

LOS COMPUESTOS VERBALES, COMPLEJOS VERBALES Y EXPRESIONES IDIOMÁTICAS DEL ZAPOTECO MIAHUATECO DE SAN BARTOLOMÉ LOXICHA

Rosemary Grace Beam de Azcona¹
ENAH

Thabgol² Emiliano Cruz Santiago
Biblioteca de Investigación Juan de Córdova

RESUMEN. Este trabajo examina cómo los verbos flexionados interactúan con raíces adicionales que pueden o no ser parte de la misma palabra prosódica, sintáctica y/o el mismo ítem léxico en el *dí'zdéh*, la variante del zapoteco miahuateco hablado en San Bartolomé Loxicha. Se comparan tres estructuras: los compuestos verbales, los complejos verbales y las expresiones idiomáticas. Los compuestos y complejos verbales representan etapas intermedias en el proceso de la gramaticalización, mientras que las expresiones idiomáticas pueden ser precursores a los compuestos. Una comparación de estos fenómenos es de interés tipológico porque demuestran cómo una lengua particular puede definir la noción de “palabra” en diferentes dominios, pero estas cuestiones también tienen aplicaciones prácticas en el diseño de la ortografía y en la lexicografía.

PALABRAS CLAVE: zapoteco miahuateco, compuestos verbales, expresiones idiomáticas, incorporación nominal, complejo verbal, palabra prosódica

1. INTRODUCCIÓN

Desde 2005 hasta la muerte repentina del segundo autor en 2015, colaboramos en la documentación del *dí'zdéh* [di^ʰi'zɪdæːʎ] o zapoteco miahuateco de San Bartolomé Loxicha. Dos proyectos centrales de nuestra colaboración han sido un diccionario (Cruz Santiago y Beam de Azcona, en preparación) y una colección de cuentos glosados (Cruz Santiago, en preparación). Tanto en la lexicografía como en la transcripción ortográfica de los textos, uno de los desafíos más interesantes ha sido distinguir entre frases y palabras que consisten en más de una raíz. El meollo del asunto es una pregunta de interés universal en la lingüística: ¿qué es una palabra? El famoso aforismo de Talmy Givón de que *la morfología de hoy es la sintaxis de ayer*, es especialmente apto para las lenguas zapotecanas. En este trabajo nos enfocamos en los predicados verbales y cómo un verbo flexionado se puede combinar con otras raíces en diferentes estructuras morfosintácticas.

Una palabra prototípica en una lengua zapotecana consiste en una raíz. Si la palabra es un verbo también tendrá un prefijo flexivo. Podríamos tomar como un ejemplo una forma

¹ Este trabajo fue presentado por los autores en CILLA en 2011. La primera autora ha convertido el volante en el presente texto. Tras la muerte del segundo autor, se han checado dudas con su hermana, Zenaida Cruz Sebastián. Otras personas que han nutrido este trabajo incluyen Abdón Cruz Cortés, Thabgol Nemecio Gaspar, Ambrocio Gutiérrez Lorenzo, Thabgol Terry Kaufman, Mario Luna, Carolyn O'Meara, Thabgol Thom Smith Stark y Norma Leticia Vásquez Martínez. Los datos fueron recopilados gracias al apoyo del Endangered Languages Documentation Programme y difundidos aquí gracias a los editores de este volumen. Cualquier error es responsabilidad de la primera autora.

² Este término en el *dí'zdéh* de San Bartolomé Loxicha se usa para indicar un finado adulto. Se usa este término aquí para reconocer de manera respetuosa y no colonial que ésta es una publicación póstuma.

como *dbáan* ‘se despertará’ que se compone del prefijo futuro definido *d-* más la raíz verbal *báan*. La raíz consiste en una sílaba tónica y podemos considerar que *dbáan* cuenta como una palabra prosódica. No es posible agregar otros morfemas dentro de la secuencia indisociable *dbáan* y podemos suponer que esta forma es una sola palabra también desde una perspectiva sintáctica. ‘Despertarse’ es un sentido léxico particular y este verbo se debe alistar en el diccionario de esta lengua. Si definimos la “palabra” según criterios fonológicos, sintácticos o semánticos, bajo cualquier definición ‘despertarse’ se va a considerar como una palabra. Las palabras de más de una raíz, que son el enfoque de este capítulo, no son palabras prototípicas, pero sí son comunes en lenguas zapotecanas. Al comparar palabras prototípicas con las distintas combinaciones de raíces verbales con otras raíces, percibimos por lo menos tres grandes categorías: compuestos, expresiones idiomáticas y complejos verbales.

En este capítulo comparamos las estructuras ejemplificadas en (1-3). La raíz verbal ‘meter’ se combina con el sustantivo ‘aire’ en un compuesto en (1), donde esta combinación resulta en el nuevo concepto léxico de ‘inflar’. Este compuesto cuenta como una sola palabra prosódica y es sintácticamente indisociable porque el sujeto *mén* viene después de toda la secuencia *kombü* y no después del núcleo verbal *ko*. La expresión idiomática en (2) también incluye ‘meter’ y ‘aire’ pero como palabras independientes en una estructura sintáctica que cuenta como un ítem léxico, donde aquí el significado es ‘abanicar’. En (3) hay una combinación de los verbos ‘ir’ y ‘pedir’ que vamos a denominar “complejo verbal”. El complejo verbal, al igual que el compuesto en (1), cuenta como una palabra prosódica y sintáctica, pero esta combinación de verbos no es un ítem léxico con un significado particular.

(1) Compuesto verbal

<i>Kombü</i>	<i>mén</i>	<i>pelot</i> ³ .
k-o+m-b-î	mén	pelòt
POT ⁴ \R1-meter+CL-DER-aire	gente	pelota

‘Alguien va a inflar la pelota.’

(2) Expresión idiomática

<i>Ko</i>	<i>Cheén</i>	<i>mbü</i>	<i>loó</i>	<i>Yovani</i> .
k-ò	tʃën	m-b-î	lõ	jobani
POT\R1-meter	Arsenio	CL-DER-aire	cara/a	Yovani

‘Arsenio va a abanicar a Yovani.’

³ En todo el trabajo, los ejemplos incluyen una representación ortográfica en cursivas en una primera línea, seguida por una representación en AFI con cortes morfémicos en la segunda.

⁴ Las abreviaciones usadas aquí son las siguientes: ADV = adverbio, APL = aplicativo, CAUS = causativo, CL = término de clase, DEM = demostrativo, DER = prefijo derivativo, DIST = distal, EST = estativo, HAB = habitual, INF = infinitivo, INTER = interrogativo, IPFV = imperfectivo, LOC = locativo, M = forma que se combina con un verbo de movimiento, MED = medial, NEG = negación, OBJ = objeto, PART = participio, PFV = perfectivo, PL = plural, POS = posesión, POT = potencial, PROG = progresivo, PROX = proximal, R = realis, REL = relativizador, R1 = reemplacivo 1 (prefijo derivacional para formar el tema habitual), R2 = reemplacivo 2 (prefijo derivacional para formar el tema perfectivo), SG = singular, SR = sustantivo relacional, SUJ = sujeto, TERM = terminativo, X = morfema de glosa desconocida, 1EXCL = primera persona exclusiva, 1SG = primera persona singular, 2FAM = segunda persona familiar, 3H = tercera persona humana, 3HR = tercera persona humana de respeto. Para indicar la concatenación de afijos usamos + para separar raíces o ítems léxicos en compuestos y complejos verbales, = para clínicos, \ para indicar morfología fusional, . para morfemas *portmanteau* y / para indicar dos posibles glosas.

(3) Complejo verbal

<i>Ndaána'b</i>	<i>xi'y</i>	<i>dīb.</i>
nd-ǎ+nà²b	ǰi²=j	dīb
EST-ir+M\pedir	3H=3INAN	otro
'Iba a pedirlo otra vez.'		

Para escribir todos estos ejemplos ortográficamente en la colección de textos (Cruz Santiago, en preparación), y para saber qué enlistar en el diccionario (Cruz Santiago y Beam de Azcona, en preparación) como entradas léxicas, tenemos que determinar no solamente que se considere como una “palabra” sino también determinar qué tipo de palabra es. La Tabla 1 resume qué tipos de palabras son las estructuras en (1-3), según nuestro análisis.

	Una palabra prosódica	Una palabra sintáctica	Un ítem léxico
Compuestos: <i>kombíi</i>	✓	✓	✓
Expresiones idiomáticas: <i>ko...mbíi</i>	--	--	✓
Complejo verbal: <i>ndaána'b</i>	✓	✓	--

Tabla 1: Tres tipos de "palabra".

El trabajo está organizado de la siguiente manera: §2 contiene información general sobre la lengua, §3 trata los compuestos verbales, §4 los complejos verbales y §5 las expresiones idiomáticas. La sección §6 aplica los análisis a las cuestiones prácticas de la ortografía y lexicografía. Finalmente, §7 considera la evolución diacrónica de estas estructuras, antes de concluir en §8.

2. LA LENGUA

La Figura 1 muestra el zapoteco miahuateco y sus vecinos. Esta lengua cuenta con alrededor de 200,000 hablantes y se puede dividir en las zonas dialectales indicadas con las formas más oscuras. La variante de San Bartolomé Loxicha (marcada) tiende a ser conservadora en cuanto a su fonología segmental, aunque hay innovación en el sistema tonal entre los hablantes nacidos a partir de los 1980. Hablantes jóvenes de la variante vecina de San Agustín Loxicha reportan dificultades en la comprensión del habla de San Bartolomé, aunque la gente grande reporta mayor inteligibilidad. La variante de Santo Domingo de Morelos (Hernández Luna, este volumen) pertenece a la misma zona dialectal de San Agustín Loxicha. Nosotros, incluyendo el autor nativo hablante, consideramos que la variante de San Agustín, Santo Domingo, Candelaria, Magdalena y otros pueblos, y la variante de San Bartolomé que se describe aquí, son variantes de la misma lengua, aunque también se podrían considerar como lenguas emergentes debido a la dificultad de comprensión que reportan algunos hablantes jóvenes de la variante más extensa geográficamente. También consideramos que San Agustín Mixtepec (Beam de Azcona, 2004; Hernández Luna, 2019) es una variante de la misma lengua, debido a la habilidad del segundo autor y su familia de comprenderla, aunque no hemos hecho la

prueba inversa para ver si hablantes de San Agustín Mixtepec comprenden esta variante.

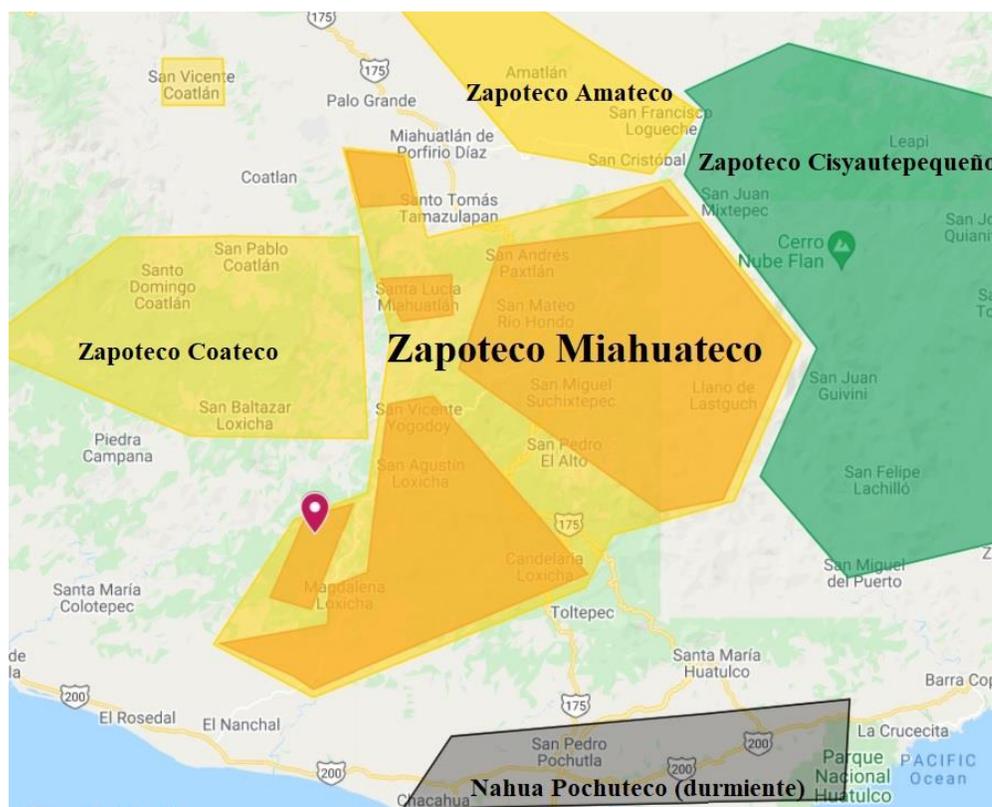


Figura 1. Zonas dialectales del zapoteco miahuateco y sus vecinos (Datos de mapa: Google)

Entre los datos generales de la lengua más relevantes para el problema de definir las palabras están algunos detalles sobre la prosodia (§2.1), el orden sintáctico y el templete morfológico de los verbos (§2.2), que veremos ahora.

