



CENTRO DE ESTUDIOS LINGÜÍSTICOS Y LITERARIOS

La gestión del número en los miembros de dos comunidades
bilingües español-otomí

Tesis que para obtener al grado de
Doctora en Lingüística

Presenta

Rosnataly Avelino Sierra

Asesor: Dr. Pedro Martín Butragueño

Ciudad de México, 2022

A mi pueblo, San Andrés Cuexcontlán

Agradecimientos institucionales

- Al Centro de Estudios Lingüísticos y Literarios de El Colegio de México por darme las herramientas necesarias para poder formarme como Doctora en Lingüística.
- Al Consejo Nacional de Ciencia y Tecnología (CONACYT) por otorgarme la beca de doctorado, sin la cual no habría sido posible realizar esta investigación.
- A Aeroméxico por su apoyo para realizar una estancia de investigación en la Universidad de Illinois en Urbana-Champaign, Estados Unidos.
- A la Fundación BBVA-COLMEX por la beca de manutención que me permitió culminar mi tesis doctoral.
- Al Grupo de Investigación “Cambio lingüístico en situaciones de contacto” y al Proyecto de Investigación “El español en contacto con otras lenguas II: variación y cambio lingüístico.”

Agradecimientos personales

- En primer lugar, quisiera agradecer a mis padres por haberme dado la vida, por su amor infinito, por todo lo que han hecho y lo que no, porque eso me ha convertido en el ser humano que hoy soy. Para los dos, sólo tengo profundos sentimientos de agradecimiento y admiración.
- A mi papá, por tomar mi mano cuando el miedo me visita y ayudarme a ser valiente. De las partes más bellas de este proceso fue haber pasado más tiempo juntos y que me hayas compartido tu conocimiento.
- A mi mamá, por ser un ejemplo de bondad, fortaleza y trabajo. Gracias, porque a tu manera siempre me demuestras tu cariño y me cuidas. Tu apoyo ha sido fundamental en este proceso.
- A mi hermano, por ser mi cómplice y enseñarme que no es malo equivocarse, que siempre es posible rectificar el camino.
- A mi Dani, por todo su amor, por ser fuente de inspiración y ejemplo a seguir, por todo su apoyo, cariño y cuidados, porque su amor y compañía me motiva a ser una mejor persona.
- A todos los miembros de la comunidad de San Andrés Cuexcontitlán y Jiquipilco el Viejo, por permitirme entrar a sus hogares, por brindarme su tiempo y compartir conmigo su lengua y un poquito de la vida.
- Al Dr. Pedro Martín Butragueño, por todas sus enseñanzas, por ser bondadoso con su tiempo y compartir conmigo su conocimiento en cada clase y asesoría, por escucharme y leer mi trabajo.

- A los miembros del comité lector: Violeta Vázquez Rojas Maldonado, Alonso Guerrero Galván y Marcela San Giacomo Trinidad, por su atenta lectura y consejos. Gracias por compartirme su conocimiento en semántica, otomí y lenguas en contacto, por haberse tomado el tiempo de escucharme, de leerme, y por sus atinados comentarios que ayudaron a mejorar este trabajo. Por supuesto, cualquier error u omisión es responsabilidad mía.
- A mis profesores del Doctorado en Lingüística, por sus enseñanzas y conocimiento compartido dentro y fuera del salón de clases: las Dras. Graciela Fernández Ruiz, Niktelol Palacios Cuahtecotzi, Esther Herrera Zendejas, Rebeca Barriga Villanueva, Julia Pozas Loyo, María Eugenia Vázquez Laslop y los Drs. Erik Franco Trujillo, Rodrigo Gutiérrez Bravo, Alfonso Medina Urrea. Especialmente y, con mucho cariño, a la Dra. Violeta Vázquez Rojas Maldonado y los Drs. Sergio Borgard Sierra y Pedro Martín Butragueño, por ser ejemplos de vida para mí.
- A Nadi, por ser como una hermana mayor en los estudios de contacto, por compartir conmigo su conocimiento y experiencia, por apoyarme y ayudarme a crecer, por contagiarme su entusiasmo y pasión por el contacto.
- A las Dras. Leonor Orozco y Érika Mendoza por haber contribuido en mi formación y por guiarme en el proceso de admisión al programa de Doctorado en Lingüística.
- A la Dra. Anna Maria Escobar, por haberme recibido en su institución y en su hogar. Por su tiempo, por compartir conmigo su conocimiento y experiencia en los estudios de contacto.
- A Eduardo y Miguel, por su amistad constante, por estar siempre, aún en la distancia.

- A Liz, Jessy y Rubí, por su amistad y cariño, que ha sido fundamental en esta última etapa.
- A David, Kasep, Eli, Lalito, Roge y Naye, por todo su cariño, por ser mi familia en la Ciudad de México.
- A mis amigos del doctorado: Clarita, Alessia, Rafa, Pau, Hugo, Claudia y Paola, por su valiosa amistad y compañía, por todas las horas y risas en el aula, la biblioteca y los jardines del Colegio.
- A Laura, por su atenta escucha y valiosa interpretación, por haber sido un gran apoyo en estos casi cinco años.
- Por último, pero no menos importante, a mis gatitos: T'axi, Louis y P'axi por ser fuente de amor y ternura.

RESUMEN

El objetivo general de este trabajo fue revisar el concepto de bidireccionalidad en la situación de contacto entre el otomí del centro y el español. Para ello, analizamos dos procesos de variación en la marcación de número del español y del otomí en contacto en dos comunidades bilingües *n̄ható*-español, San Andrés Cuexcontitlán y Jiquipilco el Viejo. Estos fenómenos fueron la concordancia plural en el español de contacto y las marcas de número en los nombres patrimoniales y los lexemas nominales del español que han ingresado al otomí.

Los resultados indican que la variación en la concordancia plural de la frase nominal y en el nivel oracional del español hablado en SAC y JV sí está relacionado con el contacto con la lengua otomí. Asimismo, advierten la existencia de diferentes sistemas de concordancia de número en español, relacionados con el tipo de bilingüismo de los hablantes. Por otro lado, la influencia del español en la marcación de número en otomí sólo se hizo evidente al realizar el análisis por grupo lingüístico. Los datos mostraron un aumento en el uso de las marcas de número del otomí, tanto en los nombres patrimoniales como en los préstamos, conforme incrementaba el conocimiento del español de los bilingües y decrecía su dominio de la lengua otomí. Esto debido al debilitamiento de algunas restricciones que determinaban la presencia / ausencia de las marcas de número en la lengua otomí.

Con base en esto, pudimos ver que es posible que el contacto afecte un mismo dominio lingüístico, en este caso la marcación de número en el otomí y el español. El otomí influye en la marcación de número del español, mediante la importación de su patrón de marcación al español de contacto. Por otro lado, el español influye en la marcación de número del otomí, al debilitar algunas restricciones sobre la presencia / ausencia de las marcas de número en otomí.

ÍNDICE GENERAL

Presentación de los datos	xx
Introducción	xxii
Capítulo 1 Marco teórico	1
1.1 El estudio del contacto de lenguas.....	1
1.1.1 El término interferencia.....	1
1.1.2 Los términos interferencia y préstamo	4
1.1.3 El término copia de código.....	9
1.1.4 El término réplica gramatical	11
1.2 El estudio del bilingüismo en los estudios del contacto	13
1.2.1. Definición	13
1.2.2 Clasificación	15
1.3 El estudio del contacto lingüístico en esta investigación	18
1.4 La marcación de número en otomí	25
1.4.1 Características generales de la lengua otomí.....	25
1.4.1.1 Introducción.....	25
1.4.1.2 Filiación lingüística de la lengua otomí.....	28
1.4.1.3. Variantes de la lengua otomí	30
1.4.1.4 El sistema fonológico de la lengua otomí.....	36
1.4.1.5 Características morfosintácticas generales de la lengua otomí	39
1.4.2 El número en el nivel oracional.....	42
1.4.2.1 La estructura del verbo en otomí	42
1.4.2.2 La expresión del número en el verbo.....	46
1.4.3 El número en el nivel frástico.....	52
1.4.3.1 La estructura de la frase nominal.....	52
1.4.3.2 Las marcas de número	54
1.4.3.3 El número en las construcciones posesivas	55
1.4.3.4 Los numerales y los cuantificadores.....	56
1.4.3.5 La expresión del número en los demostrativos	57
1.4.3.6 La expresión del número en los pronombres personales	58

1.4.3.7	La expresión del número en los adjetivos	58
Capítulo 2. Descripción de la metodología y la realidad sociolingüística de las comunidades bilingües otomí-español		
63		
2.1	Descripción de la metodología	63
2.1.1	Población y muestra.....	64
2.1.2	Materiales para la recolección de material lingüístico: la entrevista.....	73
2.1.2.1	Presentación.....	74
2.1.2.2	Historia tradicional y de vida	74
2.1.2.3	Pruebas de habilidades lingüísticas	74
2.1.2.3.1	Prueba de habilidad lingüística en español.....	77
2.1.2.3.2	Prueba de habilidad lingüística en otomí.....	79
2.1.2.3.3	Prueba de habilidad lingüística pasiva en otomí	81
2.1.2.4	Cuestionario sociolingüístico	82
2.1.3	La corpora y algunas herramientas de análisis	88
2.1.3.1	La prueba de Chi-cuadrado de Pearson	91
2.1.3.2	V de Cramer	92
2.1.3.3	La medida de Goodman y Kruskal	92
2.1.3.4	La regresión logística escalonada	94
2.1.3.5	Prueba de inferencia multimodelo	96
2.2	Descripción de la realidad sociolingüística de las comunidades bilingües otomí-español	98
2.2.1	Brevísima historia de las dos comunidades	98
2.2.2	Hablantes de otomí en las dos comunidades	104
2.2.3	Características lingüísticas de los hablantes de otomí en las dos comunidades	107
2.2.4	El uso del otomí y el español en las dos comunidades	114
Capítulo 3. El español de contacto: el caso de la concordancia plural		
135		
3.1	Introducción.....	135
3.2	Antecedentes.....	140
3.2.1	La concordancia de número en variedades de español monolingüe.....	140
3.2.2	La concordancia de número en el español en contacto con sikuani, wayuu y tikuna	143

3.2.3	La concordancia de número en el español en contacto con quechua en Perú ...	144
3.2.4	La concordancia de número en el español en contacto con quechua en Argentina	146
3.2.5	La concordancia de número en el español en contacto con mapuzugun	148
3.2.6	La concordancia de número en el español en contacto con tepehuano del sur .	150
3.2.7	La concordancia de número en el español en contacto con el otomí de Santiago Mexquititlán	151
3.3	Decisiones metodológicas para el análisis de la concordancia plural	154
3.3.1	Conformación del corpus.....	154
3.3.2	Algunas precisiones conceptuales	155
3.3.3	Factores de análisis.....	156
3.3.3.1	Factores lingüísticos generales	157
3.3.3.2	Factores lingüísticos relacionados con la marcación de número en otomí.	158
3.3.3.3	Factores sociales	163
3.4	Modelo general de la concordancia plural en el español en contacto con otomí	167
3.4.1	Modelo general: análisis descriptivo	167
3.4.1.1	Factores lingüísticos generales	168
3.4.1.2	Factores lingüísticos relacionados con el contacto.....	174
3.4.1.3	Factores sociales	177
3.4.2	Modelo general: análisis inferencial.....	185
3.4.3	Discusión del modelo general.....	189
3.5	La concordancia de número en la frase nominal	192
3.5.1	Introducción.....	192
3.5.2	El número en la frase nominal del español y del otomí	193
3.5.3	Resultados de la concordancia de número en la FN	198
3.5.4	Discusión en torno a la concordancia de número en la FN	203
3.5.5	Conclusiones en torno a la concordancia de número en la FN.....	206
3.6	La concordancia de número en el nivel oracional	209
3.6.1	Introducción.....	209
3.6.2	La concordancia de número en español y en otomí en el nivel oracional	209
3.6.3	Resultados de la concordancia plural en el nivel oracional.....	211

3.6.4	Discusión en torno a la concordancia de número en el nivel oracional	216
3.6.5	Conclusiones en torno a la concordancia de número en el nivel oracional	219
3.7	La concordancia de número en San Andrés Cuexcontitlán y Jiquipilco el Viejo: ¿una o dos comunidades?.....	221
3.7.1	La concordancia plural en la frase nominal del español de San Andrés Cuexcontitlán y Jiquipilco el Viejo	222
3.7.2	La concordancia de número en el nivel oracional del español de San Andrés Cuexcontitlán y Jiquipilco el Viejo	226
3.8	La concordancia de número en los tres tipos de bilingüismo: ¿una o dos gramáticas?.....	229
3.8.1	La concordancia de número en la frase nominal según el tipo de bilingüismo.....	231
3.8.2	La concordancia de número en el nivel oracional según el tipo de bilingüismo	235
3.9	Conclusiones del tercer capítulo.....	240
	Capítulo 4. El otomí de contacto: el caso de la marcación de número	245
4.1	Introducción.....	245
4.2	Antecedentes.....	247
4.2.1	El uso del término préstamo en los estudios de contacto	247
4.2.2	El cambio de código monoléxico y su relación con el préstamo léxico.....	252
4.2.3	Las escalas de préstamo léxico y gramatical	256
4.2.4	Los préstamos léxicos en el náhuatl	262
4.2.5	Los préstamos léxicos en el maya yucateco	266
4.2.6	El estudio del préstamo léxico y gramatical en la lengua otomí	269
4.3	Descripción general del corpus.....	286
4.4	La marcación de número en los nombres de las narraciones otomíes.....	297
4.4.1	La distribución de los nombres en las narraciones otomíes	298
4.4.2	Los patrones de marcación de número en los nombres de las narraciones otomíes	301
4.4.3	La distribución de los patrones de marcación de número en los nombres de las narraciones otomíes	309
4.4.4	Análisis descriptivo de la marcación de número en los nombres de las narraciones otomíes	317

4.4.5 Análisis inferencial de la marcación de número en los nombres de las narraciones otomíes	330
4.5 La marcación de número según el origen histórico de los lexemas	335
4.5.1 La marcación de número en los nombres patrimoniales del otomí	335
4.5.2 La marcación de número en los nombres del español que han ingresado al otomí	341
4.5.3 Conclusiones en torno a la marcación de número según el origen histórico de los lexemas	347
4.6 Análisis en tiempo aparente: el tipo de bilingüismo.....	350
4.6.1 El uso de las marcas de número en los nombres patrimoniales según el tipo de bilingüismo	351
4.6.2 El uso de las marcas de número en los préstamos según el tipo de bilingüismo	364
4.7 Análisis en tiempo aparente: índice de integración léxica	376
4.7.1 La construcción del índice de integración léxica	377
4.7.2 Clasificación categórica del léxico	381
4.7.3 El uso de las marcas de número ra y ya a través del índice de integración léxica	382
4.7.4 El uso de las marcas de número ra y ya a través de la clasificación categórica del léxico	384
4.7.5 El efecto de la frecuencia léxica en el uso de ra y ya	389
4.8 Conclusiones del cuarto capítulo	392
Capítulo 5. Conclusiones generales.....	405
Apéndice capítulo 2	434
Apéndice 2a. Cuestionario de habilidades lingüísticas en español	434
Apéndice 2b Cuestionario de habilidades lingüísticas en otomí	439
Apéndice 2c. Cuestionario de habilidades lingüísticas pasivas en otomí	442
Apéndice 2d. Cuestionario sociolingüístico	444
Apéndice capítulo 3	455

Apéndice 3a. Distribución de los patrones de marcación en la FN de acuerdo con los factores lingüísticos y extralingüísticos	455
Apéndice 3b. Resultados de la medición Goodman y Kruskal. Modelo frase nominal	457
Apéndice 3c. La concordancia de número en el nivel oracional de acuerdo con los factores lingüísticos y extralingüísticos	458
Apéndice 3d. Resultados de la medición Goodman y Kruskal. Modelo oracional	460
Apéndice 3e. La distribución de los tipos de marcación según los factores lingüísticos y extralingüísticos por comunidad.....	461
Apéndice 3f. Resultados de la medición Goodman y Kruskal. Modelo frástico para JV y SAC	463
Apéndice 3g. Distribución de la concordancia plural en el nivel oracional de acuerdo con los factores lingüísticos y extralingüísticos por comunidad.....	464
Apéndice 3h. Resultados de la medición Goodman y Kruskal. Modelo oracional para JV y SAC	466
Apéndice 3i. Resultados de la medición Goodman y Kruskal. Modelo frástico por tipo de bilingüismo: BO, BS, BE	467
Apéndice 3j. Resultados de la prueba multimodelo. Modelo frástico por tipo de bilingüismo	468
Apéndice 3k. Resultados de la medición Goodman y Kruskal. Modelo oracional por tipo de bilingüismo: BO, BS, BE.	469
Apéndice 3l. Resultados de la prueba multimodelo. Modelo oracional por tipo de bilingüismo	470
Apéndice capítulo 4	471
Apéndice 4a. Distribución de los patrones de marcación en los nombres patrimoniales según las variables lingüísticas y sociolingüísticas	471
Apéndice 4b. Resultados de la medición Goodman y Kruskal. Nombres patrimoniales...	473
Apéndice 4c. Distribución de los patrones de marcación en los préstamos según los factores lingüísticos y sociales	474
Apéndice 4d. Resultados de la medición Goodman y Kruskal. Préstamos nominales	476
Apéndice 4e. Resultados de la prueba multimodelo. Nombres patrimoniales y Préstamos	477

Apéndice 4f. Resultados de la medición Goodman y Kruskal. Índice de integración léxica	478
Apéndice 4g. Resultados de la medición Goodman y Kruskal. Clasificación categórica del léxico	479

ÍNDICE DE FIGURAS

Figura 1.1 Resultados del contacto lingüístico según Weinreich (1953)	2
Figura 1.2 Resultados del contacto lingüístico según Thomason (2001: 60).....	5
Figura 1.3 Copia global y copia selectiva adaptado de Johanson (2002).....	10
Figura 1.4 Escala de bilingüismo (adaptado de Palacios 2005: 87)	16
Figura 1.5 Marco conceptual de esta investigación.....	19
Figura 1.6 Escala de bilingüismo	22
Figura 1.7 Tronco lingüístico otomangue.....	29
Figura 1.8 Áreas geográficas del otomí, adaptado de Palancar (2013)	36
Figura 1.9 Sistema de sujeto escindido, adaptado de (Dixon, 1994: 72)	41
Figura 1.10 Sistema de objeto primario, adaptado de (Dryer, 1986: 814)	41
Figura 2.1 Ubicación de San Andrés Cuexcontitlán y Jiquipilco el Viejo.....	65
Figura 2.2 Red de colaboradores en San Andrés Cuexcontitlán	68
Figura 2.3 Red de colaboradores en Jiquipilco el Viejo.....	71
Figura 2.4 Imágenes de las pruebas de competencia.....	81
Figura 2.5 Muestra de los resultados de la medida de Goodman y Kruskal (R. Pearson, 2020)	94
Figura 2.6 Muestra de los resultados de la prueba de inferencia multimodelo	97
Figura 2.7 Lienzo de Jucutacato	100
Figura 2.8 Batallas acaecidas en Xiquipilco (Códice Telleriano-Remensis, f. 33v, 37v) ..	102
Figura 2.9 Porcentaje de hablantes de otomí en SAC y JV	105
Figura 2.10 Resultados de las pruebas de habilidades lingüísticas en SAC y JV	109
Figura 2.11 Resultados de la prueba de habilidad lingüística pasiva en otomí en SAC y JV	111
Figura 2.12 Clasificación de bilingüismo de los colaboradores de SAC	112

Figura 2.13 Clasificación de bilingüismo de los colaboradores de JV.....	113
Figura 2.14 Uso de las lenguas con diferentes interlocutores. Grupos lingüísticos de SAC	117
Figura 2.15 Uso de las lenguas con diferentes interlocutores. Grupos lingüísticos de JV. 120	
Figura 2.16 Uso de las lenguas en diferentes dominios sociales. Grupos lingüísticos de SAC	125
Figura 2.17 Uso de las lenguas en diferentes dominios sociales. Grupos lingüísticos de JV	131
Figura 3.1 Resultados de la medición Goodman y Kruskal. Modelo completo.....	187
Figura 3.2 Tipos de marcación de número	197
Figura 3.3 Distribución de los tipos de marcación según el grado de bilingüismo.....	206
Figura 3.4 Promedio de concordancia por grupo lingüístico	231
Figura 3.5 Resultados de la prueba inferencia multimodelo para la alternancia en los tipos de marcación de número en la FN por tipo de hablante.....	234
Figura 3.6 Resultados de la prueba inferencia multimodelo para la alternancia en los tipos de marcación de número en el nivel oracional por tipo de hablante	238
Figura 4.1 Comparaciones en el discurso bilingüe (adaptado de Poplack y Meechan, 1998: 131).....	255
Figura 4.2 Escala de prestabilidad de Whitney (1881: 19-20)	258
Figura 4.3 Escala de prestabilidad de Haugen (1950: 224).....	258
Figura 4.4 Escala de prestabilidad de Muysken (1981).....	259
Figura 4.5 Escala de prestabilidad de Matras (2008: 61)	260
Figura 4.6 Jerarquía de los principales rasgos semánticos de los nombres	311
Figura 4.7 Resultados de la medición Goodman y Kruskal. Nombres en las narraciones otomías	331

Figura 4.8 Resultados de la prueba inferencia multimodelo para la alternancia en los tipos de marcación de número por origen histórico	348
Figura 4.9 Pauta para el análisis por grupo lingüístico.....	351
Figura 4.10 Resultados de la prueba de inferencia multimodelo del uso de las marcas de número en los nombres patrimoniales de los tres grupos lingüísticos	353
Figura 4.11 Resultados de la prueba inferencia multimodelo del uso de las marcas de número en los préstamos de los tres grupos lingüísticos	365
Figura 4.12 Estratificación del léxico. Adaptado de (Ito & Mester, 1995: 182)	377
Figura 4.13 Evaluación del índice léxico	382
Figura 4.14 La marcación de número en los tres grupos léxicos	387
Figura 4.15 Estratificación por frecuencia	390
Figura 4.16 El efecto de la frecuencia en la marcación de número de los nombres patrimoniales y los préstamos del español	391

ÍNDICE DE TABLAS

Tabla 1.1 Número de hablantes por entidad federativa	30
Tabla 1.2 Grupos otomíes según Soustelle (1993: 217-223)	31
Tabla 1.3 Agrupaciones otomíes según su inteligibilidad (Egland 1978: 46).....	33
Tabla 1.4 Variantes del otomí según el INALI (2008: 41-57)	34
Tabla 1.5 Sistema consonántico del otomí de SAC, adaptado de Lastra (1992: 16-17)	37
Tabla 1.6 Sistema vocálico del otomí de SAC, adaptado de Lastra (1992: 16-17).....	38
Tabla 1.7 Sistemas de marcación en el dependiente y en el núcleo (Nichols, 1986: 59-64)	40
Tabla 1.8 Plantilla verbal del otomí de SAC, adaptado de Hernández Green, 2015: 29)	42
Tabla 1.9 Conjugación del verbo ver en el otomí de SAC, adaptado de Lastra (1992: 22) .	47
Tabla 1.10 Las marcas de número posibles en el sujeto y objeto primario del otomí de SAC	52
Tabla 1.11 Posición de los determinantes de la FN en el otomí de SAC	53
Tabla 1.12 Paradigma de demostrativos en el otomí de SAC, adaptado de (Lastra, 1992: 20)	58
Tabla 1.13 Pronombres personales en el otomí de SAC, adaptado de (Lastra, 1992: 19) ...	58
Tabla 1.14 Conjugación del concepto de propiedad ‘alto’ en el otomí de SAC, adaptado de (Lastra, 1992: 21)	61
Tabla 2.1 Preguntas sociodemográficas	83
Tabla 2.2 Preguntas socioeconómicas	84
Tabla 2.3 Preguntas sobre adquisición y competencia lingüística	85
Tabla 2.4 Preguntas sobre dominios lingüísticos	86
Tabla 2.5 Actitudes lingüísticas	87
Tabla 2.6 Preguntas sobre variación lingüística	88
Tabla 2.7 Ejemplo de una tabla de distribución de frecuencias	90

Tabla 2.8 Ejemplo de los resultados de la regresión logística escalonada	95
Tabla 2.9 Ejemplo de los resultados de la prueba de inferencia multimodelo	98
Tabla 3.1 Variables de análisis	166
Tabla 3.2. Concordancia de número en el corpus.....	167
Tabla 3.3 Distribución de los casos de concordancia y discordancia por variables lingüísticas generales	170
Tabla 3.4 Distribución de los casos de concordancia y discordancia por variables lingüísticas asociadas al contacto	175
Tabla 3.5 Distribución de los casos de concordancia y discordancia por variables sociales	178
Tabla 3.6 Resultados de la regresión logística escalonada. Modelo completo.....	188
Tabla 3.7 Marcación de número en la FN del español y el otomí.....	194
Tabla 3.8 La marcación de número en diferentes estructuras nominal del español y el otomí	195
Tabla 3.9 Concordancia de número en la frase nominal	198
Tabla 3.10 Resultados de la regresión logística escalonada de la variación en el tipo de marcación en la FN.....	201
Tabla 3.11 Características de la concordancia de número en el nivel oracional del español y el otomí.....	210
Tabla 3.12 Resultados de la regresión logística escalonada de la concordancia oracional	213
Tabla 3.13 Factores que determinan la concordancia de número en el nivel frástico y oracional	217
Tabla 3.14 La distribución de los tipos de marcación en SAC y JV	222
Tabla 3.15 Resultados de la regresión logística escalonada de la concordancia de número en la FN por comunidad	224

Tabla 3.16 Resultados de la regresión logística escalonada sobre la variación en la concordancia de número en el nivel oracional por comunidad	228
Tabla 4.1 Tipos de transferencia de acuerdo con Heine y Kuteva (2005) y Weinreich (1953)	250
Tabla 4.2 Identificación del cambio de código de acuerdo con el tipo de integración a la lengua base (adaptado de Poplack 1980: 584).....	253
Tabla 4.3 Porcentaje de préstamos léxicos y gramaticales en las lenguas otomí, quichua y guaraní (adaptado de Hekking & Bakker, 2007c: 444).....	282
Tabla 4.4 Porcentaje de préstamos gramaticales en las lenguas otomí, quichua y guaraní (adaptado de Hekking & Bakker, 2007c: 447).....	284
Tabla 4.5 Categoría de los préstamos del español en SAC y JV.....	287
Tabla 4.6 Adaptación fonológica de los préstamos nominales en SAC y JV	289
Tabla 4.7 Porcentaje de palabras de origen otomí y español.....	298
Tabla 4.8 Porcentaje de nombres de origen otomí y español de acuerdo con las variables sociolingüísticas	300
Tabla 4.9 Frecuencia de los patrones de marcación de número en los nombres de las narraciones otomíes	308
Tabla 4.10 Categorías semánticas de una FN (adaptado de Jackendoff, 1991 apud Corbett, 2001: 80).....	313
Tabla 4.11 Factores de análisis y predicciones.....	316
Tabla 4.12 Distribución de los patrones de marcación de acuerdo con los factores lingüísticos	318
Tabla 4.13 Distribución de los patrones de marcación de acuerdo con los factores sociolingüísticos	326
Tabla 4.14 Resultados de la regresión logística escalonada. Nombres en las narraciones otomíes	332
Tabla 4.15 Distribución de los patrones de marcación en los nombres patrimoniales.....	335

Tabla 4.16 Resultados de la regresión logística escalonada. Nombres patrimoniales	337
Tabla 4.17 Los patrones de marcación de número en los préstamos nominales del español	341
Tabla 4.18 Resultados de la regresión logística escalonada. Préstamos	343
Tabla 4.19 Uso de las marcas de número en los nombres patrimoniales. Tipo de bilingüismo	352
Tabla 4.20 Uso de las marcas de número en los préstamos. Tipo de bilingüismo	364
Tabla 4.21 Factores para construir el índice de integración léxica	380
Tabla 4.22 Resultados de la regresión logística escalonada. Índice léxico	383
Tabla 4.23 Resultados de la regresión logística escalonada. Grupos léxicos: clasificación categórica.....	385

PRESENTACIÓN DE LOS DATOS

Los ejemplos de esta investigación siguen el siguiente formato. En la primera línea se presenta la transcripción ortográfica, de acuerdo con las normas de escritura de la Lengua Hñähñu (INALI 2014). En la segunda línea se proporciona la transcripción fonológica del ejemplo con separación morfémica (el signo – indica afixo y el signo = clítico). La tercera línea corresponde a la glosa según las reglas de Leipzig (Comrie *et al.* 2008). Por último, en la cuarta línea, se ofrece la traducción libre entre comillas simples. A continuación, se muestra las abreviaturas y las equivalencias entre las grafías de la norma de escritura de la Lengua Hñähñu (INALI 2014) y el Alfabeto Fonético Internacional (AFI).

Abreviaturas

1	1a persona	INTENS	intensificador
2	2a persona	INTERR	interrogativo
3	3a persona	LIM	limitativo
CIT	citativo	LOC	preposición locativa
CONT	continuativo	NEG	negación
COP	cópula	OBJ	objeto
DEF	definido	PL	Plural
DEM	demonstrativo	POS	poseedor
DIM	diminutivo	PRFT	perfecto
DU	dual	PRS	presente
ENFT	enfático	PRON	pronombre
EXCL	exclusivo	PSD	pasado
FUT	futuro	REFL	reflexivo
IMP	imperativo	S	sujeto
IMPF	imperfecto	SG	singular
INCL	inclusivo	SIM	simultativo
INDEF	indefinido		

Equivalencias de las grafías de la norma de escritura de la Lengua Hñähñu (INALI 2014) y el

Alfabeto Fonético Internacional (AFI)

Grafía	Fonema	Grafía	Fonema	Grafía	Fonema	Grafía	Fonema	
B	b	[b]	M	m	[m]	Y	y	[j]
CH	ch	[tʃ]	N	n	[n]	W	w	[w]
D	d	[d]	Ñ	ñ	[ɲ]	Z	z	[z]
F	f	[f]	P	p	[p]	‘	‘	[ʔ]
G	g	[g]	R	r	[r]	A	a	[a]
H	h	[h]	S	s	[s]	Ä	ä	[ã]
J	j	[x]	T	t	[t]	À	à	[ɔ]
K	k	[k]	TH	th	[tʰ]	E	e	[e]
KJ	kj	[kʰ]	TS	ts	[ts]	Ë	ë	[ɛ̃]
L	l	[L]	X	x	[ʃ]	Ē	ē	[ɛ̄]
						I	i	[i]
						İ	ï	[ĩ]
						O	o	[o]
						Ö	ö	[õ]
						Ū	ū	[ɹ]
						U	u	[u]
						Ü	ü	[ũ]

INTRODUCCIÓN

Algunos autores como Granda (1996) y (Palacios Alcaine, 2011) proponen repensar la distinción clásica entre préstamo e interferencia, que se centra en la influencia unidireccional del conocimiento de la lengua nativa o materna en el uso de una segunda lengua. Y plantean un análisis de los fenómenos de contacto como procesos dinámicos y bidireccionales, en donde las lenguas en contacto se influyen mutuamente, en los diferentes niveles de la lengua. En ese sentido, apelan a un análisis de los procesos de variación y cambio que se producen en un mismo dominio lingüístico en las lenguas implicadas en una situación de contacto.

Con respecto a la situación de contacto entre el otomí y el español, contamos con un número importante de investigaciones que se podrían clasificar en dos tipos. Por un lado, están los trabajos que se centran en la influencia del otomí en el español de los bilingües (Avelino Sierra, 2017; Guerrero, 2006; Guerrero Galván & San Giacomo Trinidad, 2014; Lizárraga Navarro, 2014), a partir del análisis de fenómenos diversos, como la concordancia de género y/o número o del sistema pronominal átono de 3ª persona. En el otro extremo, tenemos algunos estudios sobre los préstamos léxicos y gramaticales del español en el otomí (Gómez Rendón, 2008; Hekking & Bakker, 2010), pero todavía son pocos los trabajos que se han dedicado a estudiar un mismo fenómeno en las dos lenguas de contacto. En este contexto, vale la pena preguntarse si el contacto puede afectar un mismo dominio lingüístico en el otomí y el español de forma simultánea.

Con esta pregunta en mente, el objetivo general de este trabajo será revisar el concepto de bidireccionalidad en una situación de contacto en México, específicamente en el contacto entre el otomí del centro y el español. Para ello, analizaré dos procesos de variación en la marcación de número de las frases nominales del español y del otomí en contacto en dos comunidades bilingües *ñható*-español, San Andrés Cuexcontitlán y Jiquipilco el Viejo. Estos

fenómenos son las relaciones de concordancia en el español de contacto y las marcas de número en los nombres patrimoniales y los lexemas nominales del español que han ingresado al otomí.

Decidí trabajar con fenómenos variables asociados al número por tres razones. Primero, porque los sistemas de marcación de número son interesantes en sí mismos, debido a su complejidad semántica y sintáctica, y a la relación que guardan con otros fenómenos lingüísticos (Corbett, 2000). En segundo lugar, porque los sistemas de marcación de número han demostrado ser un área propicia para la manifestación de procesos de variación y cambio en situaciones de contacto (Palacios Alcaine, 2011), como los que se documentan en el español andino (Martínez, 2017), el pidgin nigeriano (Tagliamonte et al., 1997) o el maya yucateco (Pfeiler, 2009; Uth & Gutiérrez Bravo, 2018), por mencionar algunos. En tercer lugar, resolví estudiar estos fenómenos porque pensamos que podrían ser buenos indicadores del efecto que el contacto ha tenido en estas lenguas.

De manera general, busco determinar si la variación observada en estos dos fenómenos tiene su origen en el contacto entre el español y el otomí, y cuál fue el mecanismo que los generó. Asimismo, pretendo definir su distribución y estabilidad en los individuos y en las comunidades. Esta investigación consta de cuatro capítulos: “Marco teórico”, “Metodología”, “El español de contacto: el caso de la concordancia plural”, “El otomí de contacto: el caso de la marcación de número”.

