

Selección referencial en textos escritos en hñöhñö (otomí de Tolimán)

Néstor Hernández-Green

Centro de Investigaciones y Estudios Superiores
en Antropología Social CDMX,
México

<https://orcid.org/0000-0001-8553-8245>

ABSTRACT: At a cross-linguistic level, referential selection (that is, the decision to use one or another linguistic form to mention a referent in the discourse) is correlated, among other factors, with referential distance (that is, the number of clauses between a mention of a referent and the prior mention of said referent; see Givón 1983: 13). This paper analyzes and describes the correlation between referential distance and referential strategies (that is, the forms selected to mention a referent) within a *Hñöhñö* written corpus, by means of two statistical methods: Kruskal-Wallis's test (Kruskal & Wallis 1952) and correspondence analysis (Benzécri 1973). The procedures used are based on the methods used in Kibrik (2011).

KEYWORDS: Otomanguean; Otopamean; Discourse; Referential distance; Morphosyntax

RESUMEN: A nivel tipológico, la selección referencial (esto es, la decisión de utilizar una u otra forma lingüística para mencionar un referente en el discurso) está correlacionada, entre otros factores, con la distancia referencial (es decir, el número de cláusulas entre una mención de un referente y la mención anterior a dicho referente; ver Givón 1983: 13). En este trabajo se analiza y describe la correlación entre distancia referencial y las estrategias referenciales (es decir, las formas seleccionadas para hacer mención de un referente) dentro de un corpus escrito en hñöhñö, por medio de dos métodos estadísticos: prueba de Kruskal-Wallis (Kruskal y Wallis 1952) y análisis de correspondencias (Benzécri 1973). La metodología de trabajo se basa en los métodos utilizados en Kibrik (2011).

PALABRAS CLAVE: Otomangue; Otopame; Discurso; Distancia referencial; Morfosintaxis

1. Introducción

La lengua hñöhñö (llamada comúnmente “otomí”) se habla en numerosas comunidades del municipio de Tolimán (Querétaro, México). Las lenguas otomíes como el hñöhñö pertenecen a la rama otopame del macrofilo otomangue (centro y sur de México). Como todas las lenguas otomíes, y la mayoría de las lenguas otomangues, el hñöhñö es tonal, de marcación en el núcleo y con orden básico VO. Existen algunas publicaciones de narraciones en hñöhñö, entre las cuales se encuentra *Ya Hñö Ma Mengu. Las Voces De Mis Familiares* (González Morales; González, L.; Martínez 2019). En estos textos pueden encontrarse distintas estrategias referenciales (Kibrik 2011: xxvi, 37) para hacer mención de los referentes en el discurso. Por ejemplo, en (1a) se hace mención de una locación por medio de una frase nominal, en (1b) se utiliza una proforma tónica para referir a una persona, en (1c) el sufijo *-bi* ‘3OBJ’ hace referencia a un receptor, y la anáfora cero (\emptyset) en (1d) refiere a un participante paciente.

- (1) a. 'na=r pa da=ma[=r tohø]_{FN}¹
 uno=SG día 1.PFV=TS\ir=SG cerro²
 'un día fui al cerro'
- b. go bi=mefi [no='o]_{PRO}
 IDEN (3)PFV=TS\hacer DEF=DEM.SG
 '(fue él quien lo hizo)'³
- c. him=ba t-um[-bi]_{SUF=r} dethö
 NEG=PFV.VEN IMPRS-dar-3OBJ=SG maíz
 'no le dieron maíz'
- d. da=ma da=heti [Ø] ne dá=juts'i [Ø]
 1.PFV=TS\ir 1.PFV=ver también 1.PFV=alzar
 'fui a verlo (al chivito) y lo levanté'

A la selección de la estrategia referencial en el discurso se le conoce como selección referencial (*referential choice*; Kibrik 2011: xxvi). A nivel tipológico, la selección referencial está correlacionada, entre otros factores, con la distancia referencial, es decir, el número de cláusulas entre una mención de un referente y la mención anterior a dicho referente (Givón 1983: 13). En este trabajo se analiza y describe esta correlación dentro de un corpus escrito en hñöhñö, por medio de dos métodos estadísticos: prueba de Kruskal-Wallis (Kruskal y Wallis 1952) y análisis de correspondencias (Benzécri 1973).

Los términos y marco conceptual utilizados en este estudio se introducen en la sección "Terminología". En la sección "Antecedentes" se hace un recuento de investigaciones sobre lenguas otomías relacionadas con el tema de este trabajo. Las estrategias referenciales de la lengua de estudio se describen en la sección "Estrategias referenciales en hñöhñö". Los procedimientos seguidos en esta investigación se describen en la sección "Metodología". El análisis de los datos y su interpretación se presentan en las secciones "Resultados" y "Discusión", respectivamente. Finalmente, la sección "Conclusiones" contiene un resumen y comentarios finales sobre la investigación.

2. Terminología

Al describir la selección referencial, Kibrik (2011: xxvi) apela a dos tipos de expresiones referenciales: estrategias referenciales plenas y estrategias referenciales reducidas. Las plenas pueden consistir en nombres propios (e.g. *María, José*), sustantivos comunes (con

¹ Ortografía práctica (cuando difiere del AFI): <ä> = /ã/, <a> = /ɔ/, <ë> = /ẽ/, <e> = /ɛ/, <ɸ> = /p^h/, <ɔ> = /ʔ, ʔ/, <ɨ> = /i/, <ɟ> = /k^h/, <ñ> = /ɲ/, <ö> = /õ/, <ɔ> = /ɔ/, <r> = /r/, <ü> = /ũ/, <u> = /i/, <w> = /w, ʷ/, <x> = /ʃ/, <y> = /j/. La tilde representa tono alto (á) y la vocal doble indica tono ascendente (aa); el tono bajo no se marca ortográficamente.

² Abreviaturas: 1 = primera persona, 2 = segunda persona, 3 = tercera persona, AUM = aumentativo, COP = cópula, DEF = definido, DEM = demostrativo, DEP = dependiente, DET = determinante, DIM = diminutivo, DU = dual, ENF = enfático, EXCL = exclusivo, FF = formativo flexivo, FN = frase nominal, FP = frase preposicional, IDEN = identificativo, IMPRF = imperfecto (pasado imperfectivo), IMPRS = impersonal, IPFV = imperfectivo, IRR = irrealis, LOC = locativo, NEG = negación, OBJ = objeto, ORD = ordinal, PFV = perfectivo, PL = plural, POS = poseedor, PPRF = plu-perfecto, PRF = perfecto, PRO = proforma, PSD = pasado, SG = singular, SUF = sufijo, TAM = Tiempo-Aspecto-Modo, TS = tema secundario, VEN = venitivo.

³ Las palabras entre paréntesis en la línea de traducción libre de los ejemplos glosados no están en la traducción original de González Morales et al. (2019), sino que se han insertado para adecuar la traducción a su presentación fuera de contexto.

o sin modificadores, e.g. *perros, los perros, los perros de la calle, los perros que viste*) o descripciones (e.g. *lo que viste ayer*); las reducidas corresponden a pronombres (ya sean ligados o libres) y formas cero (Kibrik 2011: 37). La selección referencial consiste entonces en la estrategia referencial que elige el hablante a la hora de referirse a un referente dado. Esta elección está condicionada por varios factores de activación: animacidad (humano, no-humano), protagonismo, rol sintáctico (sujeto, no sujeto), rol semántico (principal,⁴ no-principal), distancia lineal del antecedente, distancia retórica⁵ del antecedente, tipo de estrategia referencial del antecedente, rol sintáctico o semántico del antecedente, etc. En este trabajo solo nos centraremos en la distancia retórica del antecedente, la cual es uno de los factores de activación de medición objetiva y escalar, además de ser un predictor confiable de la selección referencial. En adelante se le llamará simplemente “distancia referencial”.

La distancia referencial se mide tomando en cuenta relaciones de subordinación (semántica o discursiva) entre ellas, y no en un orden lineal (Kibrik 2011: 404). Este conteo sigue la Teoría de Estructura Retórica (Fox 1987 y Mann y Thompson 1988, citados en Kibrik 2011: 403), y se explicará con más detalle en la sección “Metodología”. Se asume que la distancia referencial y la estrategia referencial se correlacionan con la activación del referente: a mayor activación, menor distancia referencial y estrategias referenciales reducidas; a menor activación, mayor distancia referencial y estrategias referenciales plenas (ver Givón 1983: 17). Véase el gradiente de estrategias referenciales en función de la continuidad de tópico en (2).

