TECNOLOGÍAS DEL LENGUAJE APLICADAS AL PROCESAMIENTO DE LENGUAS INDÍGENAS EN MÉXICO: UNA VISIÓN GENERAL[1]

César Antonio Aguilar Santiago
Universidad Veracruzana (México)
ceaguilar@uv.mx

Hamlet Antonio García Zúñiga
Instituto Nacional de Antropología e Historia - INAH (México)
hamlet_garcia@inah.gob.mx

Recibido: 16/10/2022 — **Aprobado:** 02/02/2023 — **Publicado:** 31/08/2023
DOI: doi.org/10.17533/udea.lyl.n84a04

Resumen: Este artículo ofrece una primera aproximación al estado actual de las tecnologías lingüísticas existentes en México para las lenguas indígenas. Si bien no llega a ser exhaustiva, dada la falta de difusión de éstas, su intención en esta primera cala es mostrar un panorama general, capaz de brindar una idea respecto al estado del arte de tales aplicaciones, vislumbrando su potencial para ayudar no sólo a la conservación digital de tales idiomas, sino también a proyectar una solución viable al retraso informático que sufren estas comunidades de hablantes.

Palabras claves: tecnología lingüística; procesamiento del lenguaje natural; lenguas indígenas mexicanas; brecha tecnológica; política lingüística.

LANGUAGE TECHNOLOGIES APPLIED TO THE PROCESSING OF INDIGENOUS LANGUAGES IN MEXICO: AN OVERVIEW

Abstract: This article exposes a first overview to the current situation of existing linguistic technologies in Mexico for indigenous languages. Although it is not exhaustive, because of the lack of dissemination of these, the goal is to show a general landscape, capable of providing an idea regarding the state of the art of such applications, showing their potential to help not only to the digital preservation of such languages, but also to project a viable solution to the computer delay suffered by these communities of speakers.

Keywords: linguistic technology; natural language processing; Mexican indigenous languages; digital divide; language policy.

[1] Artículo producto de ponencia presentada en el v Congreso Internacional de Investigación Lingüística, realizado entre el 7 y 9 de septiembre de 2022 en la Universidad de Antioquia.

1. Introducción

Desde la década del 2000 hasta hoy —coincidiendo con el auge de las investigaciones en procesamiento del lenguaje natural (o PLN) en América Latina— se han hecho varias propuestas por promover la conservación y revitalización de lenguas indígenas. Ejemplo de ello son los trabajos de Hinton & Hale (2001), Dyson, Hendriks & Grant (2007), así como Dyson, Grant & Hendriks (2016), quienes esbozan proyectos enfocados en el diseño e implementación de corpus lingüísticos, sistemas anotados, diccionarios electrónicos y otros recursos similares para idiomas nativos tanto en Canadá como en Estados Unidos.

Por otra parte, desde un enfoque propio de la lingüística computacional, se han generado varias herramientas electrónicas focalizadas en resolver el procesamiento acústico, léxico y sintáctico de varias lenguas indígenas, entre ellas el náhuatl, el otomí, el mixteco, el mixe y el zapoteco —las cuales se hablan en México—, expuestas en foros internacionales organizados por asociaciones relevantes al área, como la Association for Computational Linguistics (Mager, Oncevay, Ríos, Meza, Palmer, Neubig & Kann, 2021).

Cabe destacar aquí trabajos focalizados en el tratamiento morfológico, como es el caso de Medina (2007; 2008) quien ha llevado a efecto experimentos relacionados con el descubrimiento automático de afijos para el chuj —lengua perteneciente a la familia maya— y rarámuri —lengua miembro de la familia yuto-azteca—; junto con los de Mager (2017), Mager, Carrillo & Meza (2018), Mager & Meza (2018), los cuales hacen uso de modelos probabilísticos para analizar la morfología del wixarika o huichol, aprovechando sus resultados para la implementación de un sistema de traducción automática de tal lengua al español.

En México, desde el 2005, el Instituto Nacional de Lenguas Indígenas (INALI) se ha enfocado en la documentación y conservación de las 68 agrupaciones lingüísticas reconocidas en el país, integrando así una enorme colección de datos orales, escritos y multimedios que dan cuenta de su realidad actual.

Asimismo, destacan varias propuestas gestadas desde universidades y centros de investigación, tales como corpus paralelos náhuatl-español (Gutiérrez-Vázquez, Sierra & Hernández, 2016; Gutiérrez-Vázquez & Mijangos, 2018), así como en ch'ol y maya —lenguas pertenecientes a la misma familia—, mazateco, mixteco y otomí con el español (Gutiérrez-Vázquez, 2015) desarrollados por el Grupo de Ingeniería Lingüística (2003); el Acervo Digital de Lenguas Indígenas, dirigido por José Antonio Flores en el Centro de Investigaciones y Estudios Superiores en Antropología Social (2014); o la colección de estudios descriptivos que integran el Archivo de Lenguas Indígenas de México, coordinado por El Colegio de México.

De la misma manera, Aguilar y Acosta (2021) ofrecen una panorámica de varios proyectos gestionados por iniciativas independientes, como Mozilla Nativo (2018), el cual se enfoca en crear

ISSN 0120-5587
E-ISSN 2422 3174
N.º 84, 2023

versiones de este buscador web para lenguas tales como el náhuatl, el maya, el mixteco o el zapoteco, por mencionar algunas. En línea con esto, Mager, Gutiérrez-Vázquez, Sierra & Meza (2018) brindan un estado del arte bastante detallado en donde se exponen los principales desafíos a considerar para la generación de recursos computacionales para las lenguas indígenas americanas, poniendo especial énfasis en su tratamiento morfológico, así como en su traducción automática.

Finalmente, la Universidad Michoacana de San Nicolás de Hidalgo ha venido organizando de forma anual desde el 2020 un congreso internacional de PLN para lenguas indígenas, el cual ha tenido un impacto positivo en el medio académico nacional con el fin de fomentar el desarrollo de modelos y recursos computacionales para el tratamiento de idiomas nativos (García, Saenz, López & Hurtado, 2021).

Dicho lo anterior, en este trabajo se continúa con dicha revisión del panorama, con miras a llevar a cabo posteriormente un censo que ayude a determinar cuántos recursos tecnológicos existen hoy para las lenguas nativas mexicanas, siguiendo en esto el reporte que hacen Littel y sus colegas respecto a las tecnologías del lenguaje disponibles en Canadá para sus idiomas nativos (2018). A grandes rasgos, los objetivos que sustentan esta propuesta son: —Identificar la mayor parte de los recursos computacionales disponibles hasta ahora (2020) que ejecuten algún tipo de procesamiento para una lengua nativa mexicana.—Clasificarlos conforme a los criterios que plantea Littel, Kazantseva, Kuhn, Pine, Arppe, Cox & Junker (2018), considerando si tales herramientas se enfocan en el procesamiento oral, textual o mixto. De igual manera, se indicará si se tratan de recursos de consulta (p. ej., corpus lingüísticos), o si ejecutan alguna clase de análisis (p. ej., gramáticas computacionales).—Reconocer y vincular tales recursos con la familia lingüística a la cual pertenece la lengua que procesa, así como ilustrar su localización geográfica.

La metodología a seguir en esta revisión es de tipo documental, por lo que se da prioridad a una exploración de todas las fuentes bibliográficas y digitales posibles. Dadas las limitaciones de tiempo, en esta primera aproximación se descarta encuestar a grupos de nativo-hablantes, enfatizando más bien el acceso a tales recursos vía Internet. El periodo límite que se dedicó a hacer tal exploración cubre hasta el 2020 —es necesario precisar que, gracias al interés que hay en el campo, las cifras pueden incrementarse sustancialmente después de dicho año—, con miras a obtener la información básica sobre dichas herramientas y recursos. Posteriormente, se harán actualizaciones del censo y se verificará si el esquema propuesto es útil para el registro completo de las tecnologías o si requiere de algún tipo de ajuste.

