

Niktelol Palacios. *Corpus orales y lexicografía. A propósito de PRESEEA-Puebla*. El Colegio de México, México, 2023; 178 pp.

ANAYELI HERNÁNDEZ CRUZ

Escuela de Antropología e Historia del Norte de México

anayeli.h.cruz@gmail.com

El español mexicano es una de las variedades hispánicas con mayor número de hablantes, y mayor complejidad por su diversidad —producto de contactos lingüísticos históricos y recientes—; por sus discursos ideológicos construidos a lo largo de la historia por sus hablantes, y por sus grandes movimientos culturales y contraculturales —urbanos y no urbanos— que han tenido un innegable impacto en lo lingüístico. Es una variedad que aún necesita seguir siendo documentada, descrita y analizada en sus dimensiones históricas, geográficas, sociales y estilísticas. Al hacerlo, podremos entender de una mejor manera su expansión por cada rincón del territorio; la influencia que las lenguas indígenas y otras lenguas han ejercido en ella; así como la conformación de los valores que sustentan sus tradiciones verbales. Así, la compleja historia mexicana y sus dinámicas sociopolíticas revelan la creación de centros de difusión cultural, de los agentes normativos, y de la configuración de identidades locales que han dado forma a nuestro mosaico cultural y lingüístico.

Por ejemplo, el proyecto del Atlas Lingüístico de México (ALM) lo inició Juan M. Lope Blanch en los años setenta al notar la poca atención que había recibido el español en las investigaciones lingüísticas. El ALM ofrecía un panorama bastante completo de la variación en los diversos niveles de descripción lingüística. Con el surgimiento de la Sociolingüística, el interés se renueva; específicamente, en la descripción y el análisis de la variación y el cambio lingüístico en toda su diversidad (culto, popular, en contacto, en ámbitos rurales, en ámbitos urbanos, etc.), todo en contextos de uso reales y más o menos espontáneos.

En los centros de estudio hispánicos comienzan, entonces, a plantearse grandes proyectos. El proyecto para el estudio sociolingüístico del español de España y América, coordinado por Francisco Moreno Fernández y Silva Corvalán, responde a la necesidad de conocer el español en los principales núcleos urbanos teniendo una metodología que permita la comparación de los datos. En 1993 inicia el estudio de las hablas urbanas con el Proyecto para el estudio sociolingüístico del español de España y de América (PRESEEA), planteado en el congreso de la Asociación de Lingüística y Filología de

América Latina (ALFAL), celebrado en Veracruz¹. Desde entonces distintos académicos y centros de investigación se han unido a este megaproyecto.

En México, Pedro Martín Butragueño —uno de los pioneros de los estudios sociolingüísticos—, participó en las discusiones metodológicas del corpus PRESEEA. La Universidad de Alcalá de Henares, la Universidad de Valencia y El Colegio de México fueron de los primeros centros de investigación que comenzaron a recopilar corpus orales con base en el modelo PRESEEA. En 1997 el equipo coordinado por Pedro Martín Butragueño y Yolanda Lastra comienza las grabaciones que conformarían el Corpus Sociolingüístico de la Ciudad de México (CSCM)², el cual toma como base los parámetros metodológicos del PRESEEA, pero agrega un cuestionario lingüístico con preguntas de actitudes y creencias relacionadas con lo que los hablantes consideran el mejor español —desde el punto de vista de la variación regional y guiados por su idea de normatividad—, y pruebas lingüísticas que incluyen un cuestionario fónico, cuyo objetivo es caracterizar la variedad documentada desde este nivel de análisis, así como una prueba de seguridad léxica y la lectura de un texto.

