

**Tipo:** Artículo original - **Sección:** Dossier: Lenguas Indígenas: puentes entre la Formación, la Documentación y el Análisis

## **Documentación y descripción lingüística de lenguas zapotecas (otomangue)**

Language documentation and linguistic description  
in Zapotec languages (Otomanguean)

**Oscar López Nicolás**

*Centro de Investigaciones y Estudios Superiores en Antropología Social,  
Unidad Pacífico Sur, México.  
<https://orcid.org/0009-0009-5302-0543>*

e-mail: [oscar\\_n21@hotmail.com](mailto:oscar_n21@hotmail.com)

Recibido: 17/11/2024

Aprobado: 9/4/2025

### **RESUMEN**

Este trabajo tiene un doble propósito. Por un lado, dar cuenta de la diversidad lingüística en México, los desafíos para determinar cuántas lenguas se hablan en el país y la propuesta institucional del gobierno mexicano a esta realidad. Por otro lado, destacar la importancia de la documentación y descripción lingüística ante un escenario de lenguas en riesgo, con énfasis en las lenguas zapotecas del tronco otomangue. Los estudios en lenguas zapotecas han abordado varios temas de la fonología, morfología, sintaxis y lexicografía, con un importante trabajo de documentación lingüística.

**Palabras clave:** documentación lingüística; descripción gramatical; lenguas en riesgo; diversidad lingüística.

### **ABSTRACT**

The goal of this paper is twofold. On the one hand, it accounts for the linguistic diversity in Mexico, for the challenges in determining how many languages are spoken in the territory and the institutional proposals from the Mexican government vis-à-vis the actual facts. On the other hand, it highlights the importance of linguistic documentation and linguistic description in the context of endangered languages, with special emphasis on the Zapotec languages of the Oto-Manguean stock. Studies on Zapotec languages have focused on various topics in phonology, morphology, syntax, and lexicography, on the basis of robust linguistic documentation.

**Keywords:** language documentation; linguistic description; endangered languages; linguistic diversity.

---

**Conflictos de Interés:** ninguno que declarar

**Fuente de financiamiento:** sin fuente de financiamiento.

**DOI:** <https://doi.org/10.47133/NEMITYRA20250101c-A11>

**BIBID:** 2707-1642, 7, 1, pp. 113-126

**Editor responsable:** Mirtha Lugo (<https://orcid.org/0000-0003-1941-3852>) y Celeste Escobar (<https://orcid.org/0000-0002-7704-543X>).  
Universidad Nacional de Itapúa, Facultad de Humanidades, Ciencias Sociales y Cultura Guaraní, Paraguay.

## Introducción

México se sitúa entre los países con mayor diversidad lingüística y cultural. Según algunas estimaciones, actualmente ocupa el quinto rango como país multilingüe, después de Nueva Guinea, Indonesia, Nigeria e India, y el primer rango dentro del continente americano (Chamoreau, 2013). Debido a la pluralidad lingüística del país, el gobierno mexicano, a través del Instituto Nacional de Lenguas Indígenas (INALI), emprendió a principios de este siglo una investigación científica a partir de criterios de la genealogía lingüística, la dialectología y la sociolingüística para proponer una clasificación de las más de 200 lenguas indígenas habladas en México; el resultado es el Catálogo de las Lenguas Indígenas Nacionales (CLIN) donde se detallan las variantes lingüísticas de México con sus autodenominaciones y referencias geoestadísticas. En este Catálogo se propone una taxonomía lingüística con base en tres categorías que estipulan la presencia de 11 familias lingüísticas, 68 agrupaciones lingüísticas y 364 variantes lingüísticas.

El primer nivel de la clasificación reconoce las relaciones de parentesco, con variada profundidad histórica, entre las lenguas de determinados grupos étnicos asentados en el territorio nacional. El segundo nivel toma como punto de referencia las denominaciones que históricamente se les ha dado a las lenguas indígenas en el país, pero se opta por la noción agrupación lingüística ya que la etiqueta lengua subsume, en estos casos, sistemas lingüísticos independientes y formas de habla recíprocamente ininteligibles. El tercer nivel de la clasificación compagina dos criterios, a saber, la ausencia de inteligibilidad mutua entre hablantes de lenguas estructural y socialmente distantes pero llamadas con el mismo nombre, y la existencia de autodenominaciones (i.e., la expresión con la cual los hablantes nombran su lengua en su propia variante) para cada una de estas formas de habla distanciada (INALI, 2009, p. 10) y no es posible la comunicación entre los hablantes. A este respecto, el Instituto Nacional de Lenguas Indígenas señala que:

La categoría lengua o idioma, intermedia en términos de inclusión entre agrupación lingüística y variante lingüística, se define como un sistema de comunicación socializado mediante el cual dos o más individuos que se identifican como o con miembros de una comunidad lingüística pueden codificar y decodificar, en un plano de mutua inteligibilidad, los mensajes orales o escritos que llegasen a intercambiar. Para la aplicación de esta categoría, en el presente Catálogo, las variantes lingüísticas deben ser tratadas como lenguas. (INALI, 2009, p. 29)