En este sentido, cabe justificar la pertinencia de un censo de este tipo, considerando su impacto positivo en la realización de tareas tales como la documentación de lenguas, la planificación y el diseño de políticas lingüísticas, o bien, la reducción de la brecha digital dentro de las comunidades indígenas mexicanas. Para concluir, y de forma breve, los resultados que se muestran en este trabajo son: (i) un listado sobre recursos y herramientas de PLNaplicables a las familias lingüísticas de México:

ISSN 0120-5587
E-ISSN 2422 3174
N.º 84, 2023

yuto-azteca, maya, totonaco-tepehua, otomangue, álgica, mixe-zoque, seri, purépecha, chontal de Oaxaca, huave y yumana; (ii) una clasificación respecto a los procesos que ejecuten tales recursos y herramientas; (iii) una planeación esquematizada de las fases a seguir en la realización del censo, con miras a establecer colaboraciones con universidades, centros de investigación e iniciativas privadas que estén interesados en desarrollo y uso de estos recursos.

2. Lenguas indígenas mexicanas: una panorámica general

Desde la época prehispánica México ha sido un crisol de lenguas y culturas, comenzando por la olmeca, de la cual se tienen datos de sus inicios a partir del año 3000 a.C., a la cual se atribuye un rol civilizatorio que tuvo un gran impacto en las zonas centro y sureste del país, principalmente, a la que se le ha denominado Mesoamérica, en especial en la zona maya (Península de Yucatán), totonaca (estados de Puebla y Veracruz) e incluso en la esfera tolteca (valle central de México) (Suárez, 1983; Soustelle, 1979; 1982; Dihel, 2005). Hoy en día se identifican alrededor de 68 agrupaciones lingüísticas habladas en el país. El Instituto Nacional de Estadística y Geografía (INEGI) en el año 2010 —tras realizarse el censo poblacional— identificó que alrededor de 6 millones de personas hablaban alguna de las variantes de dichas agrupaciones. Por su parte, la antigua Comisión Nacional para el Desarrollo de los Pueblos Indígenas (CDI), hoy Instituto Nacional de los Pueblos Indígenas, reportó que el total de población indígena en el país gira en torno a los 12,7 millones de personas, de las cuales unos 7 millones hablaban al menos una.

Entrando en mayores detalles, en la siguiente tabla se brindan cifras respecto a las 10 agrupaciones lingüísticas con mayor número de hablantes, los cuales son:

Grupo indígena	Población (2020)
Náhuatl	1,651,958
Maya yucateco	774,755
Tseltal	589,144
Tsotsil	550,274
Mixteco	526,593
Zapoteco	490,854
Otomí	298,861
Totonaco	256,344
Ch'ol	254,715
Mazateco	237,212

Tabla 1. Lista de las diez lenguas nativas más habladas en México

ISSN 0120-5587
E-ISSN 2422 3174
N.º 84, 2023

Como se puede observar en la Tabla 1, la agrupación que cuenta con mayor número de hablantes es el náhuatl, con un poco más de un millón y medio de personas; en tanto la siguiente agrupación con más hablantes es la maya yucateca, con cerca de ochocientos mil hablantes. Posteriormente, las siguientes agrupaciones (tseltal, tsotsil, mixteco y zapoteco) se ubican en un rango de medio millón de hablantes, en tanto que los últimos se ubican en un promedio de un cuarto de millón. Aquí cabe hacer algunas observaciones:

—De acuerdo con Barriga Villanueva (1995[a]; 1995[b]; 2018), así como Lara & Vázquez Laslop (2015), desde los tiempos de la Colonia el reconocimiento y aceptación de una realidad plurilingüe en México ha dado lugar a una disputa que ha tenido un impacto directo en la implementación de las políticas lingüísticas nacionales: (i) la imposición del español —vía un sistema educativo monolingüe— como un elemento unificador en los planos social, político, cultural y étnico, poniendo en riesgo la existencia de las lenguas indígenas, y (b) en contraparte, la aceptación de tal plurilingüismo —asumido en el marco constitucional mexicano como un rasgo definitorio del país—, a pesar de que ello pusiese en crisis el concebir a México como una nación unitaria, haciendo patente la necesidad de fijar un marco legal pertinente que acepte y defienda tal diversidad (Navarrete, 2016; Stahler-Sholk y Baronnet, 2017).

—Relacionado con el reconocimiento del multilingüismo y multiculturalismo en México, cabe mencionar que tras el levantamiento del Ejército Zapatista de Liberación Nacional (EZLN) en 1994, junto con la redacción y firma de los *Acuerdos de San Andrés Sakamch'en/Larraizar* en 1996, se ha dado lugar a una reflexión en torno a la conservación y difusión de las lenguas indígenas, concibiendo ambas acciones como derechos fundamentales para los pueblos originarios (Pellicer, 1997; Izquierdo, 2005; Barriga Villanueva, 2018). Así, desde la década del 2000, el Gobierno de México ha emprendido una serie de reformas jurídicas que han dado lugar a la *Ley General De Derechos Lingüísticos De Los Pueblos Indígenas* (2003), la cual enfatiza el papel que juega la educación bilingüe como un elemento clave para fomentar la actualidad de las lenguas indígenas. Ejemplo de esta revalorización lingüística por parte del Gobierno es —justo— la puesta en operaciones del INALI en 2005, cuyas labores principales son el fortalecimiento, la preservación y desarrollo de las lenguas indígenas mexicanas.

—Teniendo en cuenta dichas acciones por parte del gobierno mexicano para apoyar la conservación y difusión de las lenguas indígenas, muchas comunidades se han involucrado en proyectos de revitalización y difusión que, poco a poco, han ido fomentando una percepción más positiva respecto a éstas. En el caso de la lengua náhuatl, de acuerdo con Olko & Sullivan (2013; 2014), se han ido implementado políticas lingüísticas que involucran la participación de hablantes y gobierno, las cuales han tenido un impacto positivo en su difusión, especialmente en el contexto educativo, tanto al nivel básico (Gomashie, 2021; Melton-Villanueva, de la Cruz & Cruz, 2022) como universitario (Figueroa, Alarcón, Bernal & Hernández, 2014; Bernal & Figueroa, 2019). Con todo, sigue habiendo todavía un considerable grado de discriminación hacia esta lengua, a pesar de su propagación incluso en Estados Unidos, como observa Villareal (2011) en las comunidades chicanas de Los Ángeles.

—Respecto a la familia maya, ésta conjunta varios idiomas, entre los cuales se cuenta al tseltal, al tsotsil y al ch'ol. La familia maya es una de las que tiene un mayor número de hablantes e, igualmente, es una de las que más territorio cubre (Península de Yucatán y sur de México, Belice, Guatemala, y algunas partes de El Salvador y Honduras). Un aspecto para añadir aquí es que —al igual que el náhuatl— sus hablantes se han involucrado en procesos de creación y gestión de políticas lingüísticas cuyo impacto ha sido positivo para revitalizar su uso y aprendizaje vía la educación bilingüe (Craveri, 2011; Briseño, 2020; 2021), junto con el uso de medios digitales (Montejo, Bastiani & Orantes, 2019).

ISSN 0120-5587
E-ISSN 2422 3174
N.º 84, 2023

—Finalmente, el mayor número de hablantes de mixteco, zapoteco y mazateco se ubican hacia el oeste de México, una zona cercana al Pacífico, en tanto que el otomí se localiza entre la zona centro del país y el Bajío, y el totonaco se sitúa hacia la costa del Golfo de México.

Tal distribución la podemos observar en el siguiente mapa:

Figura 1. Distribución poblacional de hablantes de alguna lengua indígena mexicana ubicados por estados
Nota. Mapa extraído de la encuesta intercensal INEGI (2015).