- (2) + anáfora cero
 pronombres átonos/dependientes o concordancia gramatical
 pronombres tónicos/independientes
 FNs definidas dislocadas a la derecha
 FNs definidas de orden neutro
 FNs definidas dislocadas a la izquierda
 topicalización contrastiva
 construcciones escindidas/de foco
 - FNs referenciales indefinidas (Givón 1983: 17)

Givón (1983: 13) mide la distancia referencial (en su sentido de distancia lineal del antecedente) en un rango de entre 1 (mínimo) y 20 cláusulas en la porción de texto que antecede; distancias referenciales mayores a 20 se registran como 20 en su estudio. Kibrik (2011: 408), por su parte, pone el tope superior de la distancia referencial en 6 cláusulas. Este último tope es el que se utiliza en el presente estudio.

3. Antecedentes

Las estrategias referenciales de un buen número de lenguas otomías han sido descritas, de una manera más o menos detallada, en numerosos trabajos (Andrews 1993; de Cárceres 1907[1580]; Ecker 2012; Hekking y Andrés de Jesús 1984; Hernández-Green 2015; Hess 1968; Lastra 1992; 1997; Voigtlander y Echevoyen 1979). Las estrategias correspondientes al hñöhñö (otomí de Tolimán) no se han descrito en trabajos publicados, pero guardan muchas semejanzas con el otomí de San Ildefonso Tultepec (Amealco, Querétaro; Palancar 2009); estas se describen más adelante en la sección “Estrategias referenciales en hñöhñö”.

⁴ Correspondiente a los roles temáticos A y S (Kibrik 2011: 28).

⁵ El término “distancia retórica” refiere a la distancia entre una mención de un referente y su antecedente, en términos de nodos en una red jerárquica, conectados por las mismas relaciones discursivo-semánticas (esto es, retóricas; Kibrik 2011: 14).

En cuanto a la distancia referencial en lenguas otomíes, no se ha encontrado ningún antecedente especializado en la literatura, si bien Enrique L. Palancar (p.c.) se encuentra preparando trabajos de investigación al respecto. En descripciones sobre lenguas otomíes de Querétaro y el estado de México pueden encontrarse referencias aisladas sobre alternancias entre estrategias referenciales plenas y reducidas (Andrews 1993: 97; Hernández-Green 2015: 197), anáfora por medio de morfemas ligados y morfemas cero (Hernández-Green 2015: 447; Palancar 2009: 233-234), anáfora por medio de marcas de poseedor (Hernández-Green 2015: 63), anáfora por medio de pronombres demostrativos (Hernández-Green 2015: 45), el papel de la “distancia anafórica” en el uso de demostrativos con función pronominal (Palancar 2009: 117), y sobre continuidad de tópico mediante una construcción impersonal (Hernández-Green 2015: 320).

4. Estrategias referenciales en hñöhñö

En las siguientes secciones se describen las estrategias referenciales disponibles en hñöhñö, tanto plenas como reducidas. Solo se incluyen las estrategias que se usan para referir a participantes de 3ª persona, ya que la 1ª y 2ª personas tienen un estatus discursivo especial: siempre están activados, rara vez suelen expresarse mediante estrategias referenciales plenas y no se expresan mediante formas anafóricas (Kibrik 2011: 43). Estas características las hacen poco comparables con la 3ª persona en términos de la distinción pleno/reducido y distancia referencial.

4.1 Estrategias plenas

Entre las estrategias referenciales plenas en hñöhñö se encuentran las frases nominales y los nombres propios. En esta sección también se han incluido las frases preposicionales, ya que en ocasiones estas refieren a participantes antes mencionados en el discurso, tanto animados como inanimados. Además, contienen frases nominales de por sí.

4.1.1 Frase nominal

Las frases nominales en hñöhñö pueden consistir en sustantivos escuetos, o bien en sustantivos con modificadores tanto prenominales como posnominales. Los modificadores prenominales guardan el orden relativo de ocho posiciones mostrado en la Tabla 1. La posición -8 no ha podido corroborarse en el corpus consultado, sino que se ha extrapolado de la estructura de la frase nominal en otomí de San Ildefonso Tultepec (Palancar 2009: 96). Los modificadores de las posiciones -5 y -6 se representan como clíticos porque solo se observan cuando las posiciones -4 y -7 están ocupadas, respectivamente. Ejemplos de estas posiciones (con excepción de los afectivos de la posición -2) se muestran en (3a-c); en (3d) se ilustra una frase nominal escueta.

Tabla 1. Plantilla (*template*) de posiciones prenominales

cuantificador	<i>ndunthi</i> ‘mucho(s)’ <i>gatho</i> ‘todo(s)’	-8
	<i>nu~no</i> ‘DEF’	-7
demonstrativo	= <i>na</i> ‘DEM.SG’ = <i>o</i> ‘DEM.SG’ = <i>yu</i> ‘DEM.PL’ = <i>u</i> ‘DEM.PL’	-6
	<i>ma=</i> ‘otro’	-5
numerales	<i>'ra</i> ‘(alg)unos’ <i>'na</i> ‘uno’ ⁶ <i>yoho</i> ‘dos’ <i>hñu</i> ‘tres’...	-4
número /poseedor	<i>ar</i> ‘SG’ <i>ya</i> ‘PL’ <i>ma</i> ‘1POS’ <i>ri</i> ‘2POS’ <i>rá</i> ‘SG3POS’ <i>yá</i> ‘PL3POS’	-3
afectivo	<i>tsi</i> ‘DIM’ <i>nda</i> ‘AUM’	-2
adjetivo	<i>dötä</i> ‘grande’ <i>nts'o</i> ‘malo’ <i>theni</i> ‘rojo’...	-1
sustantivo	N	0

- (3) a. *ma=gi=thogi*, [**nu=na=r** **dötä** *hñe*]
AUX:FUT=2.IRR=pasar NU=DEM.SG=SG grande río
‘(vas a pasar) un gran río’
- b. [**gatho ya** *jö'i*]_{FN} *da=tsi-the*
todo PL persona IRR=ingerir-agua
‘que toda la gente la beba’
- c. *ja gi=tini* [**no ma='na=r** *ndö*]_{FN}
justo 2.IRR=encontrar NU otro=uno=SG varón
‘ahí encontrarás al otro hombre’

⁶ Si bien este morfema es glosado como indefinido por algunos autores cuando aparece en frases nominales indefinidas (ver, por ejemplo, la glosa IND.SG en Palancar 2009), en este trabajo se glosa simplemente como el numeral ‘uno’ porque también puede aparecer en frases nominales interpretadas como definidas (por ejemplo, *no ma 'nar ndö* ‘el otro hombre (previamente mencionado)’).

- d. gi='ra-ki [mefi]_{FN}
 2.IRR=dar-1OBJ trabajo
 'para que me des trabajo'

Los modificadores posnominales pueden ser de diversos tipos. Los tipos más comunes son oraciones relativas, frases nominales y frases preposicionales. Ejemplos de cada tipo se muestran a continuación en (4).

- (4) a. no ma='na=r hñe [gi=thogi]_{OR}
 DEF otro=uno=SG río 2.IRR=pasar
 'el otro río que pasarás'
- b. em-ba=bi no rá döme [no=r dönxu]_{FN}
 decir-3OBJ=3OBJ DEF SG3POS esposo DEF=SG mujer
 'y le dijo al esposo de la mujer'
- c. bi=thogi 'na=r döthe [de ga ji]_{FP}
 (3)PFV=pasar uno=SG río.grande de de sangre
 'pasó (por) un río de sangre'

Los roles temáticos centrales de tipo S (= único argumento de predicado monovalente), A (= argumento más agentivo de predicado bivalente), R (= argumento más paciente de predicado bivalente), P (= argumento más paciente de predicado bivalente) y T (= argumento ni más agentivo ni más paciente de predicado trivalente; ver Bickel 2010) se expresan comúnmente por medio de frases nominales plenas sin una marcación especial; en unos pocos casos se ha observado el uso de la preposición *ha* 'LOC' en frases que refieren al rol R. Esto se ilustra a continuación en (5a). En (5b) se muestra una frase con semántica de tipo locativo introducida por esa misma preposición. Además, en unos pocos casos algunos roles semánticos de tipo locativo pueden expresarse por medio de frases nominales plenas, en lugar de una frase preposicional, como en el ejemplo de (5c).