La Tabla 1 esquematiza la clasificación de las lenguas indígenas habladas en México. Cada columna corresponde a una familia lingüística (indicada en versalitas y negrita en el encabezado), mientras que en cada celda de la Tabla aparecen las agrupaciones lingüísticas con un número delante que corresponde a la variante o variantes lingüísticas consignadas; así por ejemplo, una de las agrupaciones lingüísticas de la familia otomangue es el otomí (columna 2, fila 2) y el número 9 delante corresponde al total de sus variantes: 1) *ñuhu* ~ *yihu* (otomí de la Sierra), 2) *hñäñho* (otomí bajo del noroeste), 3) *ñathó* (otomí de oeste), 4) *ñöhñö* ~ *ñähñä* (otomí del oeste del Valle del Mezquital), 5) *hñähnú* (otomí del Valle del Mezquital), 6) *yihmu* (otomí de Ixtenco), 7) *ñü' hü* (otomí de Tilapia o del sur), 8) *hñöñho* (otomí del noroeste), y 9) *ñohnño* (otomí del centro). Al final de esta segunda columna aparece el número total de agrupaciones y el número total de variantes, 18 y 220, respectivamente, que conforman la familia lingüística otomangue.

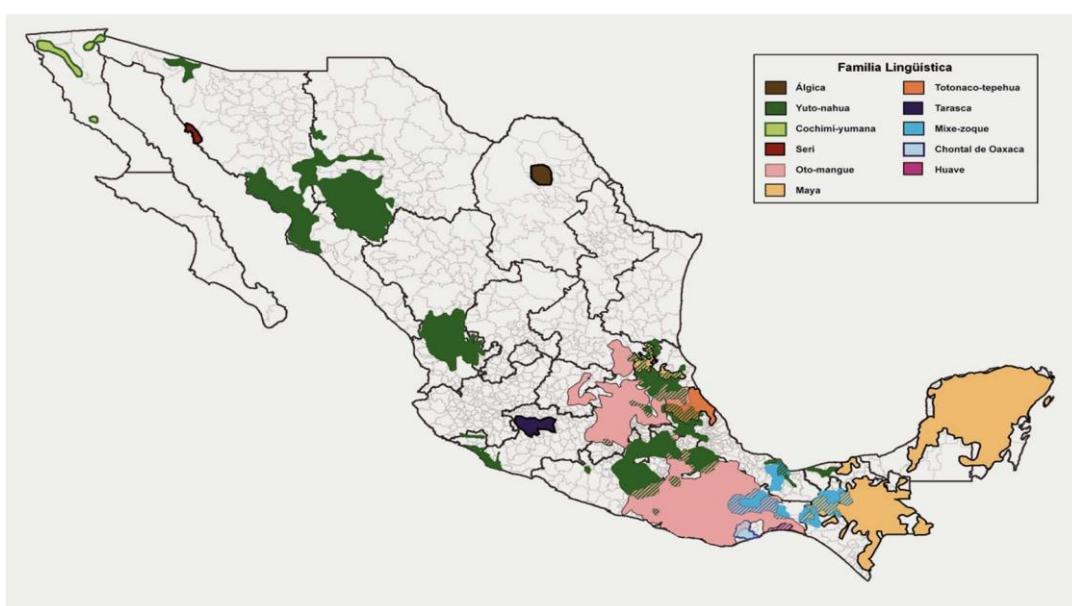
**Tabla 1.** Diversidad de las lenguas indoamericanas en México

MAYA	OTOMANGUE	YUTONAHUA	MIXEZOQUE	COCHI MÍ YUMAN	TOTONAC O-TEPEHUA	CHONTAL DE OAXACA	HUAVE	PURÉPEC HA	SERI	ÁLGICA
huasteco 3	otomí 9	pápago 1	mixe 6	paipai 1	totonaco 7	chontal de Oaxaca 3	huave 2	purépech a 1	seri 1	kikapo o 1
maya 1	mazahua 2	pima 3	sayultec o 1	ku'ahl 1	tepehua 3	<b>1</b> <b>3</b>	<b>1</b> <b>2</b>	<b>1</b> <b>1</b>	<b>1</b> <b>1</b>	<b>1</b> <b>1</b>
lacandón 1	matlatzinca 1	tepehuan o del norte 1	oluteco 1	cucapá 1	<b>2</b> <b>10</b>					
ch'ol 2	tlahuica 1	tepehuan o del sur 3	popoluca de la sierra 1	kumiai 1						
chontal de tabasco 4	pame 2	tarahuma -ra 5	texiste-pequeño 1	kiliwa 1						
tzeltal 4	chichimeco Jonaz 1	guarajío 2	ayapanec o 1	<b>5</b> <b>5</b>						
tzotzil 7	chinanteco 11	yaqui 1	zoque 8							
q'anjob'al 1	tlapaneco 9	mayo 1	<b>7</b> <b>19</b>							
akateco 1	mazateco 16	cora 8								
jakalteco 1	ixcateco 1	huichol 4								
qato'k 2	chocholtec o 3	náhuatl 30								
chuj 1	popoloca 4	<b>11</b> <b>59</b>								
tojolabal 1	zapoteco 62									
q'eqchi 1	chatino 6									
k'iche' 3	amuzgo 4									
kaqchikel 1	mixteco 81									
teko 1	cuicateco 3									
mam 5	triqui 4									
awakatek o 1	<b>18</b> <b>220</b>									
ixil 2										
<b>20</b>										
<b>43</b>										

Fuente: INALI, 2009; Chamoreau, 2013.