De acuerdo con este mapa y considerando las observaciones hechas a la Tabla 1, podemos observar que es en el sur de México en donde se localizan las mayores poblaciones de grupos indígenas, destacando particularmente Oaxaca, Chiapas (suroeste del Pacífico) y Yucatán (Península de Yucatán), en donde casi el 40 % de sus habitantes puede hablar y comprender al menos un idioma indígena. De ahí, estados como Quintana Roo y Campeche (península de Yucatán), Guerrero, Nayarit (suroeste y centro del Pacífico), San Luis Potosí, Hidalgo y Puebla (noreste y centro de México) cuentan con una densidad de hablantes que va de un 30 % a casi un 15 % de su población. Por su parte, el estado de Veracruz (Golfo de México), muestra una distribución de hablantes que gira en torno al 10 y 15 %, en donde se destacan principalmente las lenguas náhuatl y totonaca. Finalmente, en el resto del país, de manera especial en la región norte, son pocas las personas que se reconocen como hablantes de alguna

ISSN 0120-5587
E-ISSN 2422 3174
N.º 84, 2023

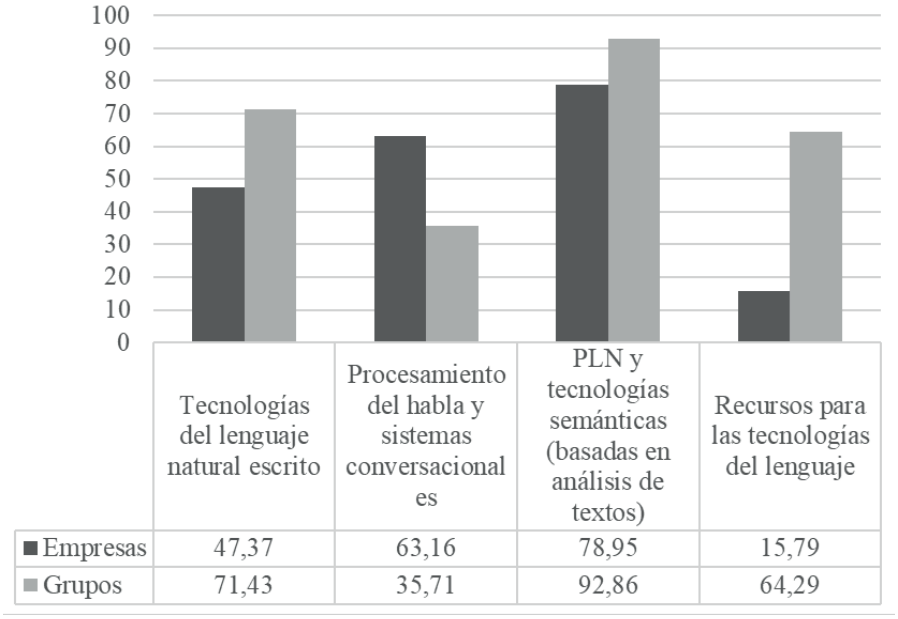
lengua indígena, ya que su distribución llega hasta un 5 % entre sus habitantes.

Teniendo en cuenta tanto el número de hablantes como su distribución a lo largo del territorio, muchas de las tecnologías que se desarrollan se orientan sobre todo hacia los grupos más amplios, como el náhuatl y el maya —aunque también se han elaborado al interior de la familia con el mismo nombre, maya, recursos como corpus lingüísticos para el tseltal, el tsotsil y el ch'ol—, en tanto que otros idiomas reciben menor atención al respecto.

Ahora bien, en México la lengua que cuenta con la mayor cantidad de recursos computacionales para su procesamiento automático es, sin duda, el español. Veamos a continuación entonces el estado actual de las tecnologías lingüísticas existentes para las lenguas indígenas mexicanas.

3. Tecnologías del lenguaje en México

En octubre de 2018 la empresa Track Global Solutions (TGS), en colaboración con la Sociedad Española de Procesamiento del Lenguaje Natural (SEPLN), así como el Grupo de Ingeniería Lingüística de la Universidad Nacional Autónoma de México (GIL-UNAM), realizó un censo preliminar, el cual les permitiría fijar un panorama amplio sobre el desarrollo y uso de tecnologías lingüísticas en México, considerando que se trata de un país estratégico para el desarrollo y comercialización de tales tecnologías. De acuerdo con dicho censo, los productos más comercializados para procesamiento del idioma son los siguientes:



	Tecnologías del lenguaje natural escrito	Procesamiento del habla y sistemas conversacionales	PLN y tecnologías semánticas (basadas en análisis de textos)	Recursos para las tecnologías del lenguaje
Empresas	47,37	63,16	78,95	15,79
Grupos	71,43	35,71	92,86	64,29

Figura 2. Recursos de PLN que se comercializan en México
Nota. Estadística tomada de TGS (2018).

ISSN 0120-5587
E-ISSN 2422 3174
N.º 84, 2023

De acuerdo con la Figura 2, la mayoría de las herramientas de PLN que se adquieren en México se enfocan en el procesamiento semántico, lo que cubre tareas de recuperación y extracción de información, minería de textos, ontologías, así como la generación de vectores de palabras (ing. *word embedding*). De ahí, los demás productos se orientan hacia el procesamiento de texto —p. ej., motores de búsqueda, reconocedores de caracteres, correctores ortográficos, etc.-, el procesamiento de habla y sistemas conversacionales —mayoritariamente, *chatbots*—, así como la generación de recursos e insumos, como es el caso de los corpus lingüísticos —orales y escritos—. Finalmente, existen dos claros proveedores de estos recursos: consultoras y empresas tecnológicas —identificadas como *empresas*—, junto con laboratorios y grupos de investigación universitarios —etiquetados como *grupos*—.

Complementario a esto, también se reconoce a los principales clientes que adquieren esta clase de tecnología, lo cual se puede ver en el siguiente gráfico:

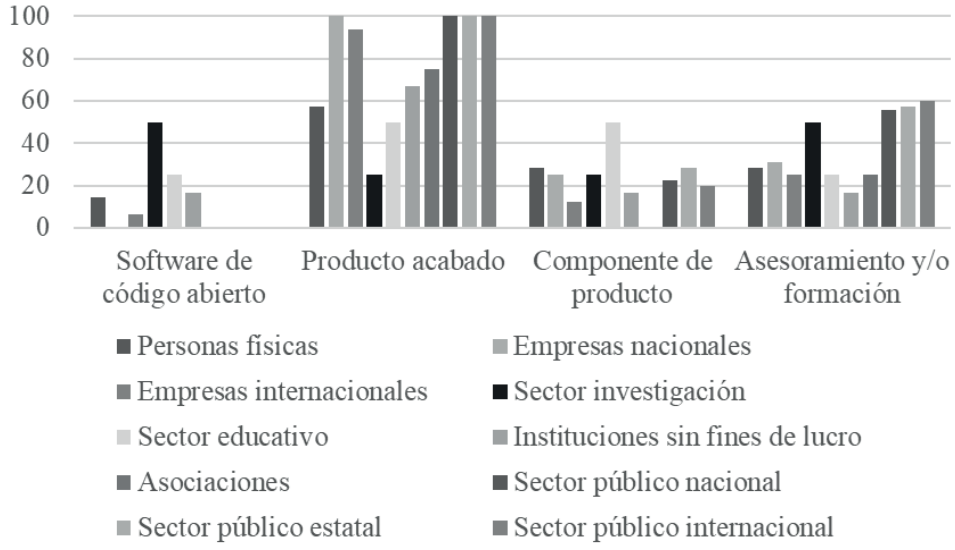


Figura 3. Principales consumidores de herramientas de PLN
Nota. *Estadística tomada de* TGS (2018).

Con base en la información de la Figura 3, el principal consumidor de estos recursos es el sector público gubernamental, cuyo mecanismo de adquisición son los fondos para el fomento de la investigación y creación de nuevas tecnologías. En este sentido, también hay una participación importante de empresas nacionales e internacionales. Todas ellas apuntan su interés a la compra de productos concluidos, ya sean prototipos, ya sean sistemas validados listos para su manufactura y distribución. Por otro lado, en rubros como el asesoramiento tecnológico, la formación de recursos humanos, la creación de componentes o la generación de software libre son del interés de sectores como la academia, asociaciones sin fines de lucro y personas físicas que requieren de herramientas

concretas para resolver sus necesidades.

