champurrado*, *chilate*, *chileatole*, *chilposo*, *chilpozontle*, *chimislán*, *guaxmole*, *molote*, *muégano*, *palanqueta*,

paste, tinga, tlalitos, tlayoyo, tlayuda, xonacate y *yolixpa*. Ejemplificaremos la complejidad del estudio de estas voces en el capítulo siguiente.

- 2) Palabras de la lengua culta que corresponden a posibles tecnicismos y términos: *background, cianótico, fisioterapia, homoparental, lumbociatalgia, retinopatía* y *rinopatía*.
- 3) Palabras relacionadas con distintos deportes: *aerobics, chuza, dojo, hapkido, punching bag, rapel, rugby, taekwondo, triatlón*.
- 4) Palabras de actividades cotidianas; por ejemplo, papelería para la elaboración de tareas escolares infantiles: (*papel*) *albanene, cúter, gioser, paspartú* y *resistol* ‘pegamento’.
- 5) Préstamos asociados con la música y las fiestas: *after, karaoke, lambada, rave, reggaetón, table dance, twist* y *tugurio*.
- 6) Préstamos cuyo uso parece una interferencia léxica del inglés, pues conservan el significado de la lengua de origen: *Hí!* ‘saludo’, *right!* ‘de acuerdo’, *speechless* ‘sin palabras’ (en este último se crea un híbrido en la expresión *quedarse speechless*, o *y yo, speechless*).
- 7) Voces relacionadas con la personalidad, los gustos y las actividades personales, en su mayoría anglicismos: *cool, dark, heavy, hípster, mirrey, relax* y *tagger*.
- 8) Interjecciones y onomatopeyas: *jguácatelas!* o *jguáchatelas!*, *linguililinguili* ‘platicar’ o *punchis punchis* ‘música electrónica que se percibe sin variaciones’.

5. NOTAS LÉXICAS DEL ESPAÑOL DE PUEBLA

Para concluir este primer acercamiento al léxico registrado en Preseea-Puebla, vamos a centrarnos en la descripción y la documentación de algunas voces propias de la cultura local cuya selección parte de nuestro juicio como hablantes de esta variedad, es decir, se trata de unidades que hemos considerado emblemáticas sin que esto las catalogue como “poblanismos”, pues el estudio dialectal requiere otro método de trabajo. No obstante, algunas de ellas son denominaciones de ámbito objetivo que aparecen, principalmente, en la última parte de la entrevista, cuando se habla de la ciudad y sus costumbres.

Con el mismo objetivo fundamental del libro —explicitar el método de trabajo con el fin de que otros equipos puedan replicarlo—, en este capítulo se aborda la documentación, o sea, la búsqueda de estas palabras en fuentes adicionales al corpus a fin de corroborar su significado y fijación dentro de la comunidad de habla.

Presentaremos el análisis de estas palabras organizado en tres grupos: 1) nombres de prácticas o conocimiento de la cultura local (*huehue*, *mosaicos*, *murales* y *pitufo*); 2) palabras de la gastronomía tradicional (*cemita*, *chalupa*, *chileatole*, *molote* y *tlalitos*), y 3) voces con variación semántica o sintáctica (*acusarse* y *arrimarse*).

Además de presentar la documentación en fuentes primarias (diccionarios) y secundarias (periódicos o fuentes especializadas), ejemplificamos con las concordancias de uso de Preseea-Puebla. En nuestra investigación, éste es el primer acercamiento al estudio del significado. Se trata, pues, de la antesala para la redacción del artículo lexicográfico.

VOCES DE PRÁCTICAS
SOCIOCULTURALES

Huehue o huéhuere

Al igual que en muchas regiones del país, la celebración del carnaval en Puebla muestra el sincretismo entre los ritos religiosos católicos y las festividades paganas con elementos indígenas. En la zona de Huejotzingo, Puebla, una de las regiones con mayor influencia indígena, el carnaval tiene una peculiaridad: el inicio de la Cuaresma, fiesta ritual católica, es el escenario para la representación de tres argumentos históricos importantes para la región. En él se representa la intervención francesa del siglo XIX, “el rapto de la hija del corregidor” y un casamiento indígena tradicional. Como parte de esta teatralidad propia de la religiosidad popular (Azor, 2006), los batallones que dan vida a los zacapoaxtlas y otros grupos indígenas, a los zuavos e incluso al batallón turco se mezclan con personajes vestidos de diablos —que, paralelamente a la representación, recorren las calles y les hacen maldades a los espectadores— y con personajes de rostros cubiertos por máscaras de ancianos de barba blanca y larga. Tanto en el carnaval de Huejotzingo como en los de otros municipios, barrios, juntas auxiliares y colonias de Puebla, la modernidad permite observar la manera en la que la caracterización teatral convive con una fiesta moderna de disfraces (Licona y Pérez, 2018: 13). Todos estos danzantes disfrazados reciben el nombre de *huehues*.

Carlos Montemayor, en su *Diccionario del náhuatl en el español de México*, define *huehues* y *huehuenches* como los danzantes que se disfrazan y actúan como viejitos en un baile tradicional. La palabra tiene su etimología precisamente en el náhuatl *huehue* ‘viejo’. Así pues, el nombre de los danzantes de un baile específico se generalizó para los danzantes del carnaval sin importar el personaje que representen.

En Preseaa-Puebla hemos documentado 11 concordancias de *huehues*, 2 de *huéhuere* y 9 de *huéhueres*:

(16) I: ya son mamarrachos éstos porque eso ya no y eso que hacen como negocio / y antes por un decir había música que realmente es específica para **huéhuere** / y ahora no / ahora bailan que La Chona y que no sé qué

[PUEB-H31-078]

(17) I: y no hay como pongamos / no sé en San Felipe en San Baltazar donde eh / allí hay **huéhueres** de a de veras / le enseñan a sus hijos desde chiquitos / tres cuatro años ya andan con su / haciendo sus pininos

[PUEB-H31-078]

(18) I: es del barrio de La Luz y sus hijos forman parte de la cuadrilla de / de los **huehues** de La Luz y este / tiene una máscara / que ojalá te la pudiera enseñar porque la tiene ahí súper guardada que tiene creo que ciento cincuenta años

[PUEB-M13-034]

Mosaicos y murales

Dentro de las prácticas socioculturales características de la ciudad de Puebla destacan los *mosaicos* y los *murales* vivientes, una actividad escolar en la que participaron miles de alumnos de educación pública de las décadas de 1960 a 1990. Ambos, *mosaicos* y *murales*, nombran figuras de gran tamaño que se aprecian a la distancia y cuya composición se logra gracias al trabajo coordinado de un grupo de personas. En el caso de los mosaicos, se eleva un trozo de tela de color para formar la imagen previamente diseñada, mientras que, en los murales, el material mostrado es cartón. Los mosaicos también pueden formarse al adoptarse cierta postura corporal, pues cada participante usa un uniforme integrado por un chaleco abierto por los lados cuyas cuatro caras son de diferente color, además de un par de mangas de doble vista. Un ejemplo mundialmente conocido de murales se dio en los Juegos Olímpicos de Moscú, 1980, donde la imagen de Misha, la mascota oficial, se formó mediante esta técnica en las ceremonias de inauguración y clausura, y, en esta

última, gracias al movimiento coordinado, derramó una lágrima de despedida.

Estas palabras aparecen en las anécdotas adolescentes de los entrevistados, como puede verse en los ejemplos (19) a (21):

(19) I: en segundo de preparatoria eh / fui abanderada / en el desfile del 5 de mayo / sí sí desfilé participé en muchísimas tablas rítmicas / en **murales** ¡por supuesto!

[PUEB-M23-070]

(20) I: en **mosaicos** y **murales** / no / eso bueno (risa) ¡lo vuelvo a hacer! (risa) / no sí me fascinó cuando fuimos a / no sé si te tocó también yo creo que sí / o bueno quién sabe eh mm hicimos tantas cosas / este cuando fuimos al Colegio Militar

[PUEB-H23-049]

(21) I: me divertí mucho / y los **murales** fascinantes ¿no? es el ver para arriba ver para debajo / de izquierda a derecha y moverte y / el color y luego ya lo ponías ¡ay! de arriba para abajo con ruido / sin ruido / grito / no /eran divertidísimos

[PUEB-M23-072]

Estas palabras no tienen una marca de uso restringido: se trata de la denominación de la actividad, por lo que su uso se documenta también en la prensa, como podemos leer en (22) y (23):

(22) Dijo que su padre juntaba la parte del desarrollo educativo con lo deportivo. «Todas las actividades extraescolares, los **mosaicos**, las tablas rítmicas, las escoltas, eso te da una formación integral que quizá los libros no te dan» (Sánchez, 2018).

(23) En el aniversario 51 del Centro Escolar Niños Héroes de Chapultepec se llevará a cabo una ceremonia con **mosaicos**, banda de guerra y coreografías (Martínez, 2008).

Durante al menos tres décadas, la presentación de estas figuras en ceremonias cívicas y militares formó parte central de la vida estudiantil de la ciudad y dio origen a una jerga asociada con estas prácticas; por ejemplo, *tarjetero* ‘joven que entrega y recoge una bolsa con el material para hacer murales y mosaicos, incluida la tarjeta con las instrucciones precisas para que cada participante forme las figuras prediseñadas’. De manera festiva, a estos jóvenes se les conocía como *tarjegatos*, cruce léxico entre *tarjeta* o *tarjetero* y *gato* ‘criado’. Como parte del mismo trabajo coordinado, los estudiantes que participaban en murales y mosaicos, generalmente sólo mujeres, aprendían coreografías; no obstante, para facilitar la tarea había *monitores* ‘personas que, mediante el ejemplo, les muestran a las demás la coreografía que deben seguir’.