Más allá de lo problemático que ha resultado determinar el número preciso de lenguas indoamericanas habladas en México; basta revisar algunos datos estadísticos para advertir esta falta de consenso (por ejemplo, el Instituto Nacional de Estadística, Geografía e Informática (INEGI) en su censo de 2020 reconoce 68 lenguas, la UNESCO da cuenta de 143 lenguas, el Instituto Lingüístico de Verano (ILV) propone 267 lenguas), vale la pena destacar que la relevancia de la diversidad lingüística en México no es solo por el alto número de variantes incluso de agrupaciones, sino fundamentalmente por el importante número de familias lingüísticas que tiene (Mapa 1). Dicho sea de paso, nótese que la mayoría de estas familias de lenguas se concentra en la región centro-sur del país que, en cierto sentido, representa la zona nuclear del área lingüística mesoamericana (Campbell, Kaufman y Smith-Stark, 1986). Algunas lenguas de la familia maya se hablan en Guatemala y Belice o lenguas de la familia yutonahua que se hablan en Estados Unidos, Guatemala, El Salvador, Honduras y Nicaragua.

**Mapa 1.** Ubicación geográfica de las familias lingüísticas en México



Fuente: INALI, 2009.

No cabe duda que para muchos la diversidad lingüística provoca una fascinación, sea porque cada lengua representa una visión propia del mundo y, por ende, establece una correlación con la diversidad cognitiva, sea por la incidencia recíproca entre estructura lingüística y comunidad de habla o incluso por las convergencias tipológicas que se descubren entre lenguas sin parentesco; lo cierto es que en la actualidad, particularmente las lenguas no hegemónicas o minoritarias enfrentan una serie de circunstancias de diversa índole como la implementación de políticas lingüísticas homogeneizantes, la estigmatización de sus hablantes, la migración, entre otras, que las vulneran y ponen en riesgo de desaparecer en muchas regiones del mundo (Moseley, 2010).

Una de las principales causas de la pérdida de una lengua es la interrupción en la transmisión intergeneracional, pues a partir de tal condición se reducen progresivamente los contextos de uso y, en consecuencia, el léxico y determinadas estructuras gramaticales dejan de ser funcionales. Atendiendo a este factor, las lenguas en peligro de desaparecer se agrupan en un continuo descendente hasta su extinción:

**Tabla 2.** Transmisión intergeneracional de la lengua

Índice de vitalidad	Grado	Población de hablantes
<i>No corre peligro</i>	5	La lengua es utilizada por todos los grupos de edad, incluidos los niños.
<i>Vulnerable</i>	4	La lengua es utilizada por algunos niños en todos los ámbitos, y por todos los niños en ámbitos restringidos.
<i>Claramente en peligro</i>	3	La lengua es utilizada sobre todo por la generación parental para arriba.
<i>Seriamente en peligro</i>	2	La lengua es utilizada sobre todo por la generación de los abuelos para arriba.
<i>En situación crítica</i>	1	La lengua es utilizada por muy pocos hablantes, fundamentalmente por la generación de los bisabuelos.
<i>Extinta</i>	0	Ya no quedan hablantes.

Fuente: UNESCO, 2003, p. 7.

La gran mayoría de las variantes lingüísticas de México se encuentran entre los niveles vulnerable y seriamente en peligro; por tanto, se torna necesario documentar, describir y revitalizar las lenguas indoamericanas en éste y en otros países.

## Lenguas en riesgo y documentación lingüística

Como alternativa a la amenaza que enfrenta la diversidad lingüística se plantean esfuerzos de documentación de lenguas con el propósito principal de “preservar el patrimonio de cada lengua como sistema simbólico y expresivo único, para su uso por generaciones actuales y venideras” (Golluscio, Pacor, Ciccone y Krasan, 2019, p. 16).

La documentación lingüística es un área de investigación y práctica de la disciplina lingüística dedicada a compilar, analizar y preservar un registro multipropósito, representativo y duradero de una lengua o cualquiera de sus variedades. Supone una tarea multifacética y fundamentalmente inductiva, cuyo fin es organizar un corpus exhaustivo de datos primarios, particularmente las muestras de comportamiento lingüístico observable que incluye, por ejemplo, las conversaciones cotidianas, el habla de los cuidadores al recién nacido, los rituales discursivamente más complejos, etc., (Himmelmann, 2007; Woodbury, 2011). Este corpus lingüístico es un acervo de datos accesibles –no solo para los interesados en algún aspecto de la lengua, sino principalmente para la comunidad de hablantes– que en suma se pueden caracterizar de la siguiente manera: datos lo más ricos posibles en términos de géneros de habla, contextos de habla y variación sociolingüística; datos perdurables y de fácil acceso, es decir, no dependientes de una tecnología en particular; datos registrados y utilizados bajo consideraciones éticas y legales; datos transcritos, anotados y glosados que permitan entender el funcionamiento de la lengua, este hecho a su vez permite la contribución a distintas ramas de la lingüística a la vez que a disciplinas afines como el campo educativo, el antropológico o la traducción; datos registrados idealmente por un equipo heterogéneo formado por lingüistas, hablantes, antropólogos, educadores, entre otros interesados.

Un aspecto relevante en el campo de la documentación lingüística que se va consolidando es el del reconocimiento e impulso de la agencia y el rol decisivo que tienen los hablantes de las lenguas documentadas (Grinevald, 2003), no solo porque se trata del factor humano más importante y generoso en esta actividad colectiva, sino porque los hablantes representan la salvaguardia, el mantenimiento y la recreación del patrimonio cultural inmaterial (UNESCO, 2003); por tanto, es la comunidad de hablantes la que tendrá un rol central al determinar qué documentar, por qué documentar, para qué documentar, para quién documentar, con qué medios documentar, en qué contextos documentar, quiénes específicamente deben documentar (Quatra, 2011); sin descuidar, desde luego elementos metodológicos en su implementación como las técnicas y estrategias de la documentación, el equipo e insumos necesarios, los *software* requeridos, los acervos lingüísticos para la salvaguardia del registro, que finalmente guían la investigación lingüística básica y aplicada (Ashmore, 2008; Montaño, Arellano y Sánchez, 2018; Pérez y Hernández, 2020).