A partir de esto, comenzaron a elaborarse más corpus orales especializados en documentar las hablas urbanas de las diferentes regiones hispanomexicanas, como el corpus *El habla de Monterrey-PRESSEA*³ o el *Corpus PRESSEA-Guadalajara*⁴. Sin duda estos proyectos son una fuente importante de registro y preservación de conversaciones que revelan usos lingüísticos condicionados por las características de los hablantes, como su edad, su género, su grado de instrucción y sus discursos construidos en torno a la configuración de identidades colectivas. También son materiales valiosos que amplían el conocimiento sobre el español mexicano en todos los niveles de descripción lingüística, porque nos permiten emprender estudios de variación fónica, de entonación, de morfología derivativa, y alimentan los repositorios que pueden ser un punto de partida para la creación de repertorios lexicográficos. A nivel de análisis léxico, los corpus nos proporcionan información valiosa relacionada con fenómenos, como el tabú lingüístico; los préstamos léxicos y su adaptación; las formas derivativas más productivas; la variación léxica, y la relación entre léxico e identidad, entre otros.

¹ Para conocer más del tema pueden consultarse los siguientes trabajos: Moreno Fernández, Francisco. 2005. “Corpus para el estudio del español en su variación geográfica y social. El corpus PRESEEA”, *Oralia* 8, pp. 7-33, y Proyecto para el Estudio Sociolingüístico del Español de España y América (PRESEEA). Francisco Moreno Fernández (coord.), <<https://preseea.uah.es/>> [consultado el 5 de mayo de 2023].

² Martín Butragueño, Pedro y Yolanda Lastra (coords.). 2011-2015. *Corpus Sociolingüístico de la Ciudad de México (CSCM)*, 1a. ed. Ciudad de México: El Colegio de México.

³ Rodríguez Alfano, Lidia, María Eugenia Flores Treviño y Tzitel Pérez Aguirre (comps.). 2010. *Corpus El habla de Monterrey-PRESEEA*. Monterrey: Universidad Autónoma de Nuevo León-Facultad de Filosofía y Letras.

⁴ Córdova Abundis, Patricia y Daniel Barragán Trejo. 2021. *El español hablado en Guadalajara. Corpus PRESEEA-Guadalajara*, Guadalajara, Universidad de Guadalajara-Centro Universitario de Ciencias Sociales y Humanidades.

Teniendo esto como antecedente, en 2011 Niktelol Palacios reúne a un equipo de jóvenes universitarios —del que yo formé parte— con diversos intereses metodológicos y disciplinarios, para la conformación de un corpus oral que sentaría las bases para los estudios del español de Puebla; este grupo de alumnos estaba ávido de registrar y analizar la variante de su español porque deseaba demostrar que había algo que lo hacía distinto, principalmente de la variante de la Ciudad de México, con la cual el español poblano comparte muchas características, ya que ambas pertenecen al área del Altiplano central. La idea original era hacer un diccionario del español de Puebla que diera cuenta de los elementos léxicos que lo caracterizan. Sin embargo, pronto nos dimos cuenta de que no contábamos con suficientes documentos de los que obtendríamos el leuario que alimentaría nuestro diccionario. Así, pese a nuestra urgencia de tener un diccionario propio, Niktelol nos hizo caer en la cuenta de que antes era necesario reunir una gran variedad de documentos que representaran las diversas tradiciones discursivas del español de Puebla, y la creación de un corpus oral era el comienzo.

Lo que inició como una inquietud, e incluso como una queja, se vio materializado en largas horas de trabajo en las que los jóvenes de este grupo comenzamos a formarnos como investigadores. Aunque nos basamos en los lineamientos del modelo PRESEEA y del CSCM, aún quedaban por tomarse algunas decisiones metodológicas y reflexionar sobre nuestro quehacer como lingüistas en formación. La recolección del corpus representó un enriquecimiento a nivel académico y humano para todos los miembros del equipo. Las temáticas tratadas en las conversaciones abarcaban historias de vida; de la Puebla de antaño y la contemporánea; de la política nacional y local, entre otras. A través de ellas se podrían obtener diversos registros y géneros textuales, como descripciones, narraciones y argumentaciones que guardan correlatos con los elementos lingüísticos que pueden aparecer en ellos. Los entrevistados en todo momento se mostraron muy entusiasmados y cooperativos; les interesaba que conociéramos la ciudad de Puebla a través de sus historias, sus sentires y sus maneras de pensar. Gracias a este trabajo pudimos conocer de manera más profunda nuestra ciudad y su diversidad.