2.1 La prosodia

Esta variante de miahuateco tiene tres tipos de contrastes suprasegmentales. La nasalización resulta de la enclitización histórica o subyacente, y la indicamos ortográficamente con <-n> después de la vocal. El rasgo glotal se escribe <'> y se puede realizar como una oclusiva glotal plena, como laringización, rearticulación o una combinación de estas, pero estas diferencias no son contrastivas. Los tonos existentes en esta variante se enlistan en la Tabla 2. La transcripción de ejemplos en este trabajo representa la pronunciación de Cruz Santiago. Cuando el tono alto se encuentra en un enclítico, se realiza con un alótono alto de nivel sin descenso. Antes de la muerte de Cruz Santiago, estuvimos contemplando si esta alotonía predecible se debería representarse ortográficamente o no, y esta decisión se quedó pendiente. En este trabajo se sigue la propuesta de Cruz Santiago, en la que se indica el alótono alto en enclíticos con dos vocales acentuadas <áá>. Noten que una vocal sin diacrítico puede indicar el tono bajo en una sílaba tónica o, si es una sílaba átona, puede indicar la falta de un tono contrastivo.

Etiqueta	Alótono principal para Cruz Santiago	Representación ortográfica	Representación en AFI
“alto”	ʌ	á	á
(alto en enclítico)	↑	áá	á
“bajo”	↓	a	à
(σ átona sin tono)	↓ ~↑	a	a
“descendente”	∨	áa	â
“ascendente”	∧	aá	ǎ

Tabla 2. Los tonos del zapoteco miahuateco de San Bartolomé Loxicha.

Todas las palabras prosódicas en esta lengua tienen una sílaba prominente en la que hay un contraste tonal. Las raíces en esta lengua consisten en una sola sílaba, como en (4). Con la excepción de préstamos como *skerban* ‘escribano’, cuando hay una palabra con más de una sílaba es siempre resulta de complejidad morfosintáctica. Los prefijos suelen ser consonantes no silábicas, aunque el prefijo estativo *na-* es excepcional por derivarse de la cópula (Beam de Azcona, en prensa). La mayoría de las palabras nativas con más de una sílaba son combinaciones de más de una raíz, como en (5), o palabras con clíticos ligados, como en (6).

- (4) *líiz*
 líz
 ‘hogar, casa de uno’
- (5) *líizgui'b*
 líz+gì'b
 hogar+fierro
 ‘la cárcel’
- (6) *líizláa*
 líz=lâ
 hogar=2FAM
 ‘tu casa’

Cuando la palabra prosódica contiene más de una raíz, como en (5), la sílaba que corresponde a la última raíz es la más prominente. Fonéticamente esta prominencia se realiza con mayor duración e intensidad que en las sílabas menos prominentes. Fonológicamente las sílabas átonas son vulnerables a la pérdida de contrastes mientras que las sílabas prominentes retienen todos sus contrastes. Hay pocos casos de reducción segmental, pero mayormente lo que ocurre es la neutralización de contrastes suprasegmentales.

Consideramos que el dí'zdéh es una lengua tonal que también tiene lo que Hyman (2006: 231) llama *stress accent*. Para Hyman, un sistema con *stress accent* tiene exactamente una sílaba acentuada por palabra, ni más ni menos. Sostenemos que en el dí'zdéh este criterio aplica y definimos la palabra prosódica como la sílaba tónica más las sílabas pretónicas que la preceden. En términos morfológicos, esto es equivalente a una raíz prominente, más cualquier prefijo o raíz no prominente que la preceda. (4) y (5) serían palabras prosódicas.

La secuencia de la palabra prosódica más los enclíticos ligados después de la raíz prominente, como en (6), son una unidad mayor (ver la noción de *clitic group* en Nespoy y Vogel, 2007: 145-163). Los clíticos están fuera del alcance de este capítulo, así que solamente nos interesa la palabra prosódica.

Aunque una sílaba pretónica puede contar con los mismos contrastes tonales que una sílaba tónica, también puede pronunciarse sin tono contrastivo, cosa que no hemos observado en ninguna sílaba tónica. No tenemos espacio en este capítulo para indagar sobre la prosodia de las frases y enunciados, pero notamos que aún palabras gramaticales con reducción segmental, como el artículo indefinido *thí*, derivado del número ‘uno’ *thíb*, tienen tonos contrastivos. Por esta razón vamos a suponer que sílabas sin tono siempre son átonas y forman parte de la misma palabra prosódica que la siguiente sílaba tónica.

2.2 El orden morfosintáctico

Para distinguir entre compuestos y frases es importante tomar en cuenta el orden de los morfemas: (7) indica el orden de argumentos con relación al verbo.

- (7) Orden canónico en la oración:
VERBO SUJETO OBJETO (frases periféricas).

Cada verbo consiste mínimamente de una raíz verbal flexionada por una categoría aspectomodal. Hay otros morfemas que pueden ir ligados. (8) muestra el templete morfosintáctico de los verbos. Aunque mostramos distintas posiciones en el templete por separado, en casos particulares dos, o inclusive tres, morfemas adyacentes se pueden fusionar (Beam de Azcona, bajo dictamen).

- (8) Templete de morfemas ligados al verbo:
(REALIS)-AM⁵-(DER)-RAÍZ-((DER)-RAÍZ)⁰⁻²=(APL)=(NEG)=(ADV)=(SUJETO)=(OBJETO)

En (7) la observación más importante es que en el orden canónico el sujeto viene después del verbo. La posición del sujeto puede así demarcar la terminación del verbo. Sin embargo, como nos expresó alguna vez Thom Smith Stark, hay ocasiones en que puede ser ambiguo si una raíz nominal funge como el sujeto o como un elemento incorporado en el verbo. Para resolver esta ambigüedad recurrimos a los clíticos adverbiales, de negación y al aplicativo en el templete de (8). Como estos clíticos ocurren antes del sujeto y después de cualquier raíz incorporada al verbo, son una prueba útil para determinar estas fronteras morfosintácticas. En (9) se muestra el verbo simple ‘comer’, de una sola raíz, con el enclítico de negación al final del verbo, mientras que en (10) un enclítico adverbial se agrega después del morfema ‘palabra’, mostrando que ‘palabra’ está incorporada y ya cuenta como parte del verbo.

- (9) Enclítico después de verbo simple
Mdáaza' *Mari guéehth.*
m-d-â=zà[?] mari gâθ

⁵ Muchos llaman TAM a estos prefijos, pero concordamos con López Nicolás (2016) que no codifican el tiempo tanto como el aspecto y el modo en lenguas zapotecanas.

R.PFV-R2-comer=NEG María tortilla
 ‘María no comió la tortilla.’

- (10) Enclítico después de verbo compuesto
Ndxóodi'zthoóz *rop xa'*
 n-ḍ̄z-ô+di'z=θōz ròp ʃã'
 R-HAB-pelear+palabra=mucho ambos 3H
 ‘Mucho/constantemente discuten verbalmente ellos.’

Una vez visto el orden en que ocurren los morfemas en verbos y oraciones, y también los patrones prosódicos de la lengua, podemos considerar tres tipos de “palabras” complejas y evaluarlas según estos criterios.

3. LOS COMPUESTOS VERBALES

Los compuestos verbales en el dí'zdéh se forman mediante la incorporación de raíces adicionales al núcleo en el verbo. En §3.1 consideramos las clases léxicas de las cuales provienen las raíces incorporadas, donde lo más común es la incorporación nominal y en §3.2 consideramos este tipo de compuesto respecto a las propiedades de los sustantivos incorporados y de los verbos con los cuales se combinan. Finalmente, en §3.3 resumimos los criterios para analizar los compuestos como palabras en sí y no secuencias sintácticas.

3.1 La incorporación de diferentes clases léxicas

Los verbos y los sustantivos son las clases léxicas mayores en el dí'zdéh. Los adjetivos y adverbios forman clases pequeñas. Raíces de estas cuatro clases léxicas se pueden incorporar para formar compuestos verbales, pero la incorporación nominal es la más común. Primero, consideramos la incorporación marginal de las otras clases antes de enfocarnos en los sustantivos en el siguiente subapartado.

La mayoría de los adverbios en la lengua ocurren como enclíticos, pero hemos analizado algunas palabras independientes que ocurren en la periferia como adverbios. De estas palabras, únicamente hemos visto a *ga'p* ‘alto, arriba’ incorporarse. En (11) ocurre independientemente al final de la oración y en (12) se incorpora en el verbo. Aunque hemos analizado esta palabra como un adverbio, otra opción sería analizarla como un sustantivo locativo que aquí se refiere al espacio donde saltó.

- (11) *Ngwáat xa' ga'p.*
 ŋ-gw-ât ʃã' gà'p
 R-PFV-saltar 3H alto
 ‘Saltó alto.’

- (12) *Ndxáatga'p xa' por náal ndyáak xa'.*
 n-ḍ̄z-ât+ga'p ʃã' por nâl n-dyâk ʃã'
 R-HAB-saltar+alto 3H por frío R-HAB\sentir 3H
 ‘Se estremece porque tiene frío’ o ‘hasta se sobresalta por el frío (que siente).’

La mayoría de los adjetivos que existen son derivados de verbos y en casos como (13)

es difícil saber si estamos viendo la incorporación de un adjetivo o la incorporación de un verbo.

- (13) *gólbil*
 g-ól+bìl
 POT-cantar+triste
 POT-cantar+entristecerse⁶
 ‘cantar entristecido.’

Otros adjetivos no parecen ser derivados tan recientemente y no existen las mismas ambigüedades. En (14) *jeh* es una palabra independiente que modifica a *maánh*, con la implicación de que el mango no estaba listo para comerse. En (15) vemos al adjetivo *jeh* pegado al verbo ‘abrir’, que da el sentido de ‘abrir forzosamente’, es decir, que sus piernas se abrieron a una posición sin estar en condiciones para hacerlo.

- (14) *Maánh yeh mdáa xa' za mbro' guéh'z ró xa'.*
 mǎŋ jǎ m-d-â jǎ' zà m-b-rò' gǎ'z ró jǎ'
 mango tierno R.PFV-R2-comer 3H entonces R-PFV-salir grano boca 3H
 ‘Mangos tiernos comió, por eso le salieron granos en la boca.’

- (15) *Ndyakaán mbi'lyeh rop xkuún por mdóbáan waáy.*
 n-djǎk=ǎ m-bi'l+jǎ ròp jkǔ-n por m-dób=ǎ wǎj
 R-HAB\sentir=1SG R-PFV\abrir+tierno ambos pierna-POS por R.PFV-sentar=1SG caballo
 ‘Siento que se me abrieron forzosamente las piernas porque monté el caballo.’