- (5) a. bi=em-ba=bi [ha no=r tsi möjō]_{FP-R}
 (3)PFV=decir-3OBJ=3OBJ LOC DEF=SG DIM sacerdote
 'le dijo al padrecito'
- b. bi=ma [ha rá nguu]_{FP-LOC}
 (3)PFV=TS\ir LOC SG3POS casa
 'se fue a su casa'
- c. da=ma [yá nguu]_{FN-LOC}
 (3)IRR=TS\ir PL3POS casa
 'y van por todas las casas'

4.1.2 Frase preposicional

Como su nombre lo indica, las frases preposicionales son introducidas por una preposición. La preposición indica la relación semántica que guarda la frase preposicional con la proposición expresada en la oración, o bien con el sustantivo modificado por dicha frase. La preposición puede estar seguida de una frase nominal, si bien cada preposición en particular se encuentra en el corpus en combinación con frases nominales de distintos grados de

complejidad: las preposiciones de (6a) se observan con frases nominales ocupadas hasta la posición -8 (cuantificadores), la preposición de (6b) se ha encontrado con modificadores hasta la posición -7 (la marca de definitud *nu/no*), las de (6c) hasta la posición -3 (número/posesión), mientras que nada puede decirse (a partir de los datos del corpus) sobre las posiciones prenominales ocupadas en frases preposicionales introducidas por las preposiciones de (6d).

- (6) a. ha/ja 'LOC'
ko(n) 'con'
- b. de 'de; desde'
- c. pa 'para'
de ga 'de'
mbo 'dentro'
- d. a(s)ta 'hasta'
dende 'desde'

Con excepción de *ha/ja* 'LOC', las preposiciones de (6) son todas prestadas del español; *de ga* 'de' se compone de dos morfemas. El primero de estos es un préstamo del español. El segundo, *ga*, es una preposición nativa que coocurre siempre con *de* en el corpus, y que en otras lenguas otomías, como el otomí de Acazulco (estado de México), puede traducirse también como 'de'.

Como se señaló en la sección "Frase nominal", las frases con roles semánticos de tipo locativo suelen estar introducidas por preposiciones. Tres ejemplos pueden verse en (7). Además, las frases preposicionales pueden tener semántica de instrumento, propósito o contenido, como se ilustra respectivamente en (8a), (8b) y (8c).

- (7) a. bi=ma [ha rá nguu]_{FP}
(3)PFV=TS\ir LOC SG3POS casa
'se fue a su casa'
- b. da=ndui da=mbu [de yá nguu]_{FP}
(3)IRR=empezar (3)IRR=volar desde PL3POS casa
'hasta que se despegan del suelo y salen volando de sus casas'
- c. nu'bya höö bi=gui rá 'ye [mbo=r hai]_{FP}
ahora sí (3)PFV=TS\hundirse SG3POS mano dentro=SG tierra
'ahora sí metió su mano dentro de la tierra'
- (8) a. ja bi=tsęki rá ñö rá 'mehñö [ko=r t'egi]_{FP}
justo (3)PFV=cortar SG3POS cabeza SG3POS esposa con=SG hacha
'y ahí le abrió la cabeza a su esposa con el hacha'
- b. ja-ka=gi 'na=r tsi zaa [pa ma bordo]_{FP}
hacer-1OBJ=1OBJ uno=SG DIM palo para 1POS bordón
'háganme un palo para mi bordón'
- c. ne höö bi=dini no=r ts'oe [de ga k'axt'a böjö]_{FP}
también sí (3)PFV=TS\encontrar DEF=SG olla de de dorado moneda
'y sí encontró aquella olla de monedas de oro'

La preposición *ko* ‘con’ de (8a) también se usa para introducir frases concomitantes (esto es, cada uno de los agentes en un evento comitativo). Frases de este tipo se ilustran en (9).

- (9) a. *bi=ma [ko ma='na=r ndo]*_{FP}
 (3)PFV=TS\ir con otro=uno=SG varón
 ‘(que) se fue con otro hombre’
- b. *mí=petsi ya bötsi [ko ndunthi ya ndo]*_{FP}
 (3)IMPRF=tener PL niño con mucho PL varón
 ‘tenía muchos hijos de muchos hombres’

Las frases preposicionales concomitantes ejemplificadas arriba se encuentran con verbos tanto monovalentes como *pa* ‘ir’ en (9a) como bivalentes como *petsi* ‘tener’ en (9b). Con verbos monovalentes, las frases concomitantes pueden tener la forma de una frase nominal (ver sección “Frase nominal”) en una construcción de coordinación escindida,⁷ la cual se mostrará más adelante en el ejemplo (18) de la sección “Afixos pronominales”.

4.2 Estrategias reducidas

Las estrategias referenciales reducidas en hñöhñö tienen una diversidad mayor que las estrategias plenas. Por un lado, tenemos proformas tónicas que pueden consistir en cuantificadores, demostrativos y otros modificadores de la frase nominal (y combinaciones de estos; ver Tabla 1) en función pronominal, o bien en combinaciones de cópula más clítico pronominal. Por otro lado, entre las proformas libres encontramos tanto adverbios locativos como marcas de persona gramatical. El término “proforma” se utiliza aquí en el sentido amplio de “un elemento en una oración que sustituye a [= desempeña la función de] otros elementos o construcciones” (Crystal 2008: 390), que abarca (pero no se limita a) pronombres y adverbios locativos como *ahí* en español. Además, encontramos afijos verbales (marcas de persona) y anáfora cero.

4.2.1 Proformas tónicas

La mayoría de las proformas libres en el corpus revisado son formas (o combinaciones de estas) que se encuentran en las posiciones -8 a -4 de la plantilla de la frase nominal en la Tabla 1. Véanse los ejemplos de (10) a continuación.

- (10) a. **gatho** *pödi*
 todo saber
 ‘todos saben’
- b. *go bi=mefi no='o*
 IDEN (3)PFV=TS\hacer DEF=DEM.SG
 ‘(fue él quien lo hizo)’
- c. **nu='u ma='ra bi=wadi bi=xo'mi no=r nt'at'i**
 DEF=DEM.PL otro=unos (3)PFV=TS\acabar(3)PFV=escarbar DEF=SG pozo
 ‘los demás acabaron de escarbar el pozo’

⁷ Esta es una construcción donde “los elementos coordinados en la estructura otomí aparecen [...] en diferentes lugares de la cláusula” (Palancar 2009: 554-571).

- d. de ge'pya mí='mɔi 'na
 desde LOC (3)PFV.DEP=TS\pararse uno
 'de ahí se detuvo uno de ellos'

Además, los clíticos demostrativos de la Tabla 1 y el clítico =se 'de por sí' pueden funcionar como pronombres en combinación con la cópula ge. Esto ocurre en contextos de focalización, como en los ejemplos de (11), o bien en frases preposicionales como en (12). El clítico locativo =ni de (12b) se verá más adelante en la sección "Proformas libres (átonas)".

- (11) a. go ge='ä bi=nja'bu
 IDEN COP=DEM.SG (3)PFV=ser.así
 '(ESO fue así)'
- b. ge=se mixki=xi-pa=bi
 COP=de.por.sí (3)PPRF=decir-3OBJ=3OBJ
 'él mismo se lo había dicho'
- (12) a. de ge='ä bi=zohɔ
 desde COP=DEM.SG (3)PFV=TS\llegar
 '(después de eso) llegó'
- b. no=r txu bi=ma de ge='ni
 DEF=SG abuela (3)PFV=TS\ir desde COP=LOC
 'aquella mujer se fue de ahí'

Finalmente, los numerales pueden entrar en construcciones pronominales en combinación con la partícula 'ra 'ORD'. Estas construcciones pueden traducirse como "el tercero" o "el cuarto", como puede verse en el ejemplo de (13).⁸

- (13) ya ata 'ra 'rato bi=ne bi=jöni ko tsi mojõ
 ya hasta ORD seis (3)PFV=querer (3)PFV=confesarse con DIM sacerdote
 'hasta el sexto (hijo) quiso confesarse con el padre'

4.2.2 Proformas libres (átonas)

El enclítico =bi '3OBJ' es un pronombre tenaz que hace referencia a participantes de tipo recipiente/benefactivo, y se ejemplifica (por conveniencia) en la sección "Afijos pronominales". Entre los textos del corpus se ha identificado solo una proforma locativa =ni 'LOC', que se usa para referir a una locación accesible en el contexto. Dos ejemplos se muestran en (14) a continuación.

- (14) ja bi=zo=ni no=r metsi bi=du=ni
 justo (3)PFV=TS\caer=LOC DEF=SG muchacho (3)PFV=TS\morir=LOC
 'cayó ahí el joven y murió ahí (en el gran agujero)'

⁸ La construcción también puede encontrarse en frases nominales (no 'ra hñõto pa 'el octavo día'), pero no se ha incluido en la sección "Frase nominal" ya que no hay suficientes datos para ubicar la construcción 'ra + numeral en la plantilla de la Tabla 1.

La mayoría de las descripciones sobre lenguas otomíes afirman que los formativos flexivos (Palancar 2011a) preverbiales indican, además de nociones de TAM, la persona gramatical del sujeto incluyendo primera, segunda y tercera persona (Voigtlander y Echevoyen 1979: 96; Andrews 1993: 6; Palancar 2009: 175; 2011b: 249; 2012; Bartholomew 2010: 503). Hernández-Green (2016: 43; 2018: 289; 2020: 4; 2021: 14), por otro lado, considera que la tercera persona no se marca explícitamente en el formativo flexivo. Esta alternativa entre dos análisis se representa con el número “3” entre paréntesis en la Tabla 2.