Es importante recalcar que estas palabras, si bien son del español de Puebla, forman parte de lo que, siguiendo a Eckert (2018: 110), llamaremos el estilo de un grupo (*group style*), esto es, unidades cuyo uso denota la identificación y pertenencia a un grupo social específico: adultos (generación 2 en Preseca-Puebla), egresados de escuela pública. Esto se confirma cuando notamos que los ejemplos de (22) y (23) son publicaciones periodísticas en las que se habla de la escuela donde se diseñaban y coordinaban estas actividades. Más aún, podemos encontrar referencia al sobrenombre burlón que recibían sus estudiantes (*cirqueros*) por participar en esas actividades, y que se volvió motivo de orgullo para ellos.¹ En (24) tenemos un ejemplo explícito de este reconocimiento identitario:

(24) Entiendo que Felipe Calderón tenga una guerra personal contra Mario Marín. Incluso que en tiempos electorales no lo tolere. Razones hay muchas. Pero, y de eso ¿qué culpa tienen los estudiantes poblanos que sufrieron su desplante? En especial, todos los que pertenecemos a la comunidad del Centro Escolar Niños Héroes de Chapultepec. Sí, los *cirqueros* estamos indignados. Rojos de ira casi (Rueda, 2010).²

¹ Fenómeno similar al ya descrito para *pipope* y *chilango*.

² El desfile cívico militar conmemorativo a la Batalla de Puebla del 5 de mayo de 1862 es otra tradición de fuerte arraigo en la ciudad. En 2010 el entonces presidente de la república se retiró de la tribuna antes de que terminaran de pasar los contingentes.

Pitufo

Analicemos ahora el uso de *pitufo*. Los pitufos son personajes creados por el dibujante belga Peyo para la historieta *La Flûte à Six Schtroumpfs* (*La flauta de los seis pitufos*), que tuvo distintas adaptaciones y se popularizó en México por la caricatura del estudio de animación estadounidense Hanna-Barbera, producida entre 1981 y 1990 (*The Smurfs* [*Los Pitufos*]). Estos personajes dieron nombre a una famosa banda delictiva en el barrio de San Antonio.

San Antonio es uno de los asentamientos indígenas fundacionales de la ciudad. Ubicado aproximadamente a un kilómetro y medio del zócalo de la ciudad (la plaza central), su constitución como centro urbano implica un complejo entramado social de migración, marginación, delincuencia y densas redes sociales en sus vecindades. Durante el siglo XIX y las primeras décadas del siglo XX fungió como zona roja. En la década de 1980 la prostitución, el narcomenudeo y la vida en las vecindades formaron el marco en el que surgió la banda de Los Pitufos, integrada por niños delincuentes, cuyos actos llenaban la sección policiaca de la prensa local³ con notas como la reproducida en (25), y que hicieron impenetrable el barrio ubicado en la zona de la 22 Poniente.

(25) Obrero asaltado por banda de “Pitufos”

Sin dinero, sin juguetes para los niños de su familia y con fuertes golpes, llegó a su domicilio un obrero la víspera de los Reyes Magos, luego de ser agredido en la 22 Poniente, entre 5 y 7 Norte por cuatro miembros de la banda de “*Los Pitufos*”, uno de los cuales no tiene más de 10 años.

Con lesiones de cuchillo en los costados y brazos, además de puños y puntapiés, permanece internado en observación en la Cruz Roja, el obrero de 21 años [...] señalando que luego de hacer las compras de

³ Tras una búsqueda en la hemeroteca de *El Sol de Puebla*, encontramos la primera mención a esta banda el 2 de enero de 1980, y entre este año y 1984 contamos 901 notas publicadas cuyo título incluye “Los Pitufos” o “pitufos(s)”, y que daban cuenta de delitos, enfrentamientos o detenciones de integrantes de esta banda delictiva.

rigor por ser Día de los Reyes, se dirigió a su domicilio por el barrio de San Antonio [...] El afectado aseguró al declarar, que cuatro sujetos le salieron al paso del pórtico de una vecindad, los cuales no son mayores de 10 años, uno de ellos ni de 10, según dijo el herido, quienes portaban cuchillos, amagando e injuriando al transeúnte (Vázquez, 1984).

El nombre de la banda se volvió tan representativo de la zona que no sólo se aplicó a sus integrantes, sino que también se extendió a los habitantes del barrio. Nos parece que los ejemplos (26) a (28) están en el límite entre el uso de la palabra como sustantivo propio o común:

(26) I: y anteriormente no era así hija / esa Puebla / de aquellos barrios que a pesar de que eran también tan pesados / había decencia / porque te digo que los chamacos de la 22 / esos putos los manejaba el puto gobierno / a los güeyes de la 22 / no me acuerdo quién estaba de gobernador / y ése era el que les daba este luz verde a todos los pinches **pitufos** / porque era el que ganaba más / y deja me voy a acordar y verás / y y ese güey fue el que los empezó a indilgar a que robaran más culero / por eso ya lo empezaron a hacer en el día / en serio

[PUEB-M21-056]

(27) I: de ahí que yo recuerde / de los barrios barrios barrios más / más peligrosos / más problemático era el / barrio de San Antonio / los famosos **pitufos** / muy muy / mucha delincuencia / mucha delincuencia ahí era / personas / la / a lo lejos se veía luego luego su fachita

[PUEB-H21-039]

(28) te digo que éramos de la vecindad éramos no sé / cuatro cinco seis no recuerdo bien / pero nos juntábamos para divertirnos no para causar destrozos desmanes / o para hacer algo malo / o sea nos juntábamos para / para tener una diversión sana / pero los **pitufos** / este / son del barrio de de San Antonio pero no andaban ahí por Analco haciendo fechorías

[PUEB-H22-046]

Pitufo es, pues, otra palabra de identidad de grupo, esta vez, de una comunidad delictiva cuya irradiación de poder se encuentra perfectamente ubicada en la geografía y la organización de la ciudad.

Como es esperable, porque se trata de prácticas muy locales, ninguna de estas tres palabras se encuentra registrada en los diccionarios consultados. Si se piensa en un diccionario de lengua general del español de Puebla, *mosaicos* y *murales* debería definirse dentro de la nomenclatura; en el caso de *pitufo* la decisión es más compleja por su transparente relación con el sustantivo propio y su uso en el límite entre éste y el nombre común. Por otra parte, *huehue* y *huehuere* podrían incluirse tanto en un diccionario integral como en uno diferencial, porque nombran una realidad cultural importante para el conocimiento de la cultura y el español de México.

VOCES DE LA GASTRONOMÍA POBLANA

En este segundo apartado se presentan palabras que pertenecen a la gastronomía local. Es esperable que formen parte del patrimonio del español de México y que se documenten en los diccionarios consultados. El interés por describirlas radica en la variación que presentan, es decir, hay tantas diferencias en el campo culinario que a un mismo nombre puede corresponder una enorme variedad de preparaciones, y cada una será representativa de la región donde se hace; incluso puede llegar a generarse una disputa identitaria sobre cuál es “la auténtica”, lo que muestra la relevancia de los alimentos para las comunidades.

Cemita

Comencemos por analizar la palabra *cemita*, que se presenta en los ejemplos (29) y (30).

- (29) I: el que viene a Puebla y no come una **cemita** ¿qué? / las **cemitas** son únicas en el mundo / de Puebla o sea las **cemitas** / y son exquisitas / o sea las **cemitas** realmente dices no / en ningún lugar del mundo hacen esas cosas / más que en Puebla

[PUEB-H22-043]

(30) E: ¿a ti qué te gusta comer cuando vas al estadio?

I: una **cemita** nomás

[PUEB-H11-003]

La *cemita* nombra un pan blanco, elaborado originalmente con acemite, que en la actualidad se hace, sobre todo, con harina de trigo. Es un pan salado, boleado (como una esfera partida por la mitad), hueco, de acabado dorado y crujiente, espolvoreado con ajonjolí. Su origen se encuentra documentado desde el siglo XVI como tributo a la Corona (Palacios, 2020a: 90). También se llama *cemita* cuando se rellena o prepara con ingredientes similares a los de cualquier torta compuesta; por ejemplo, queso, milanesa de res o de cerdo, carne enchilada o carnitas, cebolla, aguacate, pápalo y chile.

José Moreno de Alba (2003: sp) documenta el cambio lingüístico de la palabra y da cuenta incluso de la opacidad por la que se reinterpreta como un falso diminutivo:

En la evolución de la palabra se pueden observar interesantes fenómenos no sólo fonéticos sino también morfológicos. En primer lugar, hay un cambio de género en la palabra, de masculino a femenino, mediante la permutación de la *-e* final por *-a* (*el acemite*: *la acemita*). En seguida, seguramente por influencia de la *a* del artículo femenino, se elimina la *a-* inicial (*la acemita*: *la cemita*). Finalmente, el cambio más curioso: después de las modificaciones operadas en la palabra, ésta apareció a los ojos (y, sobre todo, a los oídos) de los hablantes, por el fenómeno conocido como *etimología popular*, como un *diminutivo*, y equivocadamente se interpretó la terminación *-ita* (procedente de *-ite*) como sufijo análogo al que se da en voces como *cas-ita*. Ello permite explicar que delseudiminutivo *cemita* se pase a *cema*, suprimiendo la terminación.

En (31) podemos atestiguar el uso de *cema* al que se refiere Moreno de Alba:

(31) I: ya no hay algo tan / tan / tan arraigado así como del pueblo de que te / no sé / inviten el mole o cualquier cosa / acá te vas

a comprar mole o no te van a invitar mole / te vas a tener que comprar / no sé / una **cema** / un pedazo de pizza este / unas chalupas / no sé cosas así / botanas / cosas así / pero pues ya es ya / ya con un lucro ¿no?