En el primer capítulo reviso algunas de las obras y propuestas más importantes en torno al estudio del contacto lingüístico. En *Language in Contact* de Uriel Weinreich (1953), examino el concepto de *interferencia* y sus diferentes tipos. Más adelante, continúo con la descripción de los términos *interferencia*, *préstamo* y *cambio inducido por contacto*, propuestos por Thomason y Kaufman (1988) y Thomason (2001). Posteriormente, explico

en qué consiste el modelo de *copia de código* propuesto por Johanson (2002). Por último, expongo algunos de los conceptos fundamentales de la propuesta de Heine y Kuteva (2005), específicamente la noción de *réplica gramatical*. En la sección §1.2, examino algunos conceptos y clasificaciones de bilingüismo. Después de la revisión de estos trabajos, dedicaré el apartado §1.3 a describir el marco conceptual que guía esta investigación. El capítulo culmina con la descripción de la marcación de número de la lengua otomí en el nivel frástico y oracional (§1.4).

En el segundo capítulo expongo la metodología general de este trabajo. En la primera parte, describo el tipo de muestreo que utilicé para conformar los grupos de colaboradores, así como las herramientas de elicitación. Posteriormente, ofrezco una breve descripción de la situación sociolingüística del otomí y el español en San Andrés Cuexcontitlán y Jiquipilco el Viejo, a partir de la descripción de diferentes grupos de monolingües y bilingües otomí-español, sus características lingüísticas y el uso que hacen de sus lenguas.

El tercer capítulo se centra en el español de contacto, a partir del análisis de la concordancia plural en el nivel frástico y oracional. En la primera parte reviso algunos trabajos sobre la variación en la concordancia de número en español, tanto en variedades de español sin contacto (Soler, 2012) como en el español en contacto con sikuni (Ramírez Cruz, 2009), quechua (Escobar, 2000; Martínez 2001, 2012), mapuzugun (Olate *et al.* 2019), tepehuano del sur (Torres Sánchez, 2018) y otomí de Santiago Mexquititlán (Guerrero, 2006). Con base en estos, establezco las bases para un análisis general en donde considero la concordancia plural en la frase nominal y en el nivel oracional como un solo fenómeno variable. Posteriormente, planteo un análisis de la concordancia plural en el nivel frástico y oracional de manera individual. En el apartado §3.6, analizo el comportamiento de la concordancia plural en la frase nominal y en el nivel oracional en cada comunidad. En la

siguiente sección, §3.7, planteo un análisis de la marcación plural en los dos niveles por tipo de bilingüismo. Finalmente, en §3.8, presento las conclusiones del capítulo.

El cuarto capítulo tiene como objeto de estudio el otomí de contacto, específicamente, se centra en el análisis de la marcación de número en los nombres del otomí (patrimoniales y préstamos). En primer lugar, reviso qué elementos morfológicos emplean los bilingües para marcar el número y realizo un análisis general para determinar qué factores lingüísticos y sociolingüísticos inciden en esto. En el apartado §4.5, se analiza la marcación de número según el origen histórico de los nombres: patrimonial *versus* préstamo. Más adelante, planteo un análisis de la marcación de número en los nombres patrimoniales y los préstamos por tipo de bilingüismo. En la sección §4.7, presento un análisis a partir de un índice de integración léxica y una clasificación categórica del léxico. Finalmente, en §4.8, se exponen las conclusiones de este capítulo.

Por último, en el capítulo §5, proporciono las conclusiones de esta investigación y se discuten con base en el concepto de bidireccionalidad. Además de estos capítulos, al final comparto material extra a manera de apéndices, en donde muestro los instrumentos empleados en la recolección de material lingüístico que se describieron en el segundo capítulo. Asimismo, se podrán encontrar algunas tablas y figuras que, aunque no son centrales en los análisis de los capítulos 3 y 4, sí sirvieron de base para realizarlos.

CAPÍTULO 1 MARCO TEÓRICO

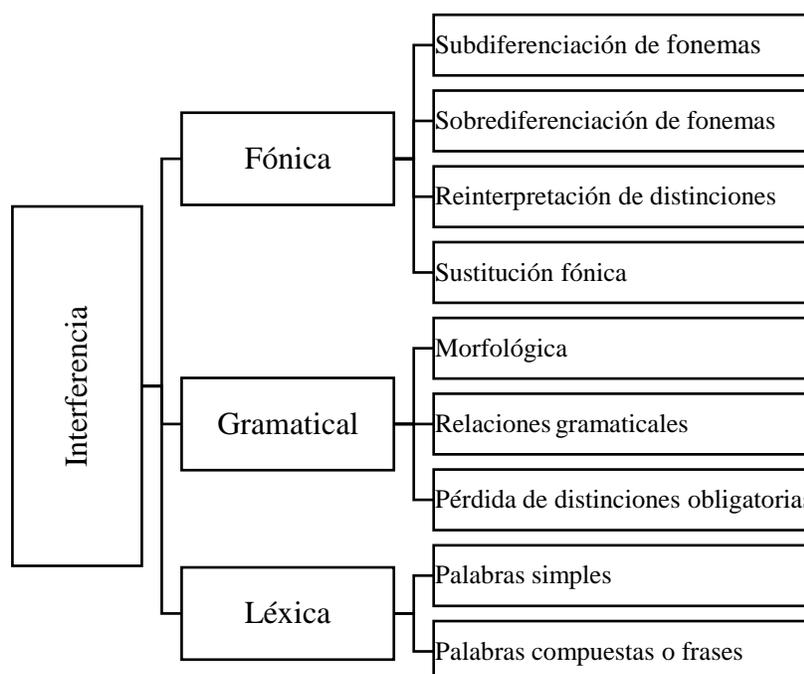
1.1 EL ESTUDIO DEL CONTACTO DE LENGUAS

En este capítulo reviso algunas de las obras más importantes en torno al estudio del contacto lingüístico. En *Language in Contact* de Uriel Weinreich, (1953), examino el concepto de *interferencia* y sus diferentes tipos. Más adelante, continúo con la descripción de los términos *interferencia*, *préstamo* y *cambio inducido por contacto*, propuestos por Thomason y Kaufman (1988) y Thomason (2001). Posteriormente, explico en qué consiste el modelo de *copia de código* propuesto por Johanson (2002). Por último, expongo algunos de los conceptos fundamentales de la propuesta de Heine y Kuteva (2005), específicamente la noción de *réplica gramatical*. En la sección §1.2, presento algunos conceptos y clasificaciones de bilingüismo. Después de la revisión de estos trabajos, dedicaré el apartado §1.3 a describir el marco conceptual que guía esta investigación. El capítulo culmina con la descripción de la marcación de número de la lengua otomí en el nivel frástico y oracional (§1.4).

1.1.1 El término interferencia

La primera obra que revisaré es *Languages in Contact* de Uriel Weinreich (1953), quien define el término *interferencia* como aquellos casos de desviación con respecto a las normas de cualquiera de las dos lenguas que ocurren en el habla de los individuos bilingües, como resultado de su familiaridad con más de una lengua (Weinreich, 1953: 1). Las interferencias pueden presentarse en el nivel fónico, morfológico, sintáctico y léxico, como se muestra en seguida.

FIGURA 1.1 *Resultados del contacto lingüístico según Weinreich (1953)*



La interferencia fónica surge cuando el bilingüe percibe y reproduce un fonema de su sistema lingüístico secundario a partir de las reglas fonéticas de su sistema lingüístico primario. Dentro de esta categoría, distingue cuatro tipos: fusión o subdiferenciación fonémica, ultradiferenciación, reinterpretación y sustitución. En el primero, se da la fusión de dos sonidos del sistema secundario en un solo fonema, debido a que en el sistema primario estos sonidos no tienen valor distintivo (Weinreich, 1953: 18). Por ejemplo, en inglés se distinguen los fonemas /ʃ/ y /tʃ/, pero en el español esos fonemas son alófonos [ʃ] [tʃ], por lo que los hablantes de español que están en proceso de adquisición del inglés suelen fusionar /ʃ/ y /tʃ/ en /tʃ/, como en el español (Silva-Corvalán, 2001: 281).

La ultradiferenciación se presenta cuando se impone una distinción fonémica del sistema primario en el secundario, que no hace esa diferenciación (Weinreich, 1953: 18). Por ejemplo, en el español [b] y [v] son alófonos y en inglés son fonemas /v, b/, por lo que los

bilingües cuya lengua dominante es el inglés suelen hacer la distinción de /b/ y /v/ en el español (Silva-Corvalán, 2001: 218). En la reinterpretación se diferencian fonemas del sistema secundario por medio de rasgos que son distintivos en el sistema primario, pero no en el secundario (Weinreich, 1953: 18-19). Por ejemplo, la palabra para ‘muchos’ en suizo alemán es [‘fil:i], fonémicamente /’fili/ pero el retorromano lo oye como /’filli/ ya que lo largo de la [i:] que en alemán suizo es redundante con su posición después de la vocal corta, se interpreta como distintivo mientras que a la brevedad, no se le presta atención, ya que la cantidad vocálica no es relevante en retorromano (Lastra, 1992: 177). Finalmente, la sustitución ocurre cuando la realización en el sistema secundario de un fonema acotado idénticamente en el sistema primario se sustituye por la realización de este fonema en el sistema primario (19). Un ejemplo de esto es lo que sucede con las oclusivas sordas /p, t, k/, que en español e inglés se definen bajo los mismos factores. Sin embargo, en inglés los alófonos más representativos son los aspirados [p^h, t^h, k^h], que no son comunes en el español, por lo que el hablante de español que está aprendiendo inglés suele sustituir [p^h, t^h, k^h] por /p, t, k/ (Silva-Corvalán, 2001: 281).

La interferencia gramatical afecta la morfología y las relaciones gramaticales.¹ En el nivel morfológico se presentan tres tipos de interferencia. La primera consiste en usar los morfemas de una lengua A en el habla de una lengua B. En el segundo tipo, se aplica una relación gramatical de la lengua A a los morfemas de la lengua B, o se descuida una relación de B porque no tiene un equivalente en A. La tercera posibilidad es la extensión o reducción de las funciones de un morfema de la lengua B, debido a la identificación de ese morfema B con un morfema de la lengua A (Weinreich, 1953: 30).

¹ El término *morfema* refiere segmentos de enunciados y el de *relación gramatical* comprende el orden y la concordancia, así como la modulación del acento y la entonación (Weinreich, 1953: 29).

Las relaciones gramaticales pueden ser objeto de interferencia en tres formas. La primera de ellas consiste en replicar la relación gramatical de una lengua en otra, donde transmite un significado no intencionado. En el segundo tipo, la reproducción de la relación gramatical de una lengua violenta el patrón de la otra lengua. La tercera forma es la imposición de una relación gramatical de un sistema en otra lengua, en la que no existe una relación obligatoria equivalente (Weinreich, 1953: 37-38).

Las interferencias léxicas ocurren cuando el vocabulario de una lengua afecta el léxico de otra, y pueden ser simples o compuestas. La primera categoría comprende tres tipos. En el primero se transfiere una palabra de una lengua a otra. El segundo tipo ocurre cuando la palabra de una lengua extiende su uso con base en el modelo de otra lengua. En el tercer tipo, la forma fónica de un signo cambia a partir de un cognado de la lengua con la que está en contacto, sin afectar su contenido (Weinreich, 1953: 47, 48, 50).

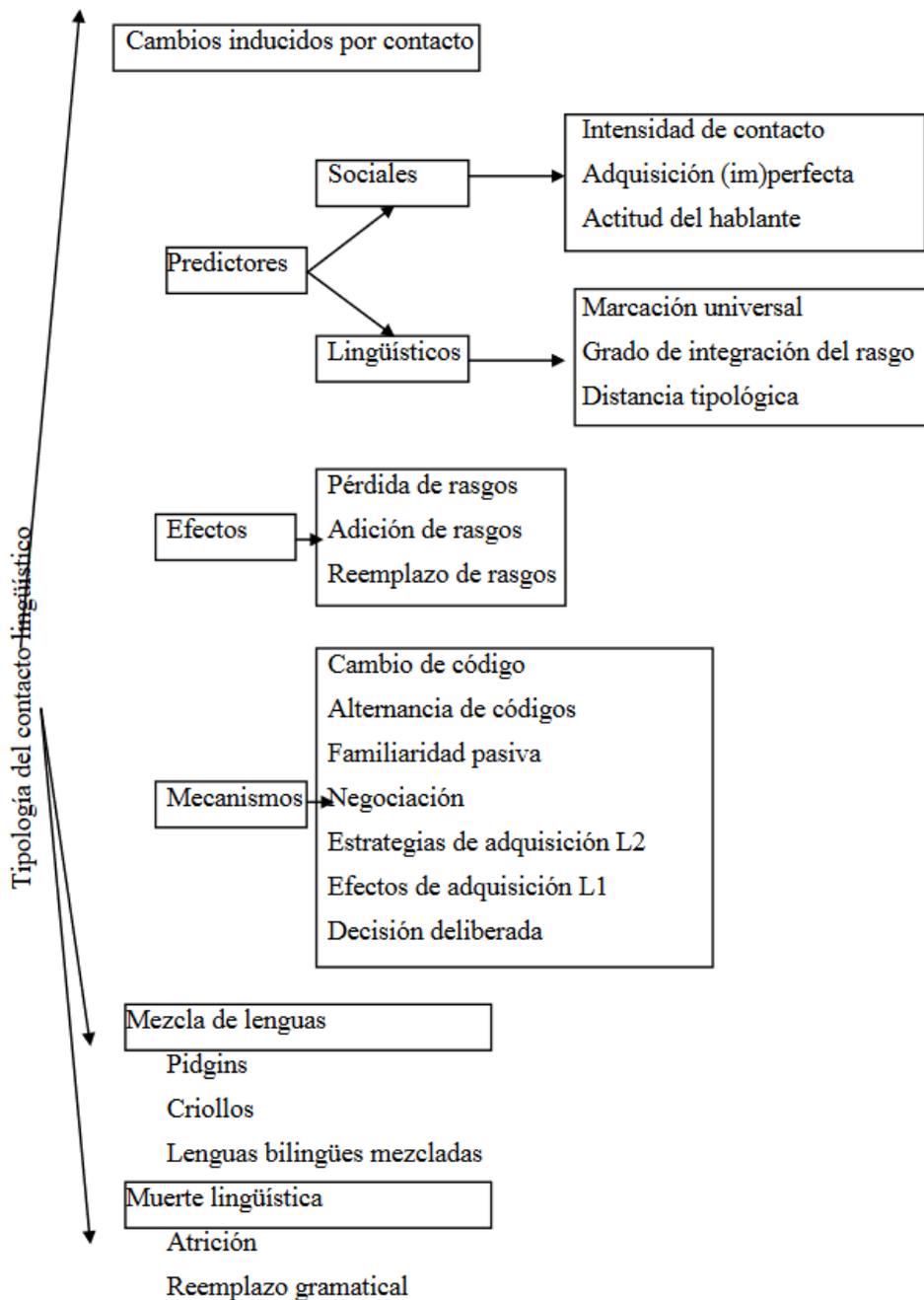
Por otro lado, las interferencias compuestas pueden ser de tres clases. Se transfieren palabras o frases compuestas y se adaptan a los patrones sintácticos de la lengua receptora. El segundo tipo es la traducción de préstamos o calcos semánticos y consiste en tomar el significado de la lengua donante y codificarlo con elementos nativos de la lengua receptora. El último implica la transferencia de algunos elementos y la reproducción de otros (Weinreich, 1953: 51).

1.1.2 Los términos interferencia y préstamo

Posteriormente, en *Language Contact, Creolization, and Genetic Linguistics* (Thomason & Kaufman, 1988) y *Language Contact* (Thomason, 2001) se reelaboran algunas ideas planteadas en Weinreich (1953). Los autores señalan que el contacto lingüístico genera tres resultados principales: los cambios inducidos por contacto, la mezcla de lenguas y la muerte

lingüística, como se muestra en el esquema de la Figura 1.2. Sin embargo, este apartado sólo se centrará en los cambios inducidos por contacto, los factores sociales y lingüísticos que los propician, los efectos que producen en las lenguas en contacto y los mecanismos que los generan.

FIGURA 1.2 *Resultados del contacto lingüístico según Thomason (2001: 60)*



En estas obras se emplea un concepto amplio de cambio inducido por contacto, y lo definen como cualquier cambio cuya aparición sería menos probable sin la presencia del contacto lingüístico, o que se debe parcialmente a este. Los cambios inducidos por contacto pueden ser directos o indirectos. El primer tipo consiste en importar de forma directa un elemento de una lengua fuente a una lengua receptora, el cual puede o no sufrir modificaciones estructurales en los rasgos de la lengua fuente durante este proceso (Thomason, 2001: 61).

Por otro lado, el cambio indirecto inducido por contacto es un cambio tardío, provocado por la importación directa de un elemento de otra lengua. Este tipo de cambio puede tener una motivación interna, pero la probabilidad de que ocurran sin la intervención del primer cambio inducido por contacto es baja. El proceso inicia con la adopción de una característica de interferencia de otra lengua, la cual desencadena otros cambios de forma sucesiva. Los cambios inducidos por contacto son fenómenos multicausales en los que intervienen tanto factores sociales como lingüísticos (Thomason, 2001: 61).

Thomason y Kaufman (1988) y Thomason (2001) sobreponen los factores sociales a los lingüísticos y consideran tres tipos: la intensidad de contacto, la presencia o ausencia de un proceso de adquisición imperfecta y la actitud de los hablantes. Con respecto al primero, señalan que mientras mayor sea la intensidad del contacto mayor diversidad de fenómenos de interferencia y préstamo serán posibles. Aunque los autores advierten la dificultad que implica determinar la intensidad de una situación de contacto, proponen considerar tres factores: la duración del contacto, el número de hablantes de cada lengua y el dominio económico (Thomason, 2001: 66).

El segundo factor social es la presencia o ausencia de un proceso de adquisición imperfecta. Con base en este y en la duración del proceso, distinguen dos fenómenos

asociados al contacto lingüístico, el préstamo y la interferencia. Las interferencias son errores o desvíos que cometen los hablantes nativos de una lengua A, al hablar una lengua B, que no han adquirido por completo. Este tipo de proceso ocurre en situaciones en donde la lengua A es desplazada, puede tomar sólo una generación y no comienza con el ingreso de vocabulario sino con cambios fónicos, sintácticos e incluso morfológicos (Thomason & Kaufman, 1988: 39). Por otro lado, el préstamo comprende rasgos o elementos de la segunda lengua (B) que los hablantes incorporan a su lengua nativa (A). Este tipo de proceso ocurre en situaciones donde la lengua A se mantiene, puede tomar cientos de años y suele iniciar con el ingreso de elementos léxicos (p. 41). El último factor social son las actitudes de los hablantes, que pueden producir excepciones para la mayoría de las generalizaciones que se hacen en torno al cambio lingüístico (Thomason, 2001: 77).

Los predictores lingüísticos de los cambios inducidos por contacto son la marcación universal, el grado de integración de un rasgo en el sistema lingüístico y la distancia tipológica entre las lenguas en contacto. La marcación determina qué rasgos de la lengua meta (B) son más difíciles de adquirir por un grupo de hablantes que están inmersos en un proceso de desplazamiento de su lengua (A), y tienen menos probabilidades de aparecer en la variante de la lengua meta de este grupo (B₁). Por otro lado, los rasgos más integrados o intrincados en un sistema lingüístico tienen menos probabilidades de ser prestados o transferidos de la lengua (A) a la lengua meta (B), en un proceso de desplazamiento lingüístico. Por último, una menor distancia tipológica entre las lenguas en contacto favorece la aparición de distintos tipos de préstamos e interferencias, independientemente de si los rasgos son marcados o tienen un nivel alto de integración (Thomason, 2001: 76-77).

Los cambios inducidos por contacto pueden producir tres efectos en la estructura de la lengua receptora: la pérdida, la adición o el reemplazo de uno de sus rasgos, debido a la

influencia de la otra lengua. Estos resultados se pueden generar a partir de siete mecanismos: el cambio de código, la alternancia de código, la familiaridad pasiva, la ‘negociación’, las estrategias de adquisición de una segunda lengua, la adquisición bilingüe de una primera lengua y la decisión deliberada (Thomason, 2001: 129).

El cambio de código es el uso de material de dos o más lenguas por un solo hablante en una misma conversación (Thomason, 2001: 132). La alternancia de código consiste en el uso de dos lenguas por el mismo hablante, cada una en entornos específicos (136). El mecanismo de familiaridad pasiva ocurre cuando un hablante adquiere un rasgo de una lengua que comprende, mas no habla (139). La negociación es un mecanismo mediante el que los hablantes modifican su lengua A para aproximarse a los patrones de su lengua B. Este proceso puede ocurrir en hablantes que dominan completa o parcialmente la lengua (142).

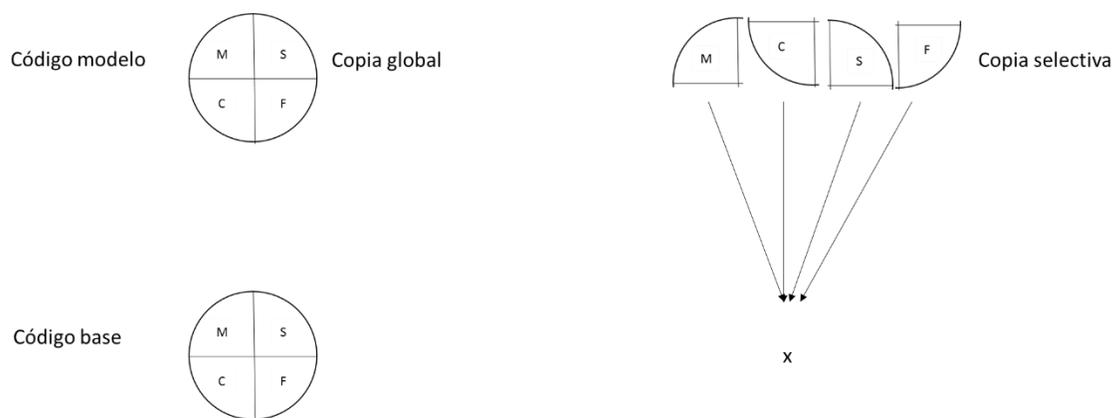
Las estrategias de adquisición de una segunda lengua comprenden el mecanismo de negociación, así como el denominado “hueco funcional”, en donde el bilingüe utiliza material de su idioma nativo mientras habla su segunda lengua para llenar los vacíos que tiene en el conocimiento de esa lengua (Thomason, 2001: 147). Otra estrategia es el mantenimiento de distinciones y patrones de la lengua materna en la gramática de la segunda lengua (148). El siguiente procedimiento consiste en omitir las distinciones que son marcadas en la segunda lengua, debido a que el bilingüe aún se encuentra en etapas tempranas de aprendizaje. Los últimos mecanismos que se mencionan son la adquisición bilingüe del primer idioma y los cambios por decisión deliberada, pero no se ahonda en ellos.

1.1.3 El término copia de código

El modelo propuesto por Johanson (2002) pretende explicar diferentes fenómenos asociados al contacto, como el préstamo, la transferencia, la convergencia y la nivelación, bajo el término *copia de código*. Este proceso consiste en insertar un elemento copiado de un código en el contexto de otro código. El concepto de código refiere tanto a lenguas como a dialectos de una misma lengua. El código del que se toma el elemento copiado se denomina código modelo (CM) y el código que sienta el marco estructural en donde las copias son insertadas es llamado código base (CB). Es importante advertir que la copia nunca es completamente idéntica al modelo (Johanson, 2002: 289).

En este enfoque, las unidades lingüísticas poseen propiedades materiales (rasgos segmentales, patrones fonotácticos, patrones acentuales), semánticas (contenido denotativo y connotativo), combinacionales (constitución interna, patrones de combinación con otras unidades) y frecuenciales. Con base en esto, se clasifican las copias en globales, selectivas o mixtas. En una copia global se reproduce un elemento del código modelo con todas sus propiedades estructurales (material, semántica, combinacional, frecuencial), como se muestra en la Figura 1.3. Estas unidades pueden ser segmentos, morfemas, lexemas de distantes clases (simples o complejas, libres o ligadas, léxicas o gramaticales, abiertas o cerradas), e inclusive frases (Johanson, 2002: 291).

FIGURA 1.3 *Copia global y copia selectiva adaptado de Johanson (2002)*



En la copia selectiva se reproducen alguna(s) de la(s) propiedades estructurales no materiales de las unidades lingüísticas. De este modo, podemos tener copias semánticas, combinatoriales y frecuenciales. En el primer tipo, se copia el contenido denotativo y connotativo de una unidad del código modelo y se inserta en un elemento del código base. En la copia combinatorial se reproducen las propiedades combinatorias internas y externas de una unidad del código modelo y se introducen en un elemento del código base. Por otro lado, en la copia frecuencial, se reproducen los patrones de frecuencia de un miembro del código modelo en una unidad del código base, lo cual produce un incremento o un decremento en su uso. Por último, en una copia mixta tiene lugar una copia selectiva (combinatorial o frecuencial) que comprende una copia global (Johanson, 2002: 292).

Además, en este enfoque, se distingue entre código primario (lengua materna) y código secundario, de acuerdo con el orden de adquisición. Asimismo, según las relaciones de dominación entre los grupos asociados a los códigos, se diferencia entre un código sociolingüísticamente dominado (A) y otro dominante (B), que goza de prestigio porque está relacionado con el poder y el estatus. Las diferencias de dominio entre los hablantes de los

diferentes códigos dan lugar a tres tipos de dinámicas lingüísticas: la adopción, la imposición y el desplazamiento (Johanson, 2002: 290).

En la adopción, los hablantes que tienen una lengua materna sociolingüísticamente dominada (A_1) insertan en ella copias de un código sociolingüísticamente dominante (B_2) y mantienen su primera lengua. En la imposición, los hablantes reproducen copias de su lengua materna sociolingüísticamente dominada (A_1) en un su código secundario dominante (B_2). Por último, en el desplazamiento, la lengua materna que es dominada (A_1) es abandonada por una variedad del código B_2 , que eventualmente se convierte en B_1 (Johanson, 2002: 290-292).

Las copias de código se acomodan a las categorías morfosintácticas del código base para poder ser insertadas. Además, sufren distintos grados de adaptación fonológica, gramatical o semántica al sistema del código básico. Posteriormente, con el tiempo y la frecuencia de uso, estos elementos son convencionalizados y aceptados en la comunidad de habla en cuestión.

1.1.4 El término réplica gramatical

Por último, revisaré algunos conceptos de la obra *Language Contact and Grammatical Change* (Heine & Kuteva, 2005). Los autores indican que la influencia de una lengua a otra se manifiesta a través de la transferencia de material lingüístico, y puede ser de cinco tipos: transferencia de una forma (sonidos o una combinación de ellos), de su significado (gramatical, funcional, o ambos), de una unidad forma-significado, de una relación sintáctica o alguna combinación de estos.

Sin embargo, sólo se centran en analizar la transferencia de significados gramaticales. Para ello, Heine y Kuteva (2005) retoman los conceptos de lengua modelo (M) y lengua réplica (R), que Weinreich (1953) empleó en el análisis de las interferencias gramaticales,

así como la noción de copia de código (Johanson, 2002) y, con base en ellos, establecen el concepto de *réplica gramatical*, entendido como un proceso mediante el que:

Speakers create a new use pattern or category in language R on the model of another language (M), where the outcome of the process is not an exact copy of what exists in M but rather a new structure that is shaped, first, by what is available in R, second, by universal constraints on conceptualization, third, by what speakers of R conceive as being pragmatically most appropriate in the situation in which language contact takes place, and, fourth, by the length and intensity of contact and -accordingly- by the relative degree to which replication is grammaticalized (Heine & Kuteva, 2005: 7).

Al igual que Johanson (2002), Heine y Kuteva (2005) señalan que el resultado de la *réplica gramatical* no es idéntico al modelo, pues se trata de un proceso creativo, regido por las estructuras de la lengua modelo y la lengua *réplica*, así como por los principios universales de la gramaticalización, cuyo comportamiento es el mismo, independientemente de la presencia o ausencia de una situación de contacto lingüístico. De este modo, la gramaticalización inducida por contacto es un fenómeno unidireccional, que va de lo léxico a lo gramatical, y puede identificarse mediante cuatro mecanismos: la extensión, la dessemantización, la decategorización y la reducción fonética (15).

Por otro lado, señalan que la *réplica gramatical* no es independiente del contexto sociolingüístico, pues para que ocurra este fenómeno se requiere cierta intensidad en la situación de contacto y un bilingüismo extensivo. Sin embargo, a diferencia de la propuesta de Thomason y Kaufman (1988) y Thomason (2001), que sobrepone los factores sociales por encima de los lingüísticos, Heine y Kuteva señalan que las variables sociales son irrelevantes para determinar el tipo de cambio inducido por contacto, pues no han encontrado parámetros sociolingüísticos significativos que se correlacionen regularmente con algún tipo de replicación gramatical. No obstante, no descartan la existencia de fuerzas sociopsicológicas

adicionales, como el deseo de usar las opciones disponibles en otra lengua para su propio beneficio, o simplemente hablar como uno de sus vecinos (2005: 28-29).

1.2 EL ESTUDIO DEL BILINGÜISMO EN LOS ESTUDIOS DEL CONTACTO

1.2.1. Definición

El contacto entre los hablantes de dos o más lenguas que conviven en una comunidad² permite el surgimiento del bilingüismo que, de manera general, se ha definido como la práctica de usar dos lenguas alternativamente (Weinreich, 1953: 1). Algunos autores distinguen entre bilingüismo social y bilingüismo individual (Appel & Muysken, 1996: 10; Silva-Corvalán, 2001: 270). Con respecto al primero, proponen tres situaciones: en la primera, las dos lenguas son habladas por dos grupos diferentes (monolingües), y sólo hay algunos bilingües que se encargan de la comunicación intergrupal. En la segunda situación todos los hablantes son bilingües;³ en la tercera, hay un grupo monolingüe y otro bilingüe, y este último suele ser minoritario, desde el punto de vista estadístico o social⁴ (Appel & Muysken, 1996: 10).

En el plano individual, se ha denominado como bilingüe a la persona que hace uso de dos lenguas (Lastra, 1992: 172; Silva-Corvalán, 2001: 269; Thomason, 2001: 3; Weinreich, 1953: 1), sin que esto implique un nivel de competencia nativa en la segunda lengua

² “Una comunidad de habla está formada por un conjunto de hablantes que comparten efectivamente, al menos, una lengua [o una variedad lingüística], pero que, además, comparten un conjunto de normas y valores de naturaleza sociolingüística: comparten unas mismas actitudes lingüísticas, unas mismas reglas de uso, un mismo criterio a la hora de valorar socialmente los hechos lingüísticos, unos mismos patrones sociolingüísticos [...] una comunidad de habla es básicamente una comunidad de consenso, de sintonía entre grupos e individuos diferentes, donde el conflicto está minimizado” (Moreno Fernández, 1998: 23-24). Para mayores detalles, véase (Gumperz, 1962).

³ Este tipo de situación es equivalente al llamado *bilingüismo o multilingüismo mutuo*, donde los hablantes de ambos grupos aprenden la lengua del otro (Thomason, 2001: 4).

⁴ Este tipo de bilingüismo es el que Thomason denomina asimétrico, pues el grupo Z aprende la lengua del grupo Y, pero el grupo Y no aprende la lengua de Z (Thomason, 2001: 4).

(Grosjean, 2004: 34; Silva-Corvalán, 2001: 270), como lo planteó Bloomfield, al describir el bilingüismo como un caso extremo de aprendizaje de una lengua extranjera, donde el hablante alcanza un nivel de competencia nativo de su segunda lengua en todas las habilidades (hablar, escuchar, leer y escribir), sin perder la competencia en su primera lengua (Bloomfield, 1956: 55-56). Contrario a esto, existen definiciones más amplias, que consideran bilingüe a cualquier persona que, además de tener una competencia nativa en una lengua, tiene conocimiento en alguna de las cuatro habilidades (hablar, entender, escribir, leer) de una segunda lengua (Appel & Muysken, 1996: 11; Grosjean, 2004: 35; Montrul, 2013: 7; Silva-Corvalán, 2001: 270; Thomason, 2001: 3).

Una de las propuestas que mejor comprende la dinamicidad y complejidad del individuo bilingüe es la de Grosjean (2004, 2012), quien define el bilingüismo como el uso de dos o más lenguas en la vida cotidiana (2012: 5), cuyos factores definatorios son la fluidez lingüística y el uso de las lenguas. Al respecto, señala que los bilingües están influidos por el principio de complementariedad, es decir, que aprenden las lenguas para distintos propósitos y las usan en diferentes dominios de su vida, con diferentes interlocutores. Por eso, los bilingües difícilmente son equilibrados en las dos lenguas y en todas las habilidades lingüísticas, pues el nivel de fluidez depende del uso que hacen de sus lenguas. Además, añade que la competencia de los bilingües cambia con el tiempo y a lo largo de su vida⁵, en función del contexto en el que interactúan con otros bilingües y monolingües, lo cual implica un proceso de adaptación, según el interlocutor (Grosjean, 2004: 34).

⁵ Esta característica concuerda con lo que Silva-Corvalán denomina *bilingüismo cíclico*, es decir, que un bilingüe puede pasar por ciclos de expansión y reducción de su competencia en una de las lenguas involucradas, lo cual no necesariamente corresponde a dicotomías fijas: bilingüe equilibrado/desequilibrado, o compuesto/coordinado, sino a una amplia gama de niveles dinámicos de competencia en la lengua subordinada que se mueve a lo largo de la vida. Esto es posible gracias a que el bilingüe tiene un nivel de competencia latente que es posible activar fácilmente (Silva-Corvalán, 2001: 270).