Aunque no es tan común como la incorporación nominal, también es posible combinar dos verbos y así obtener un nuevo concepto léxico que es más de la suma de sus partes. Es posible que los compuestos en (16) y (17) tengan un origen histórico en construcciones de verbos seriales (Gutiérrez Lorenzo, 2014) particulares, pero sincrónicamente la observación pertinente es que el verbo compuesto tiene un sentido diferente del núcleo flexionado y también diferente del verbo incorporado.

- (16) *Ñá thíb tsi'káa meétr ndxattó béhl.*
 ná θɪb tsi'=ká mǎtr n-dǔ-at+tó bǎl
 HAB\verse uno diez=como metro R-HAB-saltar+parar lumbre
 ‘El fuego se elevaba como diez metros.’

- (17) *Gaáldra' dó' páar bí'lkeh' laár.*
 g-ǎl=drǎ' dó' pǎr bí'l+kǎ' lǎr
 POT-alcanzar=NEG mecate para POT\abrirse+colgarse ropa
 ‘Ya no hay tendadero para que se tienda la ropa.’

No estamos seguros de la incorporación de raíces adverbiales, ya que el único ejemplo

⁶ Más abajo introducimos la noción de “infinitivos” (equivalentes a “gerundios modales” en la terminología de Smith Stark, 2003), pero en este verbo y también ‘parar’ en (16) y ‘colgarse’ en (17), esta forma es idéntica a la raíz no marcada. Por ahora basta reconocer que estas tres formas carecen de marcación especial.

se podría reanalizar como nominal. La incorporación de adjetivos y verbos no es tan ambigua, pero nos parece un fenómeno marginal en comparación con la incorporación nominal.

3.2 La incorporación nominal

Al examinar los verbos que se forman mediante la incorporación nominal, podemos hacer ciertas distinciones sintácticas y semánticas. Algunas tienen que ver con el nominal incorporado y con otras frases nominales en la oración que pueden o no promoverse a ser argumentos más centrales. Mithun (1984) provee una tipología adecuada para analizar los compuestos del *dí'zdéh* según estos criterios (§3.2.1). Sin profundizar mucho, podemos agregar que es posible también analizar algunos compuestos según la raíz verbal, ya que algunos compuestos parecen ser formas lexicalizadas de construcciones de verbo ligero (§3.2.2).

3.2.1 Las relaciones gramaticales y los roles semánticos

Mithun (1984) distingue entre 4 tipos de incorporación y señala que se desarrollan de manera evolutiva, al ser el tipo I el más frecuente y el tipo IV el más escaso. En el *dí'zdéh* solamente hemos observado los tipos I y II. Para Mithun, el tipo I refiere a la composición léxica en que un paciente, instrumento, o lugar se incorpora en el verbo para indicar una actividad o estado de que el referente del sustantivo incorporado es emblemático. Este tipo de incorporación puede involucrar una reducción de valencia, ya que el sustantivo incorporado deja de tener relaciones gramaticales.⁷ El tipo II difiere del tipo I en que un argumento periférico puede ocupar las relaciones gramaticales del sustantivo incorporado.

El tipo I de incorporación nominal simplemente reduce la valencia del verbo cuando deriva predicados intransitivos de otros transitivos, pero el tipo II promueve un argumento oblicuo al rol gramatical vacante por el sustantivo incorporado. Cuando un verbo transitivo incorpora su objeto directo, entonces un instrumento, lugar, o poseedor puede asumir el rol de objeto. Cuando un verbo intransitivo incorpora su sujeto, otro argumento puede tomar su estatus de sujeto. (Mithun 1984: 856).⁸

El tipo II de Mithun se ilustra bien comparando los ejemplos (18) y (19). En (18) 'nacer' es un verbo intransitivo cuyo sujeto es el paciente. Lo que nace es la pluma. Donde nacen las plumas, una expresión locativa, es en el cuerpo del pollito, que se

⁷ “[...] a V stem and a N stem are combined to form an intransitive predicate denoting a unitary concept. The compound is more than a description; it is the name of an institutionalized activity or state. The IN loses its individual salience both semantically and syntactically. It no longer refers to a specific entity; instead, it simply narrows the scope of the V. It is thus unaccompanied by markers of definiteness or number, or by demonstratives. Although it may function semantically as a patient, location, or instrument, it has no independent syntactic role in the sentence as a whole” (Mithun 1984, p.856).

⁸ “Type I NI simply lowers the valence of the V when it derives intransitive predicates from transitive ones; but Type II NI advances an oblique argument into the case position vacated by the IN. When a transitive V incorporates its direct object, then an instrument, location, or possessor may assume the vacated object role. When an intransitive V incorporates its subject, another argument may be advanced to subject status”. [Traducción propia].

expresa mediante una frase nominal poseída. En la sintaxis de (18) el ave es el poseedor del adjunto locativo. El ejemplo (19) expresa dicho evento, pero ahora ‘pluma’ está incorporado como parte del verbo y el perico ahora cuenta como sujeto.

- (18) *Ngótsláá* *dó’b* *lád* *dxuúl* *yéh*
 η-góts=lá dó’b [lád d̥zúl jǎe]_{LOC}
 R-PFV\nacer=ya pluma cuerpo pollito DEM
 ‘Ya le salieron las plumas a este pollito.’

- (19) *Ngótsdo’bláá* *mguíl*
 η-góts+dò’b=lá [m-gíl]_{SUJ}
 R- PFV\nacer+pluma=ya CL-perico
 ‘Ya emplumeció el perico.’

Según el argumento de Mithun (1984), si una lengua tiene el tipo II de incorporación, debería tener el tipo I también. En el dí’zdéh el tipo II parece ser el más común, pero sí existen ejemplos del tipo I, por ejemplo, en (20), donde la frase con *loó* es un adjunto locativo opcional. Este ejemplo tiene el mismo sujeto que tendría si la palabra ‘lengua’ fuera el objeto directo. Aquí el verbo pierde un argumento al incorporar el presunto objeto directo y no promueve ningún adjunto como argumento central.

- (20) *Mbo’loz* *mbeh’l* (*loó* *nó’*).
 m-bo’+lòz m-b-ǎ’l (lǒ nó’)
 R.HAB-sacar+lengua CL-DER-culebra (cara/a 1EXCL)
 ‘(Nos) sacó la lengua la culebra.’

Lo más común (tipológicamente y para esta lengua) es la incorporación de un paciente (o tema), sea este un presunto objeto directo como en (20) o un presunto sujeto intransitivo como en (19). Como las raíces verbales que son los núcleos de los compuestos verbales pueden ser transitivos o intransitivos, a veces también hay compuestos equivalentes, pero con diferencias de valencia. Los ejemplos (21) y (22) tienen verbos compuestos con ‘aire’ incorporados, aunque ‘globo’ (lit. ‘piel de aire’) es el objeto en (21) y el sujeto en (22). Podemos considerar que en (21) hay incorporación del presunto objeto, lo cual hace posible que un presunto locativo (donde uno mete el aire) sirva como objeto, mientras que en (22) hay incorporación de un sujeto intransitivo, y de igual manera hace posible que un presunto locativo se vuelva sujeto de un compuesto intransitivo. Así que estos también serían ejemplos del tipo II de Mithun.

- (21) *Mblómbú* *xa’* *guídmbí.*
 m-b-l-ó+m-b-î fǎ’ gíd+m-b-î
 R-PFV-R2-meter+CL-DER-aire 3H piel+CL-DER-aire
 ‘(Él/ella) infló el globo.’

- (22) *Yombú* *guídmbí.*
 j-o+m-b-î gíd+m-b-î
 POT\estar.metido+CL-DER-aire piel+CL-DER-aire
 ‘Se va a inflar el globo.’

Además de pacientes, hemos observado la incorporación de instrumentos (23-25).

- (23) *Tsaththoóz* *mbíixdo'* *má'*
 tsàθ=θōz m-bîf+dò⁹ má'
 fuerte=mucho R.PFV-envolverse+mecate animal
 'La bestia se había enredado muy fuerte.'
- (24) *Puír* *nkeníi* *waáy* *ba*
 pür ŋ-ke⁹+nî wäj bà
 puro R.HAB-pegar+pie caballo DEM
- or* *zii* *xó'b* *mén* *siil* *díits* *xa'*
 òr zĩ jó⁹b mén sīl dīts fã⁹
 hora REL POT\colocar gente silla espalda 3H¹⁰
 'Ese caballo patea cada vez que uno lo ensilla.'
- (25) *Kabródáá* *lú* *loó* *xúudláa!*
 kab+ró=dá lú lõ fūd=lâ
 POT\responder+boca=NEG 2FAM cara POS\papá=2FAM
 '¡No le rezongues a tu papá!'

3.2.2 Los compuestos con verbos ligeros

Las lenguas zapotecas tienen construcciones de verbo ligero en las que el verbo 'hacer' o una cópula 'hacerse' sirven para recibir la flexión, pero la semántica del evento o estado expresado por el predicado se indica mediante un sustantivo que funge formalmente como el objeto directo del verbo 'hacer' o como el sujeto o complemento de la cópula (Beam de Azcona 2017: 51). Abajo mostramos algunos ejemplos del dí'zdéh. En (26) el sustantivo 'trabajo' es el objeto directo de 'hacer'. 'Hacer' recibe la flexión de los prefijos realis y habitual, pero 'trabajo' provee la semántica de la acción de trabajar. En (27) la cópula 'hacerse' recibe la flexión y 'milpa' funciona como su sujeto. En este ejemplo la gente se esforzaba por regar la tierra para que se diera la milpa; de que las semillas de maíz brotaran, y que la planta creciera y que diera frutos. Esto se comunica al oyente con la palabra 'milpa' porque uno sabe la naturaleza de la milpa y cómo crece. El contenido semántico de 'hacerse' es prácticamente nulo. En (28) un infinitivo del castellano se presta al dí'zdéh, donde funciona como complemento de la cópula. En este ejemplo la cópula recibe la flexión, pero el infinitivo prestado provee el contenido semántico.

- (26) *yé'n* *xomoó* *ndlí* *xa'* *dzí'n.*
 jé⁹n fò-mō n-d-lí fã⁹ dźí⁹n
 a.ver INTER-modo R-HAB-hacer 3H trabajo
 'A ver cómo trabaja ella.'

⁹ El verbo independiente 'pegar' tiene la vocal /æ/. Probablemente la elevación es una asimilación hacia la vocal en 'pie' y se puede considerar como una forma de reducción fonológica en la sílaba átona.

¹⁰ Se puede usar el pronombre humano para referir a animales muy conocidos o para indicar empatía.

necesarias en la tarea de la documentación de esta lengua. Si los compuestos son palabras fonológicas y sintácticas, se deben escribir sin un espacio entre las dos raíces cuando aparecen en textos y otros materiales. De igual manera, si los compuestos son ítems léxicos, merecen una entrada en el diccionario que estamos preparando.

3.3.1 Reducción fonológica

Los compuestos a veces muestran reducción fonológica en la (primera) raíz verbal. Esta reducción puede incluir la pérdida de la coda. Por ejemplo, *gak* ‘hacerse’ se reduce a *ga* en algunos compuestos, pero no en todos. Más común que la pérdida de la coda es la neutralización de los contrastes tonales en sílabas átonas. Por ejemplo, el potencial del verbo ‘sacar’ es /bóʔ/, pero ocurre sin tono y con vocal modal en el compuesto en (31). La pérdida del tono es más común que la pérdida del contraste glotal, aunque ambos ocurren.