Mirando con un poco más de detalle este panorama, también es posible ubicar a los agentes que diseñan e implementan dichos productos, de acuerdo con este gráfico:

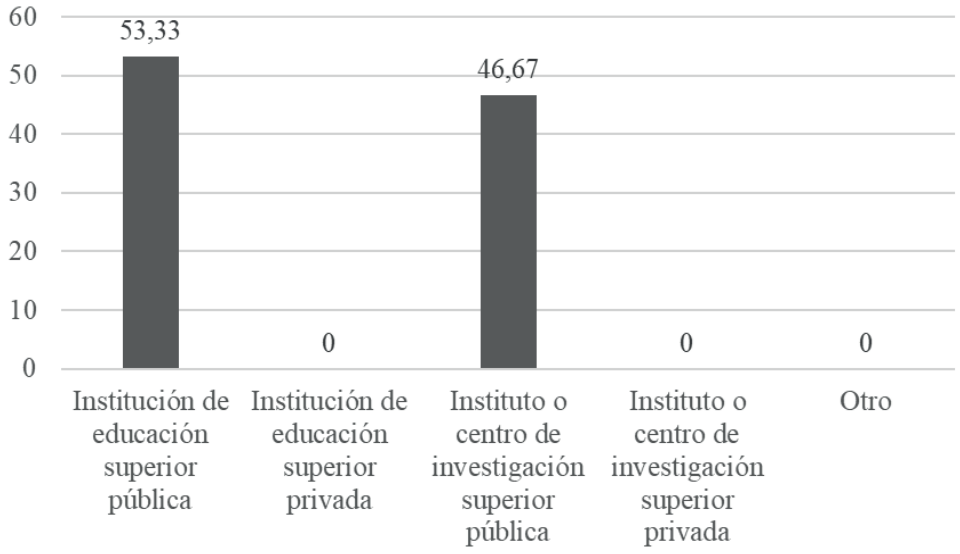


Figura 4. Agentes desarrolladores de sistemas de PLN

Nota. Estadística tomada de TGS (2018).

En concreto, la Figura 4 deja en claro que el desarrollador que juega un papel central en la creación de estos recursos electrónicos es el sector académico, ya sea desde universidades —preferentemente públicas, aunque también hay una incursión de las privadas—, ya sea desde centros de investigación financiados por el gobierno. Contrastando esta participación con la que tiene el sector privado, éste no se involucra en la fase de diseño y desarrollo, sino que, como se ha visto en la Figura 3, lo hace ya en una etapa final, cuando se tiene ya un prototipo o, en el mejor de los casos, el producto terminado.

Para concluir con esta sección, en la siguiente figura se muestra cuáles son las áreas de conocimiento con mayor actividad dentro del área de PLN. En resumen, se observa una tendencia que se ha generalizado en buena parte de América Latina, que es considerar a dicha área como un asunto para grados y posgrados en ciencias de la computación, ingeniería computacional o de *software*, informática y otras similares. En contraparte, si bien hay un interés por parte de carreras humanísticas en esta área, lo cierto es que sigue siendo una tarea preferentemente técnica. Veamos:

ISSN 0120-5587
E-ISSN 2422 3174
N.º 84, 2023

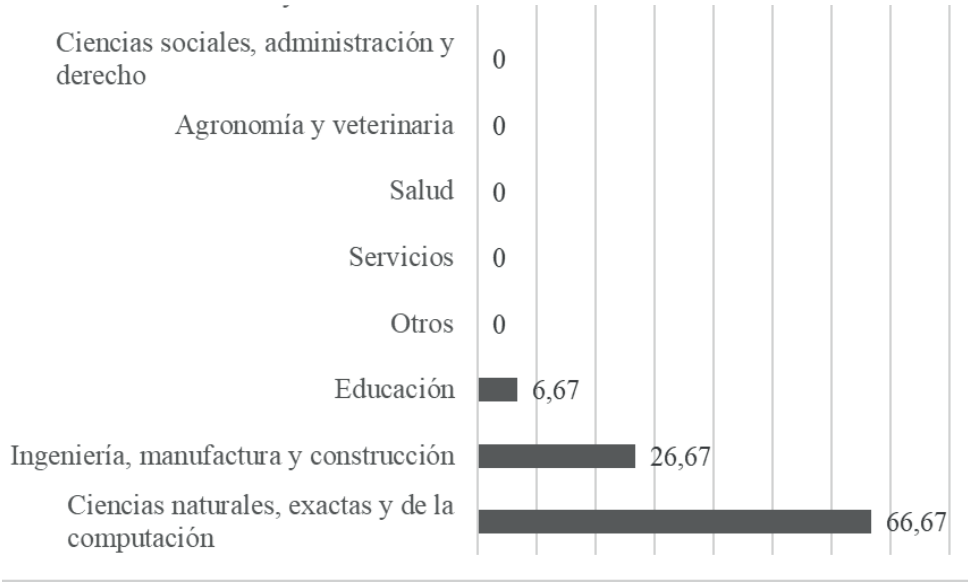


Figura 5. Principales áreas de conocimiento involucradas en tareas de PLN

Nota. Estadística tomada de TGS (2018).

Si bien es cierto que sigue existiendo todavía una mayor involucramiento de las carreras técnicas en el desarrollo del PLN en México, en los últimos años varios estudiantes, profesores e investigadores de áreas humanísticas —aparte de la lingüística teórica y aplicada— han ido incursionando en diferentes labores que han tenido un impacto positivo en la consolidación de dicha área, tales como el proceso de anotado de un corpus lingüístico, hasta desarrollar diversos análisis que son útiles para evaluar el desempeño de herramientas de análisis, como es el caso de reconocedores de habla, el diseño de *chatbots*, o el de las gramáticas computacionales.

4. Brecha digital: claroscuros en el contexto tecnológico mexicano

Teniendo en mente la información que arroja el censo hecho por TGS, se puede ver que México cuenta con un desarrollo considerable en la implementación y acceso a recursos computacionales orientados al procesamiento lingüístico, entre otras tecnologías a las cuales tiene acceso. Empero, a pesar de tal desarrollo, el país tampoco está exento de sufrir uno de los problemas que más se han generalizado en América Latina en los últimos años: la brecha digital. A grandes rasgos, este término alude a la distribución desigual en el acceso, el uso o el impacto de las tecnologías de la información y la comunicación entre grupos sociales (Maitland, 1985; Compaine, 2001; Aguilar & Acosta, 2021). Dicha desigualdad abarca desde la carencia de equipo de cómputo, la falta de acceso a Internet y otros recursos en línea, hasta el ancho de banda que brinda tal conexión. Un ejemplo de tal desigualdad se observa

ISSN 0120-5587
E-ISSN 2422 3174
N.º 84, 2023

en la distribución respecto al acceso a Internet que tuvo México durante el año 2020. De acuerdo con un informe elaborado por el Instituto Federal de Telecomunicaciones (IFT), denominado: *Encuesta Nacional sobre la Disponibilidad y Uso de Tecnologías de la Información en los Hogares* (ENDUITH, 2020), hay dos datos significativos a considerar aquí:

> —Respecto a la distribución de este servicio entre zonas urbanas y rurales, se observa que la mejor conectividad se da justo en las primeras: el 70.8 % de la población que habita las ciudades se conecta a Internet, en tanto que la que radica en el campo llega al 50.4 %. Entrando en mayores detalles, hay que tomar en cuenta que en el primer caso este porcentaje representa una población de 70.8 millones, en tanto que la segunda cubre a 13.3 millones. Al respecto, cabe mencionar que la mayoría de las comunidades indígenas mexicanas habita en las zonas rurales, aunque se dan flujos migratorios a espacios conurbanos.
>
> —Contrastando la distribución de la población que cuenta con acceso a Internet por estados, destacan sobre todo los estados ubicados en la franja norte del país (frontera con Estados Unidos), como es el caso de Nuevo León, que cuenta con la mayor cantidad de usuarios (84.5 % de toda su población). Por su parte, la franja central del país también cuenta con un amplio número de usuarios conectados, destacando la Ciudad de México (84.5 %). En contraste, las que cuentan con menos usuarios conectados son la parte suroeste del Pacífico (particularmente Michoacán, Guerrero, Oaxaca y Chiapas), varios estados de la zona norte-centro (Zacatecas, Guanajuato, Hidalgo, Puebla), y Tabasco, un estado ubicado hacia el Golfo de México. Resulta importante reconocer aquí que los estados de Oaxaca y Chiapas son los que cuentan con la mayor diversidad cultural y lingüística de México, ya que en ellos habitan un gran número de comunidades indígenas y son, también, los que menor acceso a Internet ofrecen.