Todos estos años de trabajo se describen y explican de manera magistral en el libro reseñado a continuación, *Corpus orales y lexicografía. A propósito de Preseea- Puebla*. A lo largo de los cinco capítulos que lo conforman, la autora nos presenta, en primer lugar, los trabajos que anteceden a PRESEEA-Puebla, llevados a cabo por investigadores interesados en documentar la forma de hablar de la región; luego ofrece un panorama histórico de la ciudad enfatizando la importancia sociohistórica de la urbe, que muestra el orgullo local de su población construido a lo largo de la historia. Nos presenta también un panorama sociodemográfico que permite conocer la gran diversidad lingüística y cultural de la capital poblana, pero también delinear futuras investigaciones, como el estudio de los contactos lingüísticos generados con poblaciones hablantes de diversas lenguas indígenas o con migrantes externos. Finalmente, la investigadora se centra de lleno en PRESEEA-Puebla; nos describe su metodología, los subcorpus que lo conforman, como el *Léxico de oficios, flora y fauna*, y ofrece una propuesta de lematización y análisis cuantitativo de los vocablos registrados en él. Veamos cada capítulo con mayor detalle.

En el primer capítulo, “Antecedentes: corpus orales del español de Puebla”, la autora proporciona un panorama de los esfuerzos de investigadores y equipos de trabajo por describir la variante poblana. Los proyectos emprendidos desde la Benemérita Universidad Autónoma de Puebla, a partir de marcos teóricos y metodológicos diversos, intentaron recopilar, desde los años ochenta, corpus orales que diera cuenta del español poblano; de ellos se derivaron algunas investigaciones enfocadas en las múltiples dimensiones de esta variedad: el habla popular escuchada en las vecindades tradicionales de la ciudad, el español poblano en el ámbito religioso o estudios del habla infantil. Lamentablemente, como la autora lo señala, la mayor parte de estos materiales están perdidos, y sólo tenemos noticias de ellos por las investigaciones que resultaron de ellos, o por los libros publicados en los que se explicita la necesidad de crear corpus que dieran cuenta de las peculiaridades de la variante poblana y la planeación consecuente para llevarlos a cabo.

En el capítulo dos, “El Corpus Sociolingüístico de la Ciudad de Puebla y PRESEEA-Puebla”, se nos proporciona una descripción detallada de la planeación y elaboración del corpus poblano. Para quien desee emprender un proyecto tan complejo que involucre el levantamiento, la transcripción y el etiquetado de un repositorio de habla real, éste es un capítulo imprescindible⁵. La autora nos describe detalladamente todas las decisiones tomadas para la recopilación de las 108 entrevistas que lo conforman, comenzando por la delimitación de la zona en la que se levantarían los datos. Con ello muestra la importancia de conocer la historia de la región, la configuración demográfica de la población y la complejidad lingüística producto de los contactos con otras lenguas, para poder delimitar el área de elicitación de los datos y justificar aquella que no se tomó en cuenta, pero que es importante analizar para poder conocer a cabalidad las hablas regionales. En este capítulo se muestra la importancia histórica de la ciudad de Puebla, considerando que también es un área comercial y de paso obligado hacia la capital. Su configuración poblacional e historia marcaron pautas para la conformación del español poblano, así como de las actitudes y creencias en relación con lo propio y con lo ajeno. Es decir, esta información nos permite delinear un perfil del hablante poblano y de las ideologías que subyacen en la construcción de su identidad. Además, la autora muestra el gran mosaico lingüístico que se encuentra en todo el estado, pues en él se hablan náhuatl, totonaco, mazateco y popoloca, entre otras lenguas. Al fungir la ciudad de Puebla como un centro de irradiación cultural y comercial, las migraciones desde el interior y exterior del estado han sido constantes. La autora indica que urge comenzar a hacer estudios enfocados en estas comunidades para dar cuenta del contacto lingüístico entre el español y algunas de las lenguas de las poblaciones migrantes. Lo anterior podrá, por ejemplo, mostrar el impacto de la migración en los procesos de pérdida y desplazamiento lingüístico. Además, destaca la importancia que han tenido las migraciones externas

⁵ Actualmente están en proceso de levantamiento las entrevistas del *Corpus Sociolingüístico de la Ciudad de Chihuahua* y el *Corpus Sociolingüístico de la ciudad de Xalapa*, coordinados por Anayeli Hernández Cruz y Mariana Aguirre Corona, para el caso del primero, y por Luis David Meneses, para el segundo, que replican la metodología de PRESEEA-Puebla.