(31) <i>Boró</i>	<i>xa'</i>	<i>chi</i>	<i>wli</i>	<i>jwá'n</i>	<i>mblí</i>	<i>xa'</i>
bo+ró	ʃäʔ	tʃi	w-li	xwáʔn	m-b-lí	ʃäʔ
POT.sacar+boca	3H	INTER	PART-enderezarse	cosa	R-PFV-hacer	3H
‘Debe de confesar si es cierto lo que hizo’						

La neutralización de contrastes suprasegmentales en las raíces átonas no es totalmente consistente. En este apartado hubo reducción tonal en el núcleo verbal (la primera raíz) de los compuestos en (16), (24-25) y (29), pero no en los compuestos de (12-13), (17), (19), (21) y (23). En particular, compare el tono alto en ‘hacer’ como verbo independiente en (26) versus la misma forma habitual del verbo en (29), donde aparece ligado a un nominal incorporado y carece de su tono subyacente. Otro ejemplo semejante es *kab* en (25), que tiene tono ascendente en esta forma potencial cuando no es parte de un compuesto. En los compuestos este tipo de reducción ocurre frecuentemente en la raíz átona pero nunca en la raíz prominente, que es la última, es decir, la incorporada; sin embargo, la reducción no siempre se aplica en la raíz átona. El mismo verbo que perdió su tono en el compuesto en (16) lo retuvo en el compuesto en (12).

La pérdida de la coda y la neutralización de los tonos son reducciones fonológicas que solamente toman lugar en sílabas átonas. Aún en sílabas pretónicas que no han perdido sus contrastes fonológicos, la menor duración e intensidad son pistas fonéticas de que estas sílabas son átonas. Como explicamos en §2.1, cada palabra prosódica tiene una sola sílaba tónica. Si las pistas fonéticas y la reducción fonológica indican que la primera raíz en los compuestos es átona, no puede formar su propia palabra prosódica diferente a la que incluye la última raíz, que es tónica, de otra forma, estas dos raíces deben ser parte de la misma palabra prosódica.

3.3.2 La impenetrabilidad sintáctica

En todos los ejemplos analizados como verbos compuestos en este apartado §3, el sujeto nunca interviene entre la raíz verbal y la raíz que, argumentamos, es incorporada. Al contrario, siempre ocurre después de la raíz incorporada o después de cualquier enclítico ligado a ella. De igual forma, el enclítico adverbial en (19) y los enclíticos de negación en (25) y (30) siempre vienen después de la raíz incorporada y nunca después del núcleo verbal. Los compuestos verbales son unidades indisolubles. Vimos que muchos verbos compuestos en el dízdéh son del tipo II de acuerdo con Mithun (1984), con una subcategorización diferente a la del mismo núcleo verbal cuando ocurre de manera independiente. Considerando

estos criterios, concluimos que los verbos compuestos son verbos. Es decir, la combinación de una raíz verbal flexionada con otra raíz ligada es un constituyente sintáctico mínimo. Podemos dividir estos verbos en diferentes morfemas, pero no en diferentes palabras, porque toda la secuencia cuenta como una sola palabra respecto a la sintaxis.

3.3.3 La semántica léxica

En los compuestos con verbos ligeros el núcleo verbal tiene una semántica casi vacía y el elemento incorporado contribuye con la mayor parte del significado. En otros casos es la combinación de los dos morfemas lo que crea un nuevo significado. Tender la ropa implica abrirla para máximo contacto con el aire y también colgarla para evitar contacto con una superficie que podría impedir el flujo del aire que va a secar la ropa. ‘Tender’ involucra el sentido de ‘abrir’ y también ‘colgar’ pero es un sentido diferente que el de cada verbo por separado. Puedes abrir cosas sin tenderlas y puedes colgar algo, como un mecate en un clavo en la pared, sin tenderlo. ‘Patear’ es pegar con el pie, pero puedes pegar con otras partes del cuerpo o instrumentos y no todo lo que haces con el pie es patear. Es la combinación de estas dos raíces lo que crea el significado especial de ‘patear’. En estos dos casos el compuesto tiene cierta lógica obvia, pero en otros casos el compuesto es más metafórico. Un aprendiz de esta lengua que ya conociera las distintas raíces probablemente captaría de inmediato que pegar+pie es ‘patear’ pero no que hacerse+excremento es ‘explotar’. Todos estos compuestos son ítems léxicos únicos, distintos de los que forman sus partes. Por ende, es necesario enlistar los compuestos en el diccionario.

4. LOS COMPLEJOS VERBALES

Hay por lo menos dos sentidos con los cuales los lingüistas usan el término “complejo verbal”. Unos lo usan para referirse al conjunto formado por un verbo y todos los morfemas ligados a éste, como el templete en (8). Aquí usaremos el término con el sentido que tiene en la lingüística germánica, donde se define el complejo verbal como “una secuencia de verbos que no pueden interrumpirse por elementos no verbales” (de Haan, 2010: 166, traducción propia). En las lenguas germánicas hay ambigüedades sobre si los complejos verbales constituyen palabras complejas o construcciones sintácticas y las mismas ambigüedades están presentes en las lenguas zapotecas. En este capítulo no vamos a profundizar sobre el tema, pero introducimos este concepto para contrastarlo con los compuestos y las expresiones idiomáticas.

Los complejos verbales en lenguas zapotecas inician con un verbo flexionado que pertenece a una de las clases de verbos con funciones gramaticales, incluyendo los verbos posicionales, aspectuales y modales, de movimiento y auxiliares que han perdido su sentido léxico. Esto es una de las principales diferencias entre compuestos y complejos verbales, ya que el núcleo verbal de un compuesto no está restringido a verbos gramaticales. Los verbos gramaticales históricamente han servido como los núcleos de diversas construcciones gramaticales y los complejos verbales son versiones ligadas de estas construcciones.

En el zapoteco de Teotitlán del Valle, Gutiérrez Lorenzo (2014) identifica un tipo de construcción de verbos seriales en la que el primer verbo es un verbo de movimiento, como en (32), y donde ambos verbos comparten la misma marcación aspectual y el mismo sujeto, marcados en cada verbo por separado. También en Teotitlán del Valle, los verbos ‘ir’ y ‘venir’ se pueden prefijar al verbo, como en (33), donde frecuentemente se analizan como prefijos direccionales, “andativo” y “venitivo”. En (33) hay una sola marca de aspecto, que

última raíz en el complejo verbal y está ligado a la forma estativa de ‘sentarse’. En este caso ‘cuidar’ tiene un tono contrastivo mientras que ‘sentarse’ ha perdido su tono por su posición átona dentro de la palabra prosódica.

- (37) San Bartolomé Apóstol *ndób* hasta *nal* *yéh.*
San Bartolomé Apóstol n-dób hasta nàl jǎ
 San Bartolomé Apóstol EST-sentarse hasta ahora PROX
 ‘San Bartolomé Apóstol está (en el pueblo) hasta ahora.’

- (38) *Xná'-n mbí'd ndobnáap mbéhd.*
fná'=ã m-bí'd n-dob+nâp mbæd
 mamá=1SG R-PFV\venir EST-sentarse+INF.cuidar bebé
 ‘Mi mamá vino y estaba cuidando al bebé.’

Los verbos de movimiento también ocurren con frecuencia en combinación con otras raíces verbales como parte de construcciones gramaticales. En los ejemplos de abajo, la forma perfectiva del verbo ‘ir’ tiene tono alto cuando ocurre sin otra raíz ligada (39), pero pierde su tono al ligar otra raíz en la oración de propósito en (40).

- (39) *Chá ngwá xa' mcho' xa' dó' guéxa-n.*
tǎ́ ñ-gw-á fǎ? m-tǎ́ò? fǎ? dó? géf=ã
 inmediatamente R-PFV-ir 3H R.PFV-cortar 3H mecate red=DEM
 ‘Inmediatamente fue y cortó un mecate de la red.’

- (40) *Ngwaxí' tseh'l xi'y díib golp.*
ñ-gw-a+fǐ? tsæ'l fǐ?=j díib gòlp
 R-PFV-ir+M.traer esposo 3H=3INAN otro golpe
 ‘Lo fue a traer el esposo de nuevo.’

No obstante los ejemplos anteriores, esta reducción tonal no siempre ocurre. Hay veces que hemos observado que el primer verbo retiene su tono. El imperativo viene de la misma fuente histórica que el perfectivo y estas formas tienen el mismo tono. En (41) ‘ir’ conserva su tono a pesar de que esté ligado al verbo ‘enterrar’.

- (41) *Wáka'ts gó'y.*
w-á+kà'ts gó?=j
 IMP-ir+INF\enterrar 2PL=3INAN
 ‘Vayan a enterrarlo’

4.2 El estatus sintáctico del complejo verbal

Reconocemos tres factores que pueden ser relevantes para el análisis sintáctico de los complejos verbales: su impenetrabilidad, la ocurrencia de un solo sujeto superficial, y los requisitos morfosintácticos que el núcleo pone en el verbo ligado.

La impenetrabilidad de los complejos verbales, o incapacidad de disociarse, se evidencia con la posición de los enclíticos en los siguientes ejemplos. En (42) la negación se

enclitiza después de ‘cocer’ y no después del auxiliar causativo. En (43) el adverbio ‘ya’ también ocurre después del verbo ‘caminar’ y no después del auxiliar. También en ambos ejemplos el sujeto ocurre una sola vez y viene hasta después de los enclíticos mencionados, demostrando que los complejos verbales son unidades indisolubles que, como otros verbos, tienen un solo sujeto.

(42) *Tegue'ydá* *lú* *ná!*
 te+gè?j=dá lú ná
 POT.CAUS+cocer=NEG 2FAM 1SG
 ‘¡No me ases!’

(43) *Nkedéhlá* *xa'* *mentalmente jodiíd.*
 n-ke¹²+dê=lá *ʃã?* *mentalmente xodĩd*
 EST-AUX+INF.caminar=ya 3H *mentalmente jodido*
 ‘Ya andaba él mentalmente jodido.’

Cuando un verbo funge como complemento de otro verbo, sin formar un complejo verbal con él, el complemento verbal aparece en modo potencial, como en (44). Como en otras lenguas zapotecas, los verbos suelen sufrir cambios tonales en potencial debido a que históricamente se marcaba el potencial con un tono alto flotante, a veces produciendo un tono ascendente en la superficie. El tono “bajo de nivel” en el habla de jóvenes como Cruz Santiago corresponde al tono ascendente en el habla de sus padres y otras personas nacidas en la década de 1970 o antes. Así que el hecho de que el verbo ‘morirse’ aparezca con tono bajo en potencial refleja una antigua alternancia en que aparecía con el tono ascendente.

(44) *Nalna* *na'ká* *jwá'n* *ña* *ndlí* *gathaá-n.*
 nál=na nà?ká xwá?n jà n-d-lí g-àθ=ã
 ahora=entonces tal.vez cosa DIST R-HAB-hacer POT-morir\|POT=1SG
 ‘(Ahora) entonces tal vez eso hace que me muera.’

En el dí'zdéh hay dos formas morfológicas en que puede ocurrir el segundo verbo en un complejo verbal. Una, que glosamos M y que tiene semejanzas con el potencial, se usa en combinación con verbos de movimiento, como en (45). Otra, que glosamos INF, ocurre con otros verbos, como los posicionales (46) y los modales (47).

(45) *Yayadla* *astke* *mbdiib* *bixtiil.*
 ja+jád=la àstkè m-b-dĩb bĩʃtĩl
 POT.ir+M\bañarse=2FAM hasta.que R-PFV-desgastarse jabón
 ‘Y te vas a ir a bañar hasta que se acabe el jabón.’

(46) *Za* *róoltha'* *ndóbyáad* *xa'...*
 Zà rôl=θa? n-dób+jád ʃã?

¹² Esto es el verbo posicional ‘pegarse’ /kê/, pero aquí sin su tono y con un cambio vocálico. Se ha glosado como un auxiliar porque aquí no tiene el sentido de ‘pegarse’. No lo estamos glosando como un prefijo progresivo, como en zapoteco central, porque en esta lengua cualquier verbo posicional puede funcionar así y ‘pegarse’ no es más frecuente que otros verbos posicionales en esta construcción.

- entonces mitad=ADV R.HAB-sentarse+INF\bañar 3H
 ‘Entonces en medio de cuando se estaba bañando...’
 (47) *Mbáay, mbruxyáad xa' bgui'.*
 mbâj m-b-ruj+jâd fã² bgi²
 bueno R-PFV-terminar+INF.bañarse 3H macho
 ‘Bueno, terminó de bañarse el hombre.’