Tabla 2. Formativos flexivos de 3ª persona en hñöhñö

	realis	irrealis
perfectivo	bi= ‘(3)PFV’ mí= ‘(3)PFV.DEP’ xa= ‘(3)PRF’ (mi)xki= ‘(3)PPRF’ mixpi= ‘(3)PPRF.VEN’	da= ‘(3)IRR’ nda= ‘PSD(3)IRR’
imperfectivo	ø~i=~di= ‘(3)IPFV’ mí= ‘(3)IMPRF’ ba= ‘(3)VEN’	

Los formativos flexivos enlistados en la Tabla 2 son “pronombres (personales) libres” en la tipología de Kibrik (2011), ya que se trata de elementos pronominales clíticos (y no afijos). La caracterización de estos elementos como clíticos se basa en las siguientes propiedades: a) pueden hospedarse en otros morfemas distintos del tema verbal, típicamente el morfema de negación *hi*, b) otros morfemas considerados como palabras, como el diminutivo *tsi*, pueden ocurrir entre el formativo flexivo y el predicado. Ejemplos de estas dos propiedades se muestran respectivamente en (15a) y (15b).

- (15) a. nu=yu ma=ra him=**ba** hyegi
 DEF=DEM.PL otro=unos NEG=PFV.VEN TS\soltar
 ‘pero los otros no lo dejaron’
- b. **xa**=tsi nk'a
 PRF=DIM húmedo
 ‘y estaba fresca’

Por otro lado, los formativos flexivos son también “pronombres marcados por categorías clausales”, ya que codifican nociones de TAM, además de otras categorías como movimiento asociado (ver Hernández-Green y Palancar 2021). Los formativos flexivos son también “pronombres libres tenaces” pues ocurren en posición preverbal independientemente de la expresión abierta y del estatus discursivo de su referente (ver Kibrik 2011: 96).

Finalmente, los modificadores nominales en la posición -3 de la Tabla 1 que indican poseedor también son clíticos con rasgos morfosintácticos de persona. En particular, los morfemas *rá* ‘SG3POS’ y *yá* ‘PL3POS’ refieren a un poseedor de tercera persona, ya sea que este se exprese en la misma oración (16a) o en alguna oración previa (16b).

- (16) a. em-ba=bi no **rá** dōme [no=r dōnxu]_{poseedor}
 decir-3OBJ=3OBJ DEF SG3POS esposo DEF=SG mujer
 ‘y le dijo al esposo de la mujer’

- b. mí=tsə=pu ha rá nguú no=r ndq_i
 (3)PFV.DEF=llegar=LOC LOC SG3POS casa DEF=SG varón
 no **rái** dōnxu 'mōi ha=r juní
 DEF SG3POS mujer pararse LOC=SG metate
 ‘cuando llegó a su casa aquel hombre, su esposa estaba parada en el metate’

4.2.3 Afijos pronominales

En hñöhñö pueden observarse dos afijos pronominales que refieren a una tercera persona en la morfología verbal. Por un lado, el sufijo *-bi* ‘3OBJ’ (y alomorfos *-ba*, *-pi*, *-pa*, *-wi*, *-wa*)⁹ es la estrategia referencial típica para participantes de tipo Recipiente/Benefactivo en lenguas otomías. Véanse los ejemplos de (17). A menudo el sufijo viene seguido del clítico *=bi* ‘3OBJ’, como se muestra en (17b), en un fenómeno puramente morfológico (esto es, sin añadir ningún matiz ni semántico ni pragmático) que Palancar llama “marcación doble de objeto” (Palancar 2009: 249).

- (17) a. pe no ma='na=r dōnxu bi=tsöt-**pi** ya hme
 pero DEF otro=uno=SG mujer (3)PFV=quemar-3OBJ PL tortilla
 ‘pero aquella otra mujer le quemó sus tortillas’
- b. da=um-**ba**=bi ya 'ñi
 (3)IRR=dar-3OBJ=3OBJ PL chile
 ‘que le diera chiles secos’

El sufijo *-wi* ‘DU’ por otro lado, se usa para indicar la presencia de un participante concomitante en una construcción de coordinación escindida (Palancar 2009: 554-571). En esta construcción, los participantes de un evento comitativo se realizan como expresiones no adyacentes, mientras que el verbo recibe morfología de número dual. Por lo general, en el corpus revisado, el concomitante menos topical se expresa por medio de una frase nominal plena, como en (18a), o bien simplemente es referido anafóricamente mediante el sufijo *-wi* ‘DU’, como en (18b).

- (18) a. nu'bya bi=ma bi=hñö-**wi** [tsi mojó]_{FN}
 ahora (3)PFV=TS\ir (3)PFV=hablar-DU DIM sacerdote
 ‘ahora (ella) se fue a hablar con el padrecito’
- b. wa da=nthe-**wi** a=r 'ñu
 o (3)IRR=encontrarse-DU LOC=SG camino
 ‘o (se encuentran con ellos) en la calle’

Nótese que la frase concomitante en esta construcción es una frase nominal y no una frase preposicional, como se vio en la sección “Frase preposicional”. Ambos sufijos, *-bi* ‘3OBJ’ en (17) y *-wi* ‘DU’ en (18), son tenaces, es decir, deben estar presentes incluso cuando el referente al que están asociados se encuentra expresado en la misma cláusula mediante una frase nominal plena (compárese con Kibrik 2011: 118).

⁹ Los alomorfos con consonante /p/ ocurren en posición inmediata postónica después de vocal o de consonante sorda (excepto /k/), mientras que los alomorfos con /w/ ocurren después de /k/. Los alomorfos con vocal /a/ ocurren en posición medial de enunciado, mientras que los alomorfos con vocal /i/ ocurren, por lo general, en posición final de enunciado. Compárese con las alomorfías correspondientes en otomí de San Ildefonso Tultepec (Palancar 2009: 234, 241-243).

4.2.4 Anáfora cero

La anáfora cero se observa sobre todo en terceras personas con rol temático P y T, ya que estos no tienen ninguna expresión morfológica observable en el verbo. En los ejemplos de (19a) y (19b), la frase nominal en negritas es mencionada implícitamente en la oración siguiente con rol temático P y T, respectivamente. El referente no expresado se representa por medio del símbolo “Ø”, siguiendo a Kibrik (2011).

- (19) a. **no=r** **dönxui** bi=pura bi=got'i rá goxti
 DEF=SG mujer (3)PFV=apurarse (3)PFV=TS\cerrar SG3POS puerta
 pa hinjontho da='yode [Ø]_iP
 para nadie (3)IRR=TS\oír
 ‘aquella mujer se apresuró para cerrar su puerta para que nadie la oyera’
- b. bi=ñötsi 'na=r zaa,
 (3)PFV=pelar uno=SG palo
 ne bi=ja-pi [Ø]_iT mote ha rá xutha
 y (3)PFV=poner-3OBJ detrás LOC SG3POS espalda
 ‘peló un palo y se lo pusieron en la espalda’

Participantes con rol temático R se dejan sin codificación explícita en el verbo solo con unos pocos lexemas verbales. Un ejemplo con el verbo *'uti* ‘mostrar’ de este tipo se muestra en (20) a continuación. Los dos ceros en negritas refieren al protagonista de la narración.

- (20) bi=nzofa [Ø]_iP 'na=r ndo ne bi='ñuti [Ø]_iR
 (3)PFV=hablarle uno=SG varón y (3)PFV=mostrar
 ha'bu mí='müi 'na=r ts'oe mí=ñudi ya k'axt'a böjö
 dónde (3)IMPRF=estar uno=SG olla (3)IMPRF=llenarse PL dorado moneda
 ‘(aquel señor; borracho oyó) que (un hombre) le_i hablaba y le_i enseñó
 donde estaba una olla llena de monedas de oro’

La representación de estos ceros anafóricos fuera de la morfología verbal va en contra de la representación que se da a estos elementos no pronunciados en otras lenguas otomíes y en mazahua, una familia de lenguas estrechamente emparentada con el otomí. En otras descripciones de estas lenguas, el cero correspondiente a participantes con roles temáticos P y T se representa como un sufijo verbal (ver, por ejemplo, Palancar 2009; Mora-Bustos 2018). En esas descripciones se sigue un criterio de isomorfismo paradigmático con respecto a la marcación de primera y segunda persona mediante sufijos. La decisión tomada en este trabajo busca una representación más coherente con la morfosintaxis de la lengua. Aquí se toman en cuenta dos hechos observados en construcciones monotransitivas y bitransitivas en otras lenguas otomíes descritas: a) la primera y segunda personas se codifican mediante sufijos siempre que tengan los roles P o R, pero nunca cuando tienen el rol T, mientras que b) la tercera persona, mayoritariamente, solo se codifica con sufijo cuando tiene el rol R. La Tabla 3 representa este patrón morfológico.