[PUEB-H12-007]

En cuanto al registro lexicográfico, encontramos su primera aparición en los diccionarios académicos como acepción 2 de la entrada *cemita* en el *Diccionario histórico de la lengua española* (RAE H, 1936), con la marca de uso compartida por Argentina, Colombia, Guatemala, Honduras y México, como se ve en la imagen 8.

En el *Diccionario de mexicanismos* (DMéx, 2010), la primera acepción registra “pan de acemite”, mientras que el DEM define otro tipo de pan con el mismo nombre: “cemita s f Pan dulce hecho de salvado o de harina oscura, espolvoreado con harina por encima; semita”, el cual puede encontrarse en el norte y el occidente de México, por ejemplo, en Nuevo León (semita de anís) o en Sonora (semita de cuajada). El *Diccionario enciclopédico de la gastronomía mexicana* (DEGM) reporta, además de la *cemita* poblana, al menos otros seis panes tradicionales con el mismo nombre y con preparaciones diversas, como la de Coahuila, fermentada con pulque (cemita de pulque); la *cemita* de tuba de Colima, fermentada con esta bebida tradicional del mismo nombre, o las de Zacatecas, “pan de harina de trigo, leche, canela y azúcar; suelen estar decoradas con acitrón, coco, nuez o pasitas y espolvoreadas con azúcar; es un pan muy tradicional del estado, que antiguamente se vendía en todas sus panaderías. La masa puede contener también coco o nuez como parte del pan o como adorno”.

Chalupa

Algo similar en cuanto a la diversidad de referentes nombrados ocurre con la palabra *chalupa*, uno de los alimentos más representativos de la comida poblana. Cuando Raúl Ávila describe la complejidad de los antojitos como parte de la variedad dialectal y cultural de México, explica que

Imagen 8. Artículo lexicográfico *cemita*

RAE H 1936

CÉMITA. (Aféresis de *acemita*.) f. *Argent.* Cemite, harina gruesa de trigo. Cfr. Garzón, *Dicc. Argent.*, s. v.

2. *Argent., Colomb., Guat., Hondur. y Méj.* Acemita, pan hecho de *acemite*. ¶ «Cuando la infeliz quería pedir su limosna decía invariablemente: pues vóyeme yo, frase que repetía hasta que algún harapo caído en desuso..., alguna *cemita* redonda y sabrosa, una vela..., unos zapatos viejos... acudían a hacer cierto... el sacramental vóyeme yo.» Sarmiento, *Recuerdos de provincia*, ed. 1896, p. 150. Cfr. Cuervo, *Apunt.*, § 817; Garzón, *Dicc. Argent.*; Membreño, *Hondur.*; Icazbalceta, *Dicc. Mexic.*, s. v.

FUENTE: Nuevo Tesoro Lexicográfico de la Lengua Española (RAE H, 1936).

las chalupas se hacen con masa a la que, antes de freírla, se le da forma cóncava y ligeramente alargada. Son también duras, como los totopos, aunque algunas tienen consistencia blanda, pues no se fríen. Su forma permite que se rellenen con frijoles refritos, carne de pollo deshebrada y otros guisos. Al final se rematan con salsa, queso desmoronado, lechuga y cebolla picadas (2008: 485).

Por su parte, el *DEGM* explica, bajo la entrada *chalupa*:

Antojito de masa de maíz en forma de pequeña tortilla gruesa, con o sin orilla, en forma de óvalo o de barquito, más concretamente de *chalupa*, de donde proviene su nombre. Generalmente se fríe y luego se rellena con diversos alimentos. El tamaño, la forma y los rellenos cambian según la región del país.

El artículo continúa describiendo las chalupas de Puebla, las chiapanecas, las de Nuevo León y las de Chihuahua. Ambas obras nos permiten comprender la multiplicidad de preparaciones que a lo largo del territorio nacional tiene este platillo, y la dificultad que conlleva su definición para un diccionario de lengua. Esto se confirma en las definiciones del *DEM* y el *DMéx*:

DEM: chalupa 2 Tortilla de maíz, un poco gruesa, generalmente ovalada y abarquillada, en cuyo interior se ponen frijoles, queso, hebras de carne, salsa, lechuga, etc: “¿No se te antojan unas chalupitas poblanas?”.

DMéx: chalupa. F. Tortilla de maíz pequeña y en forma ovoide, que se cubre de frijol o habas, carne picada, queso, lechuga, salsa de chile, y otros aderezos.

Las *chalupas* poblanas se hacen con una tortilla de maíz pequeña, circular, muy delgada, sofrita en manteca de cerdo y bañada en salsa verde o roja, a la cual se le coloca cebolla picada y pocas y muy finas hebras de carne de cerdo. No admite otro ingrediente. Se trata de un platillo característico de la gastronomía poblana y puede encontrarse en distintos tipos de establecimientos: puestos callejeros, puestos de ferias y verbenas, o restaurantes de alta cocina mexicana.

(32) I: en mi punto de vista Puebla / puede ser un lugar turístico excelente ¿por qué? / por la gastronomía que tiene / o sea digo tiene una gastronomía / que a mi punto de vista es más rica / que la gastronomía yucateca y que la gastronomía oaxaqueña por ejemplo / que tienen que tiene ¡mucho! / nosotros pero si ellos tienen mucho nosotros tenemos más / nosotros / digo / conocemos / las **chalupas** son de excelencia poblana

[PUEB-H22-043]

(33) I: la feria ¡ah sí! / acá en la iglesia / sí / sí hay juegos / hay toritos / hay eh **chalupitas** / de todo / de todos los antojitos

[PUEB-M21-058]

Chileatole

El DGEM documenta que la preparación tradicional del *chileatole*, esto es, atole de chile, es parte de la cocina prehispánica. Define más de 15 variedades de platillos que se preparan en el territorio nacional y com-

parten nombre. Algunos de ellos son atoles de masa de maíz; otros, guisados con caldo de pollo; los hay preparados con carne de cerdo y con lengua de vaca. También, según la región, se usan chiles distintos; por ejemplo, chile cascabel, ancho, costeño, de árbol, chipotle, guajillo o serrano. Puede comerse como platillo principal, como sopa o como antojito. Este mismo diccionario nos dice: “En Puebla, especialmente en el valle de Puebla, se prepara con masa de maíz, granos de elote, chiles verdes o rojos, guías de calabaza y epazote, sin carne”.

El *Diccionario de mexicanismos* da dos acepciones que engloban esta variedad de platillos:

DMéx: chileatole. m. Bebida de atole y chile. || 2. m. Guiso de maíz, chile y carne de cerdo.

En Puebla, el *chileatole*, como indica el compuesto, es un atole de chile. Suele venderse por las noches en los mismos puestos donde se venden elotes preparados y esquites. El *chileatole* que conocemos es verde, se prepara con chile serrano, guías de calabaza y granos y rodajas gruesas de elote. En Preseaa-Puebla encontramos 12 concordancias de uso que lo refieren como parte de la comida tradicional de la ciudad y parte de las verbenas populares:

(34) I: por ejemplo acá de este la iglesia del Sagrado Corazón de Jesús / pues vienen todos los juegos / entonces ya vienen las vendimias / el **chileatole** / los elotes / las chalupas / gelatinas las chalupitas este / los panes de fiesta

[PUEB-H31-073]

(35) I: el **chileatole** también en las noches / el **chileatole** siempre ha sido así con con los elotes / con hoja santa / que salía verde verde / era chile verde no chile de polvo

[PUEB-H32-082]

(36) I: allí en la mera esquina / en la calle la señora vendía el **chileatole** muy sabroso / mjú / muy sabroso porque pues de toda la vida

[PUEB-M32-098]

Molote

También dentro de los platillos tradicionales, específicamente de los *antojitos* de la ciudad, hablaremos de los *molotes*, cuya tortilla, preparada con una mezcla de masa de maíz y harina de trigo, muy delgada y ovalada —con una longitud aproximada de 30 centímetros—, se dobla por la mitad y se rellena con muy diversos ingredientes como sesos, tinga, queso, papas y champiñones, y se fríe. Se pueden servir bañados en crema y salsa.

Los ingredientes de la masa (maíz y trigo) y el grosor de la tortilla (muy delgada) hacen que la consistencia (crujiente) y el sabor sean muy distintos a los de las empanadas y quesadillas fritas.

- (37) I: antes estaban en la esquina donde son los baños y ahí mi mamá me llevaba desde chiquita nos llevaba al baño y saliendo los **molotes** / pero mira qué **molotes** bien ricos / mis respetos / son los que he comido más sabrosos porque hasta mira se te desbaratan en la boca de lo doradito que están ve te los recomiendo / es mira / así angostito / es angostita la entrada y tiene unas mesitas también chiquitas ahí para si te quieres sentar a comer / sí / no este y y en Puebla hay mucha cosa de bueno / la cocina es muy buena

[PUEB-M32-099]

En la documentación lexicográfica, el *DMéx* define: “molote. m. Tortilla gruesa de maíz de pequeño tamaño, por lo general en forma ovalada rellena de frijoles u otro alimento”. En el *DEGM* se define como “antojito de masa fresca de maíz, muy parecido a una quesadilla frita; el relleno y la salsa que llevan dependen de la región. Los molotes siempre se venden en la calle o en las antojerías y por lo general se sirven por las noches”, y continúa con el detalle de la preparación de los molotes de Hidalgo, Oaxaca, los molotes de plátano de Juchitán, los de Puebla, San Luis Potosí, Tlaxcala, Xalapa, la zona norte del estado de Veracruz, y de Huatusco. Cada uno de ellos, muy diferente a los otros.