Para ciertos verbos que no tienen alternancias tonales dentro del paradigma, el infinitivo y la forma M pueden ser homófonos, pero para los demás sí hay una diferencia, ilustrada abajo en la Tabla 3. Verbos cuyas raíces son de consonante inicial, como ‘calentarse’, usan la raíz como el infinitivo. Verbos de las clases A-C con raíz de vocal inicial agregan /j-/ para formar el infinitivo. Los verbos de las clases Ch y D tienen raíces de vocal inicial pero la mayoría de las formas paradigmáticas tienen temas de consonante inicial debido al uso de prefijos derivativos que se llaman “reemplacivos” (Kaufman, 2016; Beam de Azcona, 2009), que son la causa de la supletividad parcial que se ve en estas dos clases. En las clases Ch y D, el infinitivo es equivalente al tema del contrafactual, es decir, la forma contrafactual menos el prefijo nasal. El tema contrafactual es equivalente a la raíz más un prefijo derivativo que llamamos “reemplacivo 1”, que es /d-/ en ‘defecar’ y /g-/ en ‘sobar’. En las clases A-C, la forma M es segmentalmente igual al infinitivo, pero si el verbo es uno con alternancias tonales en el paradigma, la forma M tomará el mismo tono que en el potencial. En las clases Ch y D la forma M también tiene el mismo tono que en potencial, pero hay diferencias segmentales adicionales. En ambas clases hay fortificación de la consonante inicial del tema en la forma M, pero la consonante que sufre esta fortificación es el prefijo R1 en la clase Ch pero el prefijo R2 en la clase D. El prefijo R2 es la consonante inicial del tema de perfectivo, que es /f-/ en ‘defecar’ y /d-/ en ‘sobar’. Sobre por qué la clase Ch usa R1 pero la clase D R2, no podemos indagar más en este capítulo, pero la fortificación sí es llamativa porque resulta de un prefijo histórico *k-, que fue el prefijo de potencial, antepuesto a una consonante. En verbos de la clase D como ‘sobar’, el potencial tiene una versión fuerte del prefijo R1 mientras que la forma M tiene una versión fuerte del prefijo R2. Ambas formas tienen las mismas alternancias tonales. Lo que estos patrones revelan es que la forma M resulta por aplicar la morfología de potencial, históricamente *k- [H] (prefijo velar y tono alto flotante), pero a un tema que puede ser diferente al tema de potencial.

	Clase A-C	Clase Ch	Clase D
	C inicial	V inicial	
Raíz	‘calentarse’ /zæn/	‘bañarse’ /âd/	‘defecar’ /ôn/
Potencial	zjæn	gâd	djôn
Contrafactual	nzæn	ngâd	ndôn
Perfectivo	mzjæn	ngôd	mpjôn
Infinitivo	zæn	jâd	dôn
M	zæn	jâd	ðôn

Tabla 3. Formas paradigmáticas de verbos.

El hecho de que un verbo de movimiento requiere que el segundo verbo aparezca en una forma morfológicamente diferente a que si se combinara con otro verbo, es semejante a las maneras en que los verbos pueden subcategorizar sus argumentos, y sugiere un origen sintáctico para los complejos verbales. De estas dos formas, la forma M parece tener una relación histórica con el potencial, algo que quizás se relaciona al uso del potencial en ciertas

estructuras sintácticas, como la complementación en (44) arriba.

Los diferentes requisitos morfológicos que los núcleos verbales imponen en las siguientes raíces verbales sugieren que los complejos verbales tienen un origen sintáctico, pero el hecho de que hoy en día los complejos verbales tengan un solo sujeto y que no permitan que ningún clítico intervenga entre las dos raíces verbales, indican que un complejo verbal cuenta como una sola palabra desde la perspectiva de la sintaxis sincrónica.

4.3 La semántica del complejo verbal

Consideramos que los complejos verbales son construcciones sintácticas que se están convirtiendo lentamente en morfología flexiva, una ruta de gramaticalización común en la familia zapotecana (Smith Stark, 2003; Operstein, 2014; Broadwell, 2015). A diferencia de la composición, que deriva nuevos ítems léxicos, los complejos verbales son productivos como la flexión y transparentes como la sintaxis. Si es concebible que uno se traslade de un lugar a otro con el propósito de hacer la acción indicada por un verbo, es posible combinar ese verbo con un verbo de movimiento. Si es posible hacer una acción continuamente durante un tiempo, se puede combinar el verbo indicado con un verbo posicional para formar la construcción progresiva. Los complejos verbales permiten una combinación más amplia de raíces verbales que los compuestos, y devienen significados más transparentes. Un aprendiz de la lengua tendría que aprender uno por uno cuáles compuestos existen en la lengua, pero podría usar todos los verbos ya en su repertorio para formar complejos verbales.

Así, determinamos que los complejos verbales constituyen palabras prosódicas y sintácticas, y por ende se deben escribir ortográficamente sin un espacio entre los verbos; sin embargo, los complejos verbales no son ítems léxicos y no es necesario incluirlos en un diccionario.

5. EXPRESIONES IDIOMÁTICAS

Fillmore *et al.* (1988) definen las expresiones idiomáticas como las cosas de la lengua que, por ejemplo, un extranjero, no podría predecir aun sabiendo todo el vocabulario y toda la gramática. Las expresiones idiomáticas se tienen que aprender igual que uno aprende vocabulario, pero no consisten en una sola palabra prosódica, sino son combinaciones sintácticas que a veces obedecen las reglas normales de la sintaxis en la lengua y a veces no. Tales expresiones nos interesan por sus semejanzas y diferencias a los compuestos que vimos arriba. Ambos fenómenos son signos en el sentido de Saussure pues tienen un significado y un significante que consiste no solamente en una combinación especial de elementos fonológicos, sino también en una combinación especial de raíces. La diferencia es que en los compuestos el significante forma una sola unidad sintáctica mientras que en las expresiones idiomáticas el significante consiste en más de una unidad sintáctica (o palabra) que pueden o no formar un constituyente más grande. En (48) la combinación del verbo ‘unir’ con ‘mano’ como objeto y, quizás con el requisito de que ‘mano’ sea poseído por una frase nominal correferente con el sujeto, da el sentido de ‘suplicar’.

- (48) *Tséh'l mén yá' mén.*
[tséʔl]_V [mén]_{SUJ} [jáʔ mén]_{OBJ}
unir gente mano gente
‘La gente ruega’ o ‘La gente suplica’.
[Lit. ‘la gente une sus manos.’]

5.1 Expresiones idiomáticas codificantes versus decodificantes

Una de varias distinciones que hacen Fillmore *et al.* (1988) es la que hay entre expresiones idiomáticas “codificantes” versus “decodificantes” (*encoding idioms* versus *decoding idioms*). Todas las expresiones idiomáticas son codificantes porque se debe conocer la expresión para poder producirla (o “codificarla”) correctamente. No obstante, no todas las expresiones son decodificantes, pues algunas son fáciles de comprender cuando son escuchadas por primera vez, pero otras no. La expresión *con todo y _____* en español sería una expresión idiomática codificante pero no decodificante, pues aunque es posible vislumbrar someramente el sentido al oír la por primera vez, no sería posible saber que así se forma la expresión si no se ha escuchado antes. El conocer todas las palabras individuales del español, así como todas las reglas gramaticales, no ayudaría a predecir la existencia de esta expresión, pues no sería utilizada de no saber que existe. Otras expresiones idiomáticas son igualmente difíciles de predecir para producir/codificarlas, pero también son difíciles de interpretar o “decodificar”, por ejemplo: *qué poca madre, me vale, está cañón, y a todo dar*. Es aceptado en general que las expresiones que son difíciles de decodificar se consideren “expresiones idiomáticas” pero las que solamente son difíciles de producir y no de interpretar son excluidas de la categoría por algunos lingüistas. Svensén (2009: 158) las considera como “colocaciones”. Cualesquiera que sean las etiquetas, tanto las expresiones idiomáticas codificantes (colocaciones) como las decodificantes son importantes para la tarea lexicográfica porque los diccionarios deben ser herramientas útiles tanto para la producción como la interpretación (*ibidem*: 183-204).

En zapoteco podemos fácilmente encontrar estos dos tipos de expresiones idiomáticas. En (49) vemos que para expresar la noción de ‘castrar’ se dice que uno saca los testículos del animal que esté castrando. No hay desafío en la decodificación, sino en la codificación. Al oír esta expresión por primera vez el sentido es obvio, pero para referirse a la castración por primera vez sin haber oído la expresión, uno no sabría si decir que uno ‘saca’ los testículos o que uno los ‘corta’ o los ‘quita’. Entonces (49) es solamente una expresión idiomática codificante pero las expresiones en (50-51), además de ser difíciles de codificar, también serían difíciles de decodificar.

(49) *Mbo'* *xa'* *ngíudkwé* *guch.*
 m-boʔ ʃãʔ ŋgûd+kʷé gûʃʃ
 R.PFV-sacar 3H bola+genital cerdo
 ‘Castró al cerdo.’

(50) *¿Xomoód* *mblílaa* *mblóláa* *la'z* *xa'?*
 ʃomõd m-b-lí=lâ m-b-ló=lâ làʔz ʃãʔ
 cómo R-PFV-hacer=2FAM R-PFV-meter=2FAM hígado 3H
 ‘¿Cómo lo pudiste convencer?’

(51) *M-tséh'* *xa'* *ró* *wéets* *xa'*
 m-tséʔ ʃãʔ ró wêts ʃãʔ
 R.PFV-llenar 3H boca hermano 3H
 ‘Remedó a su hermano.’

Rhodes (2002, p. 113) señala que lo que lingüistas externos clasifican como expresiones idiomáticas decodificantes pueden ser más transparentes para miembros de la cultura en la que se habla la lengua. En el contexto cultural mesoamericano se entiende que el hígado es el centro de las emociones, y así (50) tiene una lógica cultural. En cierto sentido, la dificultad de decodificar una expresión idiomática siempre se debe a diferencias culturales, incluyendo entre un hablante y sus propios antepasados. Lo mismo pasa de manera sincrónica cuando una nueva expresión idiomática emerge en un grupo de personas con chistes internos o una perspectiva particular diferente a la del resto de la sociedad. Cuando Beyoncé canta de *Becky with the good hair*, las implicaciones son obvias para mujeres africano-americanas, que experimentan la presión social de aplicar tratamientos químicos para hacer que su cabello tenga una textura más semejante a la de mujeres blancas, pero para muchos blancos *Becky* es un nombre no marcado y desconocen el racismo que subyace la noción de *good hair* para mujeres negras. Esta nueva expresión idiomática en inglés es quizás solamente codificante para norteamericanos negros, pero también decodificante para angloparlantes con poco conocimiento de la cultura africano-americana.

5.2 Elementos léxicos especificados

Fillmore *et al.* (1988) también hablan de un continuo entre expresiones idiomáticas que son *substantive* (o con especificación léxica completa) frente a formales (o léxicamente abiertas). Este es un continuo porque no solo existen las dos opciones de: 1) tener una expresión idiomática solo con palabras fijas y 2) una construcción especial en la cual todas las palabras se pueden cambiar, sino que hay expresiones idiomáticas en que unas palabras son fijas y otras flexibles. En español la expresión *con todo y X* debe tener las primeras tres palabras y el último elemento puede ser cualquier frase nominal: *con todo y [zapato]*, *con todo y [mis berrinches]*, *con todo y [eso]*, *con todo y [frijoles refritos]*, *con todo y [el duelo]*, *con todo y [una persona adentro]*. En (50) la expresión idiomática depende de que el verbo sea ‘meter’ y el objeto ‘hígado’, aunque el sujeto y el poseedor de ‘hígado’ son variables. Vamos a referirnos a los dos puntos extremos en el continuo como expresiones “léxicamente abiertas”, por un lado, y expresiones “con especificación léxica plena”, por el otro; mientras que a los puntos intermedios (del tipo *con todo y X*) como expresiones “parcialmente especificadas”.