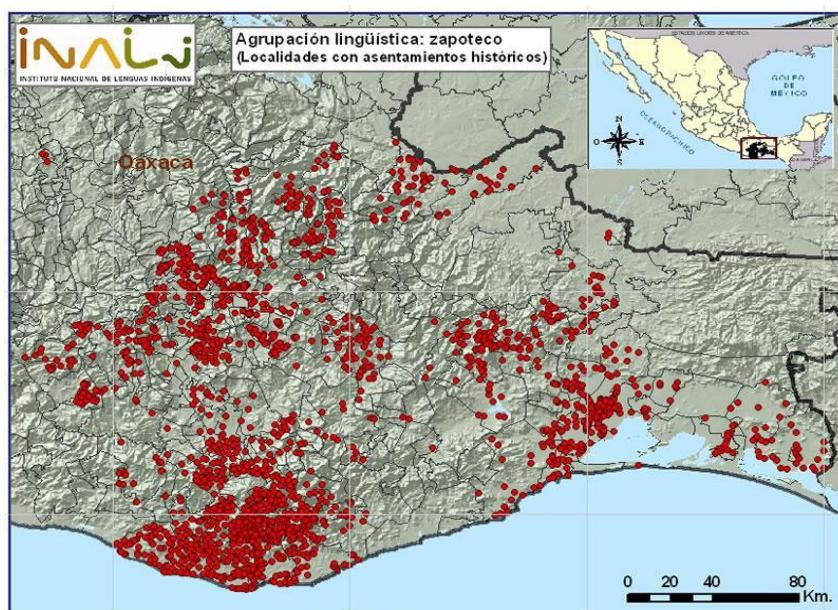
Entre las lenguas zapotecas se han llevado a cabo proyectos de documentación lingüística colectivos e individuales por parte de lingüistas hablantes y no hablantes de estas lenguas. Entre los esfuerzos colectivos están el Proyecto de Documentación de Lenguas de Mesoamérica (1993–2010) o el Proyecto de Documentación del Zapoteco del Valle del INALI (2011–2013), por citar dos ejemplos conocidos.

En la siguiente sección, por razones de espacio, se mencionan de manera general ámbitos de la descripción lingüística que, por una parte, profundizan en el conocimiento de las lenguas zapotecas y, por la otra, aportan al estado de la cuestión en la lingüística areal y tipológica.

## Descripción lingüística en lenguas zapotecas

### Preliminares

El Mapa 2 muestra la distribución geográfica de las localidades del estado de Oaxaca, México, donde se hablan las distintas variantes lingüísticas del zapoteco<sup>1</sup>, una de las agrupaciones más diversas entre las lenguas indoamericanas del país, como ya se ha señalado antes. Desde mediados del siglo pasado, Swadesh (1949, p. 420) advertía que las lenguas zapotecas exhiben tales diferencias léxicas y estructurales que “pueden llegar a diferir entre sí tan grandemente como las lenguas romances” y, más tarde, Smith-Stark (2007, p. 69) proponía incluso que “es más acertado hablar del zapoteco como una familia de lenguas estrechamente emparentadas”; conforme avanza la investigación lingüística de este grupo de lenguas se responde mejor al problema persistente de la clasificación interna (Campbell, 2017b), y se logra un mayor conocimiento de sus propiedades gramaticales en varios niveles del análisis lingüístico.

**Mapa 2.** Ubicación geográfica de la agrupación lingüística zapoteca

Fuente: INALI, 2009.

Al igual que sucedió con otras lenguas indígenas del país, las lenguas zapotecas han sido objeto de interés desde el inicio de la época colonial, claramente con fines de evangelización. Los estudios lingüísticos más tempranos en zapoteco son el *Vocabulario en lengua čapoteca* (1578) y el *Arte del idioma zapoteco* (1578), ambos escritos por el fraile dominico Juan de Córdova<sup>2</sup>; no es equivocado decir que estudios lexicográficos y gramaticales contemporáneos sobre lenguas zapotecas tienen sus antecedentes en éstos y otros textos similares producidos en la colonia, y que hoy en día son un referente descriptivo para plantear hipótesis en torno a la evolución de las lenguas modernas.

## Sobre el sistema de sonidos

Hay tres fenómenos fonológicos que han llamado la atención entre los estudiosos de las lenguas zapotecas: el tono y la fonación no modal en el ámbito vocálico, y la oposición *fortis–lenis* en el dominio consonántico. Todas las lenguas zapotecas son tonales; una lengua es tonal si la cualidad acústica (más grave o más aguda) de una palabra cambia el significado léxico de ésta (véase Yip, 2002). Pongamos por caso la forma monosílábica /ya/ que puede tener tres significados diferentes dependiendo de la cualidad acústica de la vocal. Si la vocal se pronuncia en una tonía grave (tono bajo, ortográficamente con acento grave) entonces *yà* significa ‘campana’, si la pronunciación se hace en una tonía aguda (tono alto, representado con acento agudo) *yá* significa ‘temazcal’, pero si hay una combinación de tonía aguda más tonía grave (tono descendente, en la ortografía con acento circunflejo) *yâ* significa ‘carrizo’. Los sistemas tonales de estas lenguas varían entre sistemas de 3 a 5 tonos fonológicos (Avelino, 2004; Antonio, 2015). Además de los tonos bajo, alto y descendente, se ha documentado el tono ascendente (combinación de tonía grave más tonía aguda) y tono alto ascendente (tonía aguda más tonía alta).