Nuevamente, estamos ante referentes diversos que comparten el mismo nombre y coinciden en ser antojitos o platillos tradicionales.

No obstante, el molote poblano no se refleja en la definición del diccionario de lengua general, pues ni es pequeño ni se rellena con frijol. Aunque sí se reconoce en la del diccionario de cocina mexicana.

Tlalitos

En la entrada *tlales* o *tlalitos*, el *Diccionario del náhuatl en el español de México* (Montemayor, 2007: 121) remite mediante la indicación “Véase” a “chalitos. De *tlalli*, tierra”, y bajo ésta define: “chalitos. m. pl. Residuos grasos de las frituras de carne de cerdo. En algunas regiones del país se llaman *tierritas*, *borusas*, *tlalitos* o *xalitos*. Diminutivo de *xalli*, arena. Véase *xalitos* y *jales*”.

Por su parte, el *DEGM* define:

chalitos Del náhuatl *xalli*, arena. Trozos pequeños de chicharrón muy dorados que se acostumbran consumir en el Distrito Federal en tacos con tortillas de maíz, aguacate y salsa. Se obtienen cuando se separa la piel para preparar chicharrón que se corta con algo de grasa y pedazos de carne; una vez retirada la grasa, la piel se pica, se fríe y se convierte en chalitos. Son ofrecidos por vendedores independientes dentro de una canasta o caja de cartón en los pasillos de los mercados populares. Los chalitos no se venden en las chicharronerías, ni en las carnicerías o puestos de carnitas.

En Preseaa-Puebla únicamente documentamos la forma *tlalitos* (no *tlales* ni *chalitos*). Culturalmente podemos decir que en la ciudad de Puebla sí se venden junto con el chicharrón y las carnitas, y la forma *chalitos* o *chales* no se usa. En el ejemplo (38) se narra, precisamente, una situación de contacto y variación léxica que lleva al extrañamiento de los participantes en la situación comunicativa:

- (38) I: habían dos chavos que venían de Morelos // ah pues ellos / dice “¿a dónde vamos?” los llevé al Barrio del Artista / los llevé a este / las cemitas del Carmen / y así se quedaron así “¿a poco nunca las habían visto?” / dice “no, no” dije “coman cemitas de El Carmen” / los llevé ahí / este / tiene otras cosas porque tam-

bién los llevé a una / unas gorditas que venden aquí en / qué es / la 8 / donde le echan un montón de carne / no sé qué cosa

E: [ajá]

I: [ah pues] por ahí / este / los llevé ahí / y / hablan diferente porque / dijo que le echara chales y la chava “¿qué?” / “que le echas / ¿puedes echar tantito chales?” y yo digo “qué ¿qué dice?” y eran **tlalitós** / de chicharrón / pero él le decía chales

[PUEB-M13-032]

En esta sección observamos que, si bien las palabras y sus referentes son patrimonio de la cultura culinaria nacional, la preparación específica varía por regiones, y que es precisamente esa tradición local la que carga de identidad y representatividad cultural a cada platillo y al uso del nombre en cada región.

Los diccionarios de lengua general requieren contar con información suficiente y confiable proporcionada por los estudios locales para documentarse. Los diccionarios diferenciales regionales requerirán construir definiciones culturales e incluso enciclopédicas para estas voces, mientras para los lexicógrafos de obras de lengua general el reto al definir estas palabras se halla en encontrar los aspectos semánticos relevantes y comunes para la definición. Para los hablantes, la valoración de la definición parte del detalle que permite reconocer su práctica o receta gastronómica específica, puesto que no sólo se trata de un asunto de significado, sino de identidad.⁴

Variación sintáctica y semántica

Para concluir este capítulo, analizaremos dos formas verbales que no corresponden a la realidad cultural, social o histórica de Puebla, sino que forman parte del español general y, sin embargo, presentan

⁴ Sirva de ejemplo complementario que en la actualidad resulta particularmente emblemática la discusión sobre si las quesadillas deben o no llevar queso, incluso se han difundido explicaciones presuntamente etimológicas sobre el origen náhuatl de la palabra para argumentar que no se trata de un derivado de queso. Esta disputa ha dado origen a un sinnúmero de memes y otras publicaciones en redes sociales donde se confronta la cultura capitalina con la del resto del país.

variación sintáctica o semántica en relación con la variedad estándar. En este apartado comenzaré con la documentación de los vocablos en los diccionarios para después ejemplificar su uso en el español de Puebla:

DLE: acusar. (Del lat. *accusāre*). 1. tr. Imputar a alguien algún delito, culpa, vicio o cualquier cosa vituperable. 2. tr. Denunciar, delatar. U. t. c. Prnl. 3. tr. Notar, tachar. 4. tr. Reconvenir, censurar, reprender. 5. tr. Manifestar, revelar, descubrir. 6. tr. Avisar, noticiar el recibo de cartas, oficios, etc. 7. tr. Dicho de una persona: En algunos juegos de naipes, manifestar en tiempo oportuno que tiene determinadas cartas con que por ley del juego se gana cierto número de tantos. 8. tr. Reflejar la contundencia y efectos de un golpe recibido. 9. tr. *Dep.* Dicho de un atleta o de un jugador: Mostrar inferioridad o falta de preparación física. 10. tr. *Der.* Exponer en juicio los cargos contra el acusado y las pruebas de ellos. 11. prnl. Dicho de una persona: Confesar, declarar sus culpas.

DEM: acusar v tr (Se conjuga como *amar*) 1 Denunciar a alguien como culpable de algo: “Lo *acusaron* de robo”, “Te voy a *acusar* con mi mamá” 2 Revelar algo, ponerlo de manifiesto: “Su palidez *acusa* una enfermedad”.

Ahora bien, el uso que queremos exponer es el pronominal *acusarse*, que ejemplificamos de (39) a (41):

(39) I: es la consentida de los profesores / es la que en el kinder era la ¿cómo se llama? la abanderada ¿no? / la que salía en todos los festivales una niña / muy mimada ¿no? a la que no... pues yo me acuerdo que / en sexto de primaria nosotros la molestábamos mucho porque decía “¡ay qué ñoña!” / o sea de eso que la veías y decías “¡qué horror!” ¿no? y ella iba y **se *acusaba*** / entonces es una / era una niña muy mimada por la escuela / eh que todo mundo la conocía por ser la hija de la maestra

[PUEB-M13-033]

(40)I: fue mi maestra / era buena maestra pero era muy así como que le gustaba pegar / y pues como uno no **se acusa** con los papás pues nunca reclaman porque sí había mucha violencia en ese tiempo

[PUEB-M22-062]

(41) I: a mí una pinche maestra me puso mi cachetadón en tercer año / me cae de madre / todo por decirle a mi amiga “mamadora” / que va y que **se acusa** y que se para la pinche vieja y que me pone un puto cachetadón / después hija de tu pinche madre que le robo todo lo que pude

[PUEB-M21-056]

Los ejemplos parecen vincularse con el significado definido en la segunda acepción del *DLE* y la primera del *DEM*, aunque si consideramos su forma pronominal, podríamos vincularlos con la acepción 11 del *DLE* (“Dicho de una persona: Confesar, declarar sus culpas”). En cualquiera de estos casos los ejemplos resultan poco claros. ¿Por qué la hija de la maestra se acusa a sí misma debido a que otros la molestan? (39) ¿Por qué los niños se acusarían a sí mismos con sus padres al recibir golpes de la maestra? (40). Y ¿por qué la amiga insultada se acusaría a sí misma con la maestra? (41) ¿Cuál es la culpa que confiesan estos tres hablantes?

Para entender estos usos, necesitamos saber que en el español de Puebla existe la posibilidad de que la forma pronominal cancele su valor reflexivo, y aunque se mantenga la estructura y se haga coincidir la persona con el agente, el valor del pronombre es el de un bloqueador argumental del objeto directo. Así, el uso coincide con el significado definido por el *DEM*, aunque tiene un matiz semántico más específico: quien *se acusa* necesariamente lo hace con una autoridad que puede ejercer justicia (lo que lo diferencia, por ejemplo, de *denunciar*, *rajar* o *chivatear*). La cancelación de la reflexividad y del argumento objeto hace que en el español de Puebla esta forma corresponda a la categoría gramatical del verbo *intransitivo*.

En cuanto al verbo *arrimar*, la documentación lexicográfica dice:

DLE: arrimar. (De or. inc.). 1. tr. Acercar o poner algo junto a otra cosa. U. t. c. prnl. [...] 5. prnl. Apoyarse o estribar sobre algo, como para descansar o sostenerse. 6. prnl. Agregarse, juntarse a otras personas, haciendo un cuerpo con ellas. 7. prnl. amancebarse. 8. prnl. Acogerse a la protección de alguien o de algo, valerse de él o de ello. 9. prnl. Acercarse al conocimiento de algo. *Arrimarse al punto de La dificultad* 10. prnl. *Taurom.* Torear o intentar torear en terreno próximo al toro. 11. prnl. coloq. *Ven.* Dicho de una persona: Quedarse a vivir a expensas de otra.