Con base en estas observaciones, Grosjean señala que el estudio del bilingüismo requiere tomar en cuenta la historia lingüística de los bilingües, es decir: qué lenguas adquirieron, cuándo y cómo, qué nivel de fluidez tienen en ellas y cuál es el uso que les dan en diferentes dominios lingüísticos (hogar, escuela, trabajo, calle, ciudad, etc.) y con diversos interlocutores (familiares, amigos, compañeros de trabajo) (2012: 11), pues contrario a lo que suele pensarse, la primera lengua de una persona no es necesariamente su lengua dominante.

1.2.2 Clasificación

El bilingüismo se ha caracterizado y clasificado en función de diferentes factores, como la edad, el contexto de adquisición (formal/informal), el uso y el dominio que los hablantes tienen de sus lenguas. En función de la edad en que comenzó el proceso de adquisición, se ha clasificado el bilingüismo como temprano o tardío. Se dice que el bilingüismo temprano, se da cuando una persona adquiere dos o más lenguas durante la infancia (0-12 años) y el bilingüismo tardío, cuando una de las lenguas se aprende después de los 12 años. Dentro del bilingüismo temprano, se distingue entre bilingüismo simultáneo y bilingüismo secuencial. En el primero, el hablante aprende las dos lenguas en la infancia temprana y, en el segundo, aprende primero una lengua y después la otra (Montrul, 2013: 9).

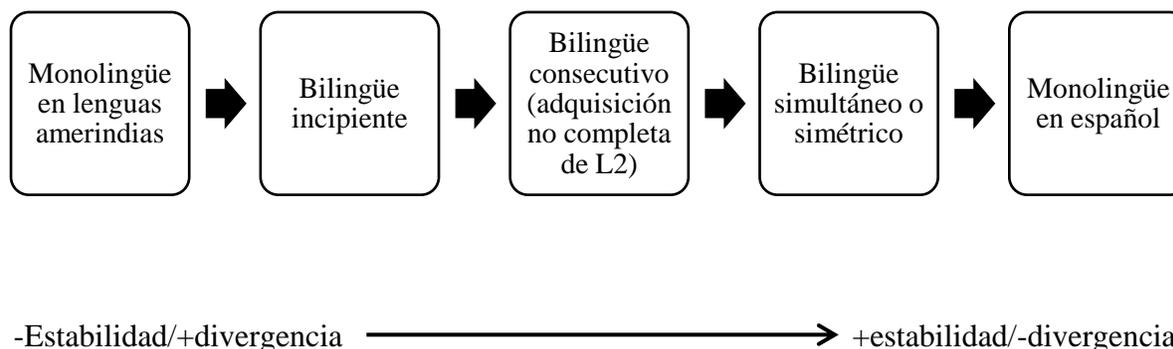
Weinreich (1953) propone otra clasificación de bilingüismo, a partir de la representación mental de las lenguas, y distingue tres tipos de bilingüe: coordinado, compuesto y subordinado. Un bilingüe coordinado es aquel que identifica palabras equivalentes en las dos lenguas y sabe que tienen significados diferentes o que se refieren a conceptos ligeramente distintos. Es decir, estos bilingües mantienen bien diferenciados los dos sistemas lingüísticos lo que, en términos de Bloomfield, sería un verdadero bilingüe. En cambio, para el bilingüe compuesto, las formas equivalentes en las dos lenguas tienen un

significado idéntico. Por último, el bilingüe subordinado tiene más dominio en una lengua, y es a través de ella que interpreta las palabras de su lengua no dominante (Appel & Muysken, 1996: 112-113; Lastra, 1992: 174-175; Moreno Fernández, 1998: 208-209; Weinreich, 1953: 10-11).

En términos sociales y educativos se ha clasificado el bilingüismo como aditivo y sustractivo. En el primero, el hablante adquiere una segunda lengua que es relevante socialmente, sin poner en riesgo el conocimiento de su primera lengua, porque esta también goza de prestigio. En contraste, en el bilingüismo sustractivo la adquisición de una segunda lengua es parte de un proceso de sustitución lingüística, en el que el hablante pierde competencia en su primera lengua (Appel & Muysken, 1996: 152).

Por último, se presenta la clasificación o escala de bilingüismo de Palacios (2005) para situaciones de contacto entre el español y lenguas amerindias. Esta clasificación conjuga diversos factores: la competencia que el bilingüe tiene en las dos lenguas, la edad y el contexto de adquisición (formal/informal), su nivel de instrucción y el ámbito (bilingüe o monolingüe) en el que se desenvuelve. A continuación, se muestra el esquema de esta escala de bilingüismo.

FIGURA 1.4 *Escala de bilingüismo (adaptado de Palacios 2005: 87)*



La escala de bilingüismo inicia con el monolingüe en lengua indígena, seguido por el bilingüe incipiente, el bilingüe consecutivo, después el bilingüe simultáneo o simétrico y, finalmente, el monolingüe en español. En términos de esta autora, el bilingüe incipiente se caracteriza por haber comenzado a aprender español en la adultez y en un contexto no escolarizado (informal), por lo que también tiene un nivel de instrucción bajo; y, en consecuencia, posee una competencia gramatical del español incompleta, que se reduce a algunos dominios; además, este bilingüe se mueve principalmente en redes monolingües en la lengua indígena y sólo tiene contacto esporádico con monolingües en español.

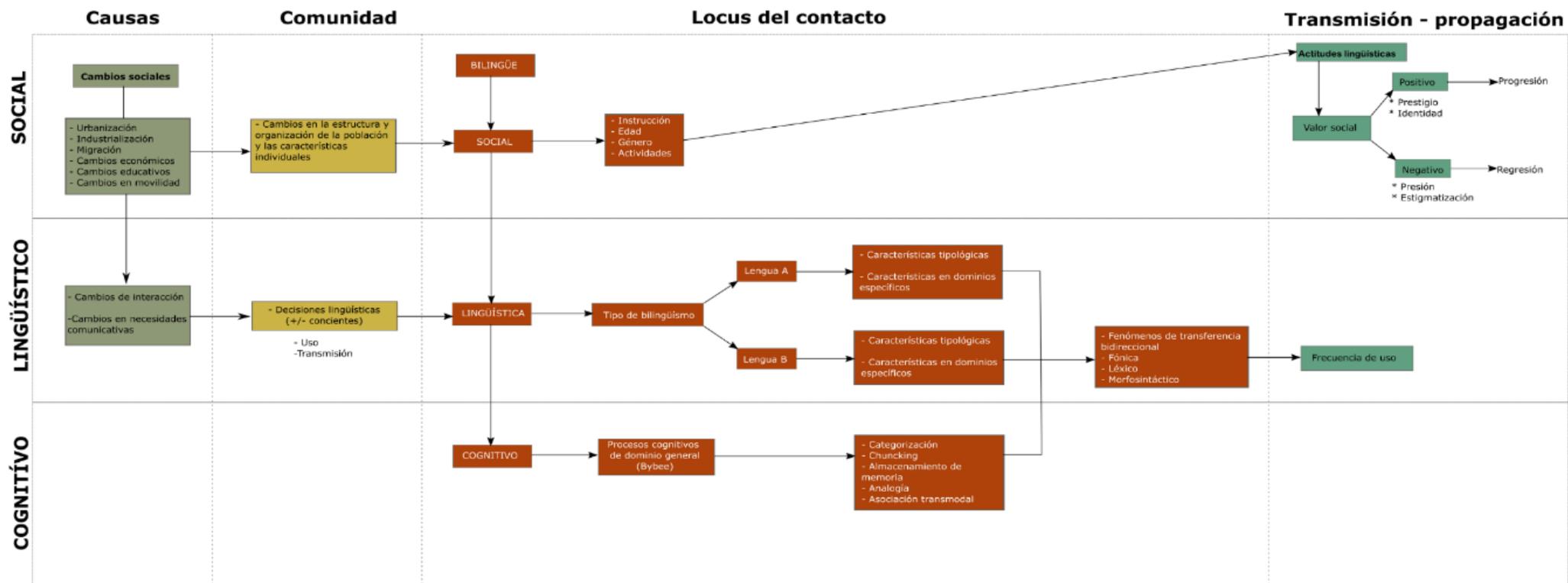
El bilingüe consecutivo se caracteriza por haber aprendido el español de manera informal, después de haber adquirido su primera lengua, tiene un nivel de instrucción bajo y, aunque posee más dominio del español que el bilingüe incipiente, su competencia en español sigue siendo incompleta. También se mueve en redes de monolingües en lengua indígena, pero, además, convive con personas cuya lengua dominante es el español. Finalmente, el bilingüe simultáneo o simétrico se caracteriza por haber adquirido las dos lenguas en la infancia, y el español en un contexto formal, por lo que tiene un nivel de instrucción medio-alto. Además, este bilingüe se mueve en redes cuyos miembros son, principalmente, monolingües en español (Palacios, 2005: 86-88).

Este modelo es novedoso porque intenta adaptar una metodología a cuestiones teóricas, que pueden llegar a favorecer la creación de un marco útil para el estudio sistemático de las diferentes situaciones de contacto (Palacios, 2005: 89), en las que existe un continuo bilingüe de gran complejidad, que se mueve del monolingüismo en una lengua indígena al monolingüismo en español (Silva-Corvalán, 2001: 270).

1.3 EL ESTUDIO DEL CONTACTO LINGÜÍSTICO EN ESTA INVESTIGACIÓN

El marco conceptual que guía el estudio del contacto en esta investigación se presenta en la Figura 1.5. Como puede observarse en la imagen, el marco conceptual tiene tres dimensiones interconectadas de forma vertical: una social, una lingüística y una cognitiva, las cuales dividen cada uno de los cuatro bloques dispuestos de forma horizontal, que simulan distintos momentos en una escala de tiempo. El primero corresponde a las causas sociales del contacto lingüístico; el segundo tiene que ver con los efectos del contacto a nivel comunidad; el tercero comprende los efectos del contacto lingüístico en el plano individual; y el último de los bloques describe el proceso de expansión o propagación de los fenómenos de contacto.

FIGURA 1.5 Marco conceptual de esta investigación



En el primer bloque están las causas del contacto lingüístico. En la dimensión social, ocurren distintos fenómenos como la urbanización, la industrialización, la migración, los cambios económicos, educativos y de movilidad. Estos hechos sociales, a su vez, generan modificaciones en la dimensión lingüística, donde se transforman las interacciones lingüísticas y las necesidades comunicativas de los hablantes. Si se avanza sobre el eje horizontal y temporal, se puede ver que los fenómenos sociales antes enunciados producen reestructuraciones en la organización (y las características) de los habitantes de las comunidades. Con base en las reestructuraciones sociales, los hablantes toman decisiones, o son obligados a tomarlas, en torno a la transmisión y el uso de las lenguas.

En el tercer bloque horizontal se encuentra la dimensión individual del contacto lingüístico. Como señala Weinreich (1953: 1), los bilingües son el verdadero *locus* del contacto y, en el marco conceptual, se conciben como individuos con una dimensión social, lingüística y cognitiva. En primera instancia, los bilingües tienen características sociales, que los hacen parte de un grupo de edad, un nivel de instrucción, una clase social y un género. De modo que, el tomar en cuenta estos factores en la conformación de una muestra también nos permite aproximarse a la estructura y organización social de una comunidad.

En el dominio lingüístico del marco conceptual, el individuo bilingüe no es la suma de dos monolingües en una persona o mente (Bloomfield, 1956: 55-56), sino una persona que tiene una competencia nativa en una lengua y alguna competencia en algunas de las cuatro habilidades (hablar, entender, escribir, leer) de otra. Asimismo, tal como lo propone Grosjean (2004, 2012), pienso que los bilingües están influidos por el principio de complementariedad, pues aprenden las lenguas para distintos propósitos y las usan en diferentes dominios de su vida, con diferentes interlocutores, lo que puede dar lugar a

especializaciones diglósicas. Aunado a esto, el nivel de competencia en las lenguas puede cambiar a lo largo del tiempo.

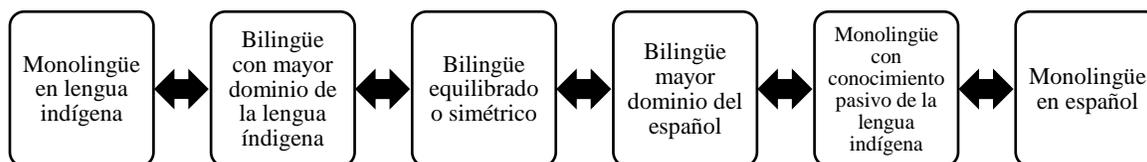
Esta postura, más abierta y flexible, permite un mayor acercamiento, tanto teórico como práctico, a bilingües y situaciones de contacto reales, en las que existe un continuo bilingüe complejo, que va desde el monolingüismo en una lengua hasta al monolingüismo en otra, pasando por una gama de bilingües con diferentes niveles de competencia en sus lenguas. Además, este enfoque considera el dinamismo que caracteriza y acompaña al bilingüismo, pues no pierde de vista que la competencia de los bilingües puede cambiar a lo largo de su vida, en función del contexto.

El bilingüe, como poseedor de dos lenguas, puede tener distintos grados de dominio de estas. La clasificación de bilingüismo que planteo en este trabajo se basa, principalmente, en el nivel de habilidades lingüísticas que tienen los hablantes en las dos lenguas y el uso que hacen de estas, pero también consideré la edad, el contexto de adquisición de las lenguas y la historia de vida de los hablantes.

Las habilidades bilingües se valorarán a partir del análisis de grabaciones de habla espontánea y de pruebas en ambas lenguas (véase §2.1.2.3). A pesar de que este tipo de pruebas conllevan problemas de tipo teórico y práctico, como hallar una norma general o estándar a partir de la cual se puedan medir las habilidades y establecer escalas en las que se establezca el nivel mínimo que se requiere para ser considerado bilingüe, van más allá de clasificaciones impresionistas y permiten una aproximación más objetiva a la competencia que los hablantes tienen de sus lenguas. Por otro lado, describiré el uso de las lenguas a partir de los datos obtenidos con un cuestionario sociolingüístico, en el que se consideran distintos dominios lingüísticos e interlocutores (véase §2.1.2.4). La Figura 1.6 muestra una escala de

bilingüismo dual y dinámica que considera los cambios y la evolución que puede tener la competencia en el español y la lengua indígena.

FIGURA 1.6 *Escala de bilingüismo*



Al igual que en la escala de bilingüismo de Palacios (2005), los extremos de esta clasificación son los monolingües. En primer lugar se encuentra el monolingüe en lengua indígena, que puede tener un conocimiento simbólico (algunos saludos y palabras) del español; después, se encuentra el bilingüe que posee una alta competencia en la lengua indígena y una competencia baja en español; enseguida está el bilingüe equilibrado o simétrico, que se caracteriza por tener una alta competencia en las dos lenguas, sin que esto implique que sean idénticas, pues incluso en este tipo de bilingües hay asimetrías, ya que la competencia lingüística no se distribuye equitativamente en los diferentes niveles de la lengua (Montrul, 2013); más adelante, se hallan los bilingües cuya lengua dominante es el español, seguidos de los monolingües con competencia pasiva de la lengua indígena; y, finalmente, los monolingües en español. Por otro lado, en la dimensión cognitiva, los hablantes tienen la capacidad de llevar a cabo procesos cognitivos de dominio general, como la categorización, la fragmentación, la memoria, la analogía y la asociación transmodal (Bybee, 2010: 6-8).

En cuanto a los fenómenos lingüísticos inducidos por contacto, la descripción de las características tipológicas de las dos lenguas en el dominio específico que se esté estudiando, en conjunción con el conocimiento de los procesos cognitivos de dominio general, permitirá

plantear hipótesis en torno a los fenómenos que pueden ocurrir en determinada situación de contacto, sin que esto implique una restricción en torno a la posibilidad o imposibilidad de que ocurra un determinado fenómeno lingüístico.

Con respecto a la terminología, hemos podido observar que los fenómenos inducidos por contacto son referidos como *interferencias*, *préstamos*, *transferencias*, *cambios inducidos por contacto*, *copia de código o réplica gramatical*. El término *interferencia* de Weinreich (1953), así como la distinción entre *préstamo* e *interferencia* que plantean Thomason y Kaufman (1988) y Thomason (2001) presentan algunos problemas. En principio, hablar de error o desviación de una norma estándar debido a una deficiencia en la adquisición de una segunda lengua muestra una postura prescriptiva que tiene implicaciones teóricas y prácticas. En lo que refiere al primer punto, dar por hecho que todas las interferencias o estructuras alternas son resultado de una adquisición deficiente de una L2, impide ver la complejidad y multicausalidad de estos fenómenos, que pueden tener una motivación cognitiva, semántica, pragmática o, incluso, interna al sistema lingüístico. Por otro lado, en términos prácticos, esta visión prescriptiva ha llevado a la estigmatización de grupos minoritarios, como los indígenas en Hispanoamérica, por considerar su español como una variedad subestándar. Este hecho tiene consecuencias en las actitudes que los hablantes tienen hacia el español, las lenguas indígenas, el bilingüismo y, en consecuencia, el devenir de las lenguas indígenas. Otro de los problemas que presenta la distinción entre *préstamo* e *interferencia* (Thomason & Kaufman, 1988) tiene que ver con el contexto social en el que ocurren estos fenómenos: el préstamo, asociado a una situación de mantenimiento, y la interferencia a una de desplazamiento, pues como veremos más adelante es posible encontrar ambos fenómenos en una situación de desplazamiento y mantenimiento. Por ello, en esta

investigación no emplearé el término *interferencia*, planteado por Weinreich (1953) y ni el de *préstamo* de Thomason y Kaufman (1988).

En su lugar, usaré el término *fenómeno variable inducido por contacto*, un concepto desprovisto de denotaciones negativas y ligado al estudio de la variación sincrónica, cuyo estatus de cambio en proceso, consolidado, o en retracción será puesto a evaluación. En la explicación de los fenómenos variables inducidos por contacto retomaré el planteamiento de Johnson (2002: 291) sobre la idea de que las unidades lingüísticas (simples o complejas, libres o ligadas, léxicas o gramaticales, abiertas o cerradas, frases) poseen propiedades materiales (rasgos segmentales, patrones fonotácticos, patrones acentuales), semánticas (contenido denotativo y connotativo), combinacionales (constitución interna, patrones de combinación con otras unidades) y frecuenciales; sin embargo, me abstendré de usar el término *copia de código*. Más bien partiré de la idea de que los hablantes perciben similitudes y diferencias entre sus lenguas y, con base en esto, establecen relaciones de equivalencia entre ellas. Además, parto de la idea de que estos fenómenos pueden ocurrir de forma bidireccional en cualquiera de los niveles de la lengua: fonológico, léxico, morfológico, sintáctico, semántico y pragmático. Asimismo, considero que los *fenómenos variables inducidos por contacto* pueden estar motivados no sólo por el contacto sino también por fuerzas internas y tipológicamente universales.

En el último bloque del eje horizontal se encuentra la expansión o propagación de los fenómenos de contacto. En el dominio social, están presentes las actitudes lingüísticas de los hablantes, la valoración o indización social del rasgo de contacto, el cual puede ser positivo o negativo y más o menos consciente. Si la evaluación del rasgo de contacto es positiva, ya sea porque esté asociado con un valor de prestigio o porque sea un rasgo identitario, es más factible que este rasgo pueda expandirse en la población monolingüe y logre formar parte de

la norma local de la comunidad. En contraste, si el rasgo de contacto tiene un valor negativo, relacionado con un grupo social estigmatizado o porque exista una fuerte presión de la norma lingüística, es posible que ese rasgo se mantenga sólo en el grupo de bilingües, que no forme parte del habla de los grupos monolingües y que no llegue a ser parte de la norma local. Sin embargo, como bien mencionan Silva-Corvalán y Enrique-Arias (2018: 316):

El destino de los préstamos y transferencias es difícil de predecir. Numerosos factores intra- y extralingüísticos entran a desempeñar un rol en el proceso de difusión de una innovación lingüística: (a) la prominencia del rasgo transferido, es decir, cuán fácilmente puede ser percibido el rasgo como transferencia o préstamo por los hablantes; (b) las actitudes subjetivas hacia la otra lengua y sus hablantes; (c) factores socioculturales de gran complejidad que determinan que las transferencias sean evaluadas como indicadores de prestigio o como rasgos estigmatizados

Además de estos factores, en el proceso de propagación es fundamental la duración y estabilidad de la situación de contacto, pues de eso depende la transmisión de las innovaciones bilingües al grupo de monolingües. A continuación, describiré cómo funciona la categoría de número en la lengua otomí con el propósito de sentar las bases que me permitirán formular hipótesis en torno a los fenómenos variables asociados al número que son objeto de esta investigación.

1.4 LA MARCACIÓN DE NÚMERO EN OTOMÍ

1.4.1 Características generales de la lengua otomí

1.4.1.1 Introducción

El objetivo de esta sección es proporcionar una descripción del funcionamiento del número gramatical en la lengua otomí en el nivel frástico y oracional con la finalidad de establecer los antecedentes necesarios para plantear hipótesis adecuadas que permitan explicar los dos fenómenos variables asociados al número que son objeto de estudio en esta investigación: la

concordancia de número plural en el español de contacto y las marcas de número no singular en los nombres (patrimoniales y préstamos) del otomí.

En la primera parte presento información general de la lengua otomí: su filiación lingüística y variación dialectal, así como algunas características de su fonología y morfosintaxis. Posteriormente, en el apartado §1.4.2 describo el funcionamiento del número en la frase nominal, específicamente la expresión de este rasgo en construcciones con determinantes (artículos, demostrativos, posesivos), cuantificadores, numerales y pronombres independientes.

En la sección §1.4.3 explico la organización del número en el nivel oracional y me enfoco en las relaciones gramaticales de sujeto y objeto. Es importante advertir que esta descripción no pretende ser exhaustiva y sólo presento la información que es pertinente para el análisis de los dos fenómenos que son objeto de esta investigación y que está disponible para la variedad de otomí de San Andrés Cuexcontitlán y Jiquipilco el Viejo.

La descripción de las características morfosintácticas generales de la lengua otomí se realiza desde una perspectiva funcional. Este enfoque parte de la idea de que la función principal de la lengua es comunicar significados semánticos y pragmáticos codificados en estructuras lingüísticas, a partir de una relación comunicativa de base cooperativa y dinámica (Dik, 1997: 5; Givón, 1984: 30-35). En ese sentido, se considera que la intención comunicativa de los participantes de la interacción verbal (emisor y receptor), así como la información pragmática que poseen, y que pueden o no compartir, determinan la manera en que los significados semánticos y pragmáticos son codificados a través de la expresión lingüística (Dik, 1997: 8-9).

Dentro de este marco, la lengua humana se compone de tres niveles unidos concéntricamente: el lexicón, la información proposicional y el discurso multiproposicional

(Givón, 2001: 8). El lexicon se puede dividir en palabras léxicas de “contenido” y palabras no léxicas “funcionales” que, a su vez, incluye los morfemas flexivos y derivativos (44). El conjunto de palabras léxicas comprende cuatro grandes clases:⁶ nombres, verbos, adjetivos y adverbios; y algunas clases menores: adposiciones, determinantes (artículos, posesivos), cuantificadores, numerales, ordinales, pronombres independientes de sujeto y objeto, conectores interoracionales (conjunciones, subordinadores) e interjecciones.

En la Gramática Funcional (GF) de Dik se reconoce las relaciones funcionales en tres niveles: semántico, sintáctico y pragmático. Las funciones semánticas especifican el papel que tienen los referentes de los términos implicados en el “estado de cosas” designado por la predicación en que estos términos aparecen (Dik, 1997: 26). En seguida defino los principales roles semánticos que serán empleados a lo largo de esta sección (Givón, 2001: 107).

- a. Agente: participante típicamente animado que actúa de forma deliberada para iniciar un evento y asume la responsabilidad de este.
- b. Paciente o Tema: participante, animado o inanimado, que se encuentra en un estado o registra un cambio de estado como resultado de un evento.
- c. Dativo: participante consciente de un evento, generalmente animado, que no es el iniciador deliberado de este.

⁶ Givón (2001: 50-54) caracteriza a los sustantivos, adjetivos y verbos a partir de un conjunto de cuatro criterios semánticos: estabilidad temporal (rango de cambio en el tiempo), complejidad (cantidad de características definitorias), concreción y compactibilidad (grado de dispersión espacial). Con base en estos criterios, un sustantivo prototípico –entendido como el miembro de la categoría que posee el mayor número de propiedades características importantes– es estable en el tiempo y tiene un número importante de propiedades definitorias con una fuerte asociación entre sí; además, es concreto y compacto. Por otro lado, un verbo prototípico, al ser un conjunto de experiencias con duración relativamente corta, es temporalmente inestable porque codifica cambios de estado o condición rápidos. Asimismo, se caracteriza por ser concreto y complejo, pues codifica actividades o cambios físicos e involucra diferentes participantes. Por último, el adjetivo prototipo es una propiedad única y concreta, analíticamente abstraída de la experiencia directa de entidades prototipo codificadas en sustantivos, por lo que se espera que sea estable, simple, concreta y compacta.

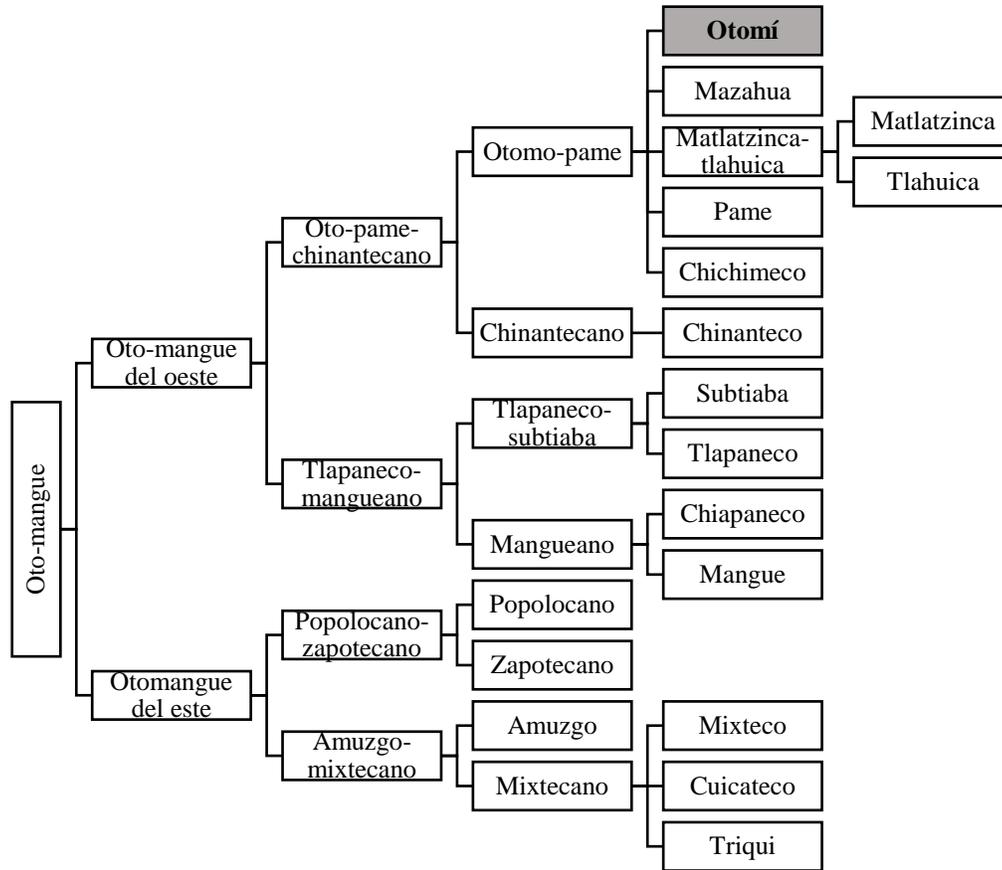
- d. Instrumento: participante, por lo regular inanimado, empleado por el agente para realizar la acción.
- e. Benefactivo o receptor: participante, típicamente animado, que se beneficia de la acción realizada.
- f. Locativo: el lugar, concreto e inanimado, donde un estado o evento ocurre, o el sitio hacia el que se desplaza algún participante.
- g. Manera: la manera en que un evento ocurre o un agente realiza la acción.

Las funciones sintácticas especifican la perspectiva desde la que se presenta el “estado de cosas” en una expresión lingüística (Dik, 1997: 26). Las principales funciones sintácticas son sujeto (S), objeto directo (OD), objeto indirecto (OI), predicativo y adverbio; el primer par ocupa una posición nuclear y las dos últimas un lugar periférico (Givón, 2001: 108). Por otro lado, las funciones pragmáticas (Tema, Tópico, Foco) especifican el estatus informativo de un constituyente dentro del marco comunicativo en el que aparecen (Dik, 1997: 26). A continuación, presento la filiación lingüística de la lengua otomí.

1.4.1.2 Filiación lingüística de la lengua otomí

La lengua otomí pertenece a la familia lingüística otópame, que comprende las lenguas chichimeca, pame, matlatzinca, tlahuica y mazahua. Esta familia, junto con los grupos tlapaneco-mangueano, popolocano-zapotecano y amuzgo-mixtecaño conforman el tronco lingüístico otomangue (INALI, 2008: 40-41; Lastra, 2010: 32-33).

FIGURA 1.7 *Tronco lingüístico otomangue*



En cuanto a su situación demográfica, la lengua otomí cuenta con un total de 239 850 hablantes (Instituto Nacional de Estadística y Geografía, 2009: 69), distribuidos en ocho estados de la República Mexicana: Puebla, Guanajuato, Querétaro, Hidalgo, Veracruz, Tlaxcala, Querétaro, Michoacán y México (INALI, 2008; Lastra, 2010: 23, 33-34). A continuación, en la Tabla 1.1 se presenta el número de hablantes por entidad federativa para el año 2005.

TABLA 1.1 *Número de hablantes por entidad federativa*

Entidad federativa	Hombres	Mujeres	Total	Distribución porcentual	Porcentaje monolingües
Hidalgo	46 017	49 040	95 057	39.7	5.1
México	39 765	43 587	83 352	34.9	1.1
Querétaro	9 060	9 873	18 933	8.0	3.5
Veracruz	8 325	8 497	16 822	7.0	13.5
C. México	5 510	6 950	12 460	5.2	0.4
Puebla	3 247	4 006	7 253	3.0	17.5
Guanajuato	395	346	741	0.3	0.1
Michoacán	248	232	480	0.2	0.0
Tlaxcala	184	204	388	0.2	0.8

La mayor cantidad de hablantes de otomí se encuentra en el estado de Hidalgo, seguido por el Estado de México, Querétaro y Veracruz. Aunque la Ciudad de México no suele considerarse como uno de los estados de origen de la población otomí, es una de las entidades con más hablantes de esta lengua, por encima de Puebla, Guanajuato, Michoacán y Tlaxcala, debido a que es uno de los destinos de migración más recurrentes de la comunidad otomí. A causa de la gran dispersión territorial de la población otomí, su lengua presenta gran variación; en el siguiente apartado presento algunas de las propuestas sobre la variación dialectal de esta lengua.

1.4.1.3. Variantes de la lengua otomí

Los estudiosos coinciden en que existe una gran variación en la lengua otomí; sin embargo, no hay un consenso en cómo agrupar esa variación. Soustelle (1993) propuso siete divisiones geográficas para el otomí a partir de la realización de segmentos linguodentales,⁷ labiales,⁸

⁷ Dentro de esta categoría revisó la realización de la oclusiva sonora, la semioclusiva sorda (oclusiva + aspiración) y la oclusiva sonora nasal en posición inicial y no inicial (Soustelle, 1993: 213).

⁸ En este grupo el autor incluye la oclusiva sonora, la semioclusiva sorda y la oclusiva sonora en posición inicial y no inicial (Soustelle, 1993: 213).

palatales,⁹ fricativos,¹⁰ semivocálicos¹¹ y vocálicos¹² de una lista de vocabulario en 33 localidades.¹³ En la Tabla 1.2 se muestra los siete grupos, las localidades encuestadas que comprenden, el tipo o representante del grupo y su distribución geográfica. De acuerdo con esta propuesta, el otomí de Jiquipilco el Viejo y el de San Andrés Cuexcontitlán formarían parte del Grupo IV.

TABLA 1.2 *Grupos otomíes según Soustelle (1993: 217-223)*

Grupo	Localidades	Tipo	Distribución geográfica
Grupo I	3, 6, 9, 13, 14, 15, 16	San Miguel, Tolimán, Querétaro	Querétaro y Guanajuato hasta el río Moctezuma al sudeste y al este hasta el valle del Laja.
Grupo II	1, 2, 3, 4, 5, 7, 8, 10, 11, 17	El Cardonal, Hidalgo	Todos los altiplanos centrales de Hidalgo y el sudeste de la Sierra Gorda, entre el río Moctezuma y el Metztlán.
Grupo III	12, 18, 19, 20, 21, 27	San Andrés Timilpan, México	La región de Jilotepec, que engloba el norte de la Sierra de las Cruces y una parte de los altiplanos. El otomí de Michoacán está vinculado a este centro.

⁹ En el conjunto de las palatales se considera la oclusiva sonora y la semioclusiva sorda en posición inicial y no inicial (Soustelle, 1993: 213).

¹⁰ En esta categoría se incluyó la fricativa sibilante /š/, la fricativa linguodental /s/, la semioclusiva linguodental /t/ en posición inicial y no inicial y la fricativa sibilante inicial delante de la oclusiva linguodental sorda /št/ (Soustelle, 1993: 213-214).

¹¹ En este grupo se consideró las semivocales /y/, /w/ y el grupo nasal-semivocal *ny-* en posición inicial y no inicial (Soustelle, 1993: 214).