Desde la perspectiva de Fillmore *et al.*, todas las construcciones especiales en la lengua son expresiones idiomáticas. Nosotros no adoptamos una postura tan radical. Si hay una construcción léxicamente abierta, no la consideramos como una expresión idiomática. Pero consideramos que las construcciones léxicamente especificadas, aun de manera parcial, sí son expresiones idiomáticas. Nuestra principal preocupación es determinar cómo representar las expresiones idiomáticas en el diccionario del dí'zdeh. Para nosotros, si hay una combinación de al menos dos palabras fijas (léxicamente especificadas) que, solas o en combinación con otras palabras léxicamente abiertas, tienen un significado particular, la consideramos como una expresión idiomática. Esto es diferente para los complejos verbales, donde el segundo verbo es léxicamente abierto y el primer verbo se restringe a un número limitado de verbos gramaticales, pero no se limita a un verbo particular, ni produce un significado léxico particular. Los compuestos son más semejantes a las expresiones idiomáticas porque hay dos raíces particulares especificadas con un significado léxico, pero forman una palabra sintáctica y no dos. Las construcciones de verbo ligero se pueden considerar como expresiones idiomáticas porque hay dos palabras, el verbo ligero y su objeto,

y es la combinación de estas dos palabras particulares la que crea un nuevo significado. Por ejemplo, en (52) la expresión ‘hacer trabajo’ tiene el sentido de ‘trabajar’.

- (52) *Tson lí'n nak kontraát líí xa' dži'n.*
 tsòn lí'n n-àk konträt líí fã² dži'n
 tres año EST-COP contrato POT\hacer 3H trabajo
 ‘De tres años es el contrato de que va a trabajar.’

Varias expresiones idiomáticas en el dí'zdéh (incluyendo, aunque no limitado a expresiones de verbo ligero) se traducen al español con verbos simples. En zapoteco su representación incluye un verbo conjugado, sin embargo, que tenga el sentido indicado depende de ciertas relaciones gramaticales con ciertos sustantivos. Estos sustantivos pueden servir como sujetos (53) o como objetos (54). Nótese que la posición del enclítico negativo en (53) hace imposible una interpretación de ‘impedir’ + ‘palabra’ como compuesto. Si fuera compuesto, la negación iría al último después de ‘palabra’. Igualmente, en (54), en que el objeto ‘vergüenza’ es una palabra clave para dar el sentido deseado, el sujeto *xa'* interviene entre verbo y objeto, haciendo imposible una interpretación de incorporación.

- (53) *Nkáza' xki'z mbyódudx bá.*
 n-ká=za² fãk-i'z m-b-jó+dùdž bá
 R.HAB-impedir=NEG POS-POS\palabra CL-AN-muchacho+pequeño MED
 ‘Ese niño no tartamudea.’

- (54) *M-bo' xa' yanthon xa' nkwi'n xa' loó mén.*
 m-bo² fã² janθò-n fã² n-kwi'n fã² lõ mén
 R.PFV-sacar 3H vergüenza-POS 3H R.HAB-mentir 3H cara gente
 ‘Se deshonra mintiendo a la gente.’

5.3 Núcleos ligeros versus núcleos pesados

Tanto para compuestos como para expresiones idiomáticas, otra distinción que podríamos hacer es si la contribución semántica se divide entre dos (o más) raíces, o si ciertas raíces son semánticamente ligeras y desempeñan funciones gramaticales más que semánticas. En las expresiones idiomáticas normalmente la combinación de un verbo con un argumento es lo importante, aunque otras palabras lexicalmente abiertas pueden ser requeridas. Por ejemplo, la expresión en (53) debe su sentido de ‘tartamudear’ al verbo ‘impedirse’ en combinación con el sujeto ‘palabra’, aunque esta expresión requiere que ‘palabra’ tenga un poseedor. Relacionando esto con los conceptos presentados en Fillmore *et al.* (1988), podemos decir que hay dos raíces con especificación plena en esta expresión y el poseedor es lexicalmente abierto. La expresión en su totalidad está parcialmente especificada entonces. Tomando en cuenta todos los compuestos formados de dos raíces y todas las expresiones idiomáticas con dos palabras lexicalmente especificadas, podemos dividirlos en un grupo donde las dos raíces o palabras hacen contribuciones semánticas semejantes versus un grupo donde una de las dos, y casi siempre la segunda, hace una mayor contribución.

Las expresiones idiomáticas como las de (48-51) tienen dos palabras, un verbo y un argumento, que son semánticamente pesadas y que juntos construyen el sentido de la expresión. Esto es cierto para casos transparentes de expresiones idiomáticas codificantes,

como ‘sacar testículo’, y para expresiones idiomáticas decodificantes que son más opacas como ‘llenar la boca de alguien’ o ‘meter el hígado de alguien’. Aún en las expresiones idiomáticas decodificantes, ya conociendo la expresión, en general se puede observar que hay algún sentido metafórico, y en casos como (50) y (51), el verbo y el argumento fijo contribuyen a la metáfora.

Podemos contrastar los compuestos y las expresiones idiomáticas que tienen dos raíces o palabras fijas de semántica pesada con compuestos y expresiones semejantes que tienen uno de los dos elementos fijos como semánticamente ligero y el otro como semánticamente pesado. Las construcciones de verbo ligero son un ejemplo obvio. Cuando en español decimos ‘robar’ y ‘trabajar’, en el dízdeh se dice ‘hacer robo’ y ‘hacer trabajo’. Verbos ligeros como ‘hacer’ y ‘hacerse’ sirven para recibir la flexión (los prefijos aspectomodales); para que los enclíticos tengan a qué ligarse y para indicar las relaciones gramaticales de la oración, por ejemplo, si es una oración transitiva o intransitiva, aunque no nos dicen mucho sobre qué pasó. Sustantivos como ‘robo’ y ‘trabajo’ informan lo que el verbo no dice. También en complejos verbales con dos raíces verbales, solo la primera raíz va flexionada y muchas veces ésta es un auxiliar u otro verbo de semántica ligera. Igualmente, en los diferentes tipos de composición que vimos, podemos encontrar casos donde el núcleo verbal tiene muy poca semántica, incluso en ocasiones se ha vuelto opaca por la reducción fonológica o por la pérdida de la raíz como verbo independiente en la lengua. En algunos compuestos la mayoría de la información la aporta el nominal incorporado.

En este trabajo solo se están considerando verbos o estructuras cuyos núcleos son verbos. Sin embargo, parece que en compuestos puramente nominales pasa algo semejante. Existen muchos compuestos nominales donde el primer sustantivo posee una semántica muy genérica e incluso se podría considerar como un término de clase, mientras que la segunda raíz nominal tiene una semántica muy específica (véase Beam de Azcona, 2012, sobre este fenómeno en los topónimos). Esto pasa frecuentemente con diferentes tipos de animales, plantas y personas: se habla del árbol-pinocote, de la fruta-aguacate, del humano-maestro. Se sabe que un maestro es un humano de igual manera que ya se sabe que trabajar es *hacer* algo.

Las lenguas zapotecas son de núcleo inicial y también son lenguas de marcación en el núcleo. El hecho de que éste reciba la marcación no dicta que el núcleo tiene que perder su semántica, pero sí puede prevenir que se pierda totalmente porque aun cuando una raíz no contribuye mucho a la semántica, lleva toda o la mayoría de la marcación gramatical. Esto le da una razón de ser. En los compuestos y en las expresiones idiomáticas, si el segundo elemento no tuviera mucho que ofrecer para construir el sentido, ni ofreciera nada en cuanto a funciones gramaticales, fácilmente se podría perder. Los compuestos se podrían reducir a palabras de una sola raíz, expresiones transitivas se podrían reducir a expresiones intransitivas, etcétera. No observamos casos de morfemas ligeros al final en estas lenguas porque la marcación va al inicio, y la marcación es lo que mantiene los morfemas ligeros. En esta lengua hay prominencia morfosintáctica al inicio y prominencia fonosemántica al final.

5.4 El estatus de verbos con *la'z*

En compuestos con sustantivos incorporados y también en muchas expresiones idiomáticas en esta y otras lenguas zapotecas, las partes del cuerpo son prominentes. La parte del cuerpo más común en estas expresiones es *la'z* ‘hígado’, el tradicional centro de emociones en Mesoamérica. Hay dos posibilidades para las relaciones gramaticales de los agentes o

experimentantes en estas expresiones: que sean sujetos de verbos compuestos con *la'z* o gramaticalmente poseedores de *la'z*, que a la vez sería el sujeto, como mostramos en (55).

- (55) a. [VERBO+là'z]_{COMPUESTO} [agente/experimentante]_{SUJ}
 b. VERBO [[là'z]_{POSEÍDO} [agente/experimentante]_{POSEEDOR}]_{SUJ}

La estructura de (55a) se ilustra en (56) y la de (55b) en (57). El enclítico de negación, siempre se agrega al final del verbo. Su respectiva posición antes y después de *la'z* demuestra que *la'z* forma parte del verbo en (56) pero no en (57). El enclítico de negación provee una prueba importante para distinguir las dos estructuras, pero si únicamente mostráramos estos ejemplos en afirmativo la estructura parecería ambigua.

- (56) *Nda'la'zza'* *xa' ché'n xa' loó mén.*
 n-da²+là²z=za² ʃà² tʃé²n ʃà² lö mén
 R.HAB-dar+hígado=NEG 3H POS 3H cara/a gente
 ‘No dona sus cosas a la gente.’

- (57) *Ndxálza'* *lázáa-n laár yéh.*
 n-d³ál=za² láz=â lâr jâc
 R-HAB-alcanzar¹³=NEG hígado\1SG=1SG ropa PROX
 ‘No me gusta esta ropa.’

Así, existen compuestos como ‘donar’ que incorporan a *la'z* y también hay expresiones idiomáticas con *la'z* como sujeto fijo de un verbo, creando un significado especial vía metáfora. Los enunciados con *la'z* son un microcosmos que ilustra el problema general de distinguir entre los compuestos y las expresiones idiomáticas. Una misma raíz, como *la'z*, puede ocurrir justo después de una raíz verbal, pero tener diferencias en su estatus sintáctico que también implica diferentes relaciones gramaticales para otros constituyentes en la oración. Por ende, es importante poder distinguir estas diferencias en el diccionario mediante alguna convención lexicográfica. Nuestra principal motivación al iniciar el análisis de estas formas fue poder tomar una decisión sobre cómo representar estas entradas en el diccionario del dí'zdéh.

6. Implicaciones para LAS REPRESENTACIONES ORTOGRÁFICA Y LEXICOGRÁFICA

Considerando que los compuestos y complejos verbales cuentan como palabras prosódicas y sintácticas, los escribimos juntos, sin un espacio entre la raíz verbal y la raíz incorporada. En cambio, los diferentes elementos de las expresiones idiomáticas cuentan como palabras independientes y por ende no se deben escribir juntos, sino separados por espacios. Escribir o no un espacio entre los diferentes morfemas que se combinan de las maneras descritas en este capítulo es una decisión ortográfica importante no solamente para el diccionario y la colección de textos de esta variante sino para todos los materiales escritos que se producen en esta lengua.

¹³ En un borrador previo teníamos este verbo glosado como ‘gustar’ pero no ocurre con este sentido cuando no aparece *la'z* en la oración. El verbo ‘alcanzar’ tiene la misma forma segmental y los mismos patrones tonales por todo el paradigma.

Otra decisión importante radica en qué formas incluir en el diccionario y cómo organizar esta información. En este capítulo determinamos que los compuestos verbales y las expresiones idiomáticas son ítems léxicos, pero los complejos verbales no lo son. Solamente los ítems léxicos deben incluirse en el diccionario. Es una práctica estándar incluir los compuestos como subentradas (Svensén, 2009: 99). Nosotros incluimos tanto los compuestos como las expresiones idiomáticas como subentradas del núcleo verbal si este es reconocible, o como una entrada independiente si no. En (58) se enlistan algunas entradas ilustrativas. En la entrada para *du* ‘esparcir’ se incluyen subentradas para el compuesto *duxníi* ‘alumbrar’ y la expresión idiomática *du xníi* ‘amanecer’. Este compuesto y esta expresión idiomática también se alistan como entradas independientes, pero en esas entradas solamente se incluye una referencia cruzada a la entrada principal para *du*. El compuesto tiene una posible pronunciación con otra vocal. Esa pronunciación, *doxníi*, también se enlista por separado con una referencia cruzada a *duxníi* y en la subentrada para *doxníi* se reconoce también la pronunciación variante. Bajo la entrada para el sustantivo *xníi* se encuentra una referencia cruzada al compuesto *duxníi* y la expresión idiomática *du xníi*, avisando al usuario que deba buscarlos bajo *du*.