DEM: arrimar v tr (Se conjuga como *amar*) 1 Acercar algo o a alguien a otra persona o cosa ya establecida o localizada en cierto lugar: “*Arrima una silla y tómate un tequilita con nosotros*”, “*Se arrimaron al fuego para calentarse*”, “*Ese remedio se unta para que no se arrimen los moscos*”, “*El niño se me arrimaba en la cama*”, “*Se arrima la tierra sobre la semilla*”, “*Arrima a ese rincón lo que vayas a llevarte*”, “*De peón comienza uno arrimándoles mezcla y ladrillo*” 2 *Arrimarse* prnl Ponerse bajo la protección de alguien, vivir en su casa sin pagar nada o a sus expensas: “*Como yo no quise arrimarme con nadie de mi familia, me puse a trabajar*” 3 *Arrimarse (Popular)* Unirse en pareja y vivir juntos sin estar casados; arrejuntarse 4 *Arrimar un golpe, arrimar una patiza, arrimar una chinga, etc (Popular)* Dar un golpe, causar un daño físico o moral: “*Si me sigues fregando te arrimo dos tres madrazos*”, “*Nomás ponerme la demanda, ya me había arrimado una chinga de las buenas*” 5 *Arrimarle a (Popular)* Consumir en exceso: “*Le arrimó durísimo al mezcal*”

Tanto en el *DEM* como en el *DLE* la forma pronominal implica un cambio de significado en relación con su forma transitiva. Veamos si los ejemplos (42) y (43) coinciden con alguna de estas acepciones:

- (42) I: **arrima** tus papeles tantito porque ahí te trae café
 E: ay gracias
 I: no te los vaya a tirar y para qué quieres / te los moja

[PUEB-M12-029]

- (43) I: luego nos hemos topado con así amigos de de la construcción que pues al momento de comer “véngase vamos a juntarnos vamos a comer” / “no” / hay unos que no se quieren **arri-mar** / <chasquido> “vente maestro vamos a comer” / “es que no traigo comida / la verdad pues no me alcanza nomás me deajo echado un taco en la mañana y ya hasta en la noche” / da tristeza da mucha tristeza

[PUEB-H11-005]

Ambos ejemplos muestran que este verbo de movimiento puede usarse para señalar el desplazamiento hacia dos posiciones opuestas: ‘alejar’ (en 42) o ‘acercar’ a un punto de referencia (43). La primera no coincide con las documentadas en las fuentes lexicográficas, por lo que la hemos considerado una variante de uso del español de Puebla.

COMENTARIOS FINALES

En los capítulos anteriores presentamos el panorama general del análisis del corpus para obtener concordancias de uso, frecuencias de aparición y dispersión del vocabulario de Preseea-Puebla. En éste hemos empleado el corpus oral para buscar, confirmar y validar que ciertas palabras se usan y tienen un significado distinto en el habla de Puebla en relación con el español general de México.

Consideramos que los corpus orales son particularmente importantes para la descripción de variedades locales de habla, y coincidimos con Briz en que se trata del “banco de prueba más eficaz y natural para analizar el lenguaje” (2012: 115). Los corpus recogidos en comunidades de habla específicas brindan a la lexicología la posibilidad de registrar este tipo de palabras, cuyo significado ha sido acuñado en la experiencia local e incluso, dentro de ella, en el habla identitaria de ciertos grupos específicos. Desde luego, este registro no permite asegurar que se trate de usos exclusivos del habla de Puebla, pero sí que se usan en el habla de la ciudad de Puebla.

En casos como los nombres de alimentos, se requiere conocer el referente para describirlo, es decir, el acercamiento debe ser semánti-

co-referencial. El registro de este vocabulario requiere de una documentación y tratamiento enciclopédico, porque se trata de palabras de ámbito objetivo; el lexicógrafo debe brindar información histórica y sociocultural que ayude al usuario externo a entender mejor las características de la realidad que desconoce. La definición de estas voces se centra en la descripción del “objeto”, y debe ser lo suficientemente preciso como para que el hablante local vea representadas sus prácticas cotidianas e identitarias.

Por otra parte, el estudio de concordancias o usos reales de habla registrados en estos corpus orales permite observar patrones sintácticos y matices semánticos que no necesariamente coinciden con los señalados en los estudios gramaticales, lo cual justifica un acercamiento léxico gramatical o de sintaxis léxica para identificar otros patrones que se manifiestan en el uso.

En cuanto a la documentación lexicográfica de palabras que denominan realidades culinarias, observamos que la misma palabra puede nombrar referentes (recetas) distintas cuyas peculiaridades otorgan identidad a la región en la que se originaron.

Las palabras que denominan prácticas socioculturales específicas no se documentan en ninguna de las obras lexicográficas, porque su registro requiere del trabajo de académicos locales. Por ejemplo, *huehue* o *huéhuere*, *mosaicos*, *murales* y *pitufo* son palabras de la realidad y de la vida cotidiana de los poblanos en una temporalidad y un contexto específicos, por lo que, incluso dentro de tal marco, pueden estar restringidas a ciertos grupos etarios o sociales.

El vocabulario local es, sin duda, la fuente lingüística más accesible y productiva de información cultural; su registro permite dar cuenta del desarrollo de hábitos, creencias sociales, así como de la historia regional (Pederson y Cassidy, 1971: 163-164). Su análisis brinda la posibilidad de acercarse al mismo tiempo a los procesos de significación de la comunidad (como unidad de denominación de prácticas culturales propias) y a su valor como símbolo de identidad.

Para concluir, queremos resaltar la importancia de seguir recogiendo corpus lingüísticos, principalmente orales y con metodologías similares, que permitan el registro sistemático de las variedades desde cada cultura local, a fin de que, pese a las limitantes, la lexicografía

se nutra de su información para documentar voces de la vida cotidiana. En el caso de este primer acercamiento descriptivo del léxico de Preseea-Puebla, pretendemos contribuir a un mejor conocimiento hacia el interior del dialecto del Altiplano central mexicano.

BIBLIOGRAFÍA

SOFTWARE Y BASES DE DATOS

- Anthony, Laurence 2020. *AntConc 3.2*. Tokio: Waseda University. Consultada por última vez en septiembre de 2001, en <<https://www.laurenceanthony.net/software/antconc/>>.
- Barriga Villanueva, Rebeca (coord.), con la colaboración de Valeria Chávez Granados 2021. *Lingmex: Bibliografía lingüística de México desde 1970*. Vigésimosexta edición. México: El Colegio de México. Consultada permanentemente en <<https://lingmex.colmex.mx/>>.
- Padró, Lluís 2020. *FreeLing 4.2*. Consultada por última vez en septiembre de 2021, en <<http://nlp.lsi.upc.edu/freeling/node/1>>.

DICCIONARIOS, CORPUS

- CEMC = Corpus del Español Mexicano Contemporáneo. Diccionario del Español de México. Consultada en <<http://www.corpus.unam.mx/cemc>>, *software* AMATE ver. 1.0
- COEM = Corpus Oral del Español de México. Coordinado por Pedro Martín Butragueño, Érika Mendoza y Leonor Orozco. México: El Colegio de México (en preparación). Consultada en <https://lef.colmex.mx/corpus_oral_del_espanol_de_mexico.html>.
- Corde = Corpus Diacrónico del Español. Real Academia Española. Consultada en <<https://www.rae.es/banco-de-datos/corde>>.
- Corpes XXI = Corpus del Español del Siglo XXI. Real Academia Española. Consultada en <<https://apps2.rae.es/CORPES/view/inicio>>.

- CREA = Corpus de Referencia del Español Actual. Real Academia Española. Consultada en <<https://corpus.rae.es/creanet.html>>.
- CSCM = Corpus Sociolingüístico de la Ciudad de México. Coordinado por Pedro Martín Butragueño y Yolanda Lastra. Tres tomos en CD. Vol. 1, *Hablantes de instrucción superior* 2011; vol. 2, *Hablantes de instrucción media* 2012, y vol. 3, *Hablantes de instrucción baja* 2015. México: El Colegio de México.
- CSCP = Corpus Sociolingüístico de la Ciudad de Puebla. Niktelol Palacios y Laura Romero Rangel (en preparación). México: El Colegio de México.
- DEGM = *Diccionario enciclopédico de la gastronomía mexicana*. Ricardo Muñoz Zurita 2020. México: Larousse. Versión en línea consultada en <<https://laroussecocina.mx/libro/diccionario-enciclopedia-de-la-gastronomia-mexicana>>.
- DEM = *Diccionario del español de México*. Dirigido por Luis Fernando Lara 2020. México: El Colegio de México. Versión actualizada en línea consultada en <<https://dem.colmex.mx>>.
- DLE = *Diccionario de la lengua española*. Real Academia Española 2020. Versión actualizada en línea consultada en <<https://www.rae.es>>.
- DMéx = *Diccionario de mexicanismos*. Academia Mexicana de la Lengua 2010. Dirigido por Concepción Company Company. México: Siglo XXI.
- Montemayor = *Diccionario del náhuatl en el español de México*. Carlos Montemayor 2007. México: Universidad Nacional Autónoma de México / Gobierno del Distrito Federal.
- RAE H 1936 = *Diccionario histórico de la lengua española*. Real Academia Española 1936. Tomo II, *B-Cevilla*. Versión en línea consultada en el Nuevo Tesoro Lexicográfico de la Lengua Española, <<https://webfrrl.rae.es/DH1936.html>>.

NOTAS PERIODÍSTICAS

- Martínez, Cecilia 2008. "El CENHCH a 51 años sigue siendo uno de los mejores del estado." *Periódicodigital.m*, 16 de febrero. Consultada en <http://periodicodigital.com.mx/notas/el_cenhch_a>

51_anos_sigue_siendo_de_los_mejores_del_estado1203141600#.UTzwFNY0iOJ>.

- Rueda, Arturo 2010. "Sobre los desplantes presidenciales (soy del CENHCH)." *Diario Cambio*, 7 de mayo. Consultada en <<http://www.quintacolumna.com.mx/columnas/tiempos/2010/mayo/colum-tiempos-070510.php>>.
- Sánchez, Francisco 2018. "Raúl Velasco de Santiago, disciplina para los jóvenes." *Intolerancia*, 11 de septiembre. Consultada en <<https://intoleranciadiario.com/articulos/2018/09/11/170461-raul-velasco-de-santiago-disciplina-para-los-jovenes.html>>.
- Vázquez, Mario 1984. "Obrero asaltado por banda de «Pitufos»." *El Sol de Puebla*, 7 de enero.