—como la libanesa o la sudamericana, y en especial la chilena y la argentina—, en la conformación de la identidad local; estos grupos han contribuido en todos los ámbitos de la vida comunitaria, como en la gastronomía poblana y en su vida intelectual.

El *Corpus Sociolingüístico de la Ciudad de Puebla* está conformado por entrevistas divididas en dos subcorpus: el que recoge datos de los oficios tradicionales y de la flora y la fauna, y el corpus PRESEEA-Puebla, el cual incluye la conversación semidirigida y un cuestionario sociológico, y se complementa con pruebas lingüísticas basadas en el CSCM. El primero resguarda conversaciones que dan cuenta de los oficios tradicionales representativos de la ciudad de Puebla, como la panadería, la siembra de durazno en Huejotzingo o el trabajo con piedra en San Salvador el Seco. En este apartado la autora pone a dialogar dos perspectivas teóricas para el estudio de este tipo de repertorios: la terminología y la lexicología. Lo anterior se debe a que las características altamente referenciales acercan a los repertorios a los léxicos de especialidad y, al conservar su carácter tradicional, los hacen muy interesantes para los estudios lexicológicos. El estar en la frontera entre lo especializado y lo popular ha hecho que los estudios etnográficos o antropológicos reciban mayor atención.

Posteriormente, tenemos una descripción detallada del levantamiento de las 108 entrevistas que conforman el corpus; se describen las variables sociales consideradas, la manera de llevar a cabo la conversación para que, en la medida de lo posible, conservara su carácter semidirigido, y la descripción de las pruebas lingüísticas elicidadas. De la prueba de actitudes y creencias la autora anota resultados comparativos interesantes entre las respuestas de los entrevistados del *Corpus Sociolingüístico de la Ciudad de México* y las de los hablantes que conformaron el corpus poblano en relación con preguntas como “¿Qué es hablar bien?”, en donde se nota que “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 [mientras que para los hablantes de la Ciudad de México] el habla de quienes tienen la autoridad por su conocimiento es relevante” (p. 59). Lo anterior es sólo un ejemplo de las muchas líneas de investigación que estos corpus nos ofrecen.

El tercer capítulo es un ejemplo del diálogo interdisciplinario entre la lexicología, la lexicografía y la ingeniería lingüística. En este apartado la autora propone una metodología para el procesamiento de textos provenientes de un corpus de habla de gran volumen. Este tipo de análisis cuantitativo hace más manejable el corpus para su lematización y el análisis de concordancias. Gracias a ellos, pudo obtener vocablos que califica como compartidos con otras variantes del español, pues aparecen registrados en el leuario del Diccionario del Español de México⁶ (DEM). En cuanto a los vocablos que no aparecen registrados en el DEM, la autora apunta a que pueden tratarse de posibles poblanismos; sin embargo, se requiere de una documentación exhaustiva para corroborar o descartar este estatus, debido a que muchos de ellos son propios de la oralidad y, por lo tanto, es difícil encontrarlos documentados en fuentes escritas; ésa podría ser una de las razones por las que no están registrados en el DEM.

⁶ Versión en línea consultada en <<https://dem.colmex.mx/>>.

En el cuarto capítulo, “Rasgos de la oralidad en PRESEEA-Puebla: reflexiones lexicográficas”, la autora analiza algunos fenómenos propios de la oralidad documentados en el corpus, como acortamientos, cruces léxicos, pseudo sufijación o juegos de palabras; y proporciona una reflexión de las decisiones lexicográficas a tomar para su tratamiento. En este apartado se muestra la complejidad que implica determinar cuándo estas unidades pueden formar parte o no de un diccionario, sea por estar integradas en el sistema, o bien, porque se trata de apariciones propias del contexto motivadas por fines lúdicos.