Tabla 3. Morfología de persona en el tema verbal

	R	P	T
1	-gi		
2	-i		
3	-bi		

El área sombreada en la Tabla 3 corresponde a celdas del paradigma donde no hay codificación morfológica explícita. El área sombreada correspondiente a primera y segunda persona es incluso imposible de expresar, como no sea en construcciones montransitivas donde el T se codifica con sufijo y el R se expresa mediante una frase preposicional (ver Hernández-Green 2015: 191; 2022: 24). Como esta área del paradigma no forma parte de la morfología verbal, en este trabajo se considera que el área sombreada correspondiente a la tercera persona tampoco forma parte de ella.

Finalmente, Hernández-Green (2016: 43; 2018: 289; 2020: 4; 2021: 14) propone que los participantes con roles temáticos A/S también son susceptibles de anáfora cero. El autor analiza los formativos flexivos con A/S de 3ª persona como meros exponentes de categorías de TAM (y otras categorías flexivas), y no como marcas de persona como se representan más arriba en la Tabla 2. En este trabajo se consideran las dos posturas en el análisis: a) los formativos flexivos como marcas de persona (esto es, pronombres libres), con el número “3” incluido en la glosa (ejemplos: *bi*= ‘3.PFV’, *mí*= ‘3.IMPRF’) y b) los formativos flexivos como meras marcas de TAM y otras categorías gramaticales, sin referencia a persona en la glosa (ejemplos: *bi*= ‘PFV’, *mí*= ‘IMPRF’).

5. Metodología

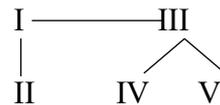
Los datos presentados y analizados en este trabajo fueron tomados de un corpus de 26 textos (5675 palabras morfosintácticas) escritos en hñöhñö por habitantes de la comunidad de Bomintzá (Tolimán, Querétaro) y alrededores. Estos textos, publicados en González Morales et al. (2019), fueron digitalizados por mí directamente en FLEx. Dentro de FLEx, se etiquetaron todas las instancias de cualquier referente mencionado más de una vez en cada texto (incluyendo las menciones sin expresión explícita, o anáfora cero). Estas etiquetas contienen: a) un número identificador único para cada referente, y b) una indicación de si se trata de la primera mención o no.

Las etiquetas descritas en el párrafo anterior se vertieron después en una hoja de Excel. En esta base de datos se indicó, por cada mención en el corpus, la estrategia referencial utilizada (esto es, FP, FN, pronombre, pronombre enfático,¹⁰ sufijo, formativo flexivo, marca de poseedor o anáfora cero), y la distancia referencial (en número de cláusulas) entre cada mención y la última mención del referente correspondiente. La distancia referencial fue medida de acuerdo con la Teoría de Estructura Retórica (Fox 1987 y Mann y Thompson 1988, citados en Kibrik 2011: 403). Siguiendo la metodología utilizada en Kibrik (2011), en el conteo se ignoran menciones en trechos de texto que corresponden a discurso directo, a menos que en estos ocurra la primera mención de un referente retomado en otro trecho que no es discurso directo; además, las cláusulas no se cuentan en un orden lineal sino tomando en cuenta relaciones de subordinación (semántica o discursiva) entre ellas. Por ejemplo, en la Figura 1 se

¹⁰ En este trabajo, la etiqueta “pronombre enfático” se ha aplicado a proformas que se utilizan en contextos de focalización (ver Crystal 2008: 192-193).

representan cinco cláusulas consecutivas en el orden indicado por los números romanos, así como las relaciones de subordinación que guardan entre ellas, en forma de líneas rectas. En esta secuencia, la distancia entre las cláusulas II y IV es 3 (líneas), y no 1 como correspondería por el orden lineal que guardan entre ellas; de igual manera, la distancia entre I y V es 2, y no 4.

Figura 1. Distancia referencial en la Teoría de Estructura Retórica



Las cinco cláusulas de (21) ilustran la estructura retórica representada en la Figura 1. Las cláusulas subordinadas se encuentran encerradas entre corchetes, seguidas de un subíndice que refiere a la cláusula a la cual están subordinadas.

- (21) a. enä
decir
'(decía)'
- b. [hin=ga hñi-he]_(21a)
NEG=1.IRR bañarse-PL.EXCL
'que no nos bañaríamos'
- c. enä
decir
'decía'
- d. [ge da=din-ga=gi ya ndogu wa ya nts'o]_(21c)
que (3)IRR=TS\encontrar-1OBJ=1OBJ PL soldado o PL malo
'que si nos encontraban los soldados o los diablos'
- e. [da=dix-ka=gi mfodi]_(21c)
(3)IRR=TS\llevar-1OBJ=1OBJ cárcel
'nos llevarían a la cárcel'

La base de datos descrita en párrafos anteriores (observaciones $n = 1105$) se sometió a una prueba de suma de rangos de Kruskal-Wallis (Kruskal y Wallis 1952) para verificar si existen diferencias significativas en la distancia referencial con respecto a cada una de las estrategias referenciales. Además, se reorganizaron los datos numéricos de distancia referencial en seis categorías nominales ("1", "2", "3", "4", "5" y ">5"), siguiendo a Kibrik (2011: 408). A estas categorías, tabuladas contra las categorías nominales de estrategia referencial, se les aplicó una prueba χ^2 y un análisis de correspondencias (Benzécri 1973) a fin de observar la asociación entre ellas. Todas estas pruebas se llevaron a cabo en RStudio 4.2.0 (RStudio Team 2015).

6. Resultados

El diagrama de caja de la Figura 2 muestra los cuartiles y las medianas de la distancia referencial (eje vertical) por estrategia referencial (eje horizontal). Las diferencias de distancia referencial dividen las estrategias referenciales en dos agrupaciones: por un lado, las estrategias reducidas DET-ENF-FF-Ø-SUF (DET = marca de poseedor, ENF = pronombre enfático, FF =

formativo flexivo, Ø = anáfora cero, SUF = sufijo) con medianas y cuartiles con valores entre 1 y 1.25, y por otro lado, las estrategias PRO-FN-FP (PRO = proforma tónica, FN = frase nominal, FP = frase preposicional), con medianas y tercer cuartil de 2 y mayores. Los resultados en que se considera nula la codificación de persona en el formativo flexivo (siguiendo a Hernández-Green 2021: 14) se muestran en la Figura 3, donde se han combinado los datos de las estrategias FF y Ø, con valores similares de medianas y cuartiles a la Figura 2.

Figura 2. Distancia referencial por estrategia referencial (con formativos flexivos)