La fonación no modal refiere a la oposición entre vocales modales y vocales no modales. Las vocales modales se pronuncian con una vibración continua de las cuerdas vocales, como sucede, por ejemplo, con las vocales del español; las vocales no modales, en cambio, se producen con una vibración irregular de las cuerdas vocales, provocado por la tensión o

distensión de la glotis (órgano cartilaginoso ubicado entre la raíz de la lengua y la faringe); así resultan varios tipos de vocal no modal en zapoteco, una vocal con una tensión total de las cuerdas vocales: la vocal ‘cortada’ (representada ortográficamente como *a'*), una vocal con la misma tensión en la mitad de la producción de la vocal pero recuperando la vibración normal: la vocal rearticulada (*a'a*), una vocal con vibración irregular sin obstruir totalmente la glotis: la vocal laringizada (*ã*), y la vocal con una distensión de la glotis: la vocal murmurada (*ã*), en algunas lenguas zapotecas cada una de estas manifestaciones de la glotis produce una vocal distinta. Todas las lenguas zapotecas tienen la oposición binaria entre vocal modal y al menos un tipo de vocal no modal, prototípicamente una vocal cortada; otras en cambio tienen en su inventario varios tipos de vocal no modal: *be* ‘frijol de mesquite’, *be'e* ‘hongo’, *beq* ‘oreja (de oveja)’ y *bee* ‘moho’ (Chávez Peón, 2010). Lo notable en las lenguas zapotecas radica en que el tono y la fonación no modal interactúan en la producción de vocales fonológicas dando como resultado el fenómeno conocido como complejidad laríngea; una propiedad fonológica poco común en las lenguas del mundo, pero compartido con otras lenguas otomangues (Arellanes, 2009; Silverman, 1997b).

Por otro lado, los inventarios consonánticos de estas lenguas se caracterizan tener pares de consonantes que en la tradición zapotequista se denomina oposición *fortis* vs., *lenis* (Nellis y Hollenbach, 1980), una distinción más allá de la diferencia sordo-sonoro que suelen presentar muchos sistemas consonánticos en las lenguas naturales, por ejemplo: /p/ vs., /b/, /t/ vs., /d/, /k/ vs., /g/, donde el primer fonema de la oposición es sordo y el segundo, sonoro. En las lenguas zapotecas la oposición fortis-lenis ocurre con todas las clases naturales de las consonantes e incide en otros contextos del sistema de sonidos: 1) en la longitud vocálica al interior de la sílaba, especialmente en la rima ya que una consonante fortis le precede una vocal corta, mientras que a una consonante lenis le precede una vocal alargada); 2) en la estructura prosódica, pues sólo las consonantes fortis en coda son moraicas<sup>3</sup>; 3) en la realización del tono de la vocal circundante, puesto que ante consonante fortis las vocales no puede llevar tono descendente o ascendente; y 4) restringe los tipos de grupos consonánticos en el inicio silábico porque, por regla general, solo permite consonantes fortis como el primer elemento del grupo. La Tabla 3 muestra un ejemplo de esta oposición consonántica en su representación ortográfica.

**Tabla 3.** Oposición *fortis* vs., *lenis* en algunas lenguas zapotecas

	con oposición			sin oposición		
<i>Fortis</i>	w	h	h			
<i>Lenis</i>	w	x	h	h	h	h

Fuente: Elaboración propia.

## Posesión inalienable y posesión alienable

La distinción inalienable-alienable en el dominio de la posesión es un rasgo areal mesoamericano (Campbell, Kaufman y Smith-Stark, 1986). La posesión inalienable hace referencia a un tipo de posesión inherente y corresponde a la posesión de partes del cuerpo y a las relaciones de parentesco; en cambio, la posesión alienable alude a la relación conferida por razones sociales o económicas y, por tanto, cubre una gran variedad de entidades disponibles de ser poseídas. Esta distinción semántica binaria de la posesión tiene una expresión variada

en el nivel gramatical. Existen lenguas, como el español, donde la distinción inalienable-alienable se mantiene en el nivel semántico y no se refleja en el dominio gramatical, de modo que la posesión se codifica mediante una sola construcción, por ejemplo: ‘la *mano* del niño’ (posesión inalienable) y ‘el *auto* de la mujer’ (posesión alienable), donde la relación entre lo poseído y el poseedor se indica mediante la preposición ‘de’. Por otro lado, hay lenguas donde tal distinción sí permea la gramática y se tienen distintas construcciones, por ejemplo, en algunas lenguas oceánicas, la posesión de partes del cuerpo y relaciones parentesco se expresa mediante aposición [poseído + poseedor], es decir, sin ningún elemento gramatical que marque la relación formal entre los dos elementos de la construcción; mientras que la expresión de la propiedad (i.e., posesión alienable) requiere un prefijo antecediendo al elemento poseído [prefijo-poseído + poseedor] (Aikhenvald, 2013, p. 3).