Otro ejemplo significativo respecto a esta diferenciación digital es el lento avance de tecnologías informáticas, tales como computadoras, teléfonos inteligentes, servidores y otros productos similares. Veamos la Figura 6

ISSN 0120-5587
E-ISSN 2422 3174
N.º 84, 2023

Indicadores sobre disponibilidad y uso de TICs	2015	2016	2017
Hogares con computadora	44.9 %	45.6 %	45.4 %
Hogares con conexión a Internet	39.2 %	47 %	50.9 %
Hogares con televisión	93.5 %	93.1 %	93.2 %
Hogares con televisión de paga	43.7 %	52.1 %	49.5 %
Usuarios de computadoras	51.3 %	47 %	45.3 %
Usuarios de Internet	57.4 %	59.5 %	63.9 %
Usuarios de Internet que realizan transacciones por esta vía	12.8 %	14.7 %	20.4 %
Usuarios de Internet que acceden desde fuera del hogar	29.1 %	20.5 %	16.7 %

Figura 6. Disponibilidad de recursos informáticos de 2015 a 2017

Nota. Estadísticas tomadas de la Encuesta ENDUTIH (2020).

Un aspecto importante para distinguir en los datos que ofrece la Figura 6 es que, al menos hasta 2017, el contar con una computadora personal o un teléfono inteligente era algo a lo que sólo podía acceder casi la mitad de los mexicanos. En contraste, la televisión —en especial, aquella cuya señal transmite canales públicos— es la tecnología de cobertura más amplia, incluyendo zonas rurales habitadas por comunidades indígenas.

Una pregunta que puede hacerse al respecto es: ¿cuál es la tecnología a la que tienen acceso los pueblos indígenas? Al respecto, el Banco Mundial realizó un estudio comparativo entre varios países latinoamericanos con población indígena —incluido México— con miras a resolver esta cuestión. Se consideraron 3 tipos de tecnologías: (i) computadoras personales, (ii) teléfonos inteligentes y (iii) acceso a puntos de Internet. Estos fueron los resultados reportados:

ISSN 0120-5587
E-ISSN 2422 3174
N.º 84, 2023

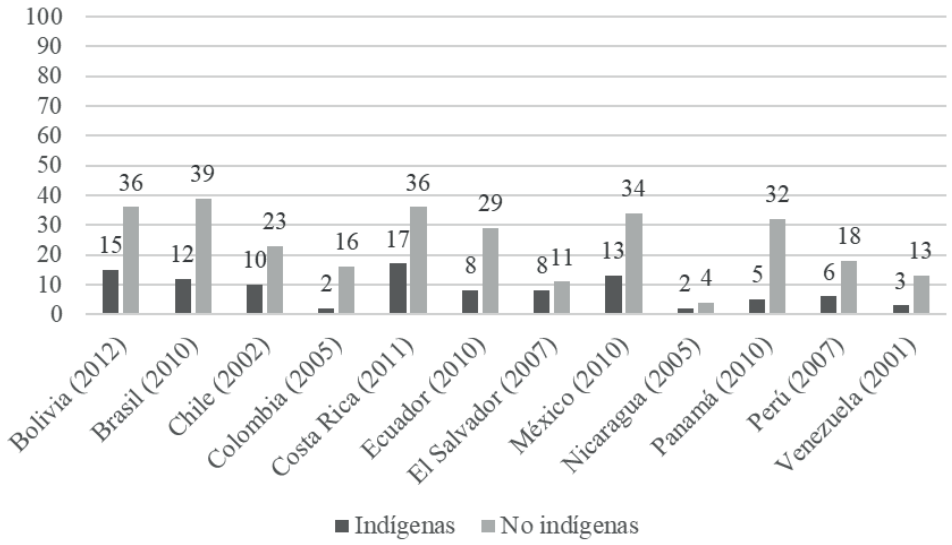


Figura 7. Acceso a computadoras por parte de comunidades indígenas en América Latina

Nota. Estadísticas tomadas del Reporte del Banco Mundial (2015).

Como se puede observar en el gráfico, la mayor parte de las comunidades indígenas de América Latina —en contraste con grupos poblacionales que habitan en medios suburbanos y rurales— no posee una computadora en sus hogares, sea de escritorio o portátil. El país en donde estos grupos tienen más posibilidades de adquirir un equipo de este tipo es Bolivia (15 %), en tanto que los países con más dificultades para acceder a uno de estos aparatos son Colombia (2 %), Nicaragua (2 %) y Venezuela (3 %), al menos del 2001 al 2005.

Con relación a la telefonía celular, los datos son los siguientes:

ISSN 0120-5587
E-ISSN 2422 3174
N.º 84, 2023

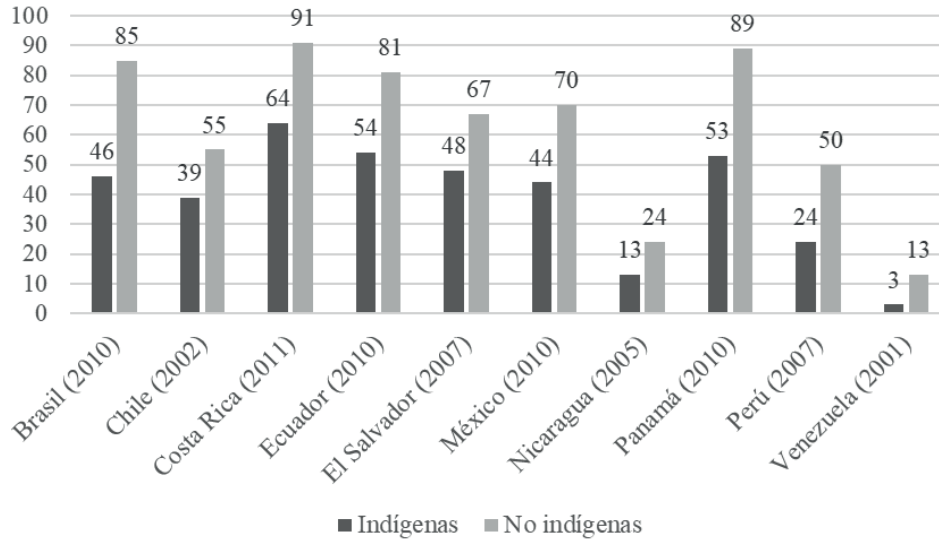


Figura 8. Acceso a teléfonos celulares por parte de comunidades indígenas en América Latina

Nota. Estadística tomada del Reporte del Banco Mundial (2015).

Se puede ver aquí un contraste notorio respecto a la posesión de computadoras: la telefonía celular —al menos hasta el 2010— es un recurso mucho más accesible, particularmente en países como Ecuador (54 %), El Salvador (48 %), Brasil (46 %) y México (44 %). En este rubro, sólo Nicaragua tiene una tasa reducida de usuarios indígenas (13 %).