¹² En este conjunto se tomó en cuenta la vocal /o/ en alternancia con /â/ en posición final y en monosílabos, las alternancias /o/, /a/ y /e/, /i/, así como la alternancia del diptongo /ey/ con las vocales simples /i/, /e/ (Soustelle, 1993: 214).

¹³ Las localidades consideradas fueron: 1. Ixmiquilpan, 2. Zimapán, 3. Tolimán, 4. Jiliapan, 5. Santa Catarina, 6. La Cañada (Querétaro), 7 (véase 4), 8. Huichapan, 9. San Miguel, Tolimán, San Pablo, San Antonio, Higuera, 10. Tecozautla, 11. San Salvador, Actopan, con límite en Pachuca, 12. San Ildefonso, 13. Amealco, San Miguel Tlaxcaltepec, Santiago Mezquititlán, San José Itó, San Bartolomé, San Ildefonso, San Pedro y Chintyú. 14. Tetillas, región de Cadereyta, Boyé, Sombrerete, 15. Dolores Hidalgo, Tequisquiapan, La Cruz del Palmar, San Juande la Vega, La Orduña, Santa Cruz, Apaseo, San Bartolo, San Pedro, Obraje de Ixtla, San Miguel de Ixtla y Mayorazgo, 16. Tierra Blanca, Cieneguillas, Corral Viejo, El Picacho, Cerro Colorado, La Cuesta y El Cano, 17. El Cardonal, 18. San Bartolo, Morelos, Sierra de las Cruces, San Marcos, San Gregorio, Malacota y San Felipe, 19. Timilpan, 20. Chapa de Mota (San Juan Tuxtepec, San Felipe Coamango, Dongú, Bodenquí) 21. La región de Jilotepec, noroeste del Estado de México hasta los límites de Querétaro, 22-23. Algunos datos de San Bartolo Oztolotepec, Santa Ana Jilotzingo, 24. Ameyalco, Atarasquillo y Xochicuautla, 25. Temoaya, Jiquipilco, Capulhuac, 26. Amanalco, 27. Siráhuato, Michoacán (Ciudad Hidalgo, San Matías y San Pedro Jarácuaro, 28. Huixquilucan, 29. Mitad meridional del Valle de Toluca (Acasulco, Atlapulco, Cholula, Tepexoyuca, Coapanoaya, Almaya, Tilapa, Xalatlaco, 30. San José del Sitio, 31. Huayacocotla, Veracruz, 32. Ixtenco, 33. Santa Ana y San Pedro (Soustelle, 1993: 203-207).

Grupo IV	25, 26, 30	San José del Sitio, México	Es un área discontinua. Comprende el altiplano de Ixtlahuaca y el enclave de Amanalco, en la Sierra que separa al Estado de México y Michoacán.
Grupo V	22, 23, 24	San Miguel Ameyalco, México	El extremo sur de la Sierra de las Cruces y la porción del altiplano circundante.
Grupo VI	28, 29, 31, 33	Santa Ana Hueytlálpam, Hidalgo	Comprende regiones alejadas. Por un lado, la Sierra, en Tierra Templada tendiendo hacia Tierra Caliente. Por otro lado, el sur del Valle de Toluca.
Grupo VII	32	Ixtenco, Tlaxcala	El único pueblo otomí del altiplano de Tlaxcala.

La siguiente propuesta se basa en pruebas de inteligibilidad¹⁴ realizadas por el Instituto Lingüístico de Verano entre los años 1968-1971 en los estados de Hidalgo, México, Guanajuato, Querétaro y Tlaxcala.¹⁵ Los resultados de las pruebas sugieren la existencia de 7 agrupaciones otomíes y una isla lingüística en Tilapa, Estado de México (Egland, 1978: 46), como se muestra en la Tabla 1.3. Aunque las comunidades de San Andrés Cuexcontitlán y Jiquipilco el Viejo no fueron consideradas en el estudio de inteligibilidad, podrían incluirse en la agrupación 4, debido a la cercanía geográfica con esas localidades y a que en otras propuestas se les ha agrupado con San Felipe, Estado de México.

¹⁴ Para la realización de estas pruebas se seleccionaron puntos geográficos que podían representar diferentes variedades dialectales, ya sea por su importancia política, religiosa o económica. En cada sitio se grabó y se transcribió una narración con la ayuda de un hablante nativo; posteriormente, se diseñaron 10 preguntas de comprensión que se pusieron a prueba con hablantes del mismo sitio. El siguiente paso consistió en aplicar la prueba a 10 hablantes representativos de cada punto de encuesta. La evaluación se realizó con promedios y se estableció como nivel crítico entre entendimiento adecuado y no adecuado el 80%: de este modo, menos de 80% correspondería a mal entendimiento y más de 80% a buen entendimiento (Egland, 1978: 1-4).

¹⁵ Las comunidades consideradas en las pruebas fueron Santa Ana, San Antonio, Anaya, Santa Clara, San Felipe, San Gregorio, Ixhuatlán, Ixmiquilpan, Ixtenco, Jiquipilco, Mexquititlán, San Nicolás, Otatitlán, San Pablo, Pueblo Nuevo, Tecozautla, Texcatepec, Tolimán, Villagrán, Zozea (Egland, 1978: 46).

TABLA 1.3 *Agrupaciones otomíes según su inteligibilidad (Egland 1978: 46)*

Agrupaciones lingüísticas	
1. Santa Ana, Pue. San Nicolás, Hgo. San Pablo, Pue.	4. San Felipe, Méx. Jiquipilco, Méx. Sta. Clara, Méx.
2. Anaya, Hgo. Ixmiquilpán, Hgo.	5. Ixtenco, Tlax.
2a. Zozea, Hgo.	6. Texcatepec, Hgo.
3. San Antonio Hgo. Ixhuatlán, Hgo. Otatitlán, Hgo. San Gregorio, Hgo.	7. Tolimán, Qro. Tecoautal, Hgo. Mexquititlán, Qro. Villagrán, Gto. Pueblo Nuevo, Méx.

Bartholomew (1994: 335-336) señala que la lengua otomí tiene cuatro áreas dialectales. El otomí del Mezquital, que se habla en la parte oeste del estado de Hidalgo, el otomí del Este, que se utiliza al este de Hidalgo, en el área que colinda con Veracruz al noroeste y con Puebla al sudeste. El otomí de Toluca, que se encuentra en el noreste del Estado de México. Y el otomí de Tula, que se habla en la parte sur de Hidalgo, en la parte noroeste del Estado de México y en partes colindantes con el estado de Querétaro. Además, advierte la existencia de dos islas dialectales, Ixtenco, en el estado de Tlaxcala y Tilapa en el Estado de México al sureste de Toluca. Según esta clasificación el otomí de San Andrés Cuexcontitlán y el de Jiquipilco el Viejo forman parte de la variante de Toluca.

Años más tarde, Lastra (2001: 24-25) propuso tres zonas dialectales: oriental, noroccidental y suroccidental. Los dialectos orientales, a su vez, se dividen en tres. El primero corresponde a la Sierra y comprende el otomí de Texcatepec, Ixhuatlán, San Antonio Huehuetla, Tutopec, San Pablito, Santa Ana Hueytlalpan. El segundo subgrupo está compuesto por el otomí de Tilapa y el de Acapulco. El otomí de Ixtenco constituye por sí

mismo un tercer subconjunto. Los dialectos noroccidentales se dividen en dos: el Mezquital y el otomí de Querétaro (sin Amealco) y Guanajuato. Por último, los dialectos suroccidentales están integrados por el otomí de Amealco, San Martín Tuchiuitlapilco (Jilotepec), San Lorenzo Nenamicoyan (Jilotepec), Dongú (Chapa de Mota), San Felipe, San Andrés, Jilotzingo. De acuerdo con esta clasificación, el otomí de San Andrés Cuexcontitlán y el de Jiquipilco el Viejo forman parte del otomí suroccidental.

Por otro lado, el INALI (2008: 37) define variante lingüística como una forma de habla que se distingue de otras variantes de la misma agrupación lingüística en su estructura y en la identidad sociolingüística de sus hablantes. Con base en los estudios disponibles sobre las diferentes agrupaciones lingüísticas y las autodenominaciones, propone para el otomí nueve variantes lingüísticas (41-57), que se muestran en la Tabla 1.4. Según esta clasificación, el otomí de San Andrés Cuexcontitlán y el de Jiquipilco el Viejo forman parte del otomí del centro, bajo la autodenominación *ñható*.

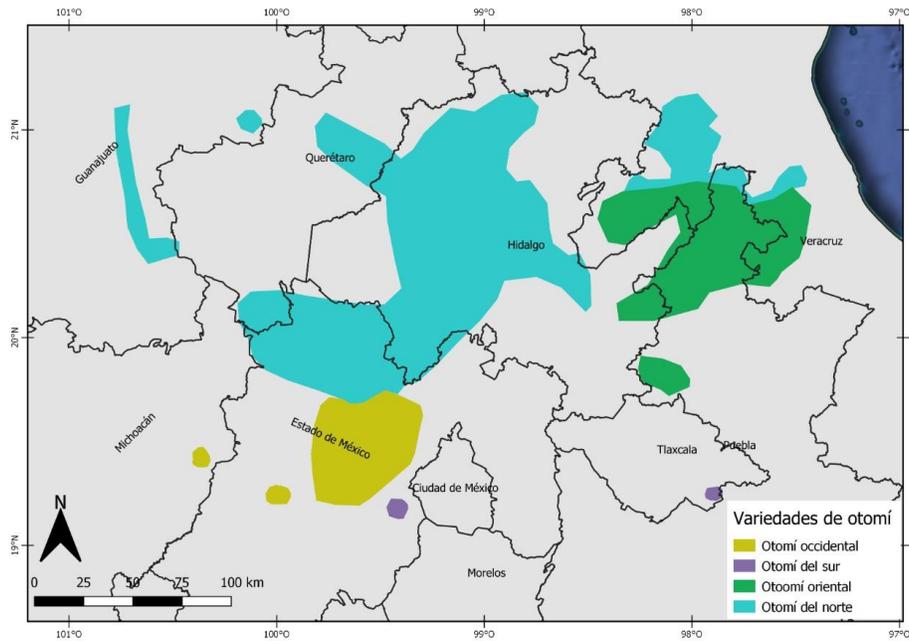
TABLA 1.4 *Variantes del otomí según el INALI (2008: 41-57)*

Autodenominación	Nombre en español
ñuju [ɲuxu] ñoju [ɲoxu] yühu [jũhu] hñãñho [hɲãɲho] ñathó [ɲathó]	otomí de la Sierra otomí bajo del oeste otomí del oeste del Valle del Mezquital
ñöhñö [ɲõhɲõ] ñähñá [ɲãhɲa] hñãhhñú [hɲãɲhú] ñänhú (del Valle del Mezquital) [ɲãnhu] ñandú [ɲãndú] ñóhnño (del Valle del Mezquital) [ɲóhnɲo] ñanhmu [ɲanhmu] yuhmu [ɲũhmu]	otomí del Valle del Mezquital otomí de Ixtenco

ñü'hü [ɲũʔhũ]	otomí de Tilapa o del sur
hñöñho [hɲõɲhõ]	otomí del noroeste
ñühú [ɲühú]	
ñanhú [ɲanhú]	
hñähñu (del centro) [hɲãɲhu]	otomí del centro
ñöthó [ɲõtho]	
ñható [ɲható]	
hñothó [hɲothó]	
ñóhnño (del centro) [ɲóhnɲo]	

A partir de los trabajos de Soustelle (1993), Egländ (1978) y Lastra (2001, 2010), Palancar (2013: 103-105) señala que las variedades actuales del otomí podrían agruparse en cuatro áreas geográficas, que comprenden al menos seis lenguas diferentes, y que, a su vez, constituyen continuos dialectales con diversas variedades lingüísticas: el otomí oriental (Huehuetla, Tenango de Coria, Texcatepec, Tulancigo, San Pablito), el otomí del norte (Valle del Mezquital, Santiago Mexquititlán, San Ildefonso Tultepec, Chantepec, Tolimán, Acambay), el otomí occidental (San Felipe, Toluca, Temoaya, San Pedro, Jiquipilco) y el otomí del sur (Ixtenco, Tilapa, Atlapulco, Acapulco), como se muestra en la Figura 1.8. Según esta propuesta, el otomí de San Andrés Cuexcontitlán y el de Jiquipilco el Viejo forman parte del otomí occidental.

FIGURA 1.8 Áreas geográficas del otomí, adaptado de Palancar (2013)



En las diferentes propuestas dialectales hasta ahora expuestas, el otomí de San Andrés Cuexcontitlán y el de Jiquipilco el Viejo forman parte de la misma variante o agrupación lingüística, lo cual permite la comparación entre las dos situaciones de contacto. En las siguientes secciones describo las características generales de la lengua otomí y la manera en que se expresa el número en el nivel frástico y oracional.

1.4.1.4 El sistema fonológico de la lengua otomí

El sistema fonológico del otomí de San Andrés Cuexcontitlán tiene 19 consonantes:¹⁶ siete fonemas oclusivos, tres nasales, cinco fricativos, dos africados, uno lateral y una rótica percusiva (Lastra, 1989: 21-23, 1992: 16-17),¹⁷ como se muestra en la Tabla 1.5. Los fonemas

¹⁶ Lastra (1989, 1992) transcribió sus datos con el alfabeto fonético americanista, pero a efectos de este trabajo los símbolos se adaptaron al alfabeto fonético internacional.

¹⁷ El otomí de Ixtenco tiene el mismo inventario consonántico: /p, t, ʃ, ʒ, ts̺, tʃ̺, k, ʔ, b, d, z, g, s, h, m, n, ɲ, r/, con excepción de /l/ (Lastra, 1997: 22). Por otro lado, el inventario consonántico de la variante de Santiago Mexquititlán difiere de los dos anteriores porque presenta sonidos consonánticos complejos aspirados,

/l/ y /r/ son de muy bajo rendimiento y se encuentran principalmente en préstamos del español (Andrews, 1949: 214).

TABLA 1.5 *Sistema consonántico del otomí de SAC, adaptado de Lastra (1992: 16-17)*

	Bilabial	Alveolar	Alveo-palatal	Velar	Glotal
Oclusivo	p b	t d		k g	ʔ
Nasal	m	n		ŋ	
Fricativo		s z	ʃ	ʒ	h
Africado		ts		tʃ	
Lateral		l			
Vibrante simple		r			

Actualmente no se cuenta con una descripción del sistema fonológico del otomí de Jiquipilco el Viejo y sólo se tiene la descripción de la fonología del otomí de San Pedro Arriba, una localidad a 6 kilómetros de distancia. El sistema fonológico de esta variedad es similar al de San Andrés Cuexcontitlán, con la única diferencia de que tiene dos consonantes más, /k^w/ y /g^w/ (Andrews, 1949: 213-214).

El sistema vocálico del otomí de San Andrés Cuexcontitlán cuenta con nueve vocales orales y dos semivocales, cuyas características articulatorias son descritas en la Tabla 1.6. Además, cuenta con tres vocales nasales: una anterior cerrada /ĩ/, otra central abierta /ã/ y una posterior cerrada /ũ/ (Lastra, 1989: 21-25, 1992: 15-16).¹⁸ El inventario vocálico de San Pedro Arriba coincide totalmente con el de San Andrés Cuexcontitlán.

glotalizados o preaspirados y preglotalizados: /p, t, k, ts, tʃ, k, ʔ, s, g, h, z, m, n, j, l, r, w, j, k^w, β, ð, g^w, p^h, k^h, t^h, k^{hw}, t^ʔ, ts^ʔ, tʃ^ʔ, k^ʔ, ^hm, ^hn, ^hj, ^hw, ^hj, ^ʔm, ^ʔn, ^ʔj, ^ʔr, ^ʔw, ^ʔj/ (Guerrero, 2015: 240-241).

¹⁸ El sistema vocálico del otomí de Santiago Mexquititlán también tiene nueve vocales orales: /i, e, ε, i, ə, ʌ, a, u, o, ɔ/, pero sólo una vocal nasal /ã/ (Guerrero, 2015: 243; Hekking & Andrés de Jesús, 1984: 12-13). El otomí de Ixtenco, por su parte, tiene ocho vocales orales, debido a que el fonema /ɔ/ se ha fusionado con /o/. Además, esta variante cuenta con cuatro vocales nasales: /ĩ/, /ũ/, /ã/ y /ẽ/ (Lastra, 1997: 27), mientras que la de San Ildefonso Tultepec, Querétaro, presenta cinco vocales nasales /ã, ẽ, ĩ, õ, ù/ (Palancar, 2009: 20).

TABLA 1.6 *Sistema vocálico del otomí de SAC, adaptado de Lastra (1992: 16-17)*

Vocales	Anterior	Central	Posterior
	No Redondeadas	No Redondeadas	Redondeadas
Semivocales	y		w
Cerrada	i	ɨ	u
Semicerrada	e		o
Semiabierta	ɛ	ʌ	ɔ
Abierta		a	

El patrón silábico no marcado del otomí es abierto CV¹⁹ y se pueden presentar diferentes configuraciones: CCV, NCV, NCCV (Andrews, 1949: 217-218). Sólo en determinados contextos se presentan sílabas cerradas del tipo CVC y CVCC, por ejemplo, en fronteras morfológicas o en la primera sílaba de un bisílabo (Guerrero, 2015: 243-244; Lastra, 1997: 24; Palancar, 2009: 11, 30-32). Las únicas consonantes que pueden aparecer en posición de coda son /ʃ/, /k/, /ʔ/ y /n/ (Lastra, 1997: 24).

En otomí se acentúan los sustantivos, los adjetivos, los verbos, los adverbios y los pronombres (Hekking & Andrés de Jesús, 1984: 19). “El acento otomí tiene una función demarcativa que permite al hablante reconocer el inicio del radical²⁰ y en esta sílaba se pueden localizar los tres contrastes tonales” (Guerrero, 2015: 236), que son uno alto (H), uno ascendente (LH) y uno bajo (L) que no se marca gráficamente (Guerrero, 2015: 236; Hekking & Andrés de Jesús, 1984: 19-20; Lastra, 1989: 25, 1992: 17, 1997: 25; Palancar, 2009: 11, 36-41). A continuación, presento algunos ejemplos.

¹⁹ La C corresponde a cualquier consonante, V a una vocal y N a una consonante nasal.

²⁰ “Las raíces léxicas generalmente son monosilábicas [...] Las excepciones se pueden entender como petrificaciones de compuestos muy antiguos o de combinaciones morfológicas que ya no son productivas (p.ej. *tsaya* ‘descansar’ y *pahni* ‘camina’). A pesar de ello, existen relativamente pocas palabras monosilábicas, ya que muchos de los lexemas de la lengua se realizan en palabras bisilábicas que morfológicamente consisten de una raíz léxica y un formativo temático sin significación léxica aparente” (Palancar, 2009: 11-12).

- (1) a. *Dí hóhki*
 dí = **hó** - hki
 1.PRES = limpiar - 3.OBJ
 ‘Yo limpio algo’
- b. *Dí hohki*
 dí = **ho** - hki
 1.PRES = afilar - 3.OBJ
 ‘Yo afilo algo’
- c. *Rá ‘yó*
 rá = **ʔyó** - Ø
 3.PRES.CONT = caminar - 3.OBJ
 ‘Él camina algo’
- d. *Dra ‘yoo*
 drá = **ʔyǒ**
 3.PRES.CONT = caminar
 ‘Él camina’
- e. *Dí hets’e*
 dí = **hetsʔe**
 1.PRES = estornudar
 ‘Yo estornudo’
- f. *Dí heets’e*
 dí = **hětsʔe**
 1.PRES = tapar
 ‘Yo tapo a alguien’

(Otomí de Temoaya, adaptado de Andrews 1949: 215).

1.4.1.5 Características morfosintácticas generales de la lengua otomí

El orden básico de constituyentes de la lengua otomí es SVO (Lastra, 1992) y se caracteriza por tener un sistema de marcación en el núcleo (Nichols, 1986),²¹ es decir, que las relaciones sintácticas se marcan morfológicamente en el núcleo de los constituyentes y no en el dependiente, como se esquematiza en la siguiente tabla.

²¹ Nichols (1986: 57) analiza las relaciones de dependencia en el nivel frástico (construcciones posesivas, con adjetivos y adposiciones), oracional (argumentos, adjuntos y verbos auxiliares) y de cláusula (relativos y oraciones subordinadas).

TABLA 1.7 *Sistemas de marcación en el dependiente y en el núcleo (Nichols, 1986: 59-64)*

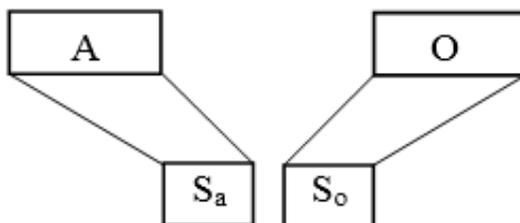
Nivel frástico		
Frase posesiva	Dependiente	Nombre ₁ + ^M Gen ^N Nombre ₂
	Núcleo	Nombre₁ ^NNombre₂ + ^MAfijo pronominal _{N1}
Frase adposicional	Dependiente	Nombre ₁ + ^M Caso ^N Adposición
	Núcleo	Nombre ₁ ^N Adposición + ^M Afij _N
Frase atributiva	Dependiente	Adjetivo + ^M Afij _N ^N Nombre
	Núcleo	Adjetivo ^NNombre + ^MAfijo_A
Nivel oracional		
Relaciones argumentales	Dependiente	Nombre ₁ + ^M Caso Nombre + ^M Caso Nombre + ^M Caso ^N Verbo
	Núcleo	Nombre₁ Nombre₂ Nombre₃ ^NVerbo + ^MAfijo₁ + ^MAfijo₂ + ^MAfijo₃
Nivel cláusula		
Relativas	Dependiente	[[^M {Ø, pro} ...] _{CR} ... ^N Nombre ...]
	Núcleo	[[... Nombre ...]_{CR} M{Ø, pro} ^N...]

Con respecto a la marcación del sujeto, sigue un sistema de sujeto escindido (Dixon, 1994),²² o activo/agentivo en términos de Mithun (1991), en el que la distinción semántica que se hace entre el sujeto (A) y el objeto (O) en una oración transitiva también se aplica al sujeto de una oración intransitiva (S), de modo que se puede tener dos tipos de sujetos intransitivos. El primero (Sa) ejerce control en los verbos de actividad y se marca como el sujeto de una oración transitiva (A). El segundo tipo (So) es afectado por la actividad o

²² Dixon (1994: 6) distingue entre oraciones intransitivas y transitivas. Las primeras constan de un verbo y una frase nominal y las oraciones transitivas, por su parte, involucran un verbo y dos o más frases nominales. La descripción de los sistemas de alineamiento se realiza en términos de tres relaciones primitivas: (S) sujeto de oración intransitiva, (A) sujeto de oración transitiva y (O) objeto de oración transitiva.

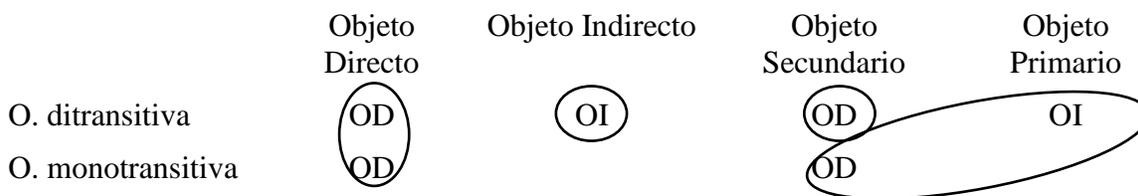
expresa un estado o condición, similar al de algunos adjetivos, y se marca como el objeto de una oración transitiva (O), como se ilustra en la Figura 1.9 (Dixon, 1994: 71-78).

FIGURA 1.9 *Sistema de sujeto escindido, adaptado de (Dixon, 1994: 72)*



En cuanto a la marcación de objeto, el otomí sigue un patrón de objeto primario (Dryer, 1986), o de objeto secundativo en términos de Haspelmath (2005), que puede ser definido en términos de las funciones de objeto directo (OD) e indirecto (OI), como se muestra en la Figura 1.10. De acuerdo con esto, un elemento nominal es objeto primario (OP) si es el objeto indirecto (OI) de una oración ditransitiva o el objeto directo (OD) de una oración transitiva, por otro lado, un objeto secundario (OS) es el objeto directo de una oración ditransitiva (Dryer, 1986: 814).

FIGURA 1.10 *Sistema de objeto primario, adaptado de (Dryer, 1986: 814)*



1.4.2 El número en el nivel oracional

1.4.2.1 La estructura del verbo en otomí

En los siguientes párrafos describo algunos aspectos básicos de la cláusula simple en el otomí de San Andrés Cuexcontitlán con base en los datos lingüísticos de Lastra (1989, 1992) y en los modelos descriptivos que plantean Palancar (2009) y Hernández Green (2015) para el otomí de San Ildefonso y de Acapulco, respectivamente.

El verbo en el otomí de San Andrés Cuexcontitlán tiene dos componentes primarios. En el primero se marca la concordancia de sujeto través de un conjunto de proclíticos²³ que codifican la persona gramatical, así como los valores de tiempo, aspecto y modo (P&TAM). El segundo componente es el tema verbal que contiene la información del predicado léxico, los sufijos de objeto (-2), los enclíticos de persona o enfáticos (=4) y los de número (=5), como se muestra a continuación.

TABLA 1.8 *Plantilla verbal del otomí de SAC, adaptado de Hernández Green, 2015: 29)*

Proclítico				Verbo						
5=	4=	3=	2=	1-		-1	-2	=3	=4	=5
Aspecto secundario	Tiempo	Modo; Persona	Aspecto; Mov. Asociado; registro	Prefijos	Tema verbal	Aplicativo	Objeto	Enclíticos adverbiales	Persona	Número

²³ En los datos de Lastra (1989, 1992) sobre el otomí de San Andrés Cuexcontitlán no se hace la distinción entre afijos y clíticos. En los datos de esta investigación haré esta distinción en la medida de lo posible con base en los criterios que empleó Palancar (2009: 50-53) para el otomí de San Ildefonso Tultepec. Este autor considera como afijos los morfemas con poco peso fonológico que siempre aparecen ligados a una unidad mayor. Desde un punto de vista fonológico, son unidades no silábicas, las sílabas con afijos no portan acento en otomí y son susceptibles a las reglas fonéticas que se aplican a la palabra fonológica de la que forman parte. Los afijos tienen una función típicamente gramatical y la categoría funcional sobre la que operan es de naturaleza léxica. Por otro lado, aunque los clíticos tampoco tienen independencia fonológica, a diferencia de los afijos, son unidades silábicas con tono inherente y no son susceptibles a las reglas fonéticas que se aplican en la palabra fonológica de la que forman parte. Además, los clíticos tienen diversas funciones en otomí y no operan dentro de un dominio categorial concreto, pues pueden ser de índole léxica, frástica e incluso el oracional.

En los ejemplos (2a)-(2c) el núcleo verbal, *ní* ‘tener’, recibe tres proclíticos diferentes que codifican la concordancia de sujeto (dependiente) con cada una de las tres personas gramaticales en singular, así como información de tiempo, aspecto y modo (P&TAM). La función sintáctica de sujeto en *ñható* también puede expresarse por medio de frases nominales y pronombres, como se muestra en los ejemplos (2d) y (2e), cuyos sujetos están codificados con la frase nominal *ra pháni* ‘el caballo’ y el pronombre de tercera persona singular *gegé*, respectivamente.

(2) a. *Di ní na ra yó*
di = ní nʌ ɾʌ yóʔ
1.PRES = tener este ART.SG perro
 ‘Tengo un perro’

b. *Gi ní na ra yó*
gi = ní nʌ ɾʌ yóʔ
2.PRES = tener este ART.SG perro
 ‘Tienes un perro’

c. *I ní na ra yó*
i = ní nʌ ɾʌ yóʔ
1.PRES = tener este ART.SG perro
 ‘Tiene un perro’

d. *Ra fáni bi het’arí*
ra pháni bi = hetʔarí
ART.SG caballo 3.CONT = correr
 ‘El caballo está corriendo’

e. *Gégé bo núk’í mándé a tói*
gegé bo = nú - kʔí mándé a tói
3.PRON 3.PRES = ver - 2.OBJ ayer LOC mercado
 ‘Él te vio ayer en el mercado’

(Otomí de SAC, adaptado de Lastra, 1989: 60)

En el segundo componente del verbo se presentan los sufijos de objeto²⁴ (-2) que pueden codificar el Paciente, el Receptor o Beneficiario y el Sujeto en verbos de marcación pacientiva. Como revisamos en el apartado §1.4.1.5, el otomí es una lengua de objeto primario o secundativo, el objeto primario (OP) corresponde al Paciente o Tema de una oración transitiva (OD) o al Benefactivo o Receptor (OI) de una construcción ditransitiva y establece una relación de concordancia con el núcleo verbal. Por su parte, el objeto secundario (OS) es el Paciente o Tema de (OD) de una oración ditransitiva y no establece ninguna relación de concordancia con el verbo. En el ejemplo (3a), el Paciente del verbo transitivo *nú* ‘ver’ (OP) está marcado con el sufijo de objeto de primera persona, *-gí* y el proclítico de número plural inclusivo *=hí* concuerda con este. Asimismo, en (3b), el Receptor o Benefactivo (OP) del verbo bitransitivo *ʔó* ‘pedir’ está codificado en el sufijo de objeto *-hkí* y el proclítico de número plural *=hí* concuerda con él. En contraste, el Tema o Paciente de la oración ditransitiva (3c) codificado en la FN *ya hmí-hí* (OS) no está marcado en el núcleo verbal.

- (3) a. *Gégé bo núgíhu mándé a tói*
 gégé bo = nú - **gí** = **hí** mándé a tói
 3.PRON.SG 3.PRET = ver - **1.OBJ** = **PL.INCL** ayer LOC mercado
 ‘Él nos vio ayer en el mercado’

²⁴ El sufijo *-gí* se emplea para marcar objetos de primera persona; además, este sufijo tiene dos alomorfos: *-kí*, que aparece después de sibilantes y *-hkí*, con verbos que llegan a perder la segunda sílaba; ambos también pueden estar determinados morfológicamente. Es importante señalar que en algunos casos la /i/ final del sufijo puede manifestarse como /e/ si la raíz del verbo tiene /e/. El objeto de segunda persona se marca con el sufijo *-kʔí*, que también presenta dos alomorfos: *-ʔé* cuando la raíz verbal tiene /e/ y *-ʔkí* como resultado de una metátesis tras vocal. La tercera persona se marca con el sufijo *-bí* y también presenta una serie de alomorfos *-pí*, *-hpí(é)*, *-Ø* y *-bí*, condicionados morfofonológicamente (Lastra, 1992: 24).

- b. *Nugége 'bídá 'óhkihu ya hmihu ngóhu*
 Nugége ʔbídá ∅ = ʔó - **hkí** = **hí**
 3.PRON.SG siempre 3.PRES = pedir - **1.OBJ** = **PL.INCL**

y Λ hmĩ = hí nugó = hí
 ART.PL tortilla = PL 1.PRON = PL
 'Él siempre nos pide tortillas'

- c. *Nugége 'bídá 'óhki ya hmihu*
 Nugége ʔbída ∅ = ʔó - **hkí** y Λ hmĩ = hí
 3.PRON-SG siempre 3.PRES = pedir - **1.OBJ** ART.PL tortilla = PL
 'Él siempre me pide tortillas'

Por otro lado, como indiqué en el apartado §1.4.1.5, el otomí tiene un sistema de alineamiento de sujeto escindido (Dixon, 1994: 71-78) que presenta dos tipos de sujeto intransitivo. El primero (S_a) ejerce control en los verbos de actividad y se codifica con los proclíticos de persona, tiempo, aspecto y modo (P&TAM). El segundo tipo (S_o) es afectado por la actividad o expresa un estado o condición y se marca con los sufijos de objeto de una oración transitiva (O). En los ejemplos de (4), el proclítico de P&TAM *dó=* codifica el sujeto (A) de la oración transitiva (4a) y al sujeto de la oración intransitiva (S_a) activa (4b). Por otro lado, el sujeto de la oración intransitiva (4c) no tiene un control de la acción (S_o) y está marcado con el sufijo de objeto *-ʔkí* que también codifica al objeto O de la oración transitiva (4a).

- (4) a. *Do núk'í mánde a tói*
dó = nú - **kʔí** mándé a tói
1.PRET = ver - **2.OBJ** ayer LOC mercado
 'Ayer te vi en el mercado'

- b. *Dó má ra ngú*
dó = má r Λ ngú
1.PRET = ir 3.POS casa
 'Fui a su casa'

c. *Ra ñoo'kí*

ra = jǒ - ʔkí
3.PRES.CONT = alto - 2.OBJ
'Eres alto'

(Otomí de SAC, adaptado de Lastra, 1989: ejemplo 5a, 93, ejemplos 5c-5e, 45)

1.4.2.2 La expresión del número en el verbo

En el segundo componente de la plantilla verbal también se alojan los enclíticos de número (=5), que pueden codificar el número del sujeto o del objeto primario (OP). Este sistema marca tres valores: singular, dual y plural, en el que el número dual refiere a dos entidades distintas del mundo real y el plural alude a tres o más entidades (Corbett, 2000: 20). Además, la primera persona dual y plural tiene dos formas que codifican los rasgos de inclusividad y exclusividad.

La Tabla 1.9 muestra la conjugación del verbo *nú* 'ver' y podemos observar que el número singular no presenta ninguna marca. El enclítico =*bé* indica el número dual exclusivo que refiere al hablante y una persona más, excluyendo al interlocutor (nosotros dos, excepto tú). En la primera persona la forma =*wí* refiere al dual inclusivo (nosotros, el hablante y el interlocutor) y en la segunda y tercera persona marca el número dual. Por otro lado, el enclítico =*hé* indica el número plural exclusivo, el cual refiere al hablante y por lo menos dos personas más, excluyendo al interlocutor (nosotros tres o más excepto tú). En la primera persona el enclítico =*hí* codifica el número plural inclusivo que refiere al hablante, al oyente y al menos a una persona más (nosotros tres o más); y en la segunda y tercera persona el número plural²⁵ (Lastra, 1992: 22).