(58) **doxníi** (var. **duxníi**)

du viA esparcir *Ndu bxíl guíts kwet mazií ndxa'p kwet*. Se sueltan chispas del cohete cuando sube. *Mdu xa' loó yú mble'k xa' gyé' ró bá'n xúud xa'*. Esparció de flores la tierra junto a la tumba de su padre.

duxníi comp. (var. **doxníi**) vtA alumbrar *Neéd dyeéhtáá lú bdoxníi*. Alumbró solamente por donde vas a caminar.

du xníi exp. *idiom.* amanecer *Goché dáak mduláá xníi*. Levántate porque ya amaneció. *Ndehdu xníi mdzín xa' noó ngen cho nhwnéy*. De madrugada (casi al amanecer) llegó y nadie lo supo.

du xníi (**du** esparcir)

duxníi (**du** esparcir)

xníi s luz *Dyenza' xká'láa-n tal nzóo xníi le'n yó*. No me agarra el sueño si hay luz en la casa. comp. **duxníi** (véase **du**), exp. *idiom.* **du xníi** (véase **du**)

En el caso de las expresiones idiomáticas parcialmente especificadas que tienen un sujeto preespecificado, las enlistamos con el verbo y el sujeto separados por un espacio, como en *du xníi*, pero cuando hay un elemento fijo que se encuentra más lejos del verbo, como un objeto directo, usamos tres puntos. Vemos que esta misma convención ha sido usada en el diccionario del zapoteco de la Sierra Juárez (Nellis y Goodner, 1983). En (59) mostramos las entradas relevantes para dos expresiones idiomáticas que aparecen como subentradas del verbo *kaá* ‘impedir’. Este verbo es ambitransitivo y puede tener o no un objeto directo. Cuando la forma inalienablemente poseída de ‘palabra’, *xki'z*, es el sujeto, tiene el significado de ‘tartamudear’, pero cuando la forma no poseída, *dí'z*, funge como objeto directo entonces significa ‘prohibir’. Las entradas para ‘palabra’ y para las expresiones idiomáticas tienen referencias cruzadas en la entrada para ‘impedir’, donde uno encuentra la información completa para las expresiones idiomáticas como subentradas.

(59) **dí'z** s palabra, lengua, idioma *Bé'y xomoód mblí xa' mte'd xa' na'r dí'z*. Quién sabe cómo le hizo para aprender varios idiomas. *Dí'z láá nakáá*. Ya es un trato. **kaá...dí'z** exp. *idiom.* (véase **kaá**), der. **xki'z**

- dí'zdéh** *comp.* zapoteco, cualquier lengua indígena *Ntakyónza' xa' dí'zdéh mbéez xa'*. Dice no entender el zapoteco.
- dí'z xtiíl** (*var.* **dixtiíl**) *comp.* español, castellano *Ndrobe'yza' ndidi'z xná'-n dixtiíl bweén-káá*. Mi mamá no puede hablar español fluidamente.
- kaá** vA impedir *Nká xa' neéd páar dyeéh mén*. Él impide el paso para caminar.
- kaá xki'z** *exp. idiom.* tartamudear *Nká xki'z mbyódudx bá no nakza' xíud xa' xna' xa' nká ró na'*. Ese niño tartamudea y sus padres no sufren de eso.
- kaá ... dí'z** *exp. idiom.* Prohibir. *Nká xa' dí'z ró' xi'n xa'*. No permite que salga su hijo/a.
- kaá xki'z** (**kaá** impedir)
- kaá ... dí'z** (**kaá** impedir)
- xki'z** s voz, palabra (posesivo) *Peén-káá ndxón xa' xki'z xna' nhwdi' nzóo xa' le'n guéht*. Apenas podía escuchar la voz de la mamá del rayo porque estaba metido dentro de la olla. *exp. idiom.* **kaá xki'z** (véase **kaá**)
- xki'zgrính** *comp.* inglés

La comparación de cómo alistamos el compuesto *duxnúi* sin espacio, las expresiones idiomáticas con un sujeto fijo, *du xnú* y *kaá xki'z*, con un espacio y la expresión idiomática con un elemento fijo más lejano, *kaá...dí'z*, con tres puntos, ilustran nuestras estrategias lexicográficas para representar las diferentes estructuras sintácticas que encontramos en estos ítems léxicos. La posibilidad de tomar estas decisiones prácticas, basándonos en criterios explícitos del análisis lingüístico, fue nuestra principal motivación para desarrollar este trabajo.

7. LA MORFOLOGÍA DE HOY (O MAÑANA) ES LA SINTAXIS DE AYER

En lenguas zapotecanas, tanto la derivación como la flexión tienen orígenes en la sintaxis. Las expresiones idiomáticas y las frases lexicalmente abiertas indiscutiblemente tienen estructura sintáctica. La diferencia es que las expresiones idiomáticas están por lo menos parcialmente especificadas, con ciertas palabras fijas y un sentido léxico particular, mientras que otras oraciones se forman libremente con una selección abierta de palabras y así producen oraciones cuyos sentidos son transparentemente la suma de sus partes. La morfología derivativa, igual que las expresiones idiomáticas, es idiosincrática y crea ítems léxicos particulares, mientras que la morfología flexiva es productiva y transparente y no crea nuevos ítems léxicos sino distintas formas paradigmáticas de la misma palabra. Las palabras complejas que hemos examinado aquí están a medio camino de dos distintas rutas de gramaticalización. Los compuestos son un punto intermedio entre las expresiones idiomáticas y la morfología derivativa. Los complejos verbales, por su parte, marcan un punto medio entre la sintaxis y la flexión.

En (1) y (2) arriba vimos el verbo 'meter' en combinación con 'aire'. En esos ejemplos observamos el compuesto con el significado de 'inflar', pero también la expresión idiomática con 'aire' como objeto directo que significa 'abanicar'. En realidad, la expresión idiomática solamente tiene el sentido de 'abanicar' si la expresión locativa contiene el sustantivo relacional *loó*, como en (60). En cambio, si es simplemente un sustantivo sencillo, también puede querer decir 'inflar' (61), igual que el compuesto (62).

(60) *Ko mén mbú loó mén.*

- | | | | |
|--------------------------------------|----------------------|------------------------|-------------------------|
| k-ò | [mén] _{SUJ} | [m-b-î] _{OBJ} | [lõ mén] _{LOC} |
| POT\R1-meter | gente | CL-AN-aire | cara/a gente |
| ‘Va a abanicarse/lo ¹⁴ .’ | | | |
- (61) *Ko mén mbúi pelot.*
k-ò [mén]_{SUJ} [m-b-î]_{OBJ} [pelòt]_{LOC}
POT\R1-meter gente CL-AN-aire pelota
‘La gente va a inflar la pelota.’
- (62) *Mblómbúi Cheén pelot.*
mb-l-ó+m-b-î [tʃɛn]_{SUJ} [pelòt]_{OBJ}
PFV-R2-meter+CL-AN-aire Chenó pelota
‘Arsenio infló la pelota.’

Al comparar (61) con (62), aunque tienen semejantes sentidos, (61) da más prominencia al aire y (62) a la pelota por ser el objeto directo un argumento central. La diferencia en la estructura de la información podría ser una razón para que existan dos formas con un significado tan similar. No obstante, también hay otra posible motivación para el compuesto en (62). Los ejemplos (60) y (61) tienen sentidos diferentes, pero la única diferencia sintáctica es la inclusión de un sustantivo relacional en (60). Las dos expresiones son tan semejantes que distinguir entre los dos sentidos sería más fácil si la estructura fuera aún más diferente, como lo es si comparamos (60) y (62). La motivación para crear el compuesto puede ser resaltar la pelota como objeto directo, o crear una distinción léxica entre ‘abanicar’ e ‘inflar’, como hay en castellano, la segunda lengua de la mayoría de los hablantes del dí’zdeh. De cualquier forma, concluimos que el sentido de ‘inflar’ originalmente se indicó con la expresión idiomática y que el compuesto es más reciente. Otras lenguas del mundo muestran semejantes patrones. Por ejemplo, Rich Rhodes habla de las propiedades “idiomáticas” de los compuestos VERBO+VERBO en Ojibwe (Rhodes, 2002: 110)

Cuando un sustantivo como *mbúi*, que fue un argumento fijo en una expresión idiomática, se incorpora en el verbo, crea una nueva palabra prosódica y sintáctica, pero, en cierto sentido, no crea un nuevo ítem léxico, ya que la combinación de ‘meter’ y ‘aire’ en la expresión idiomática ya era un ítem léxico con el significado de ‘inflar’. De esta forma, podríamos caracterizar ‘inflar’ como un ítem léxico que experimentó un cambio fonológico y morfosintáctico. Cuando un sustantivo que aparece en muchas expresiones idiomáticas, como *la’z*, se incorpora repetidas veces, existe la posibilidad de que en un futuro pueda convertirse en un sufijo derivativo. Sin embargo, debido a que la última raíz es la más prominente en lenguas zapotecanas, es más probable que los verbos ligeros, que frecuentemente ocurren en expresiones idiomáticas y que pueden reducirse en los compuestos por estar en una posición átona, se conviertan en prefijos derivativos. El prefijo causativo-agentivo *o del protozapotecano es un caso probable de un verbo auxiliar que se convirtió en un prefijo derivativo (Operstein, 2014; Foreman y Dooley, 2015). También una forma reducida de la cópula se convirtió en un prefijo estativo que deriva verbos y adjetivos (Beam de Azcona, en prensa).

Los verbos reducidos no solamente son una fuente de morfología derivativa sino

¹⁴ Es ambiguo en la oración si ambas ocurrencias de *mén* son correferenciales o si se refieren a diferentes personas.

también de morfología flexiva. Debido a su uso en la construcción progresiva, diferentes verbos posicionales se han gramaticalizado como prefijos de aspecto progresivo en diferentes lenguas zapotecanas (Smith Stark 2003; Broadwell, 2015; Beam de Azcona, bajo dictamen). En el dí'zdéh de San Bartolomé Loxicha, el aspecto progresivo no se indica con un prefijo particular sino con un complejo verbal cuyo núcleo puede ser cualquier verbo posicional. En otras lenguas zapotecas los verbos 'ir' y 'venir', o quizás versiones reducidas de ellos, se prefijan a otros verbos, donde se han analizado como "marcadores direccionales" (Gutiérrez Lorenzo, 2014: 144) o como aspectos "andativo y venitivo" o de "alejamiento y acercamiento" (López y Newberg, 1990). Las formas equivalentes en el dí'zdéh aquí se han analizado como complejos verbales cuyos núcleos son verbos de movimiento, incluyendo, pero no limitado a, los verbos 'ir' y 'venir'. Otros complejos verbales tienen núcleos como 'terminar' que pueden estar en el proceso de gramaticalizarse como nuevos prefijos aspectomodales.

Quizás el ejemplo mejor documentado de este tipo de gramaticalización en zapoteco es el prefijo de aspecto progresivo en el zapoteco central. Smith Stark (2003) fue el primero en proponer que el prefijo progresivo en zapoteco central se deriva del verbo posicional 'pegarse' que forma parte de lo que él llamó la "construcción modal", que es equivalente a la construcción progresiva con complejo verbal en el dí'zdéh. Smith Stark propuso, basándose en evidencia de la Sierra Norte, que esta construcción desarrolló históricamente de una construcción de verbos seriales, donde cada verbo tenía un sujeto explícito y donde los dos verbos no formaron una secuencia indisoluble.