En las lenguas zapotecas la distinción semántica inalienable-alienable también se refleja en la gramática a través de tres construcciones de posesión distintas (Arellanes 2020; López, 2020). Por un lado, la posesión inalienable (partes del cuerpo y relaciones de parentesco) se codifica mediante una construcción apositiva [poseído + poseedor]. Por otro lado, la posesión alienable puede expresarse mediante tres estructuras, la construcción morfoléxica que tiene dos manifestaciones, la primera donde la expresión de la entidad poseída requiere un prefijo posesivo: *x-kó 'n béné* ‘el ganado de la gente’, la segunda variante de esta construcción es una donde la entidad poseída recurre a una base léxica supletiva o alterna como en *Ihìzh nó'ólhé* ‘la casa de la mujer’, el nominal ‘casa’ es *yó'ò* (y no *Ihìzh*) cuando ocurre fuera del dominio gramatical de la posesión. La construcción perifrástica requiere de un nexo gramatical que indica la relación de posesión: *xhóá' chè yézh* ‘el maíz del pueblo’. Esta expresión variada de la posesión depende de dos factores, el primero es la distinción inalienable-alienable y el segundo es la cercanía-lejanía respecto de *ego*, las entidades cercanas de *ego* se expresan mediante la estructura con prefijo y aquellas lejanas de *ego* a través de la posesión perifrástica. ¿Cuáles entidades son cercanas y cuáles lejanas de *ego*? esa es una categorización que los hablantes hacen en función de sus interacciones con el mundo extralingüístico y que proyectan en la gramática de la lengua. En la Tabla 4 se resumen las cuatro construcciones posesivas en este grupo de lenguas.

**Tabla 4.** Tipos de construcción posesiva en lenguas zapotecas

<i>Significado inalienable</i>	<i>Significado alienable</i>	
Posesión apositiva [PDO + PSR]	Posesión morfoléxica [POS-PDO + PSR] [bases léxicas supletivas + PSR]	Posesión perifrástica [PDO GEN PSR]

Fuente: Elaboración propia.

Además de los avances en la fonología y en la interfaz semántica-sintaxis –concretamente en lo que respecta a la posesión–, el estudio de lenguas zapotecas ha logrado avances en otros niveles del análisis lingüístico y en determinados campos afines. En el nivel de la morfología, las lenguas zapotecas tienen varios prefijos para marcar las categorías de tiempo, aspecto y modo (i.e., tiempo futuro, aspecto completivo, aspecto incompletivo, aspecto progresivo, y modo irrealis) en los lexemas verbales. Cada una de estas categorías flexivas puede tener más de una expresión formal; en algunas lenguas el modo irrealis se puede codificar a través de

cuatro maneras distintas: mediante los prefijos *w-*, *g-* o mediante ajustes en el tema verbal como la sola ocurrencia del tema (*tema denudado*) o la *mutación consonántica*. Cada determinado grupo de verbos en una lengua selecciona determinadas realizaciones flexivas formando clases verbales flexivas (Corbett, 2009); la combinación de las formas flexivas en un paradigma no es predecible ni desde la semántica ni desde la fonología, sino totalmente arbitrario, es decir, cada grupo de lexemas verbales selecciona una opción que depende totalmente del léxico en cuestión. Varios estudios sobre lenguas zapotecas han dado cuenta de este fenómeno en los últimos años (Beam de Azcona, 2019; López, 2019; Antonio 2022). La Tabla 5 resume dos casos de clases flexivas en dos lenguas zapotecas y nótese que el número de clases de lengua a lengua puede variar, así como la selección y mecanismos que las lenguas tienen cada una para agrupar sus lexemas verbales.

**Tabla 5.** Clases verbales en dos lenguas zapotecas

(lengua de la rama norteña)			(lengua de la rama sureña)		
Clases verbales	incompletivo	irrealis	Clases verbales	completivo	irrealis
I	<i>sh-</i>	<i>tema denudado</i>	I	<i>b-</i>	<i>g-</i>
II	<i>dx-</i>	<i>mutación consonántica</i>	II	<i>go-, gu-</i>	<i>g-</i>
III	<i>dx-</i>	<i>g-</i>	III	<i>b-, b+mutación consonántica</i>	<i>mutación consonántica</i>
IV	<i>sh-</i>	<i>w-</i>			

Fuente: López, 2019 y Antonio, 2022.

Al momento de investigar las clases flexivas surgen varias preguntas de investigación: ¿qué factores gramaticales determinaron la evolución de las clases flexivas en estas lenguas? o ¿cómo adquieren los infantes los paradigmas flexivos en vista de la arbitrariedad de las clases flexivas verbales?

El nivel de análisis con mayor desarrollo en los últimos años es el de la morfosintaxis, ámbito donde se han logrado avances significativos tanto en el estudio de la cláusula simple como de la cláusula compleja. Sobre la cláusula simple hay consenso en torno al perfil tipológico del orden de los constituyentes en las lenguas zapotecas, considerando que el orden Verbo-Sujeto-Objeto con las correlaciones correspondientes en distintos contextos de la gramática, es el orden canónico en estas lenguas. Se reconoce también que el alineamiento nominativo-acusativo es común en todas las lenguas de esta agrupación. Los sistemas de categorización nominal sean en su expresión como clases nominales o como clasificadores nominales han revelado que estos idiomas tienen una motivación semántica que agrupa

conjuntos de sustantivos; por citar un ejemplo, las lenguas zapotecas organizan sus nominales en función del rasgo de animosidad, de modo que muchos sustantivos animados teman un prefijo b- o m-, mientras que los sustantivos inanimados toman como consonante inicial de la palabra el fonema g+ o y+. En dominio de la cláusula compleja, la atención los estudios más recientes han puesto el énfasis en la descripción y análisis de las cláusulas de relativo y las cláusulas de complemento (Gutiérrez 2021; Vásquez, 2024).