²⁵ En cambio, en el otomí de Ixtenco se marca el plural con el sufijo *-wí* y sus alomorfos *-mí* y *bí*. El plural se marca con *-mé* y no tiene variantes; *-wí* se usa para el plural inclusivo y para la segunda persona plural. Mientras que la tercera persona plural no lleva sufijo y es igual a la tercera de singular (Lastra, 1997: 36).

TABLA 1.9 *Conjugación del verbo ver en el otomí de SAC, adaptado de Lastra (1992: 22)*

Singular	Dual	Plural
	<i>Di núbé</i>	<i>Di núhé</i>
	di = nú = bé	di = nú = hé
	1.PRES = ver = DU.EXCL	1.PRES = ver = PL.EXCL
	‘Vemos’	‘Vemos’
<i>Di nú</i>	<i>Di núwí</i>	<i>Di núhú</i>
di = nú	di = nú = wí	di = nú = hí
1.PRES = ver	1.PRES = ver = DU.EXCL	1.PRES = ver = PL.INCL
‘Veó’	‘Vemos’	‘Vemos’
<i>Gi nú</i>	<i>Gi núwí</i>	<i>Gi núhú</i>
gi = nú	gi = nú = wí	gi = nú = hí
2.PRES = ver	2.PRES = ver = DU	2.PRES = ver = PL
‘Ves’	‘Ven’	‘Ven’
<i>I nú</i>	<i>I núwí</i>	<i>I núhú</i>
i = nú	i = nú = wí	i = nú = hí
3.PRES = ver	3.PRES = ver = DU	3.PRES = ver = PL
‘Veó’	‘Ven’	‘Ven’

Según Palancar (2013: 108), en la lengua otomí es obligatoria la concordancia de número no singular entre el sujeto de la oración y los enclíticos de número anclados al verbo, como se muestra en los ejemplos de (5) con sujetos en primera persona dual y plural.

- (5) a. *Nugóbé ‘bidá di ‘óhpíbé ya hmī*
 Nugó = **bé** ?bídá di = ?ó - hpí = **bé**
 PRON.1 = **DU.EXCL** siempre 1.PRES = pedir - 3.OBJ = **DU.EXCL**
- ya hmī
 PL tortilla
 ‘Nosotros dos siempre le pedimos tortillas a ella’
- b. *Nugówí ‘bidá di ‘óhpíwí ma hmīgówí gégé*
 Nugó = **wí** ?bídá di = ?ó - hpí = **wí**
 PRON.1 = **DU.INCL** siempre 1.PRES = pedir - 3.OBJ = **DU.INCL**
- ma hmí - gó = wí gégé
 1.POS tortilla - 1.ENF = DU.INCL 3.PRON
 ‘Nosotros siempre le pedidos tortillas a ella’

- c. *Nugóhé 'bidá di 'óhpíhé ma hmihé*
 Nugó = **hé** ?bidá di = ?ó - hpí = **hé**
 PRON.1 = **PL.EXCL** siempre 1.PRES = pedir - 3.OBJ = **PL.EXCL**

ma hmĩ = hé (gégé)
 1.POS tortilla = PL.EXCL 3.PRON
 'Nosotros siempre le pedimos tortillas a ella'

- d. *Nugóhu 'bidá di 'ó'hpíhu ya hmi*
 Nugó = **hí** ?bidá di = ?ó - hpí = **hí**
 PRON.1 = **PL.INCL** siempre 1.PRES = pedir - 3.OBJ = **PL.INCL**

ma hmĩ - gó = hí gégé
 1.POS tortilla - 1.ENF = PL.INCL 3.PRON
 'Nosotros siempre le pedidos tortillas a ella'

(Otomí de SAC, adaptados de Lastra, 1989: 86-90)

Sin embargo, Palancar (2013: 108-109) señala que existen al menos dos contextos en los que es posible no usar el marcador de número no singular sin que esto cause agramaticalidad, debido a que el contexto deja claro que es el mismo sujeto y que es plural. El primero es cuando hay un poseedor en otra parte de la cláusula que se interpreta como referencial con el sujeto, como se muestra en el ejemplo (6a). El segundo contexto en el que la marca de número se puede elidir es en construcciones con oraciones yuxtapuestas, como las que se presentan en (6b) y (6c). En ambos casos, el hecho de que el mismo referente esté dentro del contexto sintáctico restringido de la cláusula hace que el uso de un solo marcador de número sea suficiente.²⁶

- (6) a. *Ma dá kwáhta ma 'benihé*
 ma = dá = kwá^{ht} - a _ = ma 'beni = hé
 PURP.MOT = 1.PST.R = store[3OBJ] - B _ = 1POSS luggage = PL.EXCL
 'We went to store our luggage'

²⁶ Uno de los miembros del Comité Lector pregunta si se ha pensado en la posibilidad de que las marcas de número del otomí sean clícticos y no partículas independientes. Se ha señalado que las marcas de número en la frase nominal tienen un estatus de partícula independiente, mientras que las marcas de número pronominal que se presentan en las construcciones posesivas y en las frases verbales tienen un estatus de clíctico, véase nota 23.

b. *Ntónse dá nángi dá ñóhkhé*
 ntónse dá = náng - i [dá = ñ - hó^hk = hé]
 then 1.PSR.R = get-up - F 1.PST.R = MIDD - prepare = PL.EXCL
 ‘Then we got up to get ready’

c. *Kwándo dí né gahoohka n’aa ma tráhéhé*
 kwándo dí = né - [ga = hõ^hk - a = n’ă
 when 1.PRS.R = want[3OBJ] - 1.PRS.IRR = make[3OBJ] - B = one

ma tráhé = hé]
 1.POSS suit = PL.EXCL
 ‘When we want to make our suits’

(Otomí de San Ildefonso en Palancar (2013: 109))

Además, Palancar (2013: 114) advierte que la concordancia de número no singular entre el sujeto y los enclíticos de número anclados al verbo es obligatoria en la primera y segunda persona y opcional en la tercera. La aparición de las marcas de concordancia de número con sujetos de tercera persona está condicionada por la animacidad del nombre y la especificidad de la referencia, pues sólo se establece la relación de concordancia cuando el nombre es humano y tiene una referencia específica. En el ejemplo (7a), el enclítico plural (=hí) no está presente en el verbo cuando el sujeto de la oración es inanimado *yí cithă* ‘las mazorcas’ y animado *ya ngú* ‘los ratones’. En contraste, en el fragmento de (7b), sí está presente el enclítico de plural =hí en el verbo *ză* ‘comer’, que tiene como sujeto a una entidad humana con referencia específica (los papás del colaborador), mientras que en el verbo que tiene como sujeto a una entidad humana no específica la marca no aparece.

(7) a. *Yú tsithăä níhi bú ni(hú) ni ‘a hítsí este mah ke xo i nditsi(hú) ya ngú*
yí **tsithă** níhí Ø = bî = **(hí)** = ni ?a hítsĩ
 PL **mazorca** también 3.PRES = estar = **(PL)** = aquí LOC tapanco
 este mäh ke fo i = ndítsi = **(hí)** **ya ngú**
 este más que pues 3.PRES = comer = **PL PL ratón**
 ‘Las mazorcas están en el tapanco aunque se las comen mucho las ratas’

- b. *Asta (ma tá=hé) bi zaahú ya nata ná bi zaahú kwalkiera ya páxí ya xo mí núhú ta xo dúnthi ya khání bi bidü ná*

asta (ma tá=hé) bi = ză = hí y Λ na ta = n \acute{a}
 hasta (mis papás) 3.PRET = comer = PL PL nopal.crudo = CIT

bi = ză = hí kwalkiera y Λ páfí y Λ fo
 3.PRET = comer = PL cualquiera PL hierba REL pues

mí = nú - Ø = hí ta fo dúnthi y Λ khání
 3.IMPF = ver - 3.OBJ = PL incluso pues muchos PL persona

bi = bidũ = _ = n \acute{a}
 3.PRET = morir = _ = CIT

‘Hasta (mis papás) comieron nopales crudos, comieron cualquier yerba ya pues lo que veían, incluso pues mucha gente se murió’

(Otomí de SAC, adaptado de Lastra, 1992: 347, 360)

Según Palancar (2013), la concordancia con nombres animados sólo se presenta cuando el verbo es recíproco. En el ejemplo (7c), el verbo *ʔó* ‘estar.abajo’ no presenta el enclítico de número plural (=hí), debido a que su sujeto es un nombre animado; sin embargo, en el verbo *tɥ* ‘pelear’ que tiene como sujeto al mismo referente sí presenta la marca de plural al tratarse de una construcción recíproca. Es importante advertir que el ejemplo proviene de una narración y sólo puede ser considerado como evidencia positiva, es decir, señala que los hablantes pueden no usar las marcas de número con sujetos animados y que pueden usarlas con sujeto animados en construcciones recíprocas, pero no es una prueba de la obligatoriedad de las mismas.

(7) c. *Ra ngüü ra áni wá to ‘uutébí ya áni xi to ‘uutébík’í ma ra k’u i ‘ó(hú) ‘a haqi porke dantühú*

ɾʌ ngǔ ɾʌ áni wá to ʔíté = bí yʌ ʔáni
 SG casa SG pollo aquí 1.PRFT hacer = 3.OBJ PL pollo

ʃi = to = ʔíté - bí = kʔí ma ɾʌ kʔi
 también = 1.PRET = hacer - 3.DAT = 3PL otra IND.SG REL

i = ʔó = (hí) ʔa hʔi porke
 3.PRES = estar.abajo = (PL) LOC piso porque

da = n - tũ = hí
 3.PRES = MED - pelear = PL

‘Aquí es la casa de los pollos, les hice los nidos a los pollos y también les hice los que están abajo porque (si no) se van a pelear’

(Otomí de SAC, adaptado de Lastra, 1992: 344)

Además de las restricciones antes mencionadas, la organización de la marcación de número en el verbo presenta otro tipo de complejidades pues, como vimos en la Tabla 1.8, la estructura morfológica del verbo sólo tiene una posición designada para el enclítico de número y este puede concordar con el sujeto o el objeto primario. La Tabla 1.10 presenta las distinciones de número disponibles y las posibles interacciones entre el sujeto y el objeto primario (OP) en el otomí de San Andrés Cuexcontitlán, con base en los datos de (Lastra, 1989: 82-92) y el esquema planteado por Hernández Green, (2015: 246) para el otomí de Acazulco. En los recuadros se indica el enclítico correspondiente según el tipo de construcción. Las celdas marcadas como MED son combinaciones de sujeto y objeto primario (OP) que describen situaciones reflexivas o recíprocas. Y los recuadros sombreados corresponden a combinaciones sujeto y objeto primario (OP) que no son posibles en el otomí, debido a que no se permite que el sujeto y el objeto primario (OP) tengan el mismo referente.

TABLA 1.10 *Las marcas de número posibles en el sujeto y objeto primario del otomí de SAC*

S\OP	1SG	2SG	3SG	1DU.EXCL	1DU.INCL	2DU	3DU	1PL.EXCL	1PL.INCL	2PL	3PL
	Ø	Ø	Ø	-bé	-wí	-wí	-wí	hé	hí	-hí	hé
1 SG	MED	Ø	Ø			-wí	Ø			-hí	Ø
2 SG	Ø	MED	Ø	-bé			Ø	-hé			Ø
3 SG	Ø	Ø	Ø	-bé	-wí	-wí	-wí	-hé	-hí	-hí	-hí
1DU.EXCL		Ø _{+be}	Ø _{+be}	MED		Ø _{+be}	Ø _{+be}			Ø _{+be}	Ø _{+be}
1DU.INCL			Ø _{+wi}		MED		-wí				Ø _{+wi}
2DU	Ø _{+wi}		Ø _{+wi}	-bé		MED	-wí	-hé			Ø _{+wi}
3DU	Ø	Ø	Ø	Ø _{+wi}	-wí	-wí	-wí	Ø _{+wi}	Ø _{+wi}	Ø _{+wi}	Ø _{+wi}
1PL.EXCL		Ø _{+hé}	Ø _{+hé}			Ø _{+hé}	Ø _{+hé}	MED		Ø _{+hé}	Ø _{+hé}
1PL.INCL			Ø _{+hí}				Ø _{+hí}		MED		-hí
2PL	Ø _{+hí}		Ø	-bé			FD	-hé		MED	-hí
3PL	Ø	Ø	Ø	-bé	-wí	-wí	-wí	-hé	-hí	-hí	-hí

Cuando el sujeto y el objeto primario tienen número no singular surgen ambigüedades en la marcación de número. Al respecto, algunos autores señalan que este conflicto se resuelve en el contexto discursivo (Lastra, 1992) y otros sugieren que en esta lengua se marca el número de manera prioritaria en la primera persona por encima de la segunda y tercera persona (Hernández Green, 2015: 246); sin embargo, no hay un estudio específico sobre este tema.

1.4.3 El número en el nivel frástico

1.4.3.1 La estructura de la frase nominal

La estructura mínima de la frase nominal en otomí consta únicamente de un núcleo nominal, como se muestra en los ejemplos de (8). Sin embargo, en la mayoría de las frases el núcleo va acompañado de determinantes que delimitan su referencia: posesivos, cuantificadores, numerales, demostrativos.

- (8) a. *A hiádi ya pibuuña tsuutsíyo'*
a **hiódi** ya pi - bįná **tsítsíyo?**
LOC **besana** ya 3.INCOATIVO - existir **bagazo**
'En la besana hay bagazo'

- b. *Na ra fáni bit si 'a baathá*
 n Λ r Λ pháni bi = tsi ʔaʔ bǎthá
 DEM ART.SG caballo 3.CONT.VIS = comer LOC llano
 'El caballo come en el llano'
- c. *Na ra mésa ga zǎ*
 n Λ r Λ mésa ga zǎ
 DEM ART.SG mesa de maderas
 'La mesa de madera'

En la frase nominal los determinantes siempre preceden al núcleo. En la Tabla 1.11 presento el esquema general de la frase nominal con base en los datos del otomí de SAC (Lastra, 1989, 1992) y en la propuesta de Palancar (2009) para el otomí de San Ildefonso, y de Hernández Green (2015) para el otomí de Acapulco.

TABLA 1.11 *Posición de los determinantes de la FN en el otomí de SAC*

Posición 3	Posición 2	Posición 1	
Demostrativos	Indefinido	Marcas de número	
= n Λ DEM.1.SG	'na IND.SG	r Λ SG	
= yǐ DEM.1.PL	'ra IND.PL	y Λ PL	
	Definido	Poseivos	
= ni / kʔ Λ DEM.2.SG	nu DEF	ma 1. POS	
= kʔǐ DEM.2.PL		ri 2. POS	
		(v) 3. POS	Nombre
	Numerales		
= (n)ge / gen kʔ Λ	n'a 'uno'		
DEM.3.SG	yoho 'dos'		
	hñü 'tres'		
	Cuantificadores		
	gothó todo(s)		
	dúnthí mucho(s)		
	čǐthó mucho(s)		

En la primera posición a la izquierda del núcleo nominal están las marcas de número y de posesión, cuya distribución es complementaria. En la siguiente posición se hallan las

marcas de definitud *nu* e indefinitud *na*, así como los numerales y los cuanticadores. Es importante mencionar que estos elementos pueden coaparecer con las marcas de número de la primera posición. Por último, en la tercera posición se encuentran los demostrativos, que pueden presentarse en conjunto con algún elemento de la posición dos y las marcas de número. A continuación, describo cómo se expresa el número en cada una de estas construcciones.

1.4.3.2 Las marcas de número

En otomí los nombres no tienen flexión de número, la expresión de este rasgo se hace de manera sintáctica, mediante marcas que preceden al nombre:²⁷ $\epsilon\lambda$ para el número singular y $y\lambda$ para el plural, como se muestra en los ejemplos (9a) y (9b). Lastra (1992) señala que $y\lambda$ es el determinante tanto para plural como dual y proporciona el ejemplo (9c); sin embargo, es posible que en este caso el valor dual esté dado por el numeral *yóho* ‘dos’ y no por la marca de número como tal. En relación con esta observación, otros autores han señalado que el valor de dual en otomí sólo está presente en el número pronominal, es decir, en los pronombres personales y en las construcciones de posesivo (Palancar, 2013).

- (9) a. *Ra ngüü*
 $\epsilon\lambda$ $ng\ddot{u}$
 SG casa
 'la casa'

²⁷ El otomí del estado de Querétaro presenta esta misma característica: el artículo no expresa género, únicamente número (singular/plural). El paradigma de artículos se compone de cuatro formas, dos corresponden al artículo determinado: *ar*, para singular y *ya*, para plural; y dos artículos indeterminados: en singular, *'nar* y en plural *'ra* ya (Hekking & Andrés de Jesús, 1984: 20). En otra comunidad de ese estado, San Ildefonso Tultepec, Palancar documenta la forma *ar* como determinante singular y *ya* como plural (2009: 92), Lastra encuentra un comportamiento similar en el otomí de Ixtenco, Tlaxcala, que tiene el artículo *ra* para singular y *yi* para plural (Lastra, 1997: 28).

- b. *Ya ngüü*
yΛ ngü
PL casa
 'las casas'
- c. *Ya hóho ngüü*
yΛ yóho ngü
PL dos casa
 'las dos casas'

(Otomí de SAC, adaptado de Lastra, 1992: 18)

1.4.3.3 El número en las construcciones posesivas

En el nivel frástico, las construcciones posesivas siguen un patrón de marcación en el núcleo, como se muestra en los ejemplos de (10), donde la marca morfológica de persona gramatical modifica al núcleo de la frase nominal, que es el elemento poseído.

- (10) a. *Ma ñá*
ma ñá
1.POS cabeza
 'Mi cabeza'
- b. *Ra ñá*
rΛ ñá
3.POS cabeza
 'Su cabeza'
- c. *Ma ngüü*
ma ngü
1.POS casa
 'Mi casa'
- d. *Ma ngüü*
rΛ ngü
3.POS casa
 'Su casa'
- e. *Ma t'ú*
ma t'í
1.POS hijo
 'Mi hijo'
- f. *Ra t'ú*
rΛ t'í
3.POS hijo
 'Su hijo'

(Otomí de SAC, adaptado de Lastra, 1992: 18)

Las marcas de número en estas construcciones refieren al poseedor, no al objeto poseído, y pueden expresar tres valores: singular, dual y plural, como se muestra en los siguientes ejemplos (11). Si los objetos poseídos son dos o más esto se indica de manera perifrástica: *k'í yóho məkhñi* 'mis dos metates' *k'í məkhñi* 'mis metates'.

Singular	Dual	Plural
(11)	<i>Ma ngubé</i> mą ngú = bé 1.POS casa = DU.EXCL 'Nuestra casa'	<i>Ma ngúhé</i> mą ngú = hé 1.POS casa = PL.EXCL 'Nuestra casa'
<i>Ma ngu</i> mą ngú 1.POS casa 'Mi casa'	<i>Ma nguwí</i> mą ngú = wí 1.POS casa = DU.INCL 'Nuestra casa'	<i>Ma nguhú</i> mą ngú = hí 1.POS casa = PL.INCL 'Nuestra casa'
<i>Ri ngu</i> ri ngú 2.POS casa 'Tu casa'	<i>Ri nguwí</i> ri ngú = bé 2.POS casa = DU.EXCL 'Su casa'	<i>Ri nguhú</i> ri ngú = hé 2.POS casa = PL.EXCL 'Su casa'
<i>Rá ngu</i> ra ngú 3.POS casa 'Su casa'	<i>Yá nguwí</i> ya ngú = wí 3.POS casa = DU.INCL 'Su casa'	<i>Ya nguhú</i> ya ngú = hí 3.POS casa = PL.INCL 'Su casa'

(Otomí de SAC, adaptado de Lastra, 1992: 18)

1.4.3.4 Los numerales y los cuantificadores

Los numerales coaparecen con los determinantes de número singular y plural cuando tienen un uso cardinal. En el ejemplo (12a) se observa que el numeral *n*_Λ 'uno' precede al artículo singular *r*_Λ y en (12b) el numeral *yóhó* 'dos' precede al numeral plural *y*_Λ. Según Palancar (2009: 110), cuando los numerales son empleados para hacer referencia indefinida o aproximada en la que no es importante el número exacto, aparecen sin los determinantes de número,²⁸ como se muestra en (12c). Asimismo, los cuantificadores de la lengua otomí coaparecen con los determinantes de número, como se muestra en los ejemplos (12d) y (12e).

²⁸ En los estudios sobre cardinales existe una discusión en torno a si los cardinales tienen lecturas exactas o inexactas. Por otro lado, Ionin y Matushansky (2018) señalan que los cardinales requieren una FN semánticamente singular; no obstante, esta puede codificarse morfológicamente como singular o plural. De modo que cuando se presenta una marca plural en la FN, así como en el predicado y los determinantes se

(12) a. *Dí ní na ra fáni*

dí = ní nΛ rΛ pháni
1P.PRES = tener uno ART.SG caballo
'Tengo un caballo'

b. *Dí ní yóhó ya boo dóró*

dí = ní yóhó yΛ bö dóró
1P.PRES = tener dos ART.PL negro toro
'Tengo dos bueyes negros'

c. *Ndí p̄'sa komo 'ret'a jeya*

n - dí = p̄ʔsa - Ø - a komo ʔretʔa jeya
IMP - 1.PRES = poseer.A - 3obj - D como diez año
'Tenía como diez años'

d. *Gǎthó k'ú ya bäätsí i hníní*

Gǎthó kʔí yΛ bǎhtsí i - hníní
Todos DEM.PL ART.PL niño 3ªP.PRES - enfermo
'Todos esos niños están enfermos'

e. *Dúnthi yu ya bäätsí i hníní*

Dúnthi yí yΛ bǎhtsí i - hníní
Muchos DEM.PL ART.PL niño 3ªP.PRES - enfermo
'Muchos de estos niños están enfermos'

(Otomí de SAC, adaptado de Lastra, 1992: 133)

1.4.3.5 La expresión del número en los demostrativos

La Tabla 1.12 muestra el paradigma de demostrativos planteado por Lastra (1992), que tiene un sistema de número con dos valores, singular y plural. Sin embargo, no se ha realizado un estudio específico de estas formas para establecer el funcionamiento semántico y sintáctico de este paradigma en el otomí de SAC.

presenta un caso de concordancia semántica que refleja la pluralidad semántica de las FNs que contienen un cardinal. Con base en estas observaciones, parece ser que Palancar plantea que en otomí se presenta una relación de concordancia semántica entre los miembros de la FN cuando los cardinales tienen una lectura exacta. Sin embargo, el autor no proporciona más datos al respecto.

TABLA 1.12 *Paradigma de demostrativos en el otomí de SAC, adaptado de (Lastra, 1992: 20)*

Singular	Plural
nΛ ‘este’	
nu-nΛ ‘este’	
ni ‘ese’	yi ‘esos’
nukʔΛ ‘ese’	
kʔa ‘ese’	
(n)ge ‘aquel’	kʔi ‘aquellos’
genkʔa ‘aquel’	

1.4.3.6 La expresión del número en los pronombres personales

En los pronombres personales también se marcan tres valores, singular, dual y plural. Asimismo, en la primera persona dual y plural se tienen marcas específicas que expresan los valores de inclusividad y exclusividad, como se muestra enseguida.

TABLA 1.13 *Pronombres personales en el otomí de SAC, adaptado de (Lastra, 1992: 19)*

Persona	Singular	Dual	Plural
		nugó-bé (excl.)	nugó-he (excl.)
1 ^a	nugó	nugó-wí (incl.)	nugó-hí (incl.)
2 ^a	nukʔígé	nukʔígé-wí	nukʔígé-gí
3 ^a	gégé	nugégé-wí	nugégé-hí

1.4.3.7 La expresión del número en los adjetivos

Existe una discusión en torno a la existencia de la categoría adjetivo en otomí. La mayoría de los investigadores han descrito los conceptos de propiedad como adjetivos, pese a advertir un comportamiento sintáctico verbal (Hekking & Andrés de Jesús, 1984; Lastra, 1997; Voigtlander & Echegoyen, 1985). Por otro lado, Palancar (2006: 337) prima los criterios sintácticos y con base en estos argumenta que no existe la categoría de adjetivo en la lengua

otomí, pues los conceptos de propiedad se expresan con compuestos nominales (13a), nombres (13b) y verbos. Dentro de estos distingue entre verbos “activos” (13c) y “estativos” (13d), cuyo argumento se codifica como objeto morfológico y reciben un prefijo nasal condicionado morfológicamente.

- (13) a. **dótá** - m'asmé
ser.grande - servilleta
 ‘Mantel’
- b. nǒ = r ndǒ || Ø = ar mǎ'jǒ
 DEF = SG hombre **3.PRES** = SG cura
 ‘El hombre es un cura’
- c. Ma t'ó **már** = **dótá** / mǎ
 1POS rama **3.IMP.PC** = **ser.grande** / **ser.fuerte**
 ‘Mi rama era grande/fuerte’
- d. Xi = **hěts'i** - ‘í
 (3.)PERF = **ser.alto** - 2.OBJ
 ‘Tú eres alto’

(Palancar, 2006: 327, 334, 348)

En el otomí occidental los conceptos de propiedad también se pueden codificar con compuestos (14a),²⁹ nombres (14b) y verbos. Dentro de los verbos hay diferentes conjugaciones que parecen corresponder con las agrupaciones propuestas por Palancar (2006): en (15) hay un verbo “activo” cuyo argumento con función de sujeto es codificado con el proclítico P&TAM. Por otro lado, en (16) se halla un verbo “estativo” que se conjuga con la tercera persona del presente continuativo *ɾʌ* y su argumento con función de sujeto se codifica con los sufijos de objeto.

- (14) a. *Ra botzanga*
 Ra **bo** - tzanga
 ART.SG **negro** - lagartija
 ‘La lagartija negra’

²⁹ Lastra (1992) también señala que los adjetivos en otomí ocupan una posición prenominal: *ɾʌʔyóti zá* ‘el árbol seco’.

- b. *Dra 'behñagó*
 DrΛ = ʔbéhphΛ - gó
 1PRES.CONT = **mujer** - 1ENF
 'Yo soy mujer'
- (15) a. *Dra doot'ii*
 drΛ = dõtʔĩ
 1.PRES.CONT = ser.chaparro
 'Yo soy chaparro'
- b. *Gra doot'ii*
 grΛ = dõtʔĩ
 2.PRES.CONT = ser.chaparro
 'Tú eres chaparro'
- c. *Ra doot'ii*
 rΛ = dõtʔĩ
 3.PRES.CONT = ser.chaparro
 'Él es chaparro'
- (16) a. *Ra hyytsigígó*
 rΛ = hĩtsi - gí - gó
 3.PRES.CONT = ser.alto - 1OBJ - 1ENF
 'Yo soy alto'
- b. *Ra hyytsik'í*
 rΛ = hĩtsi - kʔí
 3.PRES.CONT = ser.alto - 2OBJ
 'Tú eres alto'
- c. *Ra hyytsi*
 rΛ = hĩtsi - Ø
 3.PRES.CONT = ser.alto - 3OBJ
 'Él es alto'

(Otomí de SAC, adaptado de Lastra, 1992: 21)

Los conceptos de propiedad que se expresan mediante estructuras verbales reciben la misma morfología de número que el resto de los verbos y distinguen tres valores: singular, dual y plural. En el número dual se agrega =bé para exclusivo y =wí para inclusivo en la 1ª

persona, =wí en la segunda y tercera persona³⁰. El número plural se forma con =hé para exclusivo y =hí para inclusivo de 1ª persona, =hí para la segunda y tercera persona, como se muestra en la Tabla 1.14 con el concepto de ‘alto’.

TABLA 1.14 *Conjugación del concepto de propiedad ‘alto’ en el otomí de SAC, adaptado de (Lastra, 1992: 21)*

Persona	Dual	Plural
1ª	ɾΛ-hĩtsi-gí-bé 3.PRES.CONT-ser. alto-1OBJ-DU.EXCL	ɾΛ-hĩtsi-gí-hé 3.PRES.CONT-ser. alto-1OBJ-PL.EXCL
	ɾΛ-hĩtsi-gí-wí 3.PRES.CONT-ser. alto-1OBJ-DU.INCL	ɾΛ-hĩtsi-gí-hí 3.PRES.CONT-ser. alto-1OBJ-PL.INCL
2ª	ɾΛ-hĩtsi-kʔi-wí 3.PRES.CONT-ser. alto-2OBJ-DU	ɾΛ-hĩtsi-kʔi-hí 3.PRES.CONT-ser. alto-2OBJ-PL
3ª	ɾΛ-hĩtsi-Ø-wí 3.PRES.CONT-ser. alto-3OBJ-DU	ɾΛ-hĩtsi-hí 3.PRES.CONT-ser. alto-3OBJ-PL

En la sección §1.4 pudimos ver que el otomí de SAC y de JV pertenecen a la misma variante en las distintas clasificaciones expuestas. El otomí es una lengua tonal, con un sistema de marcación en el núcleo, de sujeto escindido y objeto primario. Además, el ñható tiene un sistema de marcación de número nominal y otro pronominal. El número nominal se codifica mediante partículas independientes que anteceden al nombre y expresan el valor singular *ra* y plural *ya*. El número pronominal se presenta en las frases verbales y las construcciones posesivas y cuenta con una serie de enclíticos que codifican tres valores singular, dual y plural, así como de inclusividad y exclusividad.

En suma, a lo largo de este capítulo, revisé algunas de las propuestas teóricas más importantes en torno al estudio del contacto lingüístico y el bilingüismo. Posteriormente, planteé el marco conceptual que guiará el análisis del contacto lingüístico en esta

³⁰ Con otros verbos en lugar de la forma –wí y –hí para el dual y plural de tercera persona se agrega el alomorfo –mí (Lastra, 1992: 20).

investigación. Por último, proporcioné información acerca de la marcación de número en la lengua otomí, en el nivel frástico y oracional, pues esta información me permitirá formular hipótesis en torno a los fenómenos que son objeto de estudio de los capítulos tres y cuatro, a saber: la concordancia plural en el español de contacto y la marcación de número en los nombres patrimoniales y los préstamos del español en el otomí. En el siguiente capítulo, presentaré la metodología empleada en la conformación de los grupos de colaboradores y las herramientas para la obtención del material lingüístico. Asimismo, ofreceré una breve descripción de la situación sociolingüística de San Andrés Cuexcontitlán y Jiquipilco el Viejo.

CAPÍTULO 2. DESCRIPCIÓN DE LA METODOLOGÍA Y LA REALIDAD SOCIOLINGÜÍSTICA DE LAS COMUNIDADES BILINGÜES OTOMÍ-ESPAÑOL

En las páginas del segundo capítulo presento la metodología general de esta investigación, que sigue los principios de la sociolingüística variacionista. En principio, describo el tipo de muestreo que utilicé para conformar los grupos de colaboradores, así como las herramientas para la obtención de datos lingüísticos y sociales. Más adelante, explico los procedimientos que empleé para analizar los datos de forma cualitativa y cuantitativa. En la segunda parte del capítulo expongo de forma breve la situación sociolingüística del otomí y el español en San Andrés Cuexcontitlán y Jiquipilco el Viejo, a partir de la descripción de diferentes grupos de hablantes monolingües y bilingües otomí-español, sus características lingüísticas y el uso que hacen de sus lenguas.

2.1 DESCRIPCIÓN DE LA METODOLOGÍA

Debido a que los dos fenómenos lingüísticos que analizo en esta investigación son variables, empleé una metodología en tiempo aparente³¹ (Labov, 1996: 123). Elegí este método porque considero que tiene la capacidad de explicar los fenómenos de variación y cambio, tanto en comunidades monolingües como en situaciones de contacto. A esto, sumo el método que siguen Poplack & Levey (2010) y Torres Cacoullos & Travis (2018a) para determinar si un fenómeno variable tiene su origen en el contacto. Este se basa en una serie de contraposiciones sobre el comportamiento de un fenómeno lingüístico en la variedad de contacto y en otras variedades de referencia. La primera comparación se realiza entre el estado actual y uno previo de la misma variedad de contacto. En la segunda contraposición

³¹ Corresponde, en especial, a la distribución de las variables lingüísticas por niveles de edad (Labov, 1996: 99).

se revisa el comportamiento del fenómeno en variedades monolingües de las lenguas en contacto, con el objetivo de encontrar puntos de divergencia estructural hacia donde se podría dirigir un cambio. Con base en este par de oposiciones se podría afirmar que un cambio es debido al contacto si la variedad bilingüe es diferente a su contraparte sin contacto y similar al punto de referencia monolingüe de la otra lengua con la que está en contacto. En caso contrario, todavía se podría probar dicha hipótesis al realizar una tercera comparación entre las variedades habladas por los bilingües, para obtener evidencia más directa y fuerte en contra o a favor de la convergencia estructural. Nótese que la serie de contraposiciones que proponen nos remite a la conjunción del tiempo aparente y el tiempo real antes propuesta.