Yatzachi, en (63), preserva la construcción de verbos seriales y vemos que tanto el verbo posicional como el verbo 'leer' tienen sujetos explícitos y que 'persona' interviene entre ambos verbos. En (64) vemos que el dí'zdéh solamente tiene un sujeto explícito y no permite ni que el sujeto ni que el enclítico adverbial intervengan entre los dos verbos. El verbo *do* aquí se glosa como un verbo auxiliar porque ha perdido su sentido léxico en este contexto, pero es la raíz del verbo posicional 'pararse'. En el dí'zdéh se pueden sustituir otros verbos posicionales en esta posición, p.ej. véase (46) arriba. Tanto Yatzachi como el zapoteco del Istmo descienden de una lengua zapoteca que eliminó el prefijo estativo *n- antes de consonantes (Beam de Azcona, en prensa) y por esa razón ni *chi'* ni *ca-* ocurren con una nasal inicial, aunque históricamente estas formas sí fueron estativas. El verbo posicional 'pegarse' se ha reanalizado como un prefijo de aspecto progresivo en (65).

- (63) Construcción de verbo serial en Yatzachi el Bajo (reinterpretado desde Smith Stark, 2003 y Butler, 1980: 316)

<i>Bachi'</i>	<i>ben'</i>	<i>chlabe'</i>
ba=tʃiʔ	benʔ	tʃ-lab=eʔ
ya=EST.sentarse persona		IPFV-leer=3HR
'El hombre está sentado leyendo.'		

- (64) Complejo verbal en San Bartolomé Loxicha

<i>Ndozéhnlá</i>	<i>xa'</i>
n-do+zán=lá	ʃäʔ
EST-AUX+INF\calentarse=ya	3H
'Ya se estaba calentando.'	

- (65) Prefijo progresivo en zapoteco del Istmo (Pickett *et al.*, 1998: 53)

caguite *ca* *ba'du' que*
 ka-gite ka ba'du' ke
 PROG-jugar PL niño aquel
 'Aquellos niños estaban jugando.'

El proceso de gramaticalización es más avanzado en dí'zdéh que en Yatzachi, porque ya no hay dos sujetos explícitos y la secuencia de los dos verbos se ha vuelto indisoluble, pero nos parece que en esta variante del zapoteco miahuateco la gramaticalización no está tan avanzada como en el zapoteco central. La variante de San Agustín Mixtepec sí tiene el verbo 'pegarse' como un prefijo progresivo, debido a un mayor contacto con el cisyautepequeño, una lengua perteneciente al zapoteco central, pero en San Bartolomé Loxicha todavía es posible usar cualquier verbo posicional en esta construcción. El estatus verbal de las raíces posicionales también es más obvio en el zapoteco miahuateco que en el zapoteco central, debido a la retención del prefijo estativo *n-*. Su pérdida antes de consonantes da la impresión de que /ka-/ es monomorfémico en zapoteco central, aunque históricamente era bimorfémico. Los complejos verbales en el dí'zdéh están más gramaticalizados que construcciones puramente sintácticas, pero menos gramaticalizados que los afijos aspectomodales que existen en otras lenguas.

Así, las tres estructuras que examinamos en este capítulo se ubican en diferentes puntos del desarrollo diacrónico, como resumimos en la Tabla 4.

	Múltiples palabras >	Palabra prosódica y sintáctica >	Afijo
Léxico	Expresión idiomática	Compuesto verbal	Derivación
No léxico	Construcción gramatical	Complejo verbal	Flexión

Tabla 4. Dos caminos de gramaticalización.

8. CONCLUSIÓN

En este capítulo hemos examinado los compuestos verbales, complejos verbales y expresiones idiomáticas. Tres combinaciones complejas que cumplen o no con diferentes nociones de "palabra", una pregunta central para la disciplina lingüística. Determinamos que los compuestos y complejos verbales, pero no las expresiones idiomáticas, constituyen palabras prosódicas y sintácticas. Aunque algunas lenguas del mundo pueden tener palabras sintácticas que consisten en más de una palabra fonológica, otras, como el latín, el griego y el dí'zdéh tienen a la palabra fonológica como la misma unidad que la palabra sintáctica, "es decir, incluye el tema, todo afijo y ambos miembros de compuestos" (Nespor y Vogel, 2007: 110, traducción propia). También determinamos que los compuestos y las expresiones idiomáticas, pero no los complejos verbales, son ítems léxicos que se deben incluir como entradas en un diccionario. De las tres estructuras, solamente los compuestos son palabras prosódicas y sintácticas, y además ítems léxicos. Aunque son más complejos que otras palabras, los compuestos son palabras más prototípicas porque tienen estas tres cualidades, mientras que los complejos verbales y las expresiones idiomáticas son palabras fonosintácticas o léxicas, pero no ambas.

Desde una perspectiva práctica, concluimos que el análisis de una forma como una palabra prosódica y sintáctica es lo que determina dónde hay que escribir espacios ortográficos, mientras que el análisis como un ítem léxico es lo que justifica crear una entrada lexicográfica. Para el análisis sincrónico que guía nuestras decisiones ortográficas y

lexicográficas, los criterios más útiles incluyen la reducción fonética y fonológica de sílabas átonas, la (im)penetrabilidad de secuencias de raíces, y si algún significado léxico se indica con dos o más raíces especificadas).

Aunque el capítulo se realizó en su mayoría dentro del análisis sincrónico, señala ciertos desarrollos diacrónicos. Proponemos que los compuestos y complejos verbales representan etapas intermedias entre la sintaxis y los afijos, los compuestos como parte de la derivación y los complejos verbales precursores de la flexión. Los desafíos de análisis que presentan los complejos verbales y los compuestos en general en lenguas zapotecas se relacionan al hecho de que están en esta etapa intermedia entre la sintaxis que fueron ayer y la morfología que serán mañana. Esta transición es lo que dificulta el análisis sincrónico de hoy, pero también lo hace más interesante.

BIBLIOGRAFÍA

- BEAM DE AZCONA, ROSEMARY G. (2004). "Introducing San Agustín Mixtepec Zapotec", en *Proceedings of the Workshop on American Indigenous Languages*. Santa Barbara: Department of Linguistics, UCSB.
- _____. (2008). "Un estudio comparativo de los tonos del zapoteco sureño", en *Memorias del Coloquio Francisco Belmar: Las lenguas otomangués y oaxaqueñas ante el siglo XXI*. A. de Avila, A. López Cruz, V. Marcial, M. Swanton, eds. Oaxaca: Fondo Editorial IEEPO / INALI / Universidad Autónoma 'Benito Juárez' de Oaxaca / Fundación Alfredo Harp Helú; 161-188.
- _____. (2009). "Clasificación de verbos en el zapoteco miahuateco de San Bartolomé Loxicha". *Cuadernos del Sur*, vol. 14, núm. 28, pp. 59-69.
- _____. (2012). "Southern Zapotec toponyms", en Brook Lillehaugen y Aaron Sonnenschein, eds. *The Expression of Location in Zapotecan Languages*. LINCOM Europa; 261-288.
- _____. (2017). "Spanish infinitives borrowed into Zapotec light verb constructions", en Karen Dakin, Claudia Parodi y Natalie Operstein, eds. *Language Contact and Change in Mesoamerica and Beyond*. Amsterdam: John Benjamins; 55-80.
- _____. (en prensa). "The historical dialectology of stative marking in Zapotecan". *Journal of Historical Linguistics*.
- _____. (bajo dictamen). "Realis morphology and Chatino's role in the diversification of Zapotec languages". *Diachronica*.
- BROADWELL, GEORGE AARON (2015). "The historical development of the progressive aspect in Central Zapotec". *International Journal of American Linguistics*, Vol. 81, núm. 2, pp. 151-185.
- BUTLER H., INEZ M. (1980). *Gramática zapoteca: zapoteco de Yatzachi el Bajo*. México, D.F.: Instituto Lingüístico de Verano.
- CRUZ SANTIAGO, THABGOL EMILIANO (en preparación). *Kwentgox jwá'n nda' thabgol noó mén gox Guéz Xíil: Cuentos, historias y leyendas en el zapoteco de San Bartolomé Loxicha*. Rosemary G. Beam de Azcona, ed.
- _____. y ROSEMARY G. BEAM DE AZCONA (en preparación). *Diccionario del dí'zdéh, el zapoteco miahuateco de San Bartolomé Loxicha*.
- FILLMORE, CHARLES J., PAUL KAY y MARY CATHERINE O'CONNOR (1988). "Regularity and Idiomaticity in Grammatical Constructions: The Case of Let Alone". *Language* 64-3; 501-538.
- FOREMAN, JOHN y SHEILA DOOLEY (2015). "Causative morphology in Macuilianguis

- Zapotec”, en Natalie Operstein y Aaron Sonnenschein, eds. *Valence changes in Zapotec: synchrony, diachrony, typology*. Amsterdam: John Benjamins; 237-280.
- GUTIÉRREZ LORENZO, AMBROCIO (2014). *Construcciones de verbos seriales en el zapoteco de Teotitlán del Valle*. Tesis de maestría. CIESAS.
- DE HAAN, GERMEN J. (2010). *Studies in West Frisian grammar: papers by Germen J. de Haan*. Jarich Hoekstra, Willem Visser y Goffe Jensma, eds. Amsterdam: John Benjamins.
- HERNÁNDEZ LUNA, MARIO ULÍSES (2021). *Fonología del miahuateco: sincronía, diacronía y clasificación*. Tesis de doctorado, Colegio de México.
- _____. Este volumen. “El miahuateco dentro del área mesoamericana”.
- HYMAN, LARRY (2006). “Word-prosodic typology”. *Phonology* 23-2; 225-257.
- KAUFMAN, TERRENCE (2016). “Proto-Zapotec(an) Reconstructions”, ms.
- LÓPEZ L., FILEMÓN Y RONALDO NEWBERG Y. (1990). *La conjugación del verbo zapoteco: zapoteco de Yalálag*. México, D.F.: Instituto Lingüístico de Verano.
- LÓPEZ NICOLÁS (2016). *Estudios de la fonología y gramática del zapoteco de Zochina*. Tesis de doctorado, Centro de Investigaciones y Estudios Superiores en la Antropología.
- MITHUN, MARIANNE (1984). “The Evolution of Noun Incorporation”. *Language* 60-4; 847-894.
- NELLIS, NEIL Y JANE GOODNER DE NELLIS (1983). *Diccionario zapoteco de Juárez: zapoteco-español, español-zapoteco. Ca titsa' qui' ri'u*. México, D.F.: Instituto Lingüístico de Verano.
- NESPOR, MARINA E IRENE VOGEL (2007). *Prosodic phonology*. Berlin: Mouton de Gruyter.
- OPERSTEIN, NATALIE (2014). “The origin of the Zapotec causative marker *k-: a diachronic-typological perspective”. *International Journal of American Linguistics*, Vol. 80, pp. 99-114.
- _____. (2015). “Valence-altering operations in Zapotec”, en Natalie Operstein y Aaron Sonnenschein, eds. *Valence changes in Zapotec: synchrony, diachrony, typology*. Amsterdam: John Benjamins; 23-54.
- PICKETT, VELMA B., CHERYL BLACK y VICENTE MARCIAL CERQUEDA (1998). *Gramática Popular del Istmo*. Centro de Investigación y Desarrollo Binniza e Instituto Lingüístico de Verano.
- RHODES, RICHARD A. (2002). “Multiple assertions, grammatical constructions, lexical pragmatics, and the *Eastern Ojibwa-Chippewa-Ottawa Dictionary*.” En William Frawley, Kenneth C. Hill y Pamela Munro, eds. *Making dictionaries: preserving indigenous languages of the Americas*. Berkeley: University of California Press.
- SMITH STARK, THOMAS C. (2004). “El progresivo en el zapoteco”. Trabajo presentado en el VII Congreso Nacional de Lingüística, Universidad de Guadalajara.
- SVENSÉN, BO (2009). *A handbook of lexicography: the theory and practice of dictionary-making*. Cambridge: Cambridge University Press.