Finalmente, con relación al acceso al Internet, se tiene lo siguiente:

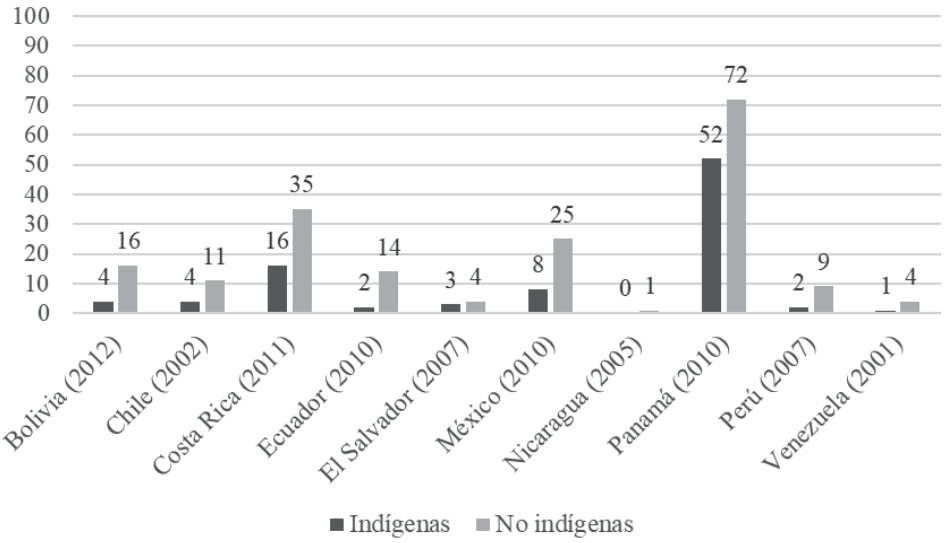


Figura 9. Acceso a Internet por parte de comunidades indígenas en América Latina

Nota. Estadística tomada del Reporte del Banco Mundial (2015).

ISSN 0120-5587
E-ISSN 2422 3174
N.º 84, 2023

Los datos que reporta la Figura 9 ayudan a corroborar que en América Latina el acceso a Internet sigue siendo un cuello de botella para superar la brecha digital: si bien es cierto que los celulares son accesibles para estos grupos, hace falta garantizar una conexión a este recurso, ya que de otra manera se reduce significativamente sus beneficios como vehículo de acceso a nuestro mundo digital.

5. PLN y lenguas indígenas: el estado actual

Después de haber presentado un panorama general respecto a las tecnologías del lenguaje en México, así como haber descrito los retos que implica la brecha digital entre las comunidades indígenas, sobre todo si se tiene en mente diseñar e implementar esta clase de tecnología para tales comunidades, veamos algunos de los proyectos que plantean soluciones a esta cuestión. Comencemos con una de las instancias que más ha fomentado el desarrollo de materiales y recursos lingüísticos —tanto impresos como electrónicos: el Instituto Nacional para las Lenguas Indígenas (INALI). Cabe destacar aquí que uno de sus grandes hitos es la labor que ha efectuado respecto a la documentación minuciosa de las agrupaciones lingüísticas del país, ayudando así a su conservación y difusión.

Gracias a esta documentación realizada por el INALI, se han podido emprender otros proyectos, en concreto aquellos que contemplan la creación de corpus lingüísticos. En este sentido, cabe destacar los experimentos realizados por el citado GIL-UNAM, quienes han implementado un corpus paralelo náhuatl-español llamado Axolotl (2015), el cual ha sido construido a partir de textos traducidos principalmente por misioneros cristianos durante los siglos XVI y XVII en México (Gutiérrez, 2015; Gutiérrez, Sierra & Hernández, 2016). El trabajo realizado en torno al Corpus Axolotl ha permitido evaluar el desempeño de herramientas automáticas para el análisis morfológico del náhuatl (Gutiérrez & Mijangos, 2018; Gutiérrez, Medina & Sierra, 2019).

Estos trabajos también han sido útiles para explorar un fenómeno como la complejidad lingüística, entendida como el grado que posee una lengua para generar combinaciones de elementos, ponderando así la capacidad de procesamiento requerida para realizar tales permutaciones (Juloa, 2008; Sampson, Gil & Trudgill, 2009; Baechler & Seiler, 2016). Si dicha ponderación se traduce en un parámetro medible, es factible realizar comparaciones entre lenguajes humanos, así como evaluar qué tan difícil es o no para un sistema informático realizar algún procesamiento lingüístico, particularmente a nivel morfológico y sintáctico.

Para concluir con este apartado, cabe destacar que el Corpus Axolotl ha tenido un impacto relevante dentro de la comunidad de PLN en México, fomentando el desarrollo de otros corpus similares, como es el caso del Corpus Paralelo de Lenguas Mexicanas (2016), el cual incluye datos de ch'ol, maya yucateco, mazateco, mixteco, náhuatl y otomí.

Otra iniciativa para considerar es la creación de portales y motores de búsqueda web para lenguas

indígenas por medio de Mozilla Firefox, en concreto bajo la guía de la iniciativa Mozilla Nativo, una comunidad que se encarga de adaptar dicho buscador a los idiomas originarios (Gómez, 2019), logrando avances importantes, en particular con el maya yucateco. Al respecto, López (2015) ofrece un ejemplo de cómo sería hacer búsquedas con Mozilla en mixteco:

Figura 11. Ejemplo de búsqueda en mixteco a través de Mozilla Nativo
Nota. Tomado de López (2015).

Entre las variantes y agrupaciones lingüísticas mexicanas que cuentan hoy con su versión de buscador Mozilla están:

—Mixteco de Mixtepec (Oaxaca)
—Mixteco de Yucuhiti (Oaxaca)
—Triqui (Oaxaca)
—Maya yucateco (Yucatán, Campeche y Quintana Roo)
—Ixil (Campeche y Quintana Roo)
—Zapoteco de Miahuatlán (Oaxaca)
—Náhuatl (Valle de México, Guerrero, Puebla, Tlaxcala, Morelos, Hidalgo, Veracruz, Tabasco, Michoacán, Jalisco)
—Purépecha (Michoacán)

Del mismo modo, existen otras lenguas indígenas latinoamericanas que cuentan con su propia versión de Mozilla Nativo, específicamente:

—Guaraní (Argentina, Bolivia, Brasil y Paraguay)
—Kaqchikel (Guatemala)

ISSN 0120-5587
E-ISSN 2422 3174
N.º 84, 2023

—K’iche’ (Guatemala)
—Kichwa (Colombia, Ecuador y Perú)
—Nahua Pipil (El Salvador)
—Quechua (Argentina, Bolivia, Chile, Colombia, Ecuador y Perú)

Un rubro relevante que se ha explorado dentro de las tecnologías lingüísticas para dichos idiomas es la generación de aplicaciones para su aprendizaje, ya sea para hablantes nativos de éstas o para personas que las adquieren como una segunda lengua. Aquí caben destacar recursos como Ko’ox App (Poot Cahun, 2019), cuya función es la de ser un asistente virtual para aprender maya yucateco, portable para teléfonos celulares. Otra iniciativa es la aplicación Vamos a aprender (2016), desarrollada por el Laboratorio de Ciudadanía Digital (LCD) y coordinada por el INALI. Entre los idiomas que cubre tal aplicación están el náhuatl, mixteco y purépecha. También existen propuestas venidas desde organizaciones civiles que han creado aplicaciones para lenguas menos conocidas, como es el caso del mixe, con Kumoontun (De Malignon, 2019), implementada por una fundación que lleva el mismo nombre.