Se han desarrollado también estudios lexicográficos (Andrade y López, 2022), de lingüística histórica (Fernández de Miranda, 1995), y sobre lingüística antropológica (Sicoli, 2021) que han contribuido al entendimiento de las lenguas zapotecas. Es evidente que falta mucho por investigar tanto en lo que respecta a fenómenos gramaticales aun no descritos como a variantes lingüísticas que no han sido exploradas aún.

## Palabras finales

Vale la pena reconocer la diversidad lingüística y promover la concientización sobre la amenaza a ésta en muchas partes del mundo. Una respuesta a esta situación dada desde la disciplina lingüística es la implementación de la documentación informada conceptual y metodológicamente, sin olvidar los principios éticos en la interacción con los hablantes y sus saberes, ni tampoco obviar las pautas legales sobre la información registrada. Cada vez se torna imprescindible atender las necesidades de las comunidades de habla, una realidad que indefectiblemente conduce a una dimensión aplicada –desde una perspectiva multidimensional y multidisciplinar– del trabajo lingüístico al campo de la revitalización lingüística o en campos afines como el de la educación intercultural, la traducción e interpretación, los medios de comunicación, en suma, al reclamo de los derechos lingüísticos de los hablantes de lenguas en riesgo de pérdida.

## Notas

<sup>1</sup> Las lenguas zapotecas se hablan también en la diáspora, es decir, en otros estados de la República Mexicana y en Estados Unidos.

<sup>2</sup> Los *Vocabularios*, las *Artes* (gramáticas) y las *Doctrinas cristianas* conforman la denominada trilogía catequística en la Nueva España (Smith-Stark, 2007). Además del zapoteco (con la *Doctrina Christiana en lengua castellana y Zapoteca* de Fray Pedro de Feria), el náhuatl (Fray Alonso de Molina), el purépecha (Fray Maturino Gilberti), el mixteco (Fray Antonio de los Reyes y Fray Francisco de Alvarado), el kaqchikel (fray Domingo de Vico), entre otras, son ejemplos de lenguas con trilogía catequística.

<sup>3</sup> En fonología, la mora ( $\mu$ ) es la unidad métrica mínima que permite contrastar entre sílabas pesadas y sílabas ligeras. Generalmente, las vocales breves equivalen a una mora y las vocales largas a 2 moras (Arellanes, 2009, p. 333).

## Referencias

- Aikhenvald, A. (2013). Possession and Ownership: A Cross-linguistic Perspective. En A. Aikhenvald y R. M. W. Dixon (Eds.), *Possession and Ownership. A Cross-linguistic Typology* (pp. 1-64). Oxford University Press.
- Antonio, P. (2015). *La fonología y morfología del zapoteco de San Pedro Mixtepec* [Tesis doctoral]. CIESAS.

- Antonio, P. (2022). Las clases verbales definidas por la alomorfía de tiempo, aspecto y modo en el zapoteco de San Pedro Mixtepec. En F. Arellanes, M. Chávez Peón, y R. Rojas (Eds.), *Estudios descriptivos de lenguas zapotecas. Fonética-fonología, morfosintaxis y semántica léxica* (pp. 227-250). UNAM.
- Arellanes, F. (2009). *El sistema fonológico y las propiedades fonéticas del zapoteco de San Pablo Güilá. Descripción y análisis formal* [Tesis doctoral]. El Colegio de México.
- Arellanes, F. (2020). De las partes del cuerpo a la pertenencia. ¿Por qué la distinción alienable-inalienable no es suficiente en zapoteco? *Amerindia*, 42, 49-73.  
<https://doi.org/10.56551/PFYB2080>
- Ashmore, L. (2008). The role of digital video in language documentation. En P. K. Austin (Ed.), *Language Documentation and Description*, 5 (pp. 77-102). SOAS.  
<https://www.lddjournal.org/article/id/1083/>
- Avelino, H. (2004). *Topics in Yalálag Zapotec, with particular reference to its phonetic structures* [Tesis doctoral]. University of California, Los Angeles.
- Beam, R. (2019). Southern Zapotec verb classes. En E. Palancar, T. Feist, y M. Bearman (Eds.), *Amerindia*, 41 (pp. 121-165). E. A. E. <https://doi.org/10.56551/NVJZ1964>
- Campbell, E. (2017). Otomanguean historical linguistics: Exploring the subgroups. *Language & Linguistics Compass*, 11. <https://doi.org/10.1111/lnc3.12244>
- Campbell, L., Kaufman, T., y Smith-Stark, T. (1986). Meso-America as linguistic area. *Language*, 62, 530-570. <http://www.jstor.org/stable/415477>
- Chamoreau, C. (2013). Diversidad lingüística en México. En C. Chamoreau (Ed.), *Amerindia*, 37(1), 3-20. A. E. A. <https://amerindia.cnrs.fr/wp-content/uploads/2021/02/Chamoreau-C.-Diversidad-lingüística-en-México-un-acercamiento.pdf>
- Chávez Peón, M. (2010). *The interaction of metrical structure, tone and phonation types in Quiaviní Zapotec* [Tesis doctoral]. University of British Columbia.
- Córdova, J. (1987a). *Arte del idioma zapoteco* [Original publicado en 1578]. Ediciones Toledo-INAH.
- Córdova, J. (1987b). *Vocabulario en lengua apoteca* [Original publicado en 1578]. Ediciones Toledo-INAH.
- Corbett, G. (2009). Canonical inflectional classes. En F. Montermini, G. Boyé, y J. Tseng (Eds.), *Décembrettes 6: morphologie et classes flexionnelles* (pp. 1-11). Somerville.
- Fernández de Miranda, T. (1995). *El protozapoteco*. El Colegio de México-INAH.
- Golluscio, L., Pacor, P., Ciccone, F., y Krasan M. (2019). *Lingüística de la documentación. Textos fundacionales y proyecciones en América del Sur*. Eudeba.
- Grinevald, C. (2003). Speakers and documentation of endangered languages. En P. K. Austin (Ed.), *Language Documentation and Description*, 1 (pp. 52-72). SOAS.  
<https://doi.org/10.25894/ldd306>