Por último, en el quinto capítulo, “Notas léxicas del español de Puebla”, la autora nos muestra algunos vocablos susceptibles de ser propios del español poblano, que califica como emblemáticos porque nombran referentes propios de la cultura poblana, lo cual los vuelve grandes señalizadores de identidad. También analiza voces que denominan prácticas socioculturales propias de la región, como *huehue* y sus variantes, o *mosaicos* y *murales*; voces de la gastronomía poblana, como *cemita*, *chalupa*, *chileatole* o *molote*, y variaciones sintácticas y semánticas en verbos como *acusarse*. La autora muestra la complejidad en el análisis semántico de este tipo de unidades, ya que, por ejemplo, en la definición de vocablos propios de la gastronomía “si bien las palabras y sus referentes son parte 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 de cada región” (p. 158); es decir, un mismo nombre puede denominar referentes distintos (modos de preparación diferentes) y cargas culturales específicas en las diversas regiones mexicanas.

En definitiva, *Corpus orales y lexicografía. A propósito de PRESEEA-Puebla*, es producto de muchos años de trabajo y de reflexiones teóricas y metodológicas que abren un gran abanico de interrogantes y de futuras investigaciones. La mirada renovada de la autora da cuenta de su apertura metodológica y teórica al poner a dialogar la lexicología, la lexicografía, la terminología, la sociolingüística y la lingüística computacional. El libro muestra con ejemplos claros y propuestas metodológicas la manera de abordar problemas centrales en la lingüística, en general, y en la de corpus, en particular, como la delimitación de la palabra para su tratamiento en los análisis lexicológicos y lexicográficos. Asimismo, da cuenta de que los análisis computacionales requieren del acompañamiento del investigador, pues el procesamiento puede producir, por ejemplo, segmentaciones morfológicas ambiguas. Los corpus orales, además, arrojan un sinfín de juegos de palabras (que, si bien no formarían parte de repertorios lexicográficos, sí pueden tratarse desde los estudios lexicológicos), los cuales se abordan de manera minuciosa en esta investigación.

Por último, el haber documentado una variedad del español que siempre se ha considerado muy parecida a la más estudiada —la de Ciudad de México— es un hecho muy loable porque, más allá de los datos lingüísticos compartidos por ambas variedades, este corpus arroja datos a nivel de ideologías lingüísticas y de construcción de normatividad que pueden explicar fenómenos pragmáticos, conversacionales, o de cambio y variación lingüística, que permiten caracterizar la comunidad lingüística de estudio como menos o más conservadora, menos o más cortés, y menos o más reticente a la adaptación del préstamo léxico en comparación con las otras comunidades hablantes del español mexicano, entre otros.

BIBLIOGRAFÍA

- CÓRDOVA ABUNDIS, Patricia y Daniel BARRAGÁN TREJO. 2021. *El español hablado en Guadalajara. Corpus PRESEEA-Guadalajara*, Guadalajara: Universidad de Guadalajara-Centro Universitario de Ciencias Sociales y Humanidades.
- MARTÍN BUTRAGUEÑO, Pedro y Yolanda LASTRA (coords.). 2011-2015. *Corpus Sociolingüístico de la Ciudad de México* (CSCM), 1a. ed. Ciudad de México: El Colegio de México.
- MORENO FERNÁNDEZ, Francisco. 2005. “Corpus para el estudio del español en su variación geográfica y social. El corpus PRESEEA”, *Oralia* 8: 129-139. <DOI: <https://dx.doi.org/10.25115/oralia.v8i1.8333>> .
- Proyecto para el Estudio Sociolingüístico del Español de España y América (PRESEEA). Francisco Moreno Fernández (coord.), <<https://presea.uah.es/>> [consultado el 5 de mayo de 2023].
- RODRÍGUEZ ALFANO, Lidia, María Eugenia FLORES TREVIÑO y Tzitel PÉREZ AGUIRRE (comps.). 2010. *Corpus El habla de Monterrey-PRESEEA*. Monterrey: Universidad Autónoma de Nuevo León-Facultad de Filosofía y Letras.