ESPECIALIZADA

- Aronoff, Mark 1985. *Word Formation in Generative Grammar*. Cambridge, Mass.: Massachusetts Institute of Technology.
- Ávila, Raúl 2008. "Diccionarios, dialectos y diferencias culturales: los antojitos mexicanos." En Dolores Azorín Fernández (ed.), *El diccionario como puente entre las lenguas y culturas del mundo: actas del II Congreso Internacional de Lexicografía Hispánica, Alicante 2006*. Alicante: Biblioteca Virtual Miguel de Cervantes, pp. 483-490.
- Azor, Ileana 2006. "Los carnavales en México: teatralidades de la fiesta popular", *América sin nombre* (8), pp. 58-67. Consultada en <<http://rua.ua.es/dspace/handle/10045/5698>>.
- Briz, Antonio (coord.) 1995. *La conversación coloquial. Materiales para su estudio*. Valencia: Universidad de Valencia.
- Briz, Antonio 2000. "La variedad coloquial en el aula de lengua extranjera. Pautas para el análisis léxico", *Carabela* 47, pp. 37-51.
- Briz, Antonio 2010. "Lo coloquial y lo formal, el eje de la variedad lingüística." En Rosa María Castañer Martín y Vicente Lagüéns Gracia (coords.), *De moneda nunca usada: Estudios dedicados a José M.^a Enguita Utrilla*. Zaragoza: Institución Fernando El Católico-Consejo Superior de Investigaciones Científicas, pp. 125-133.

- Briz, Antonio 2012. “Los déficits de los corpus orales del español (y de algunos análisis).” En Tomás Eduardo Jiménez Juliá, Belén López Meirama, Victoria Vázquez Rozas, Alexandre Veiga Rodríguez y Guillermo Rojo Sánchez (coords.), *Cum corde et in nova grammatica: estudios ofrecidos a Guillermo Rojo*. Compostela: Universidad de Santiago de Compostela, pp. 115-137.
- BUAP 2007. *Modelo Universitario Minerva. Documento de integración*. Puebla: Benemérita Universidad Autónoma de Puebla. Consultada en <https://csbiologicas.buap.mx/sites/default/files/MUM_Documentos_de_Integraci%C3%B3n.pdf>.
- BUAP. “Los exilios latinoamericanos en la BUAP.” *Crónica universitaria*. Repositorio digital consultado en <<https://cultura.buap.mx/cronicauniversitaria/exilios-latinoamericanos-buap-repositorio-digital>>.
- Bühler, Karl 1961. *Teoría del lenguaje*. Madrid: Revista de Occidente.
- Cabrera Becerra, Virginia 1995. “La fundación de Puebla: un hecho complejo”, *CIENCIA Ergo-Sum* 2 (3), pp. 287-292.
- Carcedo, Helena F. 1983. “Notas sobre un estudio lingüístico-ideológico”, *Boletín AEPE* (29), pp. 101-109. Consultada en <https://cvc.cervantes.es/ensenanza/biblioteca_ele/aepe/pdf/boletin_29_15_83/boletin_29_15_83_14.pdf>.
- Cestero, Ana M. 2020. “Uses and resources of mitigation, in contrast”, *Spanish in Context* 17 (2), pp. 362-383.
- Cestero, Ana M., y Lidia Rodríguez Alfano 2014. *Análisis de la atenuación. Guía de estudio de la atenuación*. Versión de julio de 2008 en línea consultada en <<http://presea.linguas.net>>.
- Coseriu, Eugenio 1973. *Teoría del lenguaje y lingüística general. Cinco estudios*. Tercera edición. Madrid: Gredos.
- Cruz Gómez, Sandra Rocío 2014. “Panorama sociolingüístico de las lenguas indígenas del Chiapas Actual.” En Rebeca Barriga Villanueva y Pedro Martín Butragueño (dirs.), *Historia sociolingüística de México*. Tomo 3, *Espacio, contacto y discurso político*. México: El Colegio de México, pp. 1721-1755.
- Delgado Candel, Julen 2019. *Evaluación de material didáctico en L2*. Trabajo de fin de grado en ingeniería informática de gestión de sistemas de información. Bilbao: Universidad de País Vasco. Consul-

- tada en <http://addi.ehu.es:8080/bitstream/handle/10810/37081/TFG_JulenDelgadoCandel.pdf?sequence=1&isAllowed=y>.
- Eckert, Penelope 2018. *Meaning and linguistic variation: the third wave in sociolinguistics*. Cambridge / Nueva York: Cambridge University Press.
- Faber, Pamela, Antonio Moreno Ruiz y Chantal Pérez Hernández 1999. "Lexicografía computacional y lexicografía de corpus", *Revista Española de Lingüística Aplicada* 1, pp. 175-214.
- Fajardo Rojas, Carmen Teresa 2016. *Locuciones del español de México: análisis de corpus orales*. Tesis inédita. Puebla: Benemérita Universidad Autónoma de Puebla.
- Fernández Gordillo, Luz 2006. *Tratamiento lexicográfico de los mexicanismos y pseudomexicanismos en los diccionarios académicos*. Tesis doctoral inédita. México: El Colegio de México.
- Franco Trujillo, Erik (en preparación). *Juegos verbales en la tradición popular mexicana*.
- Frenek, Margit 1953. "Designaciones de rasgos físicos en el español de México", *Nueva Revista de Filología Hispánica*, 7 (1/2), pp. 134-156.
- García Hidalgo, Ma. Isabel 1979. "La formalización del Analizador Gramatical del DEM." En Luis Fernando Lara, Roberto Ham Chande y Ma. Isabel García Hidalgo, *Investigaciones lingüísticas en lexicografía*. México: El Colegio de México, pp. 85-156.
- Garza, Gustavo 2003. *La urbanización de México en el siglo xx*. México: El Colegio de México.
- Gimate-Welsh, Adrián S. 1980a. "Lenguaje, cultura e ideología. Premisas teórico metodológicas para una investigación." En Tomás Segovia, Raúl Dorra, Noé Jitrik, Oscar del Barco, Adrián S. Gimate-Welsh y Fernando Tudela, *El lenguaje: problemas y reflexiones actuales*. Puebla: Universidad Autónoma de Puebla, pp. 93-111.
- Gimate-Welsh, Adrián S. 1980b. *Lenguaje y sociedad. Bases para una investigación regional*. Puebla: Universidad Autónoma de Puebla.
- Gimate-Welsh, Adrián S., y Enrique Marroquín 1985. *Lenguaje, ideología y clases sociales. Las vecindades en Puebla*. Puebla: Universidad Autónoma de Puebla.
- Gobierno del Estado de Puebla y H. Ayuntamiento del Municipio de Puebla 2016. *Programa Municipal de Desarrollo Urbano Sustentable*

- de Puebla, H. *Ayuntamiento del Municipio de Puebla 2014-2018*. Puebla: Secretaría de Desarrollo Urbano y Sustentabilidad / Dirección de Desarrollo Urbano / Instituto Municipal de Planeación de Puebla.
- Gobierno del Estado de Puebla y Orden Jurídico Poblano 2001. *Ley Orgánica Municipal*. Puebla: Secretaría General de Gobierno. Consultada en <http://gobiernoabierto.pueblacapital.gob.mx/transparencia_file/InformacionFiscal/norm/77.01.ley.org.mpal.pue.pdf>.
- González Avilés, Rogelio 2018. “La creación de léxico humorístico en Les Luthiers”, *Signo y Seña* (34), pp. 61-78.
- Halliday, Michael Alexander Kirkwood 1978. *El lenguaje como semiótica social. La interpretación social del lenguaje y del significado*. México: Fondo de Cultura Económica.
- Herrera Zendejas, Esther 2014. “La Babel del sur: el caso de Oaxaca.” En Rebeca Barriga Villanueva y Pedro Martín Butragueño (dirs.). *Historia sociolingüística de México*. Tomo 3, *Espacio, contacto y discurso político*. México: El Colegio de México, pp. 1691-1719.
- Hirschberg, Julia 1978. “La fundación de Puebla de los Ángeles. Mito y realidad”, *Historia Mexicana* 28 (2), pp. 185-223. Disponible en: <<https://historiamexicana.colmex.mx/index.php/RHM/article/view/2728/2238>>.
- Inclán, Rebeca, 1995. “Inmigración libanesa en México. Un caso de diversidad cultural”, *Historias. Revista de la Dirección de Estudios Históricos* (33), pp. 61-68. Consultada en <http://132.248.2.138:8991/F/B9CCEPQQYENCFCSDKH3IVDHG3KLXJK5G9DXC4H5YT18M3QGX55-01973?func=full-set-set&set_number=297623&set_entry=000016&format=999>, también disponible en <<http://www.mediateca.inah.gob.mx/repositorio/islandora/object/articulo:18462>>.
- INEGI 2010. *Censo de Población y Vivienda*. México: Instituto Nacional de Estadística y Geografía.
- INEGI 2020. *Censo de Población y Vivienda. Presentación de resultados: Puebla*. México: Instituto Nacional de Estadística y Geografía. Consultada en <https://www.inegi.org.mx/contenidos/programas/ccpv/2020/doc/cpv2020_pres_res_pue.pdf>.