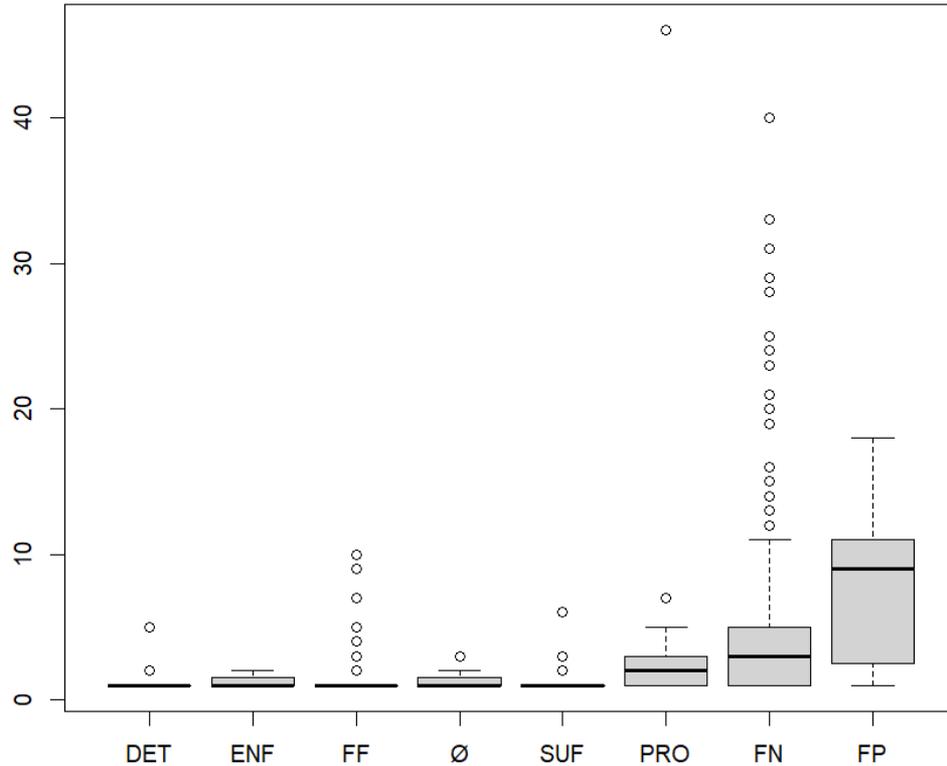
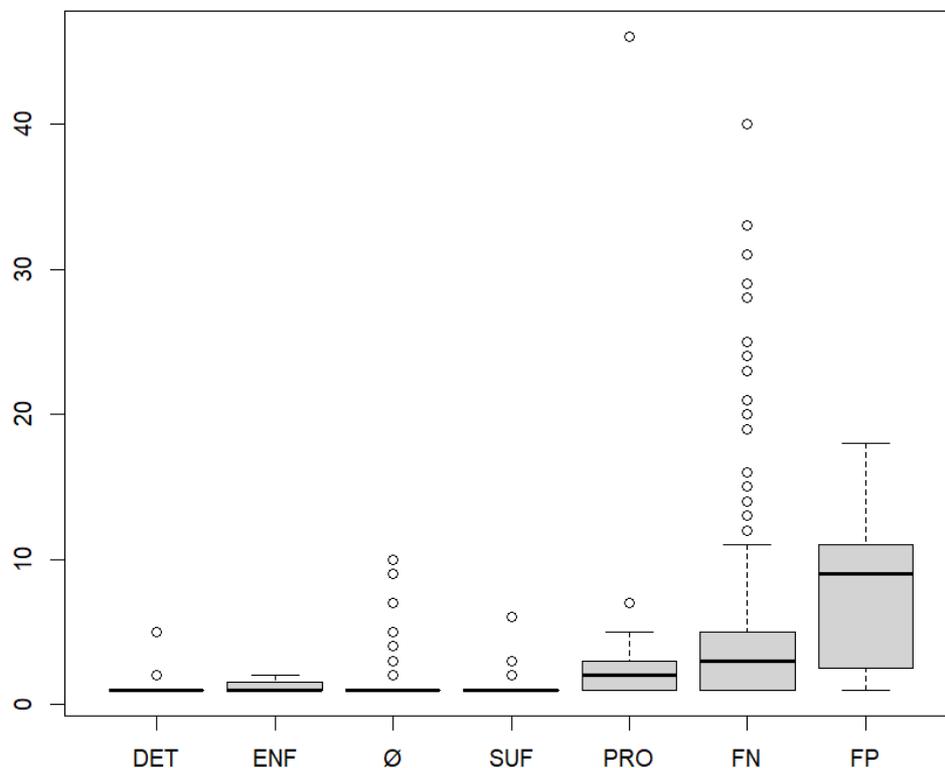


Figura 3. Distancia referencial por estrategia referencial (sin formativos flexivos)

La prueba Kruskal-Wallis ($\chi^2 = 362.07$, $df = 7$, $p < .05$) arrojó los datos resumidos en la Tabla 4. Las etiquetas de la columna “comparaciones” corresponden a cada una de las estrategias referenciales observadas (ENF = pronombre enfático, FF = formativo flexivo, FN = frase nominal, FP = frase preposicional, DET = marca de poseedor, PRO = proforma tónica, SUF = sufijo, Ø = anáfora cero). La columna “diferencia” indica si la diferencia de distancia referencial entre las estrategias referenciales es significativa (“VERDADERO”) o no (“FALSO”).

Las diferencias no significativas (“FALSO”) agrupan, por un lado, a las estrategias reducidas de anáfora cero, sufijo, formativo flexivo y marca de poseedor, y por otro lado a las proformas tónicas (estrategia reducida), las frases nominales y las frases preposicionales (estrategias plenas). Esto puede observarse en el hecho de que solo se presentan diferencias no significativas dentro de cada una de estas dos agrupaciones. Los pronombres enfáticos ($n = 4$) quedan fuera de estas dos agrupaciones, pues no muestran diferencias significativas con las demás estrategias referenciales consideradas. Los datos que combinan los formativos flexivos con la anáfora cero (siguiendo a Hernández-Green 2021: 14) agrupan exactamente las mismas estrategias (Ø-SUF-DET y PRO-FN-FP, y excluyen ENF), y no se muestran aquí. Las agrupaciones Ø-SUF-FF-DET y PRO-FN-FP se confirman en las comparaciones con diferencia significativa (“VERDADERO”): en estas encontramos siempre que las dos estrategias comparadas (columna “comparaciones”) pertenecen a agrupaciones distintas. Además, estas dos agrupaciones son casi idénticas a las observadas por comparación de medianas y cuartiles arriba en la Figura 2.

Tabla 4. Significancia de las diferencias de distancia referencial por estrategia referencial (Kruskal-Wallis, $p < .05$)

<i>comparaciones</i>	<i>diferencia observada</i>	<i>diferencia crítica</i>	<i>diferencia</i>
DET-ENF	35.953125	528.67464	FALSO
DET-FF	14.886828	181.41862	FALSO
DET-Ø	41.421875	203.48696	FALSO
DET-SUF	43.757612	209.27448	FALSO
ENF-FF	50.839953	500.29908	FALSO
ENF-FN	277.973101	501.58399	FALSO
ENF-FP	410.408333	560.97484	FALSO
ENF-Ø	5.46875	508.71741	FALSO
ENF-PRO	264.478448	531.70432	FALSO
ENF-SUF	7.804487	511.05994	FALSO
FF-Ø	56.308703	110.49545	FALSO
FF-SUF	58.64444	120.82262	FALSO
FN-FP	132.435232	263.4311	FALSO
FN-PRO	13.494653	193.42349	FALSO
FP-PRO	145.929885	317.0472	FALSO
Ø-SUF	2.335737	151.96162	FALSO
DET-FN	313.926226	184.93253	VERDADERO
DET-FP	446.361458	311.93961	VERDADERO
DET-PRO	300.431573	255.58369	VERDADERO
FF-FN	328.813055	70.72722	VERDADERO
FF-FP	461.248287	260.97628	VERDADERO
FF-PRO	315.318402	190.06663	VERDADERO
FN-Ø	272.504351	116.17473	VERDADERO
FN-SUF	270.168614	126.03741	VERDADERO
FP-Ø	404.939583	276.77219	VERDADERO
FP-SUF	402.603846	281.05464	VERDADERO

Ø-PRO	259.009698	211.23341	VERDADERO
PRO-SUF	256.673961	216.81424	VERDADERO

Las dos agrupaciones de estrategias referenciales mencionadas en los párrafos anteriores se confirman también mediante el análisis de correspondencias de los datos categóricos ($\chi^2 = 449.33$, $p < .05$). Este análisis se aplicó en la tabla de contingencia mostrada en la Tabla 5. En la Tabla 6 se muestran los datos que combinan los formativos flexivos con la anáfora cero (siguiendo a Hernández-Green 2021: 14; $\chi^2 = 439.06$, $p < .05$).

Tabla 5. Frecuencia de estrategias referenciales vs. distancia referencial (con formativos flexivos)

	1	2	3	4	5	>5
ENF	<u>3</u>	1	0	0	0	0
DET	<u>27</u>	4	0	0	1	0
Ø	<u>72</u>	19	5	0	0	0
SUF	<u>59</u>	13	4	0	0	2
FF	<u>468</u>	48	13	1	2	3
PRO	8	<u>11</u>	5	2	1	2
FN	<u>107</u>	49	36	37	17	70
FP	3	1	1	1	1	<u>8</u>

Tabla 6. Frecuencia de estrategias referenciales vs. distancia referencial (sin formativos flexivos)

	1	2	3	4	5	>5
ENF	<u>3</u>	1	0	0	0	0
DET	<u>27</u>	4	0	0	1	0
Ø	<u>540</u>	67	18	1	2	3
SUF	<u>59</u>	13	4	0	0	2
PRO	8	<u>11</u>	5	2	1	2
FN	<u>107</u>	49	36	37	17	70
FP	3	1	1	1	1	<u>8</u>

Las mayores frecuencias (en negritas y subrayadas) de las estrategias referenciales por encima de la línea punteada se encuentran en la columna de distancia referencial “1”, mientras que las segundas mayores frecuencias (en negritas, sin subrayar) se encuentran en la columna “2”. En contraste, las frecuencias mayores de las estrategias por debajo de la línea punteada tienen distribuciones dispares. Las agrupaciones de estrategias por encima y por debajo de la línea punteada coinciden (si excluimos de la estrategia de pronombre enfático “ENF”),

respectivamente, con las agrupaciones Ø-SUF-FF-DET y PRO-FN-FP de la prueba Kruskal-Wallis mostrada más arriba.