Como punto relevante hay que decir que, inclusive, se ha explorado el diseño de aplicaciones genéricas, las cuales pueden adaptarse a cualquier idioma indígena, sin necesidad de crear recursos específicos para ello. Éste es el caso de Miyotl (Álvarez, 2019), gestada por la Universidad Autónoma de Chapingo. Con la finalidad de hacer una síntesis de todos los recursos que hay hasta hoy en día para las lenguas originarias, presentamos la siguiente tabla:

Rubro	Descripción	Usuarios	Lenguas	Desarrollador	Tipo
Cursos	Desarrollo de recursos para la enseñanza-aprendizaje de lenguas indígenas nacionales	Aprendices y enseñantes de segundas lenguas	Varias (tseltal, triqui, matlatzinca, chinanteco, otomí, náhuatl, huave, mixteco, huichol, ch’ol, mixe)	Instituto Nacional de Lenguas Indígenas, Universidad Pedagógica Nacional, Universidad Intercultural del Estado de México, Universidad Autónoma de Nayarit	Institucional

ISSN 0120-5587
E-ISSN 2422 3174
N.º 84, 2023

Traductor automático	Desarrollo de sistemas que permitan la traducción automática entre lenguas indígenas y el español	Público en general	Varias (huichol, mixe, náhuatl, mexicanero, mazahua)	Instituto de Investigaciones Matemáticas Aplicadas y Sistemas, UNAM	Institucional
Corpus paralelos otomí-español (Tsunkua)		Público en general	Otomí	Elotl	Institucional
Corpus paralelos amuzgo-español		Público en general	Amuzgo	Antonio Reyes y H. Antonio García Zúñiga	Interinstitucional
Interactivos	Desarrollo de juegos para aprender lenguas	Público en general	Purépecha, Mazahua, Otomí	Centro de Investigaciones en Estudios Superiores en Antropología Social	Institucional

Tabla 2. Resumen de los actuales recursos computacionales para lenguas indígenas mexicanas

Los rubros considerados en esta tabla son: (a) cursos, ofertados principalmente por universidades y centros de investigación, especialmente de corte intercultural; (b) corpus electrónicos, considerando tanto implementaciones como el corpus paralelo otomí-español Tsunkua (2019) desarrollado por Elotl —una comunidad conformada por lingüistas, programadores y miembros de varias comunidades indígenas, sin fines de lucro—, como propuestas teóricas, p. ej., el caso de Reyes & García (2021) con su planteamiento para un corpus paralelo amuzgo-español; y (c) juegos interactivos, ya sean portables para computadoras de escritorio, teléfonos celulares, o algún otro dispositivo similar.

ISSN 0120-5587
E-ISSN 2422 3174
N.º 84, 2023

6. Consideraciones finales

La revisión que hemos hecho aquí es un paso para llevar a cabo, posteriormente, un censo mucho más detallado respecto a la totalidad de herramientas y recursos lingüísticos para las lenguas nativas de México. En esta primera cala, identificamos una variedad de recursos considerables para el tratamiento automático de dichos idiomas, aunque no cubren la totalidad.

Ahora, un aspecto positivo a considerar en varios de estos desarrollos es que ofrecen soluciones viables al problema de la brecha digital, especialmente en el terreno educativo, en especial tras el impacto negativo que tuvo la pandemia del covid-19 durante el 2020, la cual ha afectado sobre todo al sistema educativo bilingüe indígena, desde los niveles básicos hasta el universitario.

Otro aspecto para considerar de esta revisión es que varios de estos proyectos y productos todavía se encuentran al nivel de prototipos, situados en un contexto de investigación académica, sin tener un impacto real en los sectores gubernamentales y empresariales. En ese sentido, es necesario un mayor involucramiento de tales sectores, así como de la sociedad en general, especialmente transformando las ideologías lingüísticas que subyacen en ella, como el considerar que el país en conjunto es monolingüe, negando la enorme diversidad de lenguas nativas que coexisten en el territorio.

Con todo, hay que mencionar también que en años recientes se ha ido incrementando el interés por implementar herramientas electrónicas capaces de procesar automáticamente dichos idiomas, ya sea porque resultan un desafío teóricamente relevante para avanzar en el tratamiento automático del lenguaje natural, ya sea porque son proyectos patrocinados por el sector gubernamental, con el propósito de reducir el rezago tecnológico entre los pueblos originarios, o más importante aún: por un auténtico interés venido desde los hablantes de dichas lenguas, interesados en crear mecanismos digitales que les permitan conservar y difundir sus culturas y sus idiomas. En este sentido, resulta prometedor el panorama a futuro para tales pueblos.

Referencias bibliográficas

Acervo Digital de Lenguas Indígenas (2014). Centro de Investigaciones y Estudios Superiores en Antropología Social (CIESAS). Link: https://ciesas.edu.mx/investigacion/acervo-digital-de-lenguas-indigenas

Acosta, O. & Aguilar, C. (2020). A Critical Review of the Current State of Natural Language Processing in Mexico and Chile. In F. Pinarbaşi & M. Taşkiran (Eds.), *Natural Language Processing for Global and Local Business* (pp. 365-389) IGI Global.

ISSN 0120-5587
E-ISSN 2422 3174
N.º 84, 2023

Álvarez, E. (2019). Miyotl. Universidad Autónoma de Chapingo. https://proyecto-miyotl.web.app

Barriga Villanueva, R. (1995[a]): La paradoja lingüística del indígena mexicano. *Revista de Literatura Hispánica*, *1*(42), 103-112.

Barriga Villanueva, R. (1995[b]). México, un país plurilingüe. *Revista de Literatura Hispánica*, *1*(42), 115-131.

Barriga Villanueva, R. (2018). *De Babel a Pentecostés. Políticas lingüísticas y lenguas indígenas, entre historias, paradojas y testimonios*.Secretaría de Educación Pública.

Bernal, D. & Figueroa, M. (2019). Nueva oferta educativa universitaria con enfoque intercultural: el caso de la Maestría en Lengua y Cultura Nahua de la Universidad Veracruzana. *Revista Educación*, *43*(2), 16-30.

Briseño, F. (2020).Entre los derechos y las políticas lingüísticas en México. El caso de la institucionalización de la lengua maya de la Península de Yucatán", *Abralin ao Vivo-Linguists Online*, Associação Brasileira de Linguística. www.youtube.com/watch?v=hSQVl4oh7X0

Briseño, F. (2021).¿Hacia dónde va la lengua maya de la Península de Yucatán? Entre institucionalización y patrimonialización. *Maya America*, *3*(21), 135-140.

Centro de Estudios Lingüísticos y Literarios (CELL) (2022). Archivo de Lenguas Indígenas de México. El Colegio de México. https://cell.colmex.mx/proyecto/archivo-de-lenguas-indigenas-de-mexico/descripcion

Centro Cultural de España em México (2016). *Vamos a aprender*. https://ccemx.org/evento/app-nahuatl/

Compaine, B. (2001). *The Digital Divide: Facing a Crisis or Creating a Myth?*MIT Press.

Corpus paralelo otomí-español (2019). Tsunkua, Comunidad Elotl. https://tsunkua.elotl.mx

Craveri, M. (2011). La literatura maya hoy y la construcción de las identidades. Procesos constantes de afirmación y de revitalización. *Revista de Literaturas Populares*, 2, 392-409.

De Malingnon, J.-P. (2019): *Kumoontun*. Kumoontun A. C, Santa María Ocotepec, Oaxaca, México. https://kumoontun.org.

Diehl, Richard A. (2005). *The Olmecs: America's First Civilization*, London, Thames & Hudson.

Dyson, L., Grant, S., & Hendriks, M. (2016). *Indigenous People and Mobile Technology.*Routledge.

ISSN 0120-5587
E-ISSN 2422 3174
N.º 84, 2023

Dyson, L., Hendriks, M., & Grant, S. (2007). *Information Technology and Indigenous People*.IGI Global Publishing.

Figueroa, M., Alarcón, D., Bernal, D. & Hernández, J. (2014).La incorporación de las lenguas indígenas nacionales al desarrollo académico universitario: la experiencia de la Universidad Veracruzana. *Revista de la Educación Superior*, *43*(171), 67-92.