- INPI e Inali 2020. *Atlas de los pueblos indígenas de México*. México: Instituto Nacional de los Pueblos Indígenas / Instituto Nacional de Lenguas Indígenas. Consultada en <<http://atlas.inpi.gob.mx/>>.
- Lara, Luis Fernando 1996. “La cultura oral mexicana: una aproximación lexicológica.” En Thomas Kotschi, Wulf Oesterreicher y Klaus Zimmermann (eds.), *El español hablado y la cultura oral en España e Hispanoamérica*. Madrid / Fráncfort: Vervuert/Iberoamericana, pp. 463-473.
- Lara, Luis Fernando 2004. “¿Es posible una teoría de la palabra?”, *Lexis* 27 (1-2), pp. 401-427.
- Lara, Luis Fernando 2006. *Curso de lexicología*. México: El Colegio de México.
- Lara, Luis Fernando 2012a. “Hacia una tipología de las tradiciones verbales populares”, *Nueva Revista de Filología Hispánica* 60 (1), pp. 51-60.
- Lara, Luis Fernando 2012b. “La definición lexicográfica del vocabulario de germanía y jergal”, *Revista Káñina* XXXVI, pp. 13-18.
- Lara, Luis Fernando 2017. “Agradecimiento”, *Boletín Editorial El Colegio de México* (176), pp. 3-8.
- Lara, Luis Fernando, Roberto Ham Chande y Ma. Isabel García Hidalgo 1979a. “El vocabulario fundamental del español de México.” En *Investigaciones lingüísticas en lexicografía*. México: Centro de Estudios Lingüísticos y Literarios-El Colegio de México (Jornadas, 89). Versión electrónica disponible consultada en <<https://dem.colmex.mx/repository/pdfs/vocabulario-fundamental.pdf>>.
- Lara, Luis Fernando, Roberto Ham Chande y Ma. Isabel García Hidalgo 1979b. *Investigaciones lingüísticas en lexicografía*. México: El Colegio de México.
- Lastra Yolanda, y Pedro Martín Butragueño 2000. “El modo de vida como factor sociolingüístico en la Ciudad de México.” En Pedro Martín Butragueño (ed.), *Estructuras en contexto: estudios de variación lingüística*. México: El Colegio de México.
- Lázaro, Jorge, y María Guadalupe Huerta Morales 2020. “El Corpus del habla de Puebla: un corpus regional.” En Jorge Lázaro y Rafael Saldívar Arreola (eds.), *Estudios sobre léxico del español de México basados en corpus*. México: Fontamara, pp. 27-46.

- Leal Cruz, Pedro 1993. “Inestabilidad de los sonidos ‘a’ y ‘d’ a inicio de palabra en el habla palmera”, *Revista de Filología de la Universidad de La Laguna* (12), pp. 169-180.
- Licon Valencia, Ernesto, y Martha Ivett Pérez Pérez 2018. *El carnaval en la región Puebla Tlaxcala. Acercamientos etnográficos y multidisciplinares*. México: Instituto Municipal de Arte y Cultura de Puebla.
- Llisterri, Joaquim 2003. “Lingüística y tecnologías del lenguaje”, *Lynx. Panorámica de Estudios Lingüísticos* (2), pp. 9-71.
- Lope Blanch, Juan M. 1970a. *Cuestionario para la delimitación de las zonas dialectales de México*. México: El Colegio de México.
- Lope Blanch, Juan M. 1970b. “Las zonas dialectales de México: Proyecto de delimitación”, *Nueva Revista de Filología Hispánica* 19 (1), pp. 1-11.
- Lope Blanch, Juan M. 1980a. “Algunos juegos de palabras en el español de México”, *Lingüística Española Actual* 2 (2), pp. 219-243.
- Lope Blanch, Juan M. 1980b. “Un sistema de numeración festivo”, *Romanica europea et americana. Festschrift für Harri Meier*, Bonn: Bouvier Verlag, pp. 345-349.
- Lope Blanch, Juan M. 1990 [1971]. “El léxico de la zona maya en el marco de la dialectología mexicana”. En *Investigaciones sobre dialectología mexicana*. México: Universidad Nacional Autónoma de México, pp. 59-132.
- Lope Blanch, Juan M. 1990-2000. *Atlas lingüístico de México*. 6 vols. México: El Colegio de México / Universidad Nacional Autónoma de México / Fondo de Cultura Económica.
- Makri, Julie 2012. “Terminología y tipología de los procesos de la neología formal”. *Debate Terminológico* (8), pp. 3-15.
- Marín Guzmán, Roberto, y Zidane Zeraoui 2003. *Arab Immigration in Mexico in the nineteenth and twentieth Centuries. Assimilation and Arab Heritage*. Austin: Augustine Press / Instituto Tecnológico y de Estudios Superiores de Monterrey.
- Martín Butragueño, Pedro 2020. “Building the Megalopolis: dialectal leveling and language contact in Mexico City.” En Andrew Lynch (ed.), *The Routledge handbook of spanish in the global city*. Londres / Nueva York: Routledge, pp. 234-274.
- Mendoza Guerrero, Everardo 1996. *El español hablado en Sinaloa: el léxico en la conformación dialectal del Noroeste*. Tesis para obtener el

- grado de maestro en lingüística hispánica. México: Universidad Nacional Autónoma de México.
- Mendoza Guerrero, Everardo 2003. *El habla de Culiacán: fonética, morfosintaxis y léxico*. Tesis para obtener el grado de doctor en lingüística hispánica. México: Universidad Nacional Autónoma de México.
- Mendoza Guerrero, Everardo 2011. *El habla de Culiacán*. Culiacán, Sinaloa: Universidad Autónoma de Sinaloa / El Colegio de Sinaloa / Instituto Municipal de Cultura.
- Moreno de Alba, José G. 2003. "Acemite / acemita / cemita / cema." En *Suma de minucias del lenguaje*. México: Fondo de Cultura Económica. Consultada en <<https://www.fondodeculturaeconomica.com/obra/suma/r1/buscar.asp?idPALABRA=207&starts=Todos&word=acemite%20/acemita%20/%20cemita%20/%20cema>>.
- Moreno Fernández, Francisco 1993. "Proyecto para el estudio sociolingüístico del español de España y América (Preseea)", *Lingüística* 5, pp. 268-271.
- Moreno Fernández, Francisco 1997. "Metodología del Proyecto Sociolingüístico para el Estudio del Español de España y de América (Preseea)." En *Trabajos de sociolingüística hispánica*. Alcalá de Henares: Universidad de Alcalá, pp. 137-167.
- Moreno Fernández, Francisco 2003. "Metodología del Proyecto Sociolingüístico para el Estudio del Español de España y de América (Preseea). Versión revisada-octubre 2003." Consultada en <<https://preseea.linguas.net/Portals/o/Metodologia/METODOLOG%C3%8DA%20PRESEEA.pdf>>.
- Morúa, María del Carmen 2011. "Datos conversacionales como fuente para explorar formas de diferenciación sociolingüística y registros de ámbito local-regional." En Pedro Martín Butragueño (ed.), *Realismo en el análisis de corpus orales. Primer coloquio de cambio y variación lingüística*. México: Centro de Estudio Lingüísticos y Literarios-El Colegio de México, pp. 491-515.
- Morúa Leyva, María del Carmen, y Julio César Serrano Morales 2004. "Dos mil kilómetros de por medio: dialectología perceptual contrastiva del español mexicano." En María del Carmen Morúa Leyva y Rosa María Ortiz Ciscomani (eds.), *Memorias del VII*

- Encuentro Internacional de Lingüística en el Noroeste*. Tomo 2. Hermosillo: Universidad de Sonora, pp. 253-276.
- Nowikow, Waczesław 2001. “Sobre la elaboración del modelo para el estudio del español hablado de Puebla (México): esbozo de una experiencia”, *Moenia: Revista Lucense de Lingüística & Literatura* (23), pp. 451-474.
- Olivar Espinosa, Stephany 2020. “La entonación de los enunciados aseverativos del español de contacto de hablantes bilingües de San Miguel Canoa, Puebla: una variedad del español mexicano”, *Anuario de Letras. Lingüística y Filología* 8 (2), pp. 65-116. Consultada en <<https://doi.org/10.19130/iifl.adel.2020.2.24873>>.
- Olivar Espinosa, Stephany (en preparación). *Algunos aspectos en el estudio de la entonación del español en contacto con el náhuatl de San Miguel Canoa, Puebla*. Tesis doctoral. México: El Colegio de México.
- Orozco, Leonor 2014. “El empleo de ¿no?, ¿eh? y ¿verdad? en situación de entrevista sociolingüística.” En Pedro Martín Butragueño y Leonor Orozco (eds.), *Argumentos cuantitativos y cualitativos en sociolingüística: Segundo coloquio de cambio y variación lingüística*. México: El Colegio de México, pp. 643-668.
- Padró, Lluís 2011. “Analizadores multilingües en *FreeLing*”, *Linguamática* 3 (1), pp. 13-20.
- Palacios, Niktelol 2010. “El léxico de las drogas en el español de México.” En Rodrigo Pardo Fernández (coord.), *Caleidoscopio hispánico: estudios lingüísticos y literarios*. Morelia: Universidad Michoacana de San Nicolás de Hidalgo, pp. 71-87.
- Palacios, Niktelol 2014. *Estudio lexicológico del vocabulario textil común en el español de México*. Tesis de doctorado. México: El Colegio de México.
- Palacios, Niktelol 2015. “Actitudes y creencias lingüísticas en el Corpus sociolingüístico de la ciudad de Puebla (CSCP)”, XIII Congreso Nacional de Lingüística de la AMLA, Tuxtla Gutiérrez, Universidad Autónoma de Chiapas, 7-10 de octubre. Documento inédito.
- Palacios, Niktelol 2017. “Estrategias y mecanismos de atenuación en Preseea-Puebla: instrucción educativa baja”, *Anuario de Letras. Lingüística y Filología* 5 (2), pp. 149-181.