Finalmente, el diagrama mostrado en la Figura 4 muestra los resultados del análisis de correspondencias (Benzécri 1973) aplicado a la tabla de contingencia de la Tabla 5. Los números indican la distancia referencial y las letras (D = marca de poseedor, E = pronombre enfático, F = formativo flexivo, Ø = anáfora cero, S = sufijo, R = proforma tónica, N = frase nominal, P = frase preposicional) refieren a las estrategias referenciales. Una vez más, se observa una diferencia clara entre las estrategias Ø-S-F-D (elipse) y R-N-P (rectángulo). El diagrama correspondiente a la combinación de formativos flexivos y anáfora cero (siguiendo a Hernández-Green 2021: 14) puede verse en la Figura 5.

Figura 4. Análisis de correspondencias entre distancia referencial y estrategias referenciales (con formativos flexivos)

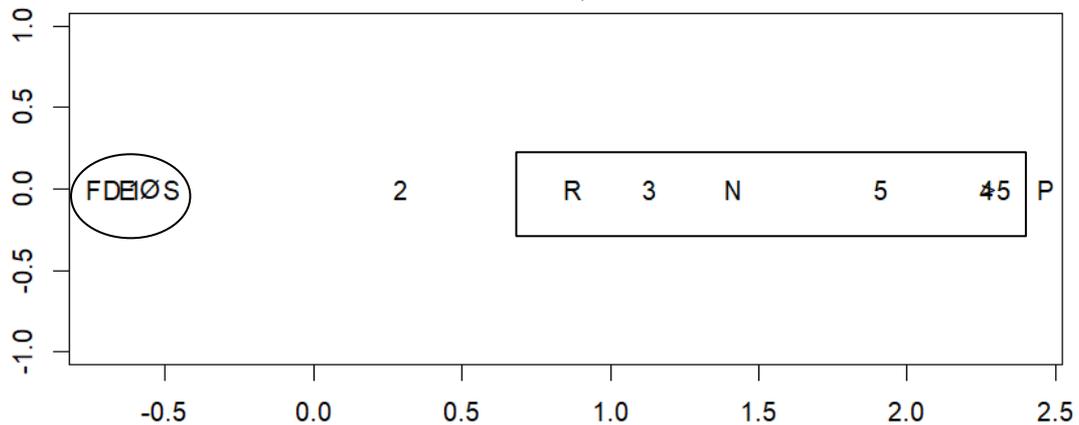
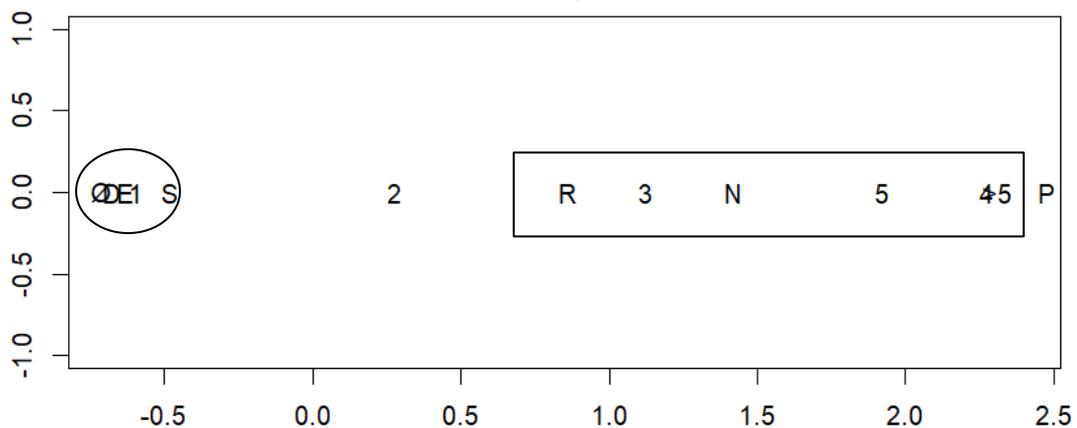


Figura 5. Análisis de correspondencias entre distancia referencial y estrategias referenciales (sin formativos flexivos)



7. Discusión

Los resultados presentados en la sección anterior muestran que existe una correlación clara entre la distancia referencial y la estrategia referencial usada en la mención de los referentes en los textos del corpus analizado. Correlaciones similares basadas en medias de distancia referencial se han encontrado en lenguas como el yagua (Payne 1985), el árabe escrito (Fahkri 1995) y el mandarín escrito (Pu 1989), todas conformes (en general) con la escala de codificación por continuidad o accesibilidad de tópico de Givón (1983: 17). En los datos de la

sección anterior (con formativos flexivos y anáfora cero combinados)¹¹ se observa también esta conformidad, como se muestra en (22). El signo + indica mayor continuidad, o bien, menor distancia referencial, mientras que el signo - indica la situación opuesta.

- (22) + anáfora cero ($\bar{x} = 1.215$)
 afijos y clíticos ($\bar{x} = 1.799$)
 proformas tónicas ($\bar{x} = 3.313$)
 - frases nominales/preposicionales ($\bar{x} = 5.865$)

La categoría de clíticos en (22) solo contiene la proforma locativa =*ni* 'LOC', la cual se ha incluido junto con las proformas tónicas en el análisis mostrado en la sección anterior. A pesar de las diferencias en la naturaleza prosódica de ambos tipos de proformas, las medias de distancia referencial son muy cercanas entre ellas: $\bar{x} = 2.2$ y $\bar{x} = 2.313$, respectivamente.

En contraste con las similitudes entre la proforma locativa =*ni* 'LOC' y las proformas tónicas, dentro del análisis que considera a los formativos flexivos como marcas de 3ª persona (Voigtlander y Echegoyen 1979: 96; Andrews 1993: 6; Palancar 2009: 175; 2011b: 249; 2012; Bartholomew 2010: 503) existen discrepancias en cuanto a distancia referencial entre las formas clíticas. Por un lado, la proforma locativa presenta una media de $\bar{x} = 2.2$, mientras que los formativos flexivos tienen tan solo $\bar{x} = 1.202$. La media de distancia referencial de los formativos flexivos es menor y más cercana a la media de la anáfora cero, la cual es de $\bar{x} = 1.289$. Estos hechos entran en conflicto con el gradiente de continuidad o accesibilidad de tópico de Givón (1983: 17), pues se trata de un sistema de estrategias referenciales con pronombres libres átonos de continuidad menor que los afijos verbales y la anáfora cero. El sistema de estrategias referenciales del hñöhñö se alinea más fielmente a las tendencias tipológicas si se asume, como lo hace Hernández-Green (2021: 14), que los formativos flexivos que se observan en verbos con A/S 3ª persona no es una marca de persona, sino solo de TAM y otras categorías gramaticales. Bajo este análisis, resumido arriba en (22), la distancia referencial de las distintas estrategias referenciales, reducidas y plenas, sigue los patrones observados en las lenguas del mundo.

La presentación de los hechos descritos en el párrafo anterior no puede ni pretende responder a la pregunta sobre si en lenguas otomíes el formativo flexivo de 3ª persona codifica morfológicamente la persona gramatical o no. Tomando en cuenta el carácter tenaz de los formativos flexivos analizados como marcas de 3ª persona, es obligatorio que cada cláusula con predicado verbal contenga por lo menos un formativo flexivo, lo cual reduce forzosamente la distancia referencial entre una y otra mención. La media tan baja de distancia referencial de los formativos flexivos es posiblemente un efecto de la morfología de la predicación en lenguas otomíes más que un efecto del uso de estos morfemas en el discurso. De este modo, el comportamiento de los formativos flexivos como estrategias referenciales no puede tomarse como un criterio para decidir si estos codifican persona gramatical o no.

Finalmente, las proformas enfáticas muestran un comportamiento distinto de las demás estrategias referenciales en términos de distancia referencial. Por un lado, el análisis de correspondencias las alinea con las estrategias reducidas DET-Ø-SUF(-FF). Por otro lado, en contraste, la prueba Kruskal-Wallis muestra que la distancia referencial de proformas enfáticas no presenta diferencias significativas con respecto a la de las otras estrategias referenciales. Estas discrepancias pueden tener, por lo menos, dos posibles explicaciones no necesariamente excluyentes la una de la otra.¹² La primera es que el número de proformas enfáticas en el corpus

¹¹ Cuando no se combinan los formativos flexivos y la anáfora cero, la media de afijos y clíticos es $\bar{x} = 1.6$, y la media de anáfora cero es $\bar{x} = 1.289$.