Gomashie, G. (2021).Nahuatl and Spanish in Contact: Language Practices in Mexico. *Languages*, *6*(135). https://doi.org/10.3390/languages6030135

Grupo de Ingeniería Lingüística, (2016). Corpus Paralelo de Lenguas Mexicanas www.corpus.unam.mx/geco/portal/index/cplm

Grupo de Ingeniería Lingüística, Instituto de Ingeniería, UNAM (2015). Corpus Axolotl. www.corpus.unam.mx/axolotl

Grupo de Ingeniería Lingüística, Instituto de Ingeniería, UNAM (2003). Instituto de Ingeniería de la Universidad Nacional Autónoma de México. www.iling.unam.mx/corpus

Gutiérrez-Vázquez, X. (2015). Bilingual lexicon extraction for a distant language pair using a small parallel corpus. In *Proceedings of the 2015 Conference of the North American Chapter of the Association for Computational Linguistics: Student Research Workshop*. Denver, CO, USA, ACL Publications, 154-160.

Gutiérrez-Vázquez, X., Sierra, G., & Hernández, I. (2016). Axolotl: a Web Accessible Parallel Corpus for Spanish-Nahuatl. In *Proceedings of the Tenth International Conference on Language Resources and Evaluation LREC 2016*. Portoro, Slovenia, European Language Resources Association (ELRA), 4210-4214.

Gutiérrez-Vázquez, X. & Mijangos, V. (2018). Comparing morphological complexity of Spanish, Otomi, and Nahuatl. In *Proceedings of the Workshop on Linguistic Complexity and Natural Language Processing (LC&NLP-2018)*. Santa Fe, NM, USA, ACL Publications, 30-37.

Hinton, L. & Hale, K. (2001). *The Green Book of Language Revitalization in Practice*.Brill.

Instituto Federal de Telecomunicaciones (2020). *Encuesta Nacional sobre la Disponibilidad y Uso de Tecnologías de la Información en los Hogares (ENDUITH, 2020)*.Instituto Federal de Telecomunicaciones. www.ift.org.mx/comunicacion-y-medios/comunicados-ift/es/encuesta-nacional-sobre-disponibilidad-y-uso-de-tecnologias-de-la-informacion-en-los-hogares-endutih

Izquierdo, M. (2005). El reconocimiento de los derechos de los pueblos indígenas en México. *Cuadernos*

ISSN 0120-5587
E-ISSN 2422 3174
N.º 84, 2023

Constitucionales de la Cátedra Fadrique Furió Ceriol,. 50-51, 109-124. Instituto Nacional de Lenguas Indígenas (INALI).www.inali.gob.mx

Lara, L. F. & Vázquez Laslop, M. E. (2015). *El estudio de las lenguas en México: avatares de dos siglos*.El Colegio Nacional.

Littel, P., Kazantseva, A., Kuhn, R., Pine, A., Arppe, A., Cox, Ch., & Junker, M. (2018).Indigenous language technologies in Canada: Assessment, challenges, and successes. In *Proceedings of the 27th International Conference on Computational Linguistics*, Santa Fe, NM, USA, ACL Publications, 2620–2632.

López, N. (2015).Localización de Firefox en lenguas indígenas. *Revista Sinfin*, L10n. www.revistasinfin.com/articulos/l10n-localizacion-de-firefox-en-lenguas-indigenas

Mager, M., Oncevay, A., Ríos, A., Meza, I., Palmer, A., Neubig, G., & Kann, K. (2021). *Proceedings of the First Workshop on Natural Language Processing for Indigenous Languages of the Americas*. Stroudsburg PA, USA, ACL Publications. https://aclanthology.org/2021.americasnlp-1.0/

Mager, M., Gutiérrez-Vázquez, X., Sierra, G., & Meza, I. (2018). Challenges of language technologies for the indigenous languages of the Americas. In *Proceedings of the 27th International Conference on Computational Linguistics COLING 2018*. Santa Fe, NM, USA, ACL Publications, 55-69.

Mager, M. & Meza, I. (2018). Hacia la traducción automática de las lenguas indígenas de México. En J. Girón & I. Galina, I. (Eds.), *Digital Humanities 2018. Puentes-Bridges* (pp. 637-639). COLMEX/UNAM/RedHD.

Mager, M., Carrillo, D., & Meza, I. (2018).Probabilistic Finite-State Morphological Segmenter for the Wixarika (Huichol) Language. *Journal of Intelligent & Fuzzy Systems*, *34*(5), 3081–3087.

Mager, M. (2017). *Traductor híbrido wixarika-español con escasos recursos bilingües* [Tesis de Maestría, Universidad Autónoma Metropolitana].

Maitland, D. (1985). *The Missing Link*.International Telecommunication Union.

Medina, A. (2007). Affix discovery by means of corpora: Experiments for Spanish, Czech, Ralámuli and Chuj. In A. Mehler, A. & R. Köhler (Eds.), *Aspects of Automatic Text Analysis* (pp. 277—299). Springer.

Medina, A. (2008). Affix discovery based on entropy and economy measurements. In N. Gaylord, A. Palmer, & E. Ponvert (Eds.), *Computational Linguistics for Less-Studied Languages* (pp. 99-112). CSLI Publications.

ISSN 0120-5587
E-ISSN 2422 3174
N.º 84, 2023

Melton-Villanueva, M., de la Cruz, A., & Morales, O. C. (2022).Práctica autóctona para revitalizar la lengua náhuatl en comunidades bilingües de México. *Lenguas Radicales*, *1*(3), 31-43.

Montejo, O., Bastiani, J. & Orantes, S. (2019). Experiencia digital en la enseñanza del Ch'ol en Chiapas, México. *Revista Senderos Pedagógicos*,10, 145-162.

Navarrete, F. (2016).México sin mestizaje: una reinterpretación de nuestra historia. *Ciclo de conferencias: El Historiador frente a la Historia 2016. Desigualdad y violencia en la historia*, Ciudad de México, Instituto de Investigaciones Históricas de la UNAM, 26 de abril.www.youtube.com/watch?v=qLCUzzLls0E

Olko, J., & Sullivan, J. (2013).Empire, colony, and globalization. a brief history of the Nahuatl language. *Colloquia Humanistica*, 2. 181-216.

Olko, J., & Sullivan, J. (2014). Toward a Comprehensive Model for Nahuatl Language Research and Revitalization. In H. Leung, Z. O'Hagan, S. Bakst, , A. Lutzross, J. Manker, N. Rolle, & K. Sardinha (Eds.), *Annual Meeting of the Berkeley Linguistics Society*, Vol. 40 (pp. 369-397).University of California at Berkeley.

Pellicer, D. (1997).Derechos lingüísticos en México: realidad y utopía. *xx International Congress of the Latin American Studies Association (LASA97)*, Guadalajara, México, 17-19 de abril.

Poot Cahun, J. M. (2019). Koox App - Aprende MaayaT'aan. https://apkcombo.com/es/koox-app-aprende-maayat-aan/com.lighthouse.koox

Proyecto Mozilla Nativo (2018). Comunidad de Mozilla Nativo, Fundación Mozilla.https://wiki.mozilla.org/Nativo

Reyes, A. & García, H. (2021). Hacia el desarrollo de un corpus oral en lengua amuzga. En M. García, M. Sáenz, A. López & A. Hurtado (Eds.), *Procesamiento de lenguaje natural para las lenguas indígenas* (pp. 132-144)Universidad Michoacana de San Nicolás de Hidalgo.

Soustelle, J. (1979). *Los olmecas* Fondo de Cultura Económica.

Soustelle, J. (1982). *Los mayas*.Fondo de Cultura Económica.

Stahler-Sholk, R. & Baronnet, B. (2018).La escuela es la comunidad: luchas indígenas y autonomía en México., En S. Plá & S. Rodríguez (Coords.), *Saberes sociales para la justicia social: educación y escuela en América Latina (pp. 99-136).*Universidad Pedagógica Nacional/La Carreta.

Suárez, J. (1983). *The Mesoamerican Indian Languages*.Cambridge University Press.

ISSN 0120-5587
E-ISSN 2422 3174
N.º 84, 2023

Villareal, B. (2011).El náhuatl en Los Ángeles: el papel de la lengua indígena en la creación de la identidad chicana. *Mester*, *40*(1), 81-100.

ISSN 0120-5587
E-ISSN 2422 3174
N.º 84, 2023