- Palacios, Niktelol 2019. “Funciones y mecanismos de la atenuación en Preseesa-Puebla: instrucción educativa alta.” En *Voces de la lingüística mexicana contemporánea*. México: El Colegio de México, pp. 235-270.
- Palacios, Niktelol 2020a. “Los nombres de los panes en la ciudad de Puebla.” En Jorge Lázaro y Rafael Saldívar Arreola (eds.), *Estudios sobre léxico del español de México basados en corpus*. México: Fontamara, pp. 81-99.
- Palacios, Niktelol 2020b. “Tejiendo palabras: estudio lexicológico de oficios.” En Alan Emmanuel Pérez Barajas y Axel Hernández Díaz (coords.), *Propuestas metodológicas para la investigación lingüística. Aplicaciones teóricas y descriptivas*. México: Universidad de Colima, pp. 343-366.
- Palacios, Niktelol 2020c. “Compuestos sintagmáticos y locuciones nominales en el español de México: criterios léxico-semánticos para su distinción.” En Esther Hernández y Pedro Martín Butragueño (eds.), *Las palabras como unidades lingüísticas*. Madrid / México: Consejo Superior de Investigaciones Científicas / El Colegio de México, pp. 443-467.
- Palacios, Niktelol (en preparación). “Funciones de la atenuación y factores situacionales en Preseesa-Puebla: Instrucción educativa baja y alta.” En Karina Fascinetto Zago y Karen Miladys Cárdenas Almanza (eds.), *Tópicos en lingüística*. México: Benemérita Universidad Autónoma de Puebla.
- Palacios, Niktelol, y Dinorah Pesqueira 2018. “Quiénes hablan mejor y quiénes hablan bien: acercamiento al estudio de actitudes y creencias lingüísticas de la Ciudad de México y Puebla”, XIII Congreso Internacional de Lingüística Xeral, Vigo, Universidad de Vigo, 13-15 de junio. Documento inédito.
- Palacios, Niktelol, y Juan Luis Serralde (en preparación). “Lematización de un corpus oral: a propósito de Preseesa-Puebla.”
- Pederson, Lee, y Frederic G. Cassidy 1971. “Chicago Words: The Regional Vocabulary”, *American Speech* 46 (3-4), pp. 163-192.
- Ramírez Carrillo, Luis Alfonso 1994. “De buhoneros a empresarios: la inmigración libanesa en el sureste de México”, *Historia Mexicana* 43 (3), pp. 451-486.

- Recalde, Montserrat, y Victoria Vázquez Rozas 2009. "Problemas metodológicos en la formación de corpusorales." En Pascual Cantos Gómez y Aquilino Sánchez Pérez (eds.), *A survey of corpus-bases-research*. Murcia: Asociación Española de Lingüística del Corpus, pp. 51-64.
- Redodem = Red de Documentación de las Organizaciones Defensoras de Migrantes 2019. *Informe. Migraciones en México: fronteras, omisiones y transgresiones*. Consultada en <<https://redodem.org/informes/>>.
- Rivera Domínguez, María Ligia de la Soledad 1984. *Léxico religioso e ideología*. Tesis de licenciatura. Puebla: Universidad Autónoma de Puebla.
- Rodríguez González, Félix 1989. "Los cruces léxicos en el ámbito político-periodístico", *Verba* 16, pp. 357-386.
- San Vicente, Félix 2017. "Hablado / Parlato: Aspectos lingüísticos y discursivos de la comunicación oral." En Félix San Vicente, Gloria Bazzocchi y Pilar Capanaga (eds.), *Oraliter. Formas de la comunicación presencial y a distancia*. Bolonia: Bologna University Press, pp. 17-72.
- Sánchez Gavi, José Luis 2016. "Movilidad humana. El fenómeno migratorio en Puebla bajo la perspectiva de la Iglesia católica", *Tla-melaua* 9 (39), pp. 109-130.
- Serrano Morales, Julio César 2019. "¿Cuántos dialectos del español existen en México? Un ensayo de dialectología perceptual." En Niktelol Palacios (ed.), *Voces de la lingüística mexicana contemporánea*. México: El Colegio de México, pp. 321-346.
- Sierra Martínez, Gerardo 2011. "Lexicografía computacional en las búsquedas onomasiológicas con lenguaje natural." En María Eugenia Vázquez Laslop, Klaus Zimmermann y Francisco Segovia (eds.), *De la lengua por sólo la extrañeza. Estudios de lexicología, norma lingüística, historia y literatura en homenaje a Luis Fernando Lara*. Vol. 1. México: El Colegio de México, pp. 445-464.
- Sierra Martínez, Gerardo, Julián Solórzano Soto y Arturo Curiel Díaz 2017. "GECO, un Gestor de Corpus colaborativo basado en web", *Linguamática* 9 (2), pp. 57-72. Consultada en <<https://linguamatica.com/index.php/linguamatica/issue/view/19/V9N2>>.

- Tavera-Cervantes, Xóchitl 2021. *Broncas, bronquillas y broncones: combinatoria, significación y usos de la palabra bronca en el español hablado en México*. Tesis inédita. Morelia: Universidad Michoacana de San Nicolás de Hidalgo.
- Torres Villaseñor, Hermila E. 1970. “Léxico relativo al vestuario infantil”, *Anuario de Letras* 8, pp. 241-253.
- Valadez, Carmen Delia 2020. “Llovizna y mejilla ‘cachete’: una nota léxica a partir del *Atlas lingüístico de México*”, *Lingüística Mexicana*. Nueva Época 2 (1), pp. 119-129. Consultada en <http://linguisticamexicana-amla.colmex.mx/index.php/Linguistica_mexicana/article/view/383/353>.
- Vázquez Zúñiga, Iliana 2020. “De los dólares a los *changarros*: Migración interna, internacional y reproducción social en el sur de Puebla, México”, *Migraciones Internacionales* 11. Consultada en <<https://migracionesinternacionales.colef.mx/index.php/migracionesinternacionales/article/view/2131>>.
- Vélez Pliego, Francisco M. 2011. “Puebla de Zaragoza, antigua Ciudad de los Ángeles. Patrimonio cultural de la humanidad”, *Revista Sociedad, Ciudad y Territorio* (1), pp. 1-47. Consultada en <<http://148.228.173.140/OUMpuebla/documents/PUEBLAFVP.pdf>>, también disponible en <<https://issuu.com/ablmartilu/docs/pueblafvp>>.
- Villena Ponsoda, Juan Andrés, Juan Antonio Moya Corral, Antonio Manuel Ávila Muñoz y Matilde Vida Castro 2003. “Proyecto de investigación de la formación de dialectos (FORDIAL)”, *ELUA* (17), pp. 607-638. Consultada en <https://rua.ua.es/dspace/bitstream/10045/6196/1/ELUA_17_34.pdf>.
- Zacarías Ponce de León, Ramón Felipe 2017. “Neología e identidad lingüística en el español de México”, *e_Scripta Romanica* 4, pp. 116-124.
- Zacarías Ponce de León, Ramón Felipe 2018. *¿Cómo se forman las palabras? Claves de combinación*. México: Universidad Nacional Autónoma de México.
- Zeraoui, Zidane 1997. “La emigración árabe en México.” En María Elena Ota Mishima (ed.), *Destino México. Un estudio de las migraciones asiáticas a México, Siglos XIX y XX*. México: El Colegio de México, pp. 257-304.

- Zeraoui, Zidane 2006. “La inmigración árabe en México: integración nacional e identidad comunitaria”, *Contra Relatos desde el Sur. Apuntes sobre África y Medio Oriente* 2 (3), pp. 11-32. Consultada en <<http://bibliotecavirtual.clacso.org.ar/ar/libros/argentina/cea/contras/3/zeraoui.pdf>>.
- Zimmermann, Klaus 1996. “Lenguaje juvenil, comunicación entre jóvenes y oralidad.” En Thomas Kotschi, Wulf Oesterreicher y Klaus Zimmermann (eds.), *El español hablado y la cultura oral en España e Hispanoamérica*. Madrid / Fráncfort: Vervuert / Iberoamericana, pp. 475-514.

Corpus orales y lexicografía. A propósito de Preseaa-Puebla
se terminó de imprimir en octubre de 2022,
en los talleres de SmartbooksMx Press,
Av. Universidad 626, col. Letrán Valle,
0350, Ciudad de México.
Portada: Pablo Reyna.
Tipografía y formación: LaCriba.
Cuidó la edición Carlos Mapes bajo la coordinación
de la Dirección de Publicaciones de El Colegio de México.
La edición consta de 350 ejemplares.

CENTRO DE ESTUDIOS LINGÜÍSTICOS Y LITERARIOS

ESTUDIOS DE LINGÜÍSTICA

XXXVIII

El español de Puebla pertenece a la zona dialectal del Altiplano central. Quizá por el predominio político, socioeconómico y cultural de la capital del país o por el hecho de que esta zona se percibe bastante homogénea, representada por los usos lingüísticos de la Ciudad de México, el español de Puebla ha sido poco investigado en la lingüística mexicana.

En este libro se presenta una propuesta de análisis que combina objetivos, métodos y herramientas de la lingüística computacional, la lexicografía, la lexicología y la sociolingüística. Se busca que la exposición detallada de las decisiones metodológicas sirva de punto de partida para la preparación de lemmarios de otros corpus del Proyecto para el Estudio Sociolingüístico del Español de España y América (Pressea) —y de otros corpus orales regionales—; asimismo, se espera que, a partir de un método común, pueda haber un acercamiento a la variación lexicológica por parcelas referenciales, por el origen etimológico o por la distribución social e identitaria de ciertos ítems.

ISBN: 978-607-564-388-5