¹² Agradezco a un dictaminador anónimo por hacerme ver que las dos explicaciones no son incompatibles entre ellas.

es demasiado bajo ($n = 4$) para que el análisis estadístico pueda arrojar resultados comparables con los de las otras estrategias referenciales. La segunda explicación es que el carácter enfático de estas proformas puede estarlas distinguiendo de las otras estrategias referenciales en cuanto al grado de complejidad funcional: además de servir de referencia anafórica, solo las proformas enfáticas tienen una función discursiva adicional (esto es, la de focalizar) con respecto al resto de las estrategias estudiadas en este trabajo.

8. Conclusiones

Con base en una colección de textos escritos en hñöhñö (otomí de Tolimán) y la aplicación de distintos métodos estadísticos, en este trabajo se ha mostrado que la distancia referencial (esto es, el número de cláusulas entre una y otra mención de un referente en particular) se correlaciona con la estrategia referencial utilizada en cada mención. En general, las estrategias referenciales reducidas de marca de poseedor, sufijos de persona, formativos flexivos y anáfora cero tienden a distancias referenciales cercanas a 1. En comparación, la estrategia reducida de proformas libres y las estrategias plenas (FNs y FPs) presentan distancias referenciales mayores.

En contraste con las estrategias referenciales mencionadas en el párrafo anterior, las proformas enfáticas presentan una distribución distinta. Esto podría deberse, por un lado, a que el número de proformas enfáticas en el corpus es muy reducido ($n = 4$), por otro lado, a que su comportamiento esté influido por su función focalizadora, o bien, a una combinación de estas dos posibilidades.

Finalmente, no se observan diferencias significativas de distancia referencial entre la anáfora cero y los formativos flexivos del verbo, los cuales son pronombres libres según la tipología de estrategias referenciales de Kibrik (2011). Si bien esto podría apoyar la postura de que los formativos flexivos no son marcas explícitas de persona (como tampoco lo es la anáfora cero), más investigaciones complementarias serían necesarias para corroborar o rechazar esta hipótesis.

Referencias

- Andrews, Henrietta (1993). *The function of verb prefixes in Southwestern Otomí*. Arlington, TX: SIL International.
- Bartholomew, Doris A. (2010). Notas gramaticales. En Luis Hernández Cruz; Moisés Victoria Torquemada; Donald Sinclair Crawford. *Diccionario del hñähñu (otomí) del Valle del Mezquital, estado de Hidalgo*, pp. 497-516. Ciudad de México: Instituto Lingüístico de Verano.
- Benzécri, Jean-Paul (1973). *L'Analyse des Données. L'Analyse des Correspondances: Vol. II*. París: Dunod.
- de Cárceres, Pedro (1907[1580]). Arte de la Lengua Othomi. *Boletín del Instituto Bibliográfico Mexicano* 6: 38-155.
- Bickel, Balthasar (2010). Grammatical relations typology. En Jae Jung Song. *The Oxford handbook of linguistic typology*, pp. 399-444. Oxford: Oxford University Press.
- Crystal, David (2008). *A dictionary of linguistics and phonetics*. Oxford: Blackwell.
- Ecker, Lawrence (2012[1952]). *Compendio de gramática otomí*. Ciudad de México: UNAM-Instituto de Investigaciones Antropológicas.
- Fahkri, Ahmed (1995). Topic continuity in Arabic narrative discourse. En Mushira Eid (ed.), *Perspectives on Arabic Linguistics: Papers from the Annual Symposium on Arabic Linguistics. Volume VII: Austin, Texas 1993*, pp. 141-155. Austin, TX: John Benjamins. <https://doi.org/10.1075/cilt.130.13fak>
- Givón, T. (1983). Introduction. En T. Givón (ed.), *Topic continuity in discourse: A quantitative cross-language study*, pp. 5-41. Amsterdam/Philadelphian: John Benjamins.

- González Morales, Victoriana; González, Luis A.; Martínez, Luis Mauricio (2019). *Ya Hñö Ma Mengu. Las voces de mis familiares*. Ciudad de México: Del Lirio.
- Hekking, Ewald; Andrés de Jesús, Severiano (1984). *Gramática Otomí*. Querétaro: Centro de Estudios Lingüísticos y Literarios, Universidad Autónoma de Querétaro.
- Hernández-Green, Néstor (2015). *Morfosintaxis verbal del otomí de Acapulco* (Disertación de doctorado en Lingüística Indoamericana). Ciudad de México: Centro de Investigaciones y Estudios Superiores en Antropología Social. Disponible en: <http://repositorio.ciesas.edu.mx/handle/123456789/349>
- Hernández-Green, Néstor (2016). Alineamiento semántico y lexicalización en el sistema de marcación de sujeto en otomí-mazahua. *Signos Lingüísticos* XII (24): 36-67. <http://orcid.org/0000-0001-8553-8245>
- Hernández-Green, Néstor (2018). El sistema aspectual del otomí de Acapulco. *Cuadernos de Lingüística de El Colegio de México* 5(2): 280-334. <https://doi.org/10.24201/clecm.v5i2.119>.
- Hernández-Green, Néstor (2020). Inflectional tone in Hñähñu (Mezquital Otomí). *Amerindia* 42: 1-21.
- Hernández-Green, Néstor (2021). Concordancia de número plural con participantes de 3ª persona en el verbo en ñätho (otomí de San Felipe). *Semas* 2(4): 9-34. Recuperado a partir de <https://semas.uaq.mx/index.php/ojs/article/view/66>
- Hernández-Green, Néstor (2022). Exploración sobre roles gramaticales en otomí de San Felipe Santiago a partir de datos de corpus. *Cuadernos de Lingüística de El Colegio de México* 9: 1-64. <https://doi.org/10.24201/clecm.v9i0.222>
- Hernández-Green, Néstor; Palancar, Enrique L. (2021). Associated motion in the Otomi family. En Antoine Guillaume; Harold Koch (eds.), *Associated Motion*, pp. 527-566. Berlín: Walter de Gruyter. <https://doi.org/10.1515/9783110692099-014>
- Hess, H. Harwood (1968). *The syntactic structure of Mezquital Otomí*. La Haya: Mouton.
- Kibrik, Andrej A. (2011). *Reference in discourse*. Nueva York: Oxford University Press.
- Kruskal, William H.; Wallis, W. Allen (1952). Use of ranks in one-criterion variance analysis. *Journal of the American Statistical Association* 47(260): 583-620.
- Lastra, Yolanda (1992). *El otomí de Toluca*. Ciudad de México: UNAM-Instituto de Investigaciones Antropológicas.
- Lastra, Yolanda (1997). *El otomí de Ixtenco*. Ciudad de México: UNAM-Instituto de Investigaciones Antropológicas.
- Mora-Bustos, Armando (2018). Omisión del proclítico de tiempo, aspecto, modo y persona en mazahua. *Lenguaje* 46(2): 191-219. <https://doi.org/10.25100/lenguaje.v46i2.6580>
- Palancar, Enrique L. (2009). *Gramática y textos del hñähño otomí de San Ildefonso Tultepec, Querétaro*. Ciudad de México: Plaza y Valdés.
- Palancar, Enrique L. (2011a). Revisiting the conjugation classes of Eastern Highlands Otomi. *STUF - Language Typology and Universals* 64(3): 213-236. <https://doi.org/10.1524/stuf.2011.0016>
- Palancar, Enrique L. (2011b). The conjugations of Colonial Otomi. *Transactions of the Philological Society* 109(3): 246-264. <https://doi.org/10.1111/j.1467-968X.2011.01275.x>
- Palancar, Enrique L. (2012). The conjugation classes of Tilapa Otomi: An approach from canonical typology. *Linguistics* 50(4): 783-832. <https://doi.org/10.1515/ling-2012-0025>
- Payne, Thomas E. (1985). Referential distance and discourse structure in Yagua. *Work Papers of the Summer Institute of Linguistics, University of North Dakota Session*, 29(7). Available at: <https://commons.und.edu/sil-work-papers/vol29/iss1/7>
- Pu, Ming-Ming (1989). Topic continuity in written Mandarin discourse. *Annual Meeting of the Berkeley Linguistics Society* 15: 256-267.
- RStudio Team. (2015). *RStudio: Integrated Development Environment for R*. Boston, MA: RStudio, Inc. Obtenido de <http://www.rstudio.com/>.

Voigtlander, Katherine; Echegoyen, Artemisa (1979). *Luces contemporáneas del otomí: Gramática del otomí de la sierra*. Ciudad de México: Instituto Lingüístico de Verano.

CRedit:

RECURSOS: Luis A. González

AGRADECIMIENTOS

Agradezco a Luis A. González, quien amablemente me facilitó el libro de narraciones en que se basa este trabajo. También agradezco a dos dictaminadores anónimos, por sus atinados comentarios. Cualquier error o carencia que se encuentre en este artículo es responsabilidad solo mía.

Recibido: 1/9/2022

Versión corregida: 11/11/2022

Aceptado: 15/11/2022

Publicado: 16/11/2022