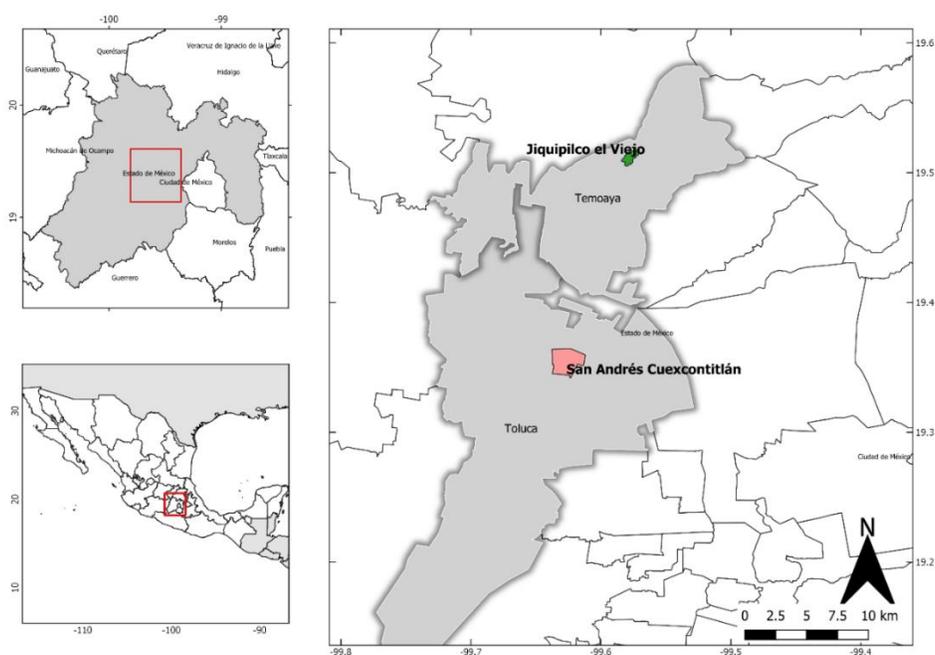
2.1.1 Población y muestra

Incorporé las observaciones en tiempo aparente en dos dimensiones: una intercomunitaria, mediante el contraste de dos comunidades con diferentes situaciones sociolingüísticas, y otra intracomunitaria, a partir de la inclusión de diferentes grupos etarios. Las observaciones de tiempo aparente intercomunitario se basan en la comparación de dos comunidades bilingües español-*n̄ható* del Estado de México con situaciones sociolingüísticas distintas. Tomé esta decisión porque esperaba que la implementación de esta oposición permitiría explorar la posibilidad de encontrar diferentes soluciones o recursos para la expresión del número en los dos fenómenos variables que son objeto de estudio o, al menos, distintos grados de difusión.³²

³² Si bien para algunos fenómenos de contacto, como el sistema pronominal del español, se ha constatado soluciones similares en diferentes situaciones de contacto, también hay casos en los que comunidades bilingües en las que están implicadas las mismas lenguas construyen soluciones diferentes para los mismos fenómenos (Palacios, 2017: 9-10).

La primera localidad en la que trabajé es San Andrés Cuexcontitlán, una comunidad de tipo interfase rural-urbana³³, situada en el municipio de Toluca (Figura 2.1). Se ubica en las siguientes coordenadas geográficas: latitud norte 19° 35'53'', longitud oeste 99° 62'62'', con una altura de 2 596 metros sobre el nivel del mar. Su perímetro es de 20 938.33 m y su extensión territorial de 134 409 km². Por su parte, la segunda comunidad, Jiquipilco el Viejo, es una población rural del municipio de Temoaya, cuyas coordenadas son: latitud norte 19° 30'59'', longitud oeste 99° 34'19'' y una altura de 2965 metros.

FIGURA 2.1 *Ubicación de San Andrés Cuexcontitlán y Jiquipilco el Viejo*



³³ El criterio que se ha empleado para considerar a una comunidad como rural es el número de habitantes. El INEGI define como rural a aquella comunidad que tiene 2 500 habitantes o menos. Por otro lado, Gustavo Garza considera como ciudad a los asentamientos poblacionales con más de 15 000 habitantes (Garza Villarreal, 2003: 33). No obstante, este criterio resulta insuficiente, pues no alcanza a comprender la complejidad de lo rural. Para esta investigación me ceñiré a los criterios de la CONAPO, que establecen dos dimensiones para definir lo rural: 1) dimensión I: predominio de los usos del suelo extensivo-suelo cultivado, suelo natural y suelo construido; 2) dimensión II: tamaño y densidad poblacional. De acuerdo con dichos parámetros, una comunidad interfase rural-urbana tiene entre 5 001 y 50 000 habitantes; una densidad poblacional de 1-150 habitantes por km²; y 10-50% del territorio construido (Larralde & González, 2013). San Andrés Cuexcontitlán tiene 35 011 habitantes y una densidad poblacional de 260 habitantes por km², índices que posicionan a dicha comunidad como una localidad de interfase rural-urbana.

Decidí trabajar en la comunidad de San Andrés Cuexcontitlán porque soy originaria y miembro de la comunidad. Mi familia paterna es originaria y vive en esta localidad y mis abuelos, tíos, padre y primos de más edad son hablantes de *ñathó*. Yo viví 17 años en San Andrés Cuexcontitlán, hasta que tuve que migrar a la Ciudad de México para continuar con mis estudios de nivel superior. Mi interés por el contacto lingüístico entre el español y el otomí en mi comunidad surgió en la infancia, constantemente me preguntaba por qué mi papá no había querido enseñarme a hablar su lengua materna y por qué cuando mis tías hablaban otomí usaban elementos del español, me preguntaba si eso pasaba porque no existía una palabra equivalente en otomí o si lo hacían con la intención de que nosotros, los niños que no hablábamos otomí, pudiéramos entender parte de la conversación.

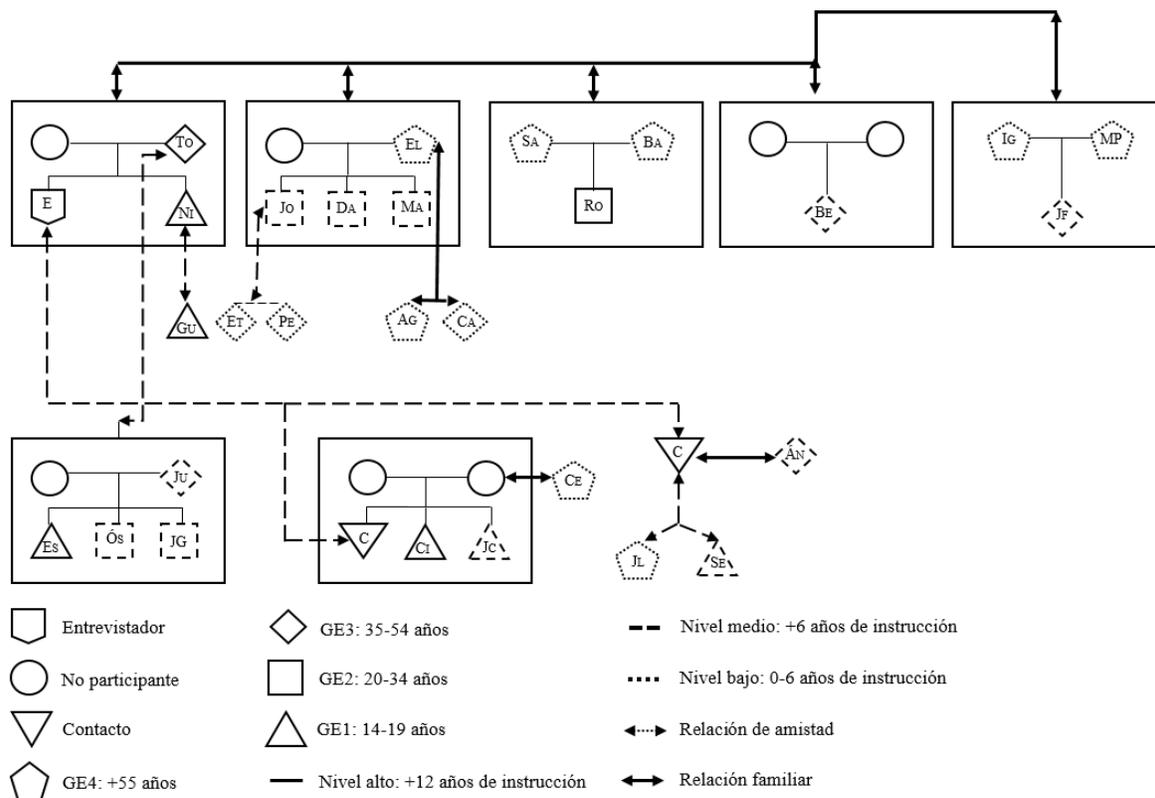
Este interés se mantuvo discreto pero constante hasta que comencé a estudiar la licenciatura en Lengua y Literaturas Hispánicas en la Universidad Nacional Autónoma de México. Durante el primer año conocí el quehacer de la lingüística y me generó una especial fascinación la sociolingüística y, por supuesto, el contacto de lenguas. Por esta razón, en 2014, comencé a desarrollar mi tesis de licenciatura con la guía del Dr. Pedro Martín Butragueño, cuya finalidad era describir la situación sociolingüística del contacto otomí-español en San Andrés Cuexcontitlán y analizar la neutralización de los rasgos de género y número en los clíticos de objeto directo del español.

Mi predilección por la comunidad de Jiquipilco el Viejo también responde a intereses e inquietudes personales. Como mi padre trabajaba en la primaria de esa comunidad, tuve la oportunidad de estudiar en ese centro educativo y de observar que, aunque en Jiquipilco el Viejo también se hablaba otomí, la lengua y su uso parecían distintos al de San Andrés Cuexcontitlán. Estas primeras anotaciones fueron las que, en 2019, me motivaron a estudiar la comunidad de Jiquipilco el Viejo como parte de mi proyecto de investigación doctoral.

Uno de los objetivos era comparar las dos comunidades a la luz del contacto lingüístico y social existente entre los hablantes de español y otomí, con la expectativa de encontrar diferentes soluciones para la expresión del número en los dos fenómenos variables que son objeto de estudio o distintos grados de difusión.

Como mencioné anteriormente, recolecté los datos de San Andrés Cuexcontitlán en 2014 como parte de mi tesis de licenciatura (Avelino Sierra, 2017). Utilicé el método de bola de nieve o redes sociales para conformar el grupo de colaboradores de mi comunidad. Este método consiste en contactar y trabajar con algunos miembros que, a su vez, te presentan con otros y estos a otros (Moreno Fernández, 1998: 312). Inicé a trabajar con los miembros de mi familia nuclear (padre y hermano), después con algunos miembros de mi familia extensa (tíos, tías, primos y primas) y, posteriormente, con vecinos, compadres, amigos y familiares de amigos hasta completar el grupo de 28 personas que se esquematiza en seguida.

FIGURA 2.2 Red de colaboradores en San Andrés Cuexcontitlán



En la imagen se observan las relaciones que guardan los colaboradores entre sí y las redes con distinta densidad (Lastra & Martín Butragueño, 2000). Cada uno de los colaboradores está representado por un par de letras al interior de una figura geométrica según su edad y nivel de instrucción. En un nivel superior, los rectángulos simbolizan familias nucleares que constituyen redes sociales densas. También se observan dos tipos de flechas: las de líneas continuas representan relaciones familiares y las discontinuas de amistad o compadrazgo.

La red principal de este estudio, considerada así por su densidad y multiplicidad, comprende cinco familias nucleares. Las primeras cuatro guardan una relación estrecha, debido a que existe una relación fraternal entre sus núcleos formales. Estas familias, a su vez,

se vinculan con una quinta por una relación de primos en segundo grado. La conformación de esta red es importante pues, como se observará más adelante, cada uno de sus integrantes presenta diferentes grados de bilingüismo y características sociales particulares, aun cuando las relaciones y el intercambio lingüístico entre ellos son estrechos. También se incluyen otras familias vinculadas con algunos miembros de la red principal mediante relaciones de amistad o compadrazgo.

Después de conformar la red de colaboradores revisé la distribución de tres factores sociolingüísticos: género³⁴, edad y nivel de instrucción³⁵. En el grupo de 28 colaboradores están representados por igual mujeres y hombres. En cuanto al nivel de instrucción, hay tres niveles: bajo, con menos de 6 años escolares, medio de 7 a 11 años, y alto con más de 12 años. Además, los participantes pertenecen a cuatro grupos etarios: el primero (GE1) está constituido por seis colaboradores de entre 14 y 19 años, el segundo (GE2) tiene seis integrantes de 20 a 34 años, el tercer conjunto (GE3) consta de ocho personas con una edad de 35 a 55 años y la cuarta agrupación (GE4) tiene ocho miembros con más de 55 años. De este modo, el tener representados cuatro grupos etarios permitirá la incorporación de observaciones en tiempo aparente en esta comunidad.

Llevé a cabo el trabajo de campo en Jiquipilco el Viejo durante el mes de julio del 2019. Decidí trabajar solo con 16 colaboradores porque el eje principal del trabajo es la

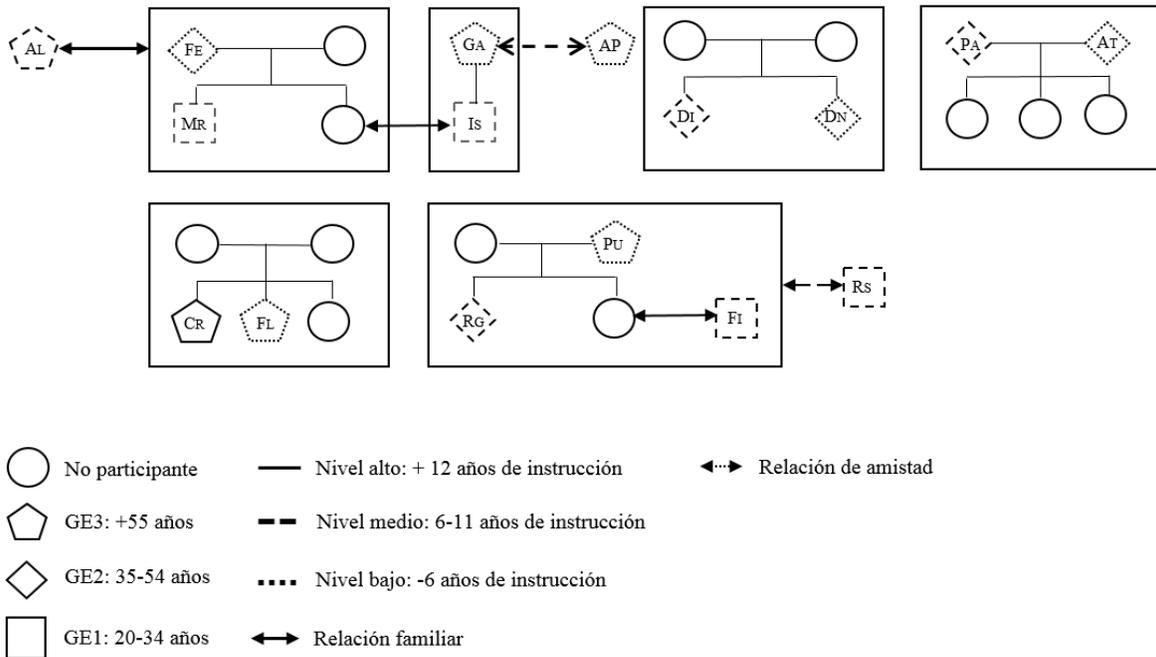
³⁴ La diferencia en el habla masculina y femenina no está basada en diferencias biológicas, sino en los roles sociales que el hombre y la mujer tienen en una comunidad de habla determinada (Labov, 2001: 263). Esta variable está relacionada con el segundo principio del cambio lingüístico “the linguistic conformity of women: For stable sociolinguistic variables, women show a lower rate of stigmatized variants and higher rate of prestige variants than men” (Labov, 2001: 266), cuya validez ha sido demostrada en numerosas investigaciones (Labov, 1996); sin embargo, también se han documentado contraejemplos en los que esta variable tiene un efecto opuesto (Labov, 2001: 166-169) o nulo.

³⁵ La clase social se ha incorporado en los estudios variacionistas a partir de varios indicadores, como la educación, los ingresos económicos y la ocupación, ya sea de manera individual o conjunta; sin embargo, en ocasiones es difícil discernir la jerarquía de estos o su traslape (Moreno Fernández, 1998: 53-54). Por esta razón decidimos incluir un solo factor: el nivel educativo, que también suele estar relacionado con el grado de bilingüismo en situaciones de contacto (Palacios, 2005).

comunidad de San Andrés Cuexcontitlán y Jiquipilco el Viejo es una comunidad más pequeña. En esta localidad también utilicé el método de bola de nieve o redes sociales para conformar el grupo de colaboradores. Mi padre fue el principal vínculo con las personas de Jiquipilco el Viejo, debido a que él laboró como profesor en la primaria de esta comunidad por 30 años. El primer paso fue establecer contacto con algunos compadres, comadres y amigos de esta localidad. Ellos aceptaron formar parte de esta investigación y me permitieron entrevistarlos. Posteriormente, me vincularon con miembros de su familia nuclear y extensa, amigos y conocidos hasta conformar el grupo de 16 colaboradores.

La Figura 2.3 muestra las relaciones que los colaboradores guardan entre sí y las redes que conforman. Al igual que en la figura de SAC, cada uno de los participantes está representado por un par de letras al interior de una figura geométrica, de acuerdo con su edad y nivel de instrucción. Los rectángulos simbolizan familias nucleares, las flechas continuas representan relaciones familiares y las discontinuas de amistad o compadrazgo.

FIGURA 2.3 Red de colaboradores en Jiquipilco el Viejo



En Jiquipilco el Viejo colaboré con los miembros de seis familias. Sin embargo, a diferencia de la red de San Andrés Cuexcontitlán, estas no guardan ninguna relación, salvo las dos familias representadas en los primeros dos recuadros, donde la hija de la colaboradora (Ga) es nuera de la familia de (Fe). Posterior a la conformación de la red revisé la distribución de los factores edad y nivel de instrucción. En la muestra están representados tres niveles de instrucción: siete colaboradores de nivel bajo (-6 años escolares), ocho personas con instrucción media (7 a 11 años) y sólo una persona de nivel alto (más de 12 años). Los participantes pertenecen a tres grupos etarios: uno (GE2) con cuatro colaboradores de entre 20 y 34 años, otro (GE3) está integrado por seis hablantes de 20 a 34 años y el último (GE4) por seis personas con más de 55 años. Esta distribución permitirá incorporar observaciones en tiempo aparente de esta comunidad.

En este punto es importante señalar que trabajar con personas de San Andrés Cuexcontitlán y de Jiquipilco el Viejo me permitió ponderar algunas de las ventajas y desventajas que puede tener el ser miembro de una de las comunidades que son parte de esta investigación. Haber crecido y vivido gran parte de mi vida en San Andrés Cuexcontitlán me brindó varias ventajas. Entre estas se halla el conocimiento de los dominios sociales relevantes y sus dinámicas, de los tipos de relaciones que se establecen entre sus miembros y el valor social que poseen. Estos antecedentes, sin duda alguna, facilitaron la planeación del muestreo y la preparación de los cuestionarios.

En segundo lugar, mi pertenencia a la comunidad y la convivencia con sus miembros me permitió observar diferencias individuales en el conocimiento y uso del español y el otomí, así como el cambio que pueden sufrir los bilingües en el dominio de sus lenguas durante el transcurso de su vida, lo cual se traduce en la existencia de un continuo bilingüe que será descrito en §2.2. Durante el proceso de muestreo y levantamiento de datos, el ser miembro de la comunidad me facilitó el acceso a redes sociales múltiples y densas, de las que además formo parte. En la fase de análisis, me brindó la posibilidad de contrastar los datos obtenidos mediante los diferentes cuestionarios con las actitudes y el comportamiento de los colaboradores en situaciones cotidianas, sin el efecto inhibitorio de la entrevista.

No obstante, las bondades, también me encontré con algunas dificultades. En primer lugar, me di cuenta de que tenía algunas ideas preconcebidas relacionadas con la transmisión de las lenguas y su uso; por ejemplo, al inicio de la investigación, pensaba que la lengua otomí se seguía transmitiendo en algunos sectores de la población y a partir de los resultados pude ver que no es así. Otro de los problemas a los que me enfrenté fue poder hallar el valor descriptivo y explicativo de la cotidianidad, más allá de los acontecimientos llamativos o anormales. Sirva de muestra mi inadvertencia en torno a la importancia de las relaciones de

compadrazgo en SAC, pues fue hasta que un miembro externo a la comunidad me hizo notar la particularidad de estos vínculos que pude ver su relevancia en la formación de redes sociales múltiples y su superposición a los vínculos familiares.

Por otro lado, mi posición como persona externa al pueblo de Jiquipilco el Viejo me trajo más desventajas que ventajas. Contrario a lo que sucedió con SAC, antes de comenzar con la investigación desconocía cuáles eran los dominios sociales relevantes para esa comunidad, sus dinámicas y valor social-cultural. Fue hasta el trabajo de campo que pude identificar algunos de estos elementos; no obstante, mi nivel de comprensión aún es limitado. También fue complicado poder constatar el dominio y el uso que los miembros de esta comunidad hacen de sus lenguas, más allá de los datos obtenidos con la entrevista y las observaciones realizadas durante el trabajo de campo. A continuación, describiré el método y los materiales empleados en la recolección de los datos lingüísticos y sociales.

2.1.2 Materiales para la recolección de material lingüístico: la entrevista

Además de emplear el mismo procedimiento para la conformación del grupo de colaboradores en las dos comunidades, también usé una metodología análoga en la recolección de datos con la finalidad de poder realizar una comparación entre ellas, a manera de estudio de tendencia. Constituí la corpora de esta investigación a partir de una entrevista que constó de seis actividades:³⁶ narraciones en otomí y español, un cuestionario de habilidades lingüísticas en español y otro en otomí, una prueba de conocimiento pasivo en otomí y un cuestionario sociolingüístico.

³⁶ Los datos de San Andrés Cuexcontitlán fueron grabados directamente en una computadora portátil con el programa Audacity y un micrófono Behringer, modelo C-1U, a una frecuencia de 44100 Hz y 16 bits. Para las grabaciones de Jiquipilco el Viejo empleamos una grabadora DR40x.

2.1.2.1 Presentación

Al inicio de cada entrevista realicé una breve presentación en la que proporcioné mis datos y el objetivo de la investigación, subrayando el interés por estudiar el otomí de San Andrés Cuexcontitlán y Jiquipilco el Viejo. Planteé un objetivo general para evitar alguna predisposición que pudiera afectar el desarrollo y los resultados de la entrevista. Asimismo, indiqué que la información personal tendría un uso único y exclusivo para fines académicos. Posteriormente, procedí a explicar cada una de las actividades-pruebas.

2.1.2.2 Historia tradicional y de vida

En la primera parte de la entrevista grabé dos narraciones en español y en otomí, una relacionada con la localidad (un mito, una leyenda, un cuento, un acontecimiento importante) y otra con la historia de vida del colaborador (un hecho sobresaliente, una anécdota de la infancia o del trabajo)³⁷. Esta actividad tenía como objetivo obtener material lingüístico para caracterizar el español y el otomí de los colaboradores.

2.1.2.3 Pruebas de habilidades lingüísticas

En el apartado §1.2 presenté algunas definiciones de bilingüismo y expresé mi afinidad con Grosjean (2012, 2014) y sus ideas en torno al estudio de este fenómeno. Si consideramos el principio de complementariedad y la idea de que la competencia de los bilingües cambia con el tiempo y a lo largo de su vida, una descripción del bilingüismo a partir de la edad y el orden de adquisición me parecía insuficiente para caracterizar el bilingüismo de los

³⁷ Cada uno de los relatos debía tener una duración aproximada de ocho o diez minutos y debía ser narrado en español y luego en otomí. Empleé dicho orden para entender lo que el informante narraba, pues esto me permitiría prolongar la conversación en caso de ser necesario.

colaboradores de esta investigación. Ante esta situación, el primer paso fue revisar las herramientas disponibles para ponderar las habilidades bilingües.

El primer instrumento disponible eran los censos poblacionales, que proveen información acerca del número de hablantes monolingües y bilingües de una lengua en un territorio determinado. Aunque son una herramienta útil que proporciona una imagen general de las situaciones de contacto, la información que ofrecen sobre bilingüismo individual y comunitario es limitada (Flores Farfán, 1999: 29-31; Gómez Rendón, 2008: 228; Romaine, 1989: 26; Terborg, 2011: 14-15). En el nivel individual, los cuestionarios censales no brindan información sobre la adquisición de las lenguas (cuándo y cómo), el grado de bilingüismo o las actitudes y las lealtades lingüísticas de los hablantes. A nivel comunidad, no dan cuenta de las relaciones que existen entre las lenguas, sus hablantes y su entorno, pues no aportan información sobre la función y el estatus de las lenguas, ni de su uso en determinados dominios.

En los estudios de bilingüismo, el uso de métodos objetivos para medir las habilidades bilingües ha sido una práctica poco frecuente (Treffers-Daller, 2018), como lo muestran los resultados de Hulstijn (2012) quien, después de analizar 140 artículos publicados en la revista *Bilingualism: Language and Cognition* entre 1998 y 2011, encontró que sólo el 19% de los artículos publicados en la mitad de este periodo usaron una prueba objetiva para medir el bilingüismo. Aunado a eso, algunas de esas herramientas presentaban sesgos teóricos y metodológicos como evaluar únicamente las habilidades en la segunda lengua o hacer comparaciones entre los bilingües y los “hablantes monolingües nativos” sin tomar en cuenta la heterogeneidad de estos últimos, dando lugar a resultados sesgados que generaban una visión negativa de las habilidades bilingües, a la que se denominó “déficit bilingüe”.

En el área de adquisición de segundas lenguas se ha hecho hincapié en que las habilidades de los bilingües no sólo comprenden el conocimiento lingüístico (fonológico, morfológico, sintáctico y léxico) en las cuatro habilidades (hablar, escuchar, leer, escribir), sino también las habilidades sociolingüísticas en cada lengua y el procesamiento de ambas. Esta complejidad exige la construcción de herramientas que midan las habilidades bilingües de forma apropiada. Aunque todavía no se cuenta con una medida que evalúe de manera integral las habilidades de un bilingüe, se ha desarrollado una gran cantidad de instrumentos y pruebas. Entre estas se encuentra el uso de reportes parentales en el estudio de niños bilingües (B. Z. Pearson et al., 1993) y el de las autoevaluaciones en los adultos (Terborg, 2011). Otro de los métodos recurrentes es medir un dominio, una habilidad o un componente que funcione como *proxy* o representante de todas las habilidades bilingües; por ejemplo, los test de vocabulario que miden el tamaño del léxico en una L2 (Bialystok et al., 2010).

Según Torres Sánchez (2019), en los estudios de contacto se han empleado al menos dos tipos de métodos para medir el bilingüismo. El primero que documenta es el etnográfico, que se basa en la (Flores Farfán, 1999; Hill & Hill, 1999). El segundo tipo se fundamenta en parámetros explícitos, tanto lingüísticos como sociales (Diebold, 1961; Escobar, 1988; Palacios, 2005). A estos se suman los cuestionarios basados en el uso y la autoevaluación (Mulík et al., 2021).

Con base en estos antecedentes, la presente investigación pretendía ir más allá de la división mono/bilingüismo y acercarse a la complejidad que hay en las comunidades bilingües de San Andrés Cuexcontitlán y Jiquipilco el Viejo. Por ello, decidí emplear una prueba de habilidad lingüística en español y otra en otomí con la finalidad de aproximarme, cualitativa y cuantitativamente, al conocimiento que los colaboradores tienen de sus lenguas.

Las pruebas de habilidades lingüísticas de esta investigación se basan en la metodología de Torres Sánchez (2018, 2019). Este cuestionario tiene como principio de ponderación la presencia o ausencia de interferencias en las habilidades orales de tres niveles lingüísticos (fónico, léxico y sintáctico-pragmático) del español en el momento de la grabación.

2.1.2.3.1 Prueba de habilidad lingüística en español

La prueba de habilidad lingüística del español consta de tres secciones que corresponden al nivel fonológico, morfosintáctico y léxico (véase Apéndice 2A). Seleccioné los ítems del cuestionario a partir de los resultados de investigaciones previas sobre el contacto español-otomí (Guerrero, 2006; Guerrero Galván & San Giacomo Trinidad, 2014; Hekking, 2001; Lastra, 1987) y de algunos fenómenos observados en las grabaciones piloto en San Andrés Cuexcontitlán (Avelino Sierra, 2017).

La sección fónica consta de doce palabras con segmentos fónicos, estructuras silábicas y configuraciones acentuales que propician la aparición de transferencias fónicas. Por ejemplo, incluí palabras como *correr* y *amarillo*, con las que pretendía evaluar la realización plena o asibilada de /r/ y la realización plena o debilitada de /j/, para confirmar la vitalidad de los fenómenos documentados por Lastra en 1987. También consideré los ítems *lombriz* y *corazón*, con el objetivo de probar la realización o elisión de /s/ y /n/ en coda, como mecanismo de recuperación de la estructura silábica CV del otomí (Guerrero, 2006; Guerrero Galván & San Giacomo Trinidad, 2014; Lastra, 1987). Con ítems como *café*, *fotógrafo*, buscaba documentar la realización bilabial oclusiva sorda /ϕ/ del fonema labiodental fricativo sordo /f/ del español (Lastra, 1987), así como posibles movimientos acentuales a la primera

sílaba. El procedimiento consistió en mostrar a los colaboradores imágenes y después pedirles que enunciaran las palabras con las que nombraban las entidades.

El apartado léxico tenía la finalidad de valorar el conocimiento léxico del español. Elicité 15 ítems de diferentes grupos semánticos que presentaban distintos grados de dificultad: cuatro verbos (*johya* ‘alegrarse’, *k’uts’i* ‘arrancar’, *hats’i* ‘amanecer’, *k’ot’i* ‘escapar’), cuatro sustantivos (*he’mi* ‘documento’, *ronjua* ‘ayate’, *ntsu* ‘miedo’, *huit’ua* ‘calzón’), tres adjetivos (*nts’o* ‘horroroso’, *mähotho* ‘hermoso’, *tsegi* ‘friolento’), dos pronombres (*nukʔígé* ‘tú’, *gégé* ‘él’) y dos adverbios (*mote* ‘afuera’, *mbo* ‘adentro’). Para la realización de este apartado grabé los ítems en otomí con la ayuda de un bilingüe; la dinámica consistió en que el informante escuchara la grabación y dijera su equivalente en español.

El propósito de la sección morfosintáctica fue acercarse al conocimiento que los colaboradores tienen en este nivel y constó de dos apartados. En el primero, mediante diez estímulos visuales evalué la producción de algunas construcciones sintácticas en las que potencialmente encontraría algunas transferencias (véase Apéndice 2A). Por ejemplo, con la imagen de tres manzanas pretendía revisar la concordancia de número entre el núcleo de la FN y el numeral; con la imagen del niño bañándose buscaba obtener oraciones con un uso alterno del *se* (omisión o hipercorrección), etcétera.

En la segunda parte, el objetivo era valorar el nivel de comprensión que los colaboradores tenían de algunas oraciones del español con cierta complejidad. Constó de cuatro oraciones, con las primeras dos, *Mis tíos, mi mamá y María fueron a misa ayer* y *La bicicleta que traía Juan, Pedro se la regaló* pretendía verificar el grado de comprensión de una oración con sujetos coordinados. En la tercera oración, *Marcos le prestó una pluma a Margarita*, valoré la comprensión de una construcción ditransitiva y en la cuarta, *Julia y Karen se peinaron*, de oraciones con *se*. El procedimiento consistió en presentarle oraciones

al colaborador y hacerle algunas preguntas (véase Apéndice 2A). Por ejemplo, en la primera oración, *Mis tíos, mi mamá y María fueron a misa ayer*, se les preguntaba: *¿quiénes fueron a misa, mis tíos o mi mamá?, ¿María iba con mis tíos?*

2.1.2.3.2 Prueba de habilidad lingüística en otomí

Este cuestionario (Apéndice 2B) tuvo como objetivo aproximarse a las habilidades que los colaboradores tienen en la lengua otomí, a partir de la presencia/ausencia de transferencias fonéticas, léxicas y morfosintácticas en elementos o contextos en los que potencialmente podrían presentarse. Esta herramienta constó de tres secciones: fonética, léxica y morfosintáctica.

La primera se compone de 15 ítems con los que evalué la presencia o ausencia de transferencias que podrían tener lugar debido a las diferencias entre los sistemas fonológicos del español y el otomí. Por ejemplo, con la producción de las palabras *dese* ‘capulín’ deseaba comprobar la producción o no de la vocal /ε/. Las palabras *dada*, *tu*, *mbəhə* ‘señor’ permitían documentar la realización de los fonemas /ʌ/ y /ɔ/, y en *mu* ‘barriga’ la de la vocal /i/. La dinámica consistió en proporcionar al informante la palabra en español y pedir su correspondiente en otomí.

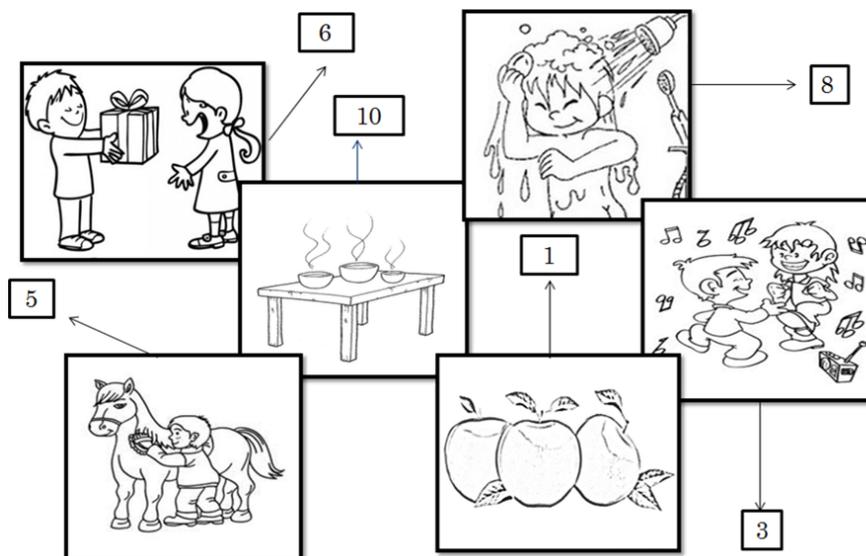
La segunda sección tenía como propósito dar cuenta del conocimiento que los colaboradores tienen del léxico en otomí³⁸. Estuvo constituida por 15 ítems de distintas clases

³⁸ Uno de los miembros del Comité Lector expresó su preocupación sobre un posible conflicto entre la sección léxica de las pruebas de habilidad lingüística y el análisis de la marcación de número en los préstamos nominales del español en otomí, pues pensó que uno de los factores (léxico) que definen la variable independiente “bilingüismo” podría estar asociado con la variable dependiente. Sin embargo, esto no es así, pues la variable dependiente en esos análisis es la presencia o ausencia de las marcas *ra* y *ya*. Otra de las preguntas que surgió fue si se había controlado el efecto que el monolingüismo en español del entrevistador podría tener en el uso de préstamos del español durante la aplicación del cuestionario de habilidades lingüísticas en otomí, pues se ha mostrado que las características lingüísticas del interlocutor pueden afectar el comportamiento lingüístico de los bilingües. Lamentablemente, esto no se controló.

léxicas y funcionales: cuatro verbos (*gä'tsi* 'suspirar', *agi* 'enterrar', *bets'i* 'embarrar' y *häm̄bi* 'arrebatar'); cuatro sustantivos (*bots'e* 'costilla', *ajüa/zidada* 'Dios', *bax'bots'e* 'chiquihuite' y *däza* 'plátano'); cuatro adjetivos (*hñu* 'pesado', *oki* 'agujerado', *seh̄e* 'solo' y *mfege* 'roto'); dos pronombres (*nda'ä* 'cuál' y *nu'u* 'aquellos') y una preposición *ha* 'en'. El procedimiento consistió en darle al colaborador la palabra en español y pedirle que proporcionara su equivalente en otomí.

Por último, con la sección morfosintáctica buscaba dar cuenta de algunos fenómenos de transferencia del español en el otomí. Para ello, empleé diez imágenes del cuestionario de competencia lingüística del español y les pedí a los colaboradores que dijeran una oración en otomí que describiera la situación. Por ejemplo, con la imagen del niño acariciando un perro y un gato, pretendía observar si la relación de coordinación entre frases nominales se realizaba mediante el elemento coordinador del otomí o del español. Con el dibujo de un niño dándole un regalo a una niña y de un niño jalándole el cabello a una niña, quería evaluar la emisión de oraciones con doble objeto (Palancar, 2009) y ver si esta ha sido influida por el español. A continuación, se muestran algunas de las imágenes que utilicé.

FIGURA 2.4 *Imágenes de las pruebas de competencia*



2.1.2.3.3 Prueba de habilidad lingüística pasiva en otomí

Durante la recolección del corpus trabajé con colaboradores que tenían un rendimiento menor a los 5.00 puntos en la prueba de competencia en otomí, por lo que resultaba difícil integrarlos en el grupo de bilingües; sin embargo, era notoria su habilidad para entender *ñható*. Por ello, pensé en otra prueba para aproximarme al grado de comprensión que esos colaboradores tienen de la lengua otomí (véase Apéndice 2C).

La prueba constó de tres secciones. En la primera ponderé la comprensión léxica, a partir de 57 ítems con distinto grado de complejidad, por ejemplo: *dehe* ‘agua’, *tsi* ‘comer’, *muza* ‘zapote’, *bospi* ‘ceniza’, *fugi* ‘espuma’. En la segunda sección examiné el nivel de comprensión de 33 construcciones sintácticas de diversa dificultad. En la tercera sección revisé la comprensión de una anécdota a partir de la petición de una paráfrasis y de diez preguntas sobre el suceso.

2.1.2.4 Cuestionario sociolingüístico

Las pruebas de habilidades lingüísticas en español y otomí y la prueba de habilidad pasiva en otomí pretendían aproximarse al dominio que los hablantes tienen de las lenguas e ir más allá de la clasificación básica monolingüe/bilingüe. El cuestionario sociolingüístico, por su parte, tenía el propósito de completar el perfil lingüístico y sociolingüístico de los colaboradores y las relaciones que establecen con las lenguas (usos, funciones, estatus, actitudes, lealtades, etc.) y los miembros de la comunidad (dominios).

El cuestionario sociolingüístico de esta investigación³⁹ fue diseñado por Alonso Guerrero Galván, en el marco del proyecto *Normatividad y variación de las lenguas otopames*⁴⁰ y se ha empleado en otras comunidades de habla otomí.⁴¹ Está dividido en siete secciones: 1) datos sociodemográficos, 2) datos socioeconómicos, 3) adquisición y competencia lingüística, 4) uso de las lenguas, 5) actitudes lingüísticas, 6) identidad y cultura y 7) variación (véase Apéndice 2D), que a continuación describo en más detalle.

Datos sociodemográficos

En esta sección pedí al colaborador que me proporcionara algunos datos personales para completar su perfil sociolingüístico. Les solicité información acerca de su edad y nivel de estudios para poder clasificarlos dentro de un grupo etario (adolescentes, jóvenes, adultos y

³⁹ Una encuesta similar fue empleada en el proyecto “La vitalidad de las lenguas indígenas de México: un estudio en tres contextos” (Terborg, 2011) que, a pesar de proporcionar más información lingüística que los censos, sigue siendo limitado, pues apenas se esbozan los dominios lingüísticos.

⁴⁰ Proyecto a cargo de Alonso Guerrero Galván, cuyo objetivo es estudiar la evolución fonológica de las lenguas otopames, en el contexto de la sistematización ortográfica; la relevancia de la variación dialectal en el registro sistemático de la lengua y la formación de comunidades textuales. Fuente: <http://linguistica.inah.gob.mx/index.php/pro/27-normatividad-y-variacion-en-lenguas-otopames>.

⁴¹ Este cuestionario ha sido aplicado en la comunidad de El Espíritu, municipio de Ixmiquilpan, Hidalgo; y en San Ildefonso Chantepec como parte de un peritaje lingüístico cuyo propósito era autenticar como indígena a esa comunidad (Guerrero Galván, 2014).

mayores) y de instrucción (bajo, medio, alto). También les pedí algunos datos sobre su situación migratoria, pues la historia de vida de los hablantes ha mostrado ser sumamente importante para determinar el nivel de fluidez de los bilingües y el uso que hacen de sus lenguas. Enseguida presenté algunas de las preguntas que conforman esta sección.

TABLA 2.1 *Preguntas sociodemográficas*

Preguntas
1. Nombre:
2. Género:
4. Lugar de nacimiento:
5. Edad:
6. Domicilio actual:
7. ¿Cuántos años ha vivido en su comunidad actual?
8. ¿Ha salido muchas veces de su comunidad?
9 ¿Cuántos años ha estado en la escuela?
9.1 ¿Actualmente estudia?
10. Estado civil:

Datos socioeconómicos

La segunda sección tenía el objetivo de obtener datos acerca de la ocupación y los ingresos económicos de los colaboradores, como se ejemplifica en la Tabla 2.2. Estos datos también son útiles para completar el perfil social de los colaboradores, porque permiten saber más sobre las redes y los grupos sociales a los que pertenecen, así como los contextos y tipo de hablantes con los que interactúan.

TABLA 2.2 *Preguntas socioeconómicas*

Preguntas
11. ¿Realiza alguna actividad económica?
12. ¿Usted trabaja para alguien o lo hace por cuenta propia?
13. ¿Qué actividad económica realiza principalmente para subsistir?
14. ¿Realiza alguna otra actividad complementaria? ¿Cuál?
15. ¿Durante el último año usted ha migrado para trabajar?
16. ¿Cuándo viaja para trabajar, se queda en una sola ciudad?
17. ¿Cuánto tiempo dedica a las actividades de subsistencia?
18. ¿Recibe ingresos por las actividades que realiza para subsistir?
19. ¿Aproximadamente de cuánto es su ingreso mensual?

Adquisición y competencia lingüística

Este apartado indaga sobre las lenguas que hablan los colaboradores y las habilidades que poseen en estas (hablar, escuchar, leer, escribir). Como se muestra en la Tabla 2.3, también solicita información acerca de la edad y el contexto de adquisición de las lenguas, pues son factores de suma importancia para la caracterización y clasificación de la competencia lingüística de los colaboradores. Además, empleé uno de los métodos que Appel y Muysken (1996: 53-154) proponen para la valoración del bilingüismo: la autoevaluación de la competencia lingüística, mediante preguntas como *¿Cuál lengua habla mejor?*

TABLA 2.3 Preguntas sobre adquisición y competencia lingüística

Preguntas

20. ¿Qué lenguas habla?
 21. ¿Dónde aprendió el otomí? ¿Quién le enseñó a hablarlo o cómo lo aprendió?
 22. ¿Dónde aprendió el español? ¿Quién le enseñó a hablarlo o cómo lo aprendió?
 23. ¿A qué edad aprendió el otomí?
 24. ¿A qué edad aprendió el español?
 27. ¿Cuál lengua habla mejor?
 28. ¿Cuál lengua entiende mejor?
 32. ¿Lee el español?
 33. ¿Lee el otomí?
 34. ¿Escribe el español?
 35. ¿Escribe el otomí?
-

Dominios lingüísticos⁴²

El propósito de esta sección fue obtener información sobre el uso de las lenguas: en qué contextos se usan y con qué interlocutores. La primera parte indaga sobre el uso que los colaboradores hacen del español y el otomí, según el interlocutor. Incluye los miembros de la familia nuclear (padre, madre, hermanos, cónyuge e hijos, según sea el caso) y algunos elementos de la familia extensa (tíos, primos, abuelos y, de ser necesario, suegros, yerno/nuera y nietos); también considera interlocutores con lazos de amistad y compadrazgo. La segunda parte presenta nueve contextos o dominios: casa, trabajo, mercado, ciudad, escuela, iglesia, fiestas, juntas, calle; y, a partir de estos, plantea diversas situaciones con potenciales interlocutores.⁴³ A continuación, presento algunos ejemplos.

⁴² En el sentido de Fishman (1972) un ámbito o dominio sería un cúmulo de contextos o situaciones en torno a un tema central que estructura la percepción que el hablante tiene de estas situaciones. En ese sentido, cada ámbito exige una elección lingüística concreta (Appel & Muysken, 1996: 40).

⁴³ Blom y Gumperz (1972: 46) plantean tres niveles de interpretación: en el primero está el contexto, que es el escenario de la interacción, generalmente, un entorno reconocido socialmente. En el siguiente nivel están las diferentes situaciones sociales que se pueden presentar en un contexto, momento y con interlocutores concretos. El tercer nivel, se encuentra el evento social que se presenta en un contexto, momento y situación social determinados.

TABLA 2.4 *Preguntas sobre dominios lingüísticos*

Preguntas
36. ¿Qué lengua habla con su familia?
37. ¿Con la mamá?
38. ¿Con el papá?
39. ¿Con los hermanos o hermanas?
40. ¿Con el abuelo?
41. ¿Con la abuela?
46. ¿Con la pareja?
46.1 ¿Con los hijos o hijas?
46.6 ¿Con su nieto o nieta?
47. ¿Con los amigos?
48. ¿Con los compadres?

En una escala de importancia, espero que el factor más determinante en la elección de las lenguas sea el entendimiento común entre el colaborador y su oyente; por ejemplo, es posible que el colaborador elija hablar español si sabe que su interlocutor no sabe hablar ni entiende otomí. El segundo factor es el tipo de relación que tienen el colaborador y su interlocutor (familiar, de amistad, etc.) y el tercero el contexto o ámbito en el que se encuentran.

Actitudes lingüísticas

Con el quinto apartado pretendía conocer las actitudes de los colaboradores hacia el español y el otomí, pues han demostrado ser de gran importancia en el mantenimiento y/o desplazamiento de las lenguas. Como se muestra en la Tabla 2.5, pregunté sobre la importancia de las lenguas, la situación del otomí en la comunidad y si los hablantes tenían interés en aprender a leer y escribir en *n̄ható*. Con las siguientes preguntas indagué sobre la importancia de incluir el otomí en la esfera pública (medios de comunicación: radio y televisión) y educativa (programas bilingües). Además, quería saber si los colaboradores han

sufrido discriminación por hablar otomí, ya que este tipo de hechos suele repercutir en las decisiones de las personas para el mantenimiento y la transmisión de las lenguas.

TABLA 2.5 *Actitudes lingüísticas*

Preguntas
61. ¿Cree que alguna lengua es más importante?
64. ¿Cree que el otomí puede perderse o dejarse de hablar?
65. ¿Se debería enseñar el español en las escuelas?
64. ¿Cree que los otomíes tienen derecho a que se les enseñe en su lengua?
66. ¿Se debería enseñar el otomí en las escuelas?
68.3 ¿Se ha sentido discriminado por hablar otomí?
68.4 ¿Se ha sentido discriminado por no hablar otomí?
69. ¿Sus padres le han dicho que el otomí es feo o no debe hablarlo?
69.8 ¿Usted cree que el español no se debe hablar?
69.7 ¿Usted cree que el otomí no se debe hablar?
69.4 ¿Personas ajenas a la comunidad le han dicho que el otomí no se debe hablar?

Identidad y cultura

Con las preguntas de esta sección pretendía conocer qué tan identificados se sienten los colaboradores con la cultura otomí, pues se ha observado que este factor es de gran importancia para el mantenimiento o desplazamiento de las lenguas. Esto se realizó a partir de preguntas como: *¿cree que las costumbres otomíes son distintas a las de las personas que no son de la comunidad?*, con las que buscaba ver si el colaborador se identificaba con la cultura otomí en contraposición con la mestiza. Mediante las siguientes cuestiones: *¿se identifica con la cultura otomí?*, *¿puede mencionar tres aspectos que lo identifiquen más con la cultura otomí?*, intentaba observar si el colaborador se reconocía como parte de la cultura otomí y si distinguía cuáles eran los lazos que lo unían.

Variación

Finalmente, en la sección de variación apelaba al conocimiento metalingüístico de los colaboradores e indagaba sobre la percepción que tienen de su variante de otomí, con respecto al de poblaciones vecinas. También, les pregunté sobre la variación que observaban dentro de su comunidad de habla, en función del género y la edad. Asimismo, indagué sobre la variación y los cambios que ha sufrido el otomí en los niveles fónico, léxico y sintáctico, y sus posibles causas, como se muestra enseguida.

TABLA 2.6 *Preguntas sobre variación lingüística*

Preguntas					
74.1 ¿En dónde cree que se habla mejor el otomí?					
0. San Andrés Cuexcontitlán	1. San Cristóbal Huichochitlán	2. San Pablo Autopan			
3. Temoaya	4. Son iguales	5. Otro lugar	6. Sin opinión		NR
74. 2 ¿Quiénes hablan mejor el otomí?					
0. los niños	1. Los jóvenes	2. Los adultos	3. Los ancianos	4. Las mujeres	
5. Es igual		6. Sin opinión			NR
75 ¿Cree que el otomí que hablan los jóvenes es distinto al que hablan los ancianos?					
0. No	1. Sí	2. Sí, un poco	3. Sí, es muy diferente	4. Es igual	5. Sin opinión
	NR				
75.1 ¿Cree que ha cambiado el otomí últimamente?					

2.1.3 *La corpora y algunas herramientas de análisis*

A partir de las herramientas antes descritas obtuve los datos para esta investigación. El corpus para el análisis de la concordancia plural en el español en contacto (Capítulo §3) se construyó a partir de las narraciones tradicionales y de vida en español de los cuarenta y tres

colaboradores. Las transcripciones de las narraciones se realizaron con base en los criterios del Corpus Sociolingüística de la Ciudad de México (Martín Butragueño & Lastra, 2011). Para la construcción de la base de datos tenía contemplado fichar cien ejemplos por colaborador: cincuenta casos de concordancia plural en el nivel frástico y cincuenta en el oracional, respetando el principio de exhaustividad; sin embargo, no fue posible alcanzar esa cifra con todos los hablantes, por lo que al final el corpus quedó conformado por 3 818 casos de concordancia plural.

El corpus del cuarto capítulo, que versa sobre la marcación de número en los nombres del otomí (patrimoniales y préstamos), se construyó a partir de las narraciones tradicionales y de vida en otomí de 26 bilingües otomí-español de San Andrés Cuexcontitlán y Jiquipilco el Viejo. Transcribí y traduje diez minutos de las narraciones y recolecté todos los elementos léxicos y funcionales del español, así como los casos de cambio de código. El corpus quedó conformado por 20 902 palabras, de las cuales 4 774 resultaron ser elementos del español (22.83%).

Los dos fenómenos que son objeto de estudio de los capítulos tres y cuatro son variables cualitativas⁴⁴ y fueron analizados de forma cualitativa y cuantitativa. En los análisis cuantitativos utilicé herramientas provenientes de la estadística descriptiva e inferencial. La estadística descriptiva o deductiva provee procedimientos para describir, organizar y presentar las características más relevantes de un conjunto de datos. Para realizar una descripción numérica de las variables lingüísticas utilicé tablas de distribución de frecuencias, como la que se muestra en la Tabla 2.7. En la primera columna se presentan las

⁴⁴ Una variable es una característica de interés y puede ser cualitativa o cuantitativa. Las variables cualitativas son propiedades, atributos o cualidades y pueden ser nominales (sin un orden) u ordinales (con un orden o jerarquía). Las variables cuantitativas son valores numéricos que se determinan por conteo o medición y estas pueden ser discretas (números naturales) o continuas (números reales) (Millones et al., 2018: 16).

categorías de las variables, en la segunda columna se codifica el conteo de las observaciones y en la tercera el porcentaje (Millones et al., 2018: 18).

TABLA 2.7 *Ejemplo de una tabla de distribución de frecuencias*

Categorías de la variable	Conteo de observaciones	Porcentaje (%)
C ¹	O ¹	P ¹
C ²	O ²	P ²
C ⁿ	O ⁿ	P ⁿ

La estadística inferencial, por su parte, proporciona técnicas para analizar e interpretar los resultados de una muestra y poder generalizarlos a una población.⁴⁵ Esto se hace a partir de la estimación de parámetros o del contraste de hipótesis. Una hipótesis es una conjetura sobre determinados fenómenos o hechos, para poder establecerla es necesario fijar el problema que será el objeto de estudio, las variables que lo definen y se expresa del siguiente modo (Mendenhall et al., 2010: 4).

$$\begin{cases} H_0: \mu = \mu_0 \\ H_1: \mu = \mu_1 \end{cases}$$

Donde H₀ representa la hipótesis nula, H₁ la hipótesis alternativa y μ un parámetro.⁴⁶ La hipótesis nula indica que un parámetro de población (tal como la media, la desviación estándar, etc.) es igual a un valor hipotético, el cual suele ser una afirmación inicial que se basa en análisis previos o en conocimiento especializado. La hipótesis alternativa indica que un parámetro de población es más pequeño, más grande o diferente del valor hipotético de la

⁴⁵ El término población “se refiere al conjunto total de unidades de análisis correspondientes al estudio que se desea realizar, de los cuales se desea describir su comportamiento y/u obtener conclusiones” (Millones et al., 2018: 16). Una muestra, por su parte, “es un subconjunto de la población, y debe ser representativa y aleatoria. La muestra es representativa si lo que se necesita conocer de la población está presente en la muestra, es decir, si los datos asociados a la muestra se asemejan a la población en estudio” (Millones et al., 2018: 16).

⁴⁶ Un parámetro “es una medida que resume y describe una característica de la población; su valor se calcula usando todos los datos de la muestra. Los parámetros se denotan usando letras griegas” (Millones et al., 2018: 16-17).

hipótesis nula y suele representar la idea que se busca probar. Además de esto, es necesario establecer un nivel razonable de confianza o seguridad (en lingüística suele usarse el 95%) y un nivel de error o significación α (en lingüística suele usarse el 5%), que es complementario al nivel de confianza. Después de haber establecido la hipótesis estadística y fijar los criterios para aceptarla o rechazarla se contrasta con los datos empíricos para determinar si es compatible o no con estos (Herrera Soler et al., 2011: 123).

2.1.3.1 La prueba de Chi-cuadrado de Pearson

El primer estadístico que utilicé en los análisis fue la prueba de Chi-cuadrado de Pearson (χ^2). Se trata de una prueba no paramétrica,⁴⁷ que sirve para determinar si un conjunto de datos (observados) se ajusta o no a una distribución de probabilidad dada (Canavos, 1988: 370-371). Existen tres tipos: la prueba de bondad de ajuste, la prueba de independencia y la prueba de homogeneidad. En esta investigación únicamente usé la prueba de independencia de Chi-Cuadrado para determinar una asociación entre dos variables categóricas. En todos los casos, la H_0 es que las dos variables categóricas son independientes y la H_1 es que la distribución de una de las variables es explicable por variaciones en la otra. El nivel de error o significación α es de 0.05 y el intervalo de confianza del 95%. La prueba de Chi-cuadrada se calcula mediante la siguiente fórmula.

$$\chi^2 = \sum_i \frac{(\text{observada}_i - \text{teórica}_i)^2}{\text{teórica}_i}$$

⁴⁷ Las pruebas paramétricas dependen en gran medida de algunas características de la población como una distribución normal, la igualdad de varianzas entre dos grupos, la escala de medida de las variables (intervalo o de razón). Por otro lado, las pruebas no paramétricas tienen menos requerimientos (Herrera Soler et al., 2011: 121).

2.1.3.2 V de Cramer

También empleé la V de Cramer para diagnosticar la intensidad de la relación de dependencia entre dos variables. Proporciona valores entre 0 y 1, en donde valores próximos a 0 indican una magnitud o tamaño del efecto muy baja y los valores cercanos a 1 muy alta (Herrera Soler et al., 2011: 184). La fórmula de este estadístico es:

$$V = \sqrt{\frac{\chi^2}{n(\min[r, c] - 1)}}$$

Establecí una escala análoga a la de (Torres Sánchez, 2018: 152) y el tamaño del efecto (ES) se interpreta del siguiente modo: si $ES \leq 0.2$ el resultado es débil y las variables categóricas sólo están débilmente asociadas; si $0.2 < ES \leq 0.6$, las variables están asociadas moderadamente; si $ES > 0.6$, las variables están fuertemente asociadas. Las tablas de contingencia, la prueba de Chi-cuadrada de Pearson y la V de Cramer fueron realizadas con la ayuda del programa SPSS (2020).

2.1.3.3 La medida de Goodman y Kruskal

Además de las herramientas descriptivas, hice uso de procedimientos inferenciales como la regresión logística escalonada. Antes de describir ese estadístico es importante advertir que cuando se construye un modelo de regresión las variables independientes deben tener una alta correlación con la variable dependiente y una baja correlación entre las variables independientes. Una correlación alta entre dos variables independientes puede generar colinealidad, es decir, que las dos variables independientes estén relacionadas de tal forma que constituyan una combinación lineal. Esto significa un problema, pues impide ver la

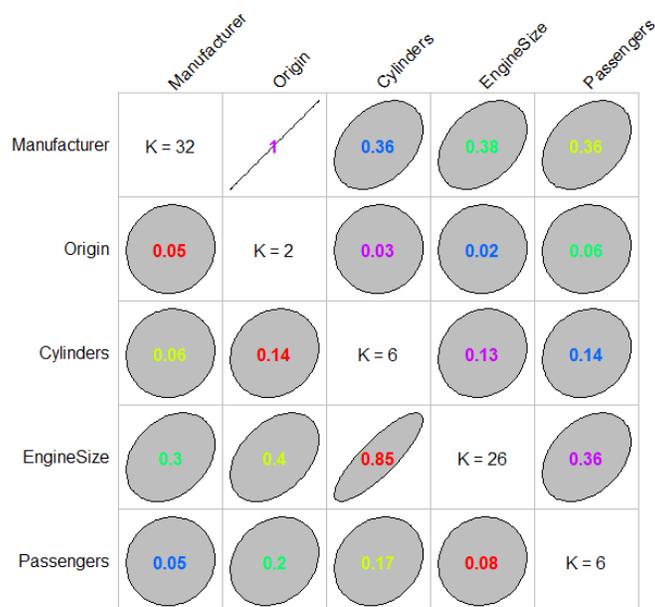
influencia que cada una de las variables tiene y es difícil alcanzar una explicación del fenómeno en cuestión, además de que los resultados no son fiables.

La Chi-cuadrado de Pearson y la V de Cramer son las medidas de asociación entre dos variables categóricas más utilizadas, pero tienen la desventaja de ser simétricas y partir del supuesto de que la correlación entre X y Y es la misma que entre Y y X (Herrera Soler et al., 2011: 257). Por esta razón, decidí utilizar la medida de Goodman y Kruskal tau del paquete “The GoodmanKruskal” en la plataforma de R. Este estadístico mide la asociación entre dos variables categóricas e indica si la variación en una variable puede explicarse a partir de la otra; no obstante, a diferencia de la Chi-cuadrado de Pearson, es asimétrica. Esto permite percibir que, en algunos casos, la fracción de variabilidad en X que es explicable por variaciones en Y puede ser muy diferente de la variabilidad en Y que es explicable por variaciones en X (R. Pearson, 2020). Por ejemplo, la variabilidad en el nivel de instrucción que es explicable por la variabilidad en la edad no necesariamente corresponde a la variabilidad en la edad que la variabilidad en la instrucción puede explicar.

Realicé esta técnica con la función `GKtauDataframe` del paquete `The GoodmanKruskal` en la plataforma de R. Este procedimiento proporciona valores que van de 0 a 1; mientras el valor esté más cercano a 1 mayor correlación habrá entre las variables y mayor riesgo de colinealidad (R. Pearson, 2020). A continuación, muestro un ejemplo del tipo de gráfico que nos brinda esta herramienta. En los ejes vertical y horizontal se encuentran cinco variables y al centro de cada recuadro se presenta la relación que existe entre estas. Los resultados muestran una relación de correlación completa entre los factores `Manufacturer-Origin` (1). Asimismo, se presenta una relación de correlación fuerte entre las variables `EngineSize-Cylinder` (0.85). Este recurso permite tener la certeza de que no exista o haya una

correlación baja entre las variables independientes; y, en caso de haberla, redefinir las variables y obtener un modelo explicativo más fiable.

FIGURA 2.5 Muestra de los resultados de la medida de Goodman y Kruskal (R. Pearson, 2020)



2.1.3.4 La regresión logística escalonada

La regresión logística escalonada es una prueba que forma parte de los modelos lineales generalizados (GLM)⁴⁸ y permite realizar análisis inferenciales sobre la relación de un conjunto de variables múltiples independientes sobre una variable dependiente categórica con un nivel de significación estadística. Un modelo de regresión puede tener fines explicativos o predictivos. En el primer caso, se busca evaluar cómo el cambio en factores determinados (variables independientes) tiene un efecto en otra característica en concreto

⁴⁸ Los modelos lineales generalizados (MLG) proporcionan un marco flexible para describir cómo una variable dependiente puede ser explicada por una serie de variables explicativas. La variable dependiente puede ser continua o discreta (de valor entero), y las variables explicativas pueden ser cuantitativas o categóricas. Además, abarcan muchas situaciones, como los ANOVA, las regresiones múltiples o la regresión logística (Calcagno & Mazancourt, 2010: 1-2).

(variable dependiente). En un modelo predictivo lo que se busca es estimar o aproximar el valor de una característica (variable dependiente) en función de los valores que pueden tomar en conjunto otros factores (variables independientes). Existen varios métodos de selección de variables: i) técnica de pasos hacia adelante: consiste en introducir las variables en el modelo sólo si cumplen una serie de condiciones hasta que se hayan integrado todas las que lo cumplan; ii) técnica de pasos hacia atrás: se introducen en el modelo todas las variables y se van eliminando si cumplen una serie de condiciones definidas hasta que no quede ninguna variable que cumpla la condición impuesta; iii) técnica por pasos: se combinan los dos métodos anteriores (Morales Peláez, 2006: 195-196).

En esta investigación empleé una regresión logística por pasos con un fin explicativo. En ese sentido, los resultados permitirán saber qué factores condicionan el fenómeno variable en cuestión, cuál es la dirección de su efecto y cuál es la jerarquía entre ellos. La prueba se realizó con la ayuda del software R (R CORE TEAM, 2021). En la regresión logística se tuvo variables de respuesta dicotómica. A continuación, presento un ejemplo del tipo de tabla en la que se exponen los resultados de esta prueba.

TABLA 2.8 *Ejemplo de los resultados de la regresión logística escalonada*

VARIABLES	ESTIMADO	ERROR STD.	VALOR Z	Pr(> z)
Intersección	Valores	Valores	Valores	Valores
Variables independientes	Valores	Valores	Valores	Valores

En la primera columna de la tabla se encuentran las variables más relevantes según el modelo aplicado. En la columna con la etiqueta “Estimado” se ofrecen los resultados sobre el sentido de la relación en una escala de log(odds). Los valores en el intervalo de 0 a $+\infty$ indican una relación positiva, es decir, un incremento en la probabilidad de que ocurra el fenómeno variable estudiado. Los valores en la escala de 0 a $-\infty$ señalan una relación

negativa, es decir, un decrecimiento en la probabilidad de que se presente el fenómeno de estudio. En las siguientes dos columnas se proporcionan los valores de error estándar y de Z. En la última columna se presentan los valores de probabilidad, que indican si la variable en cuestión es significativa, de acuerdo con un nivel de significación menor a 0.05.

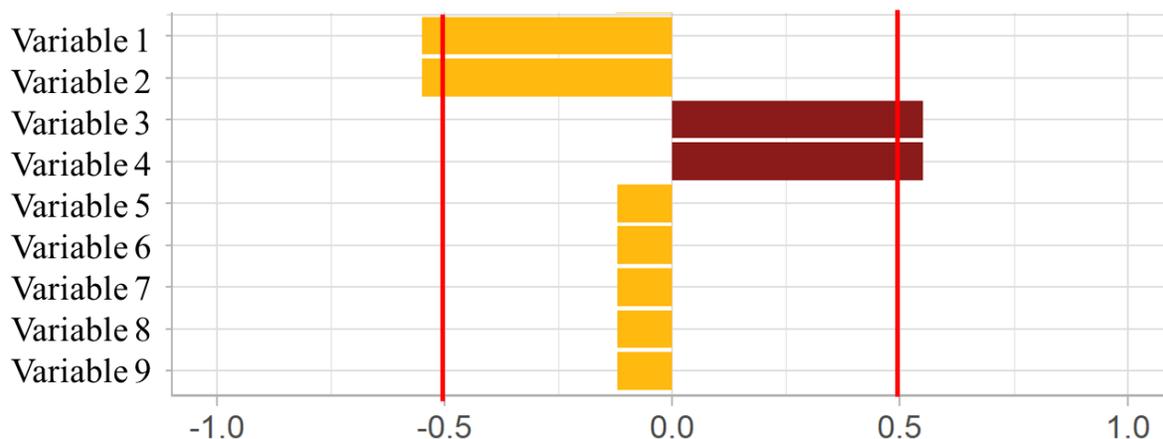
2.1.3.5 Prueba de inferencia multimodelo

En algunos análisis, específicamente en aquellos donde comparo los diferentes grupos de bilingües, utilicé una aproximación estadística diferente basada en la teoría de la información. A diferencia de las pruebas de hipótesis nulas, este tipo de enfoque compara todos los modelos candidatos, los clasifica e identifica el mejor en función de su valor del criterio de información. Este método permite realizar una inferencia multimodelo, en lugar de utilizar un único modelo (Burnham & Anderson, 2002).

Para ello, utilicé el paquete `glmulti`⁴⁹ para la versión 3.01 de R. Clasifiqué los modelos de mejor a peor según el criterio de información de Akaike. El conjunto de modelos con una diferencia $AIC < 2$ del mejor modelo fueron considerados con la misma plausibilidad (Calcagno & Mazancourt, 2010). La importancia relativa de cada predictor se calculó a partir de la suma de los pesos de Akaike ($\sum W_i$) de cada modelo en el que el predictor aparece. Los resultados de esta prueba se presentan en gráficas y tablas como las que se muestran enseguida.

⁴⁹ `glmulti` proporciona un envoltorio general para `glm` y funciones relacionadas. Se pretende que sea flexible para que el apoyo a cualquier función de este tipo, y cualquier criterio de información, sea fácil de implementar. Admite fórmulas e interacciones entre variables. Una dificultad obvia con este tipo de enfoque es que el número de modelos a considerar se vuelve fácilmente prohibitivo. Por esta razón, `glmulti` está optimizado para minimizar los requisitos de memoria y el tiempo de cálculo, lo que permite gestionar conjuntos de candidatos muy grandes. El resultado de un análisis `glmulti` es un objeto que contiene el conjunto de confianza de los modelos (es decir los n mejores modelos) y su soporte (Calcagno & Mazancourt, 2010).

FIGURA 2.6 Muestra de los resultados de la prueba de inferencia multimodelo



En la figura anterior se ejemplifica el tipo de gráfico utilizado en los capítulos de análisis. En la gráfica se presentan las variables predictoras en el conjunto de modelos $\Delta AICc < 2$ y el 95% del conjunto de modelos para el fenómeno en cuestión. La importancia de cada variable se muestra mediante la suma de los pesos de Akaike. El signo (\pm) de la estimación de los parámetros representa el efecto (positivo o negativo) de cada predictor sobre la variable de respuesta.

En la tabla inferior se ejemplifica la manera en que se reportan los resultados de la selección de modelos basados en información teórica e inferencia multimodelo para el fenómeno en cuestión. Detallo el conjunto de modelos $\Delta AICc < 2$. Con "X" indico la inclusión de la variable en cada modelo. Las variables explicativas se presentan a la cabeza de las columnas.

TABLA 2.9 Ejemplo de los resultados de la prueba de inferencia multimodelo

Modelo	Var. 1	Var. 2	Var. 3	Var. 4	Var. 5	Var. 6	AIC	Var
Wi	Valor	Valor	Valor	Valor	Valor	Valor		
1	X	X	X				Valor	Valor
2	X	X	X	X			Valor	Valor
3				X	X	X	Valor	Valor
4		X	X		X	X	Valor	Valor

Después de haber descrito las herramientas estadísticas empleadas en el análisis de los dos fenómenos variables asociados al número, en el siguiente apartado proporciono una breve descripción de la historia y la situación sociolingüística de San Andrés Cuexcontitlán y Jiquipilco el Viejo.

2.2 DESCRIPCIÓN DE LA REALIDAD SOCIOLINGÜÍSTICA DE LAS COMUNIDADES BILINGÜES

OTOMÍ-ESPAÑOL

2.2.1 Brevísima historia de las dos comunidades

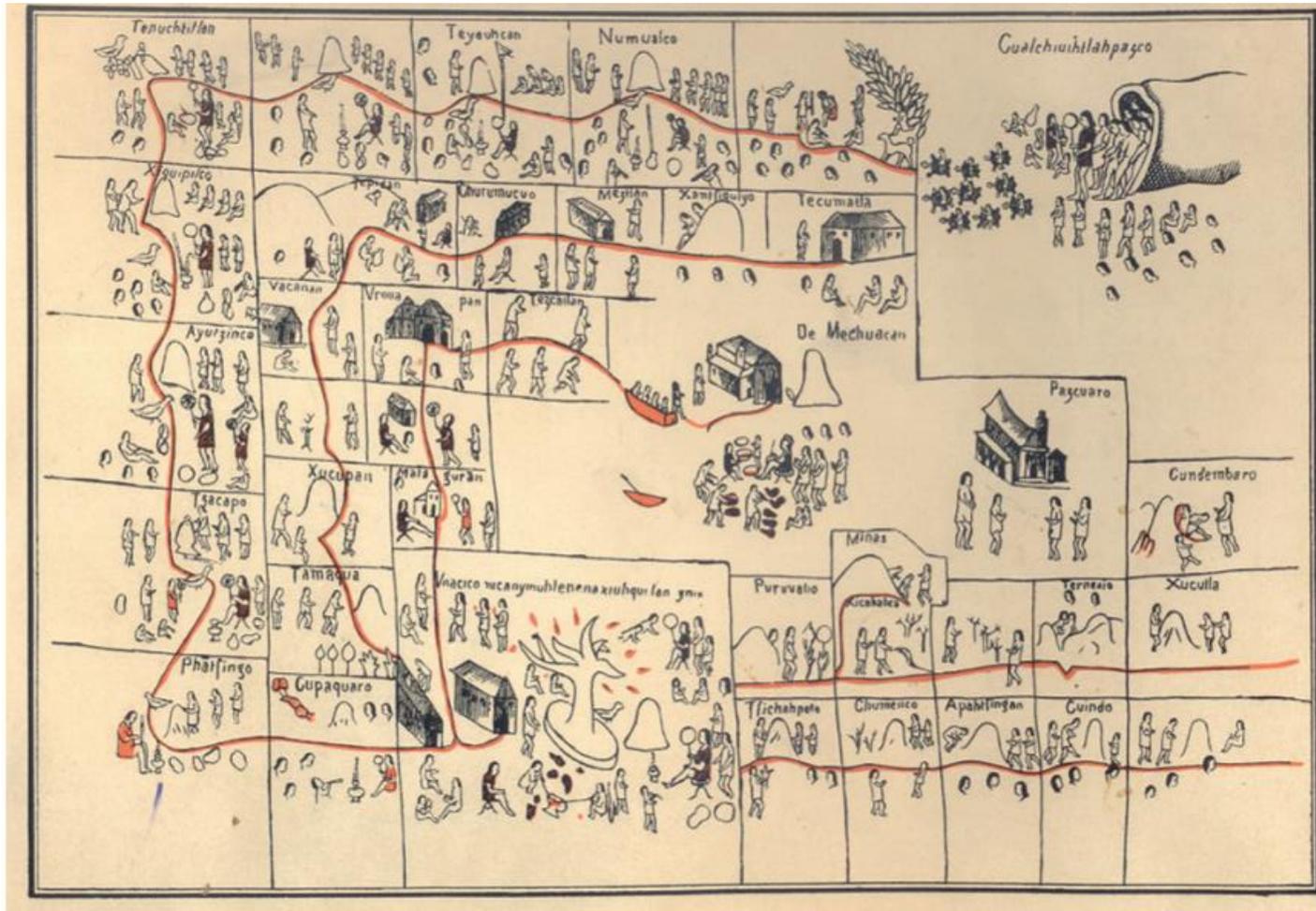
Jiquipilco el Viejo es un asentamiento otomí, cuyo origen se remonta al siglo XV,⁵⁰ según algunos documentos como el Lienzo de Jucutacato (1530-1540) y el Códice Telleriano Remensis. El Lienzo de Jucutacato describe la ruta que siguieron los purépechas desde su lugar de origen hasta el territorio que ocuparon en el estado de Michoacán (Tejeda Alvarado, 2019: 42). En esta obra, Xiquipilco⁵¹ es representado como un punto intermedio en el viaje. En la Figura 2.7 se muestra una de las versiones del Lienzo de Jucutacato. El recuadro que lleva por título *Xiquipilco* consta de dos escenas:

⁵⁰ Jiquipilco el Viejo es conocido como *Ndonguu* “Lugar de la casa antigua”, lo que indica que se trató de un asentamiento importante durante la época prehispánica (Salazar Mondragón, 2019: 122).

⁵¹ *Xiquipilco* se traduce como ‘lugar de ocho mil’, de *xiquipilli* ‘ocho mil’ y *co* ‘lugar de’. Este lugar se representa con una bolsa. Con base en eso, sugieren que el topónimo indica ‘lugar donde se hacen bolsas o costales para el cacao’ (Tejeda Alvarado, 2019: 177).

en uno de ellos las figuras que se aproximan al cerro, tomado como parámetro identificativo del lugar al que han llegado; dos figuras se dirigen a él caminando, mientras que el conjunto de cuatro, ubicados en el otro extremo de la figura jeroglífica, marcan un alto, por su posición de estar sentados. En el segundo momento, el conjunto que se representa por el personaje guía del grupo y los elementos que conforman el carácter reverente del momento en que se llega al nuevo lugar seleccionado en su camino (Tejeda Alvarado, 2019: 177)

FIGURA 2.7 Lienzo de Jucutacato



Lienzo de Jucutacato, Michoacán: paisajes, tradiciones y leyendas; Eduardo Ruiz, México, 1940.

El Códice Telleriano Remensis fue escrito a mediados del siglo XVI, aproximadamente entre 1562 y 1563. La obra consta de tres partes: un calendario con periodos de veinte días de los años solares (ff. 1r-7r), un ciclo adivinatorio de 260 días (ff. 8r-24r) y una sección de anales históricos que comprende un periodo de 1198 a 1562 (25r-49r) (Iwaniszewski, 2004: 45). En la tercera sección de esta obra se documentan dos batallas acaecidas en Xiquipilco, de acuerdo con el jeroglífico del *xiquipilli*, que es la bolsa que utilizaban los sacerdotes para guardar el copal usado en sus ceremonias. El primer combate ocurrió en 1462 y se enfrentaron los habitantes de Michoacán contra los de Xiquipilco (folio 33v). El segundo encuentro acaeció en 1473 entre los aztecas y los habitantes de Xiquipilco (folio 37v) (Arroyo Estrada, 1985: 31); ambos folios se muestran en la Figura 2.8.

Durante los primeros años de la conquista, se estableció una encomienda en el territorio de Xiquipilco y se construyó un templo dedicado a San Juan. Por ello, esta comunidad fungió como cabecera política y religiosa, y gozó de cierto estatus durante la dominación española, posiblemente por la importancia que tuvo durante la época prehispánica (Arzate Becerril, 1996: 183 apud Palma Patricio, 2018: 83). Después de que el templo se incendió, los habitantes decidieron trasladarse a otro lugar. Unos migraron al norte y, en 1593, fundaron el pueblo de San Juan Jiquipilco, el cual adquirió la categoría de municipio en 1822. Otra parte de la población emigró al sur y constituyó el pueblo de Temoaya en 1560. Estos hechos causaron que Xiquipilco poco a poco fuera perdiendo importancia como pueblo prehispánico y encomienda colonial (Arroyo Estrada, 1985: 36; Palma Patricio, 2018: 83-84).

FIGURA 2.8 Batallas acaecidas en Xiquipilco (Códice Telleriano-Remensis, f. 33v, 37v)



En cuanto a la comunidad de San Andrés Cuexcontitlán, no tenemos datos exactos acerca de su fundación, pero sí sabemos que en 1746 esta comunidad ya formaba parte de Toluca y estaba compuesta por 134 familias (Villaseñor y Sanches, 1952: 237). Pese a que no contamos con información sobre las lenguas que hablaban estos pobladores, es posible suponer que eran hablantes de otomí, pues se tiene registros de la existencia de asentamientos otomí-mazahua en esta zona desde el periodo epiclásico (Lastra, 2010: 81).

En la primera mitad del siglo XX, la producción agrícola era la principal actividad económica de los habitantes de Jiquipilco el Viejo y San Andrés Cuexcontitlán. La crianza de ganado ovino y otros animales, así como la producción de chincuetes, fajas, productos de lana (colchas, cojines, gabanes), tapetes anudados, molcajetes y sombreros complementaban la actividad económica. Sin embargo, los cambios en los precios internacionales de los productos agrícolas y la variabilidad en la demanda de los productos artesanales afectaron la economía de la población de estas comunidades. Aunado a esto, en 1940, se fundó el corredor industrial Lerma-Toluca, que abarca parte de los municipios de Toluca, Lerma, San Mateo Atenco, Zinacantepec y Huixquilucan. El proceso de industrialización y urbanización de esta zona modificó el espacio geográfico, la estructura y la organización social de las comunidades de Jiquipilco el Viejo y San Andrés Cuexcontitlán, así como las actividades económicas y la movilidad de sus habitantes (Barrientos López, 2004: 12-14).

Algunos habitantes de SAC y JV viajan diariamente a las zonas industriales del Estado de México, especialmente al corredor industrial Lerma-Toluca, en donde se desempeñan mayormente como obreros industriales. También viajan a la Ciudad de México y a diferentes puntos urbanos (Nezahualcóyotl, Naucalpan y Tlanepantla), de lunes a viernes, desempeñan trabajos relacionados con la albañilería, el trabajo doméstico, la seguridad pública y privada, y el comercio ambulante. Los fines de semana regresan a sus comunidades

en Toluca y Temoaya. También se mueven a diversos puntos de la República Mexicana, donde comercian con artesanías y otros productos. Aunado a esto, en años recientes ha aumentado la migración de la población de esta zona a los Estados Unidos, sobre todo en los municipios de Jiquipilco y Temoaya (Barrientos López, 2004: 12-14).

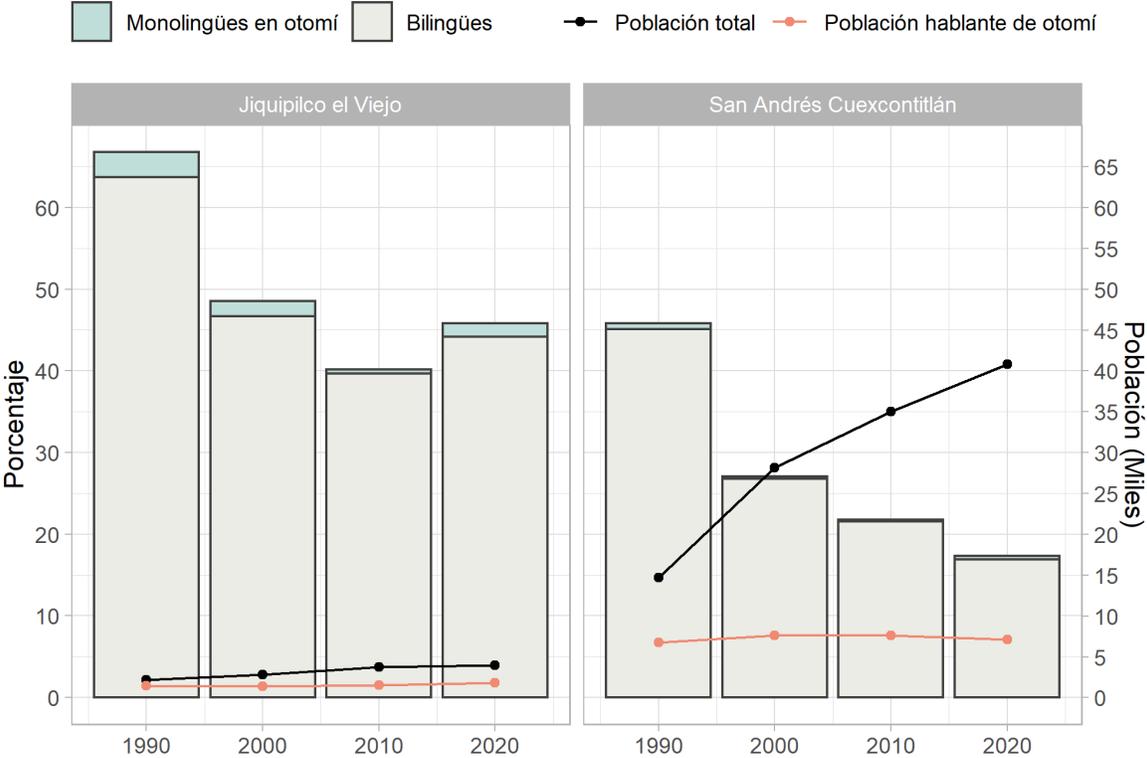
Todos estos cambios sociales y económicos incrementaron la intensidad del contacto entre el otomí y el español en San Andrés Cuexcontitlán y Jiquipilco, en especial a partir de 1950. Además, tuvo lugar una ruptura en la transmisión de la lengua otomí alrededor de los años ochenta y noventa, debido a las políticas educativas implementadas en los años setenta (González Villarreal, 2018) y al incremento de los movimientos migratorios estacionales a la Ciudad de México y a diferentes puntos de la República Mexicana. Estos hechos se constatan en las observaciones que Lastra hizo sobre SAC en 1987, en las que subraya el alto grado de bilingüismo de la comunidad, pero también deja ver algunos indicios sobre la ruptura en la transmisión del otomí, al señalar la existencia de niños que no adquirieron esa lengua y que sólo poseen un conocimiento pasivo de la misma: “Sus habitantes son casi todos bilingües, aunque hay algunos monolingües en otomí y algunos niños no tienen dominio activo, aunque sí pasivo del otomí” (Lastra, 1987: 39).

2.2.2 Hablantes de otomí en las dos comunidades

Los cambios sociales antes descritos han venido modificando la estructura y organización social de las comunidades, como lo muestran los datos censales de la Figura 2.9. En la gráfica se presenta el número de habitantes y hablantes de otomí de San Andrés Cuexcontitlán y Jiquipilco el Viejo. En el eje horizontal se muestran los resultados correspondientes al periodo de 1990 a 2020. En el eje vertical de la izquierda se codifica el porcentaje de bilingües otomí-español y monolingües en otomí, cuyos valores se grafican en las barras de color

verde. En el eje vertical de la derecha se presenta la población de las dos comunidades en miles. La línea de color negro indica el número de habitantes y la línea roja el número de hablantes de otomí.

FIGURA 2.9 *Porcentaje de hablantes de otomí en SAC y JV*



En 1990, Jiquipilco el Viejo tenía un total de 2 118 habitantes, de los cuales el 66.8% eran hablantes de otomí (bilingües = 1350, monolingües = 65) (Instituto Nacional de Estadística, 1990). Diez años después, la población de esta localidad incrementó ligeramente a 2 804 personas y el porcentaje de hablantes de otomí descendió a 48.53% (bilingües = 1308, monolingües = 53) (Instituto Nacional de Estadística, 2000). En 2010, el número de habitantes de JV ascendió a 3 724 y el porcentaje de hablantes de otomí disminuyó a 40.19% (bilingües = 1477, monolingües = 20) (Instituto Nacional de Estadística, 2010). Actualmente, la población de esta comunidad es de 3 965 personas y el porcentaje de hablantes de otomí

asciende a 45.85% (bilingües = 1753, monolingües = 65) (Instituto Nacional de Estadística, 2020).

De manera general, los datos indican que la población de Jiquipilco el Viejo no ha incrementado sustancialmente en las últimas tres décadas, posiblemente a causa de un aumento en la migración a diferentes centros urbanos y a los Estados Unidos. Por otro lado, el número de hablantes de otomí ha disminuido del 66.80% al 45.85%. Aunado a ello, las observaciones realizadas durante el trabajo de campo indican que la población más joven de la comunidad no habla la lengua otomí, sólo la entiende, lo que sugiere la existencia de una ruptura en la transmisión de la lengua otomí.

En 1990, San Andrés Cuexcontitlán tenía 14 673 habitantes y el número de hablantes de otomí comprendía el 45.84% (bilingües = 6 620, monolingües = 107) de la población total (Instituto Nacional de Estadística, 1990). Para el año 2000, el número de habitantes incrementó casi al doble 28 148 y el porcentaje de hablantes otomí se redujo a un 27% (bilingües = 7 539, monolingües = 86) (Instituto Nacional de Estadística, 2000). Diez años más tarde, la población de esta comunidad ascendió a 35 011 y el número de hablantes de *ñható* descendió al 21% (bilingües = 7 279, monolingües = 81) (Instituto Nacional de Estadística, 2010). El último censo indica que SAC tiene 40 839 habitantes, de los cuales apenas el 17.31% habla otomí (bilingües = 6 911, monolingües = 162) (Instituto Nacional de Estadística, 2020). Los datos de la Figura 2.9 nos muestran un acelerado crecimiento poblacional asociado al proceso de periurbanización de la ciudad de Toluca, que ha generado un descenso importante de los hablantes de otomí. Aunado a esto, la mayoría de las personas bilingües *ñható*-español son mayores de 40 años, lo cual también sugiere una ruptura en la transmisión de la lengua otomí. Los cambios en la estructura social de las dos comunidades

se reflejan en las características lingüísticas y sociales de los colaboradores, como analizaremos en los siguientes dos apartados.

2.2.3 Características lingüísticas de los hablantes de otomí en las dos comunidades

Como señalamos en el apartado §1.2, en este trabajo partí de un concepto amplio de bilingüismo, entendido como una persona que tiene una competencia nativa en una lengua y alguna competencia en alguna de las cuatro habilidades (hablar, entender, escribir, leer) de otra. Además, al igual que Grosjean (2004, 2012), pienso que dos de los factores determinantes para definir el bilingüismo son el nivel de fluidez o competencia, y el uso de las lenguas. Asimismo, concuerdo con la idea de que los bilingües están influidos por el principio de complementariedad; por ello, difícilmente son equilibrados en las dos lenguas y en todas las habilidades lingüísticas. De acuerdo con estos planteamientos, la clasificación de bilingüismo de esta investigación se basa en los resultados de las pruebas de habilidad lingüística, que proporcionan información acerca del conocimiento que los bilingües tienen de sus lenguas, y en los datos del cuestionario sociolingüístico, cuyos resultados brindan información acerca del uso de las lenguas con diferentes interlocutores y en distintos dominios sociales.

En primer lugar, realicé la evaluación de los cuestionarios de habilidades lingüísticas por sección (fónica, léxica y sintáctica) y, a partir de esos resultados, hice una valoración global de cada prueba. En los tres cuestionarios empleé una escala de 1 a 10, pues partí del supuesto de que los colaboradores, al ser bilingües, obtendrían un rendimiento por arriba de los 5.00 puntos. Con respecto a las habilidades lingüísticas, establecí un índice en el que consideré de los 5.00 a los 6.50 puntos como baja competencia, el intervalo que comprende

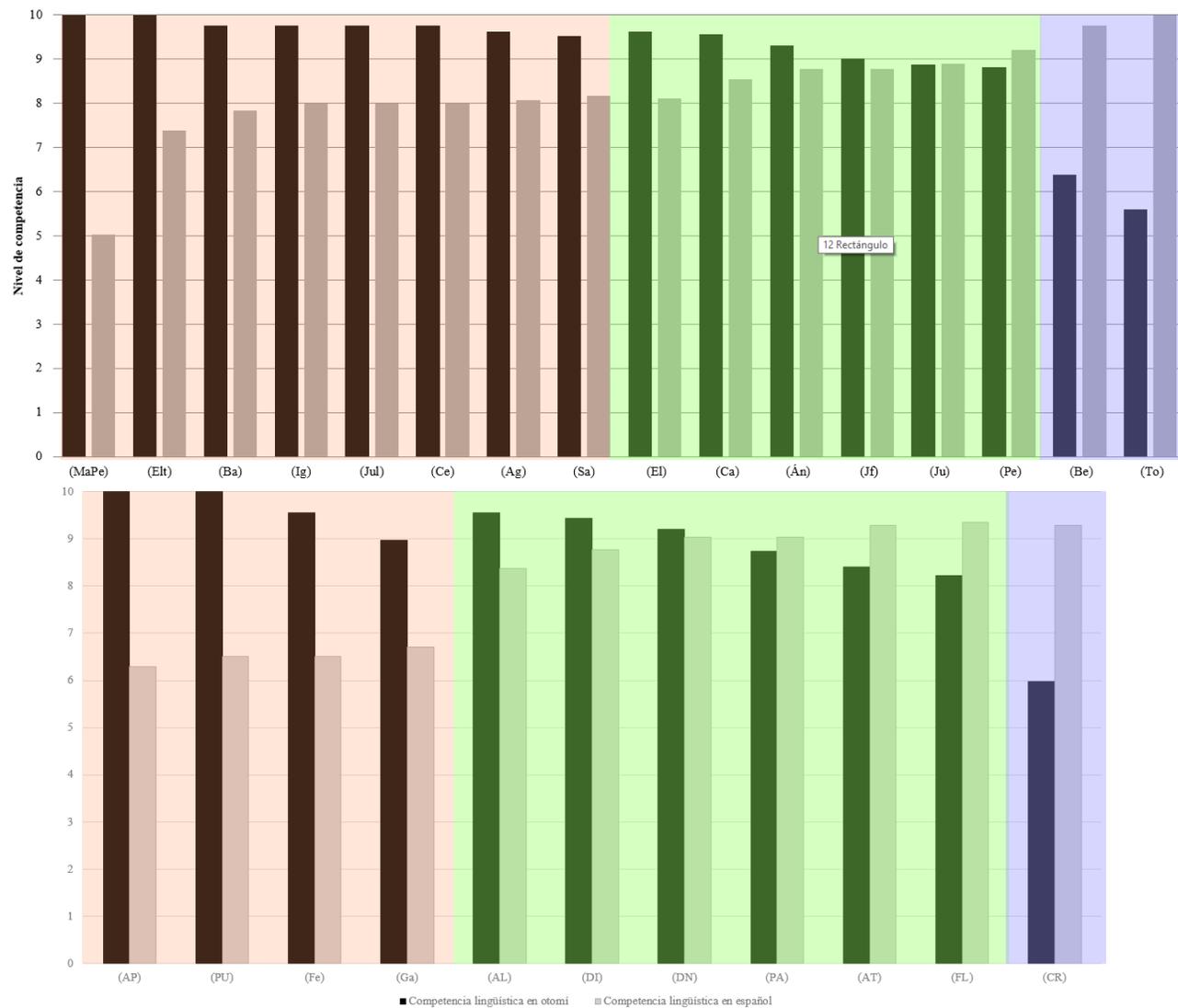
de los 6.60 a los 8.20 puntos corresponde a la mediana competencia y, por último, tomé de los 8.30 a los 10.00 como competencia alta.

Con base en esos parámetros y en la escala de bilingüismo de esta investigación (véase Figura 1.6), ponderé el conocimiento que los cuarenta y seis colaboradores tienen de sus lenguas, y constituí cinco grupos con características lingüísticas y sociolingüísticas específicas: bilingües con mayor dominio del otomí (BO), bilingües simétricos (BS), bilingües con mayor dominio del español (BE), monolingües con conocimiento pasivo del otomí (PO) y monolingües en español (ME).

En el grupo de bilingües con mayor dominio del otomí (BO) incluí a los colaboradores que presentaron un nivel alto de habilidades lingüísticas en otomí y uno bajo o mediano en español. El conjunto de bilingües simétricos (BS) está compuesto por los hablantes que alcanzaron un nivel de habilidad alto en ambas lenguas. Los bilingües con mayor dominio del español (BE) son los hablantes que tienen un nivel de habilidad alto en español y uno bajo en otomí. Por último, los monolingües en español con conocimiento pasivo del otomí (PO) fueron aquellos que no alcanzaron 5.00 pts. en la prueba de otomí y obtuvieron más de 6.00 pts. en la prueba de comprensión del otomí.

En la Figura 2.10 se presentan los resultados de las pruebas de habilidad lingüística en español y en otomí de los bilingües de San Andrés Cuexcontitlán y Jiquipilco el Viejo. Los hablantes en fondo naranja pertenecen al grupo de bilingües con mayor dominio del otomí (BO), el color verde identifica a los bilingües simétricos (BS) y el azul enmarca a los bilingües con mayor dominio del español (BE).

FIGURA 2.10 Resultados de las pruebas de habilidades lingüísticas en SAC y JV



Los bilingües con mayor dominio del otomí (BO) presentaron un nivel de habilidad bajo-medio en español (5.10-8.20 pts.) y uno alto en otomí (8.30-10.00 pts.). La mayoría de estos bilingües son del grupo etario cuatro y tienen un nivel de instrucción bajo. Su otomí tiene un número considerable de palabras del español, que han sido adaptadas fonológica y morfosintácticamente. Por otro lado, su variedad de español presenta numerosas transferencias en los niveles fonológico y morfosintáctico, como la omisión de fonemas a inicio o final de palabra, el uso alterno de preposiciones y de pronombres de objeto directo e indirecto.

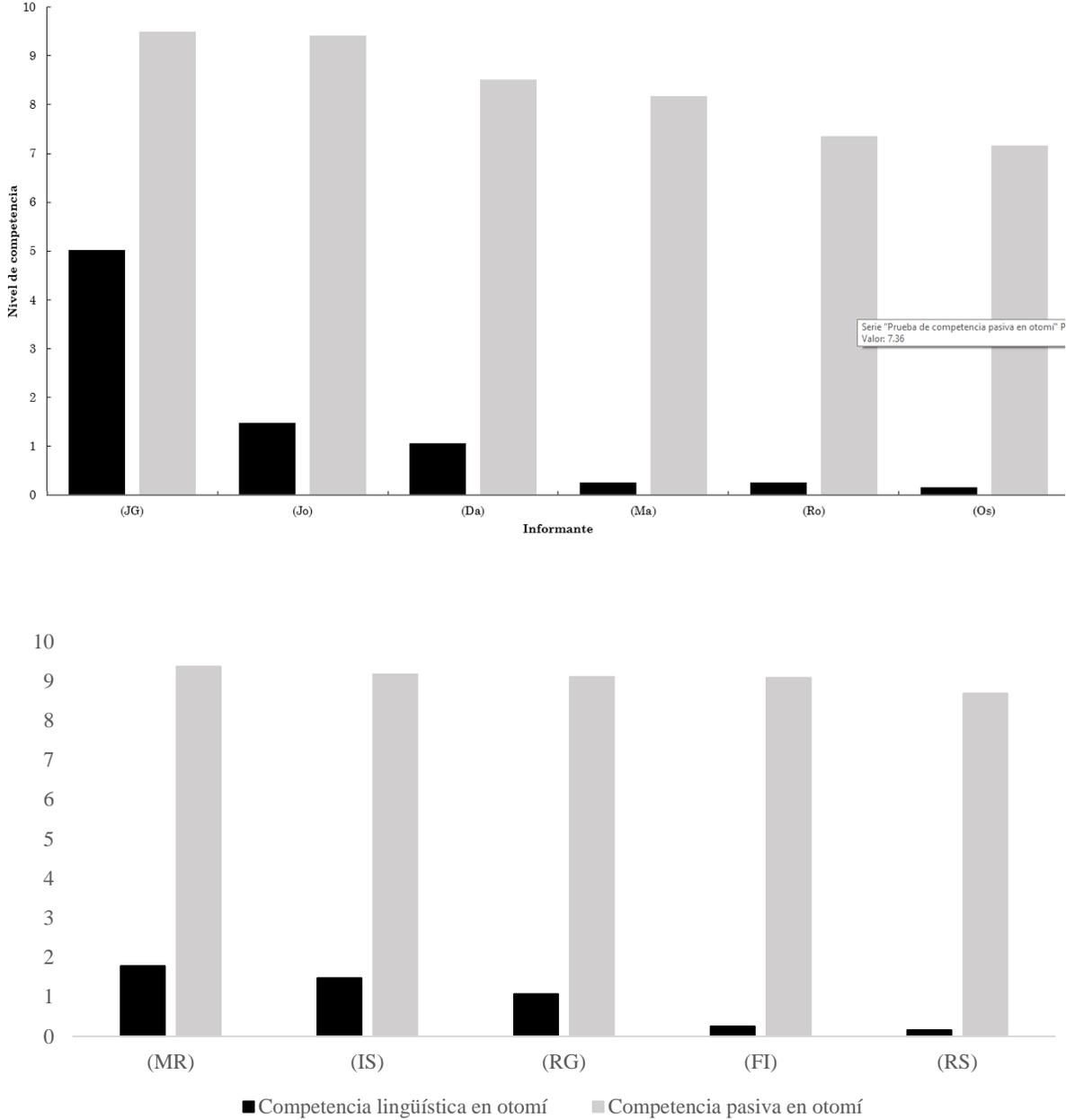
Los bilingües simétricos (BS) tienen una habilidad lingüística alta (8.30-10.00 pts.) en ambas lenguas. La mayoría son del tercer grupo de edad y tienen como máximo 6 años de instrucción. El otomí de este conjunto presenta características similares al de los bilingües con mayor dominio del otomí, pero el número de préstamos del español parece ser mayor. Su español también presenta transferencias en los niveles fonológico y morfosintáctico, aunque en menor porcentaje que los bilingües con mayor dominio del otomí.

Los bilingües con mayor dominio del español (BE) se caracterizan por tener un nivel de habilidades lingüísticas en otomí bajo (5.10-6.50 pts.) y uno alto en español (8.30-10.00 pts.). Los miembros de este conjunto son de los grupos etarios 3 y 4, y tienen un nivel de instrucción medio-alto. En comparación con los otros grupos de bilingües (BO y BS), su otomí tiene un mayor número de préstamos del español y gran parte de estos no son adaptados a la fonología y morfosintaxis del *n̄ható*. En contraste, su variante de español no presenta transferencias del otomí.

Los monolingües con conocimiento pasivo del otomí (PO) obtuvieron menos de 5.00 puntos en la prueba de habilidad lingüística en otomí, pero un promedio arriba de 6.00 pts. en la prueba de conocimiento pasivo, como se puede observar en la Figura 2.11. Los

miembros de este conjunto son del segundo grupo etario, tienen un nivel de instrucción medio-alto y su español no presenta transferencias del otomí.

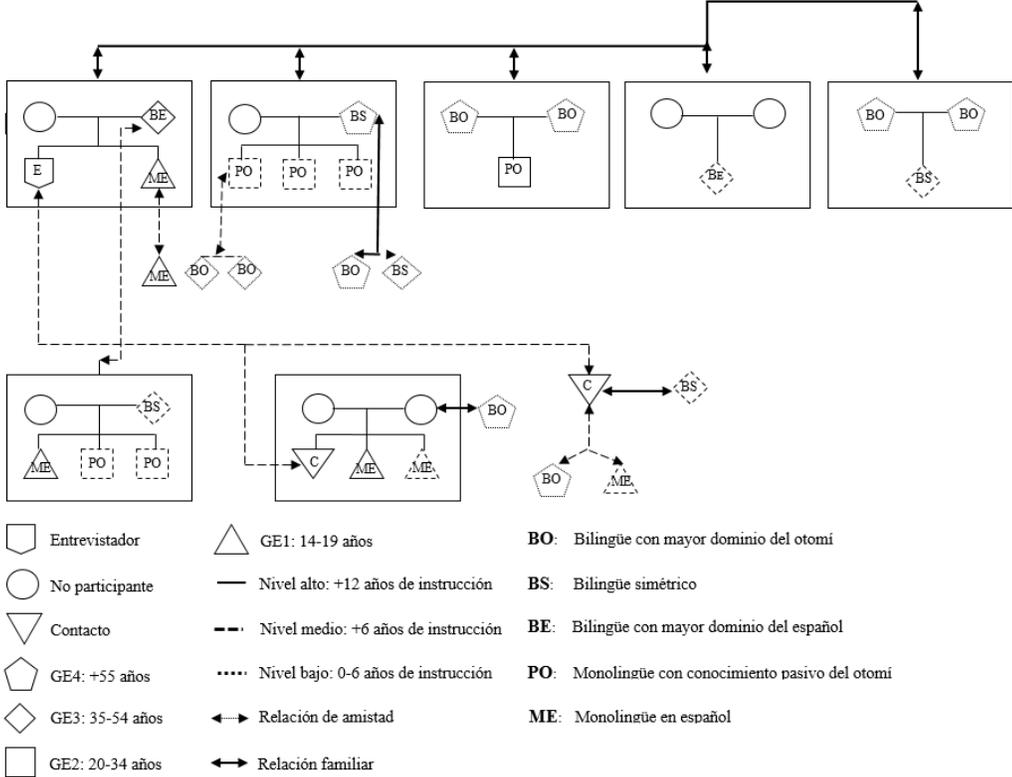
FIGURA 2.11 Resultados de la prueba de habilidad lingüística pasiva en otomí en SAC y JV



Los monolingües en español (ME) son del primer grupo etario (14-19 años) y tienen un nivel de instrucción medio-alto. Este conjunto tiene una competencia nativa del español

y sólo tiene un conocimiento simbólico de la lengua otomí, que incluye algunos saludos, nombres de plantas y animales, frases sencillas, por lo que únicamente se mueven en redes monolingües en español. A continuación, en la Figura 2.12 se muestra la distribución de estos grupos lingüísticos en la red de San Andrés Cuexcontitlán.