- Grupo especial de expertos sobre las lenguas en peligro convocado por la UNESCO. (2003). *Vitalidad y peligro de desaparición de las lenguas*. UNESCO.
- Gutiérrez, L. (2021). *A description and analysis of the syntax and functions of the subordinate clauses in Teotitlán del Valle Zapotec (with an introductory overview of TdVZ phonology and morphosyntax)* [Tesis doctoral]. The University of Texas at Austin.
- Hernández, E., y López, O. (2022). *Diccionario zapoteco (di'dz xi'dza')*. Instituto Nacional de los Pueblos Indígenas.
- Himmelmann, N. (2007). La documentación lingüística: ¿qué es y para qué sirve? En J. Haviland y J. A. Flores (Coords.), *Bases de la documentación lingüística* (pp. 15-47). INALI.
- INALI. (2009). *Catálogo de las Lenguas Indígenas Nacionales. Variantes Lingüísticas de México con sus autodenominaciones y referencias geográficas*. Instituto Nacional de Lenguas Indígenas.
- López, O. (2019). Clases flexivas verbales del zapoteco de Zochina. En E. Palancar, T. Feist, y M. Bearman (Eds.), *Amerindia, 41* (pp. 167-197). E. A. E. <https://doi.org/10.56551/XMOH2204>
- López, O. (2020). Mecanismos gramaticales de la posesión en el zapoteco. En O. Vázquez y P. Martínez (Coords.), *Patrimonio cultural de Oaxaca: investigaciones recientes* (pp. 228-251). INAH.
- Montaño, C., Arellano, I., y Sánchez, L. (2018). *¿Qué hacer para registrar y preservar una lengua? Aproximaciones a la documentación lingüística*. Ediciones del Lirio.
- Moseley, C. (2010). *Atlas de las lenguas del mundo en peligro* (2<sup>a</sup> ed.). UNESCO.
- Nellis, D., y Hollenbach, B. (1980). Fortis versus lenis in Cajonos Zapotec phonology. *International Journal of American Linguistics*, 46, 92-105. <https://doi.org/10.1086/465639>
- Pérez, A., y Hernández, A. (2020). *Propuestas metodológicas para el trabajo y la investigación lingüística. Aplicaciones teóricas y descriptivas*. Universidad de Colima.
- Quatra, M. (2011). Auto-documentación Lingüística: La experiencia de una comunidad Jodí en la Guayana Venezolana. *Language Documentation and Conservation*, 5, 134-156. <http://hdl.handle.net/10125/4495>
- Reyes, G. (1891). *Artes de las lenguas cerrana y de el valle* [Original publicado en 1700]. F. Belmar (Ed.). Imprenta del Estado.
- Sicoli, M. (2021). *Saying and Doing in Zapotec. Multimodality, Resonance, and the Language Joint Actions*. Bloomsbury Press.
- Silverman, D. (1997). Laryngeal complexity in Otomanguean vowels. *Phonology*, 14(2), 235-261. <https://doi.org/10.1017/S0952675797003412>

- Smith-Stark, T. (2007a). Algunas isoglosas zapotecas. En C. Buenrostro, J. J. Rendón, O. Schumann, L. Valiñas, y M. A. Vargas (Eds.), *Clasificación de las lenguas indígenas de México. Memorias del III Coloquio Internacional de Lingüística Mauricio Swadesh* (pp. 69-133). UNAM-INALI.
- Smith-Stark, T. (2007b). La trilogía catequística: artes, vocabularios y doctrinas en la Nueva España como instrumentos de una política lingüística de normalización. En R. Barriga y P. Butragueño (Coords.), *Historia sociolingüística de México* (Vol. 1, pp. 451-482). El Colegio de México.
- Swadesh, M. (1949). El idioma de los zapotecos. En Mendieta y Núñez (Coords.), *Los zapotecos: monografía histórica, etnográfica y económica* (pp. 417-448). UNAM.
- Vásquez, N. (2024). *Construcciones multiverbales monoclausales en el zapoteco de San Pedro Mixtepec* [Tesis doctoral]. CIESAS.
- Woodbury, A. (2011). Language documentation. En P. K. Austin y J. Sallabank (Eds.), *The Cambridge Handbook of Endangered Languages* (pp. 159-186). Cambridge University Press.
- Yip, M. (2002). *Tone*. Cambridge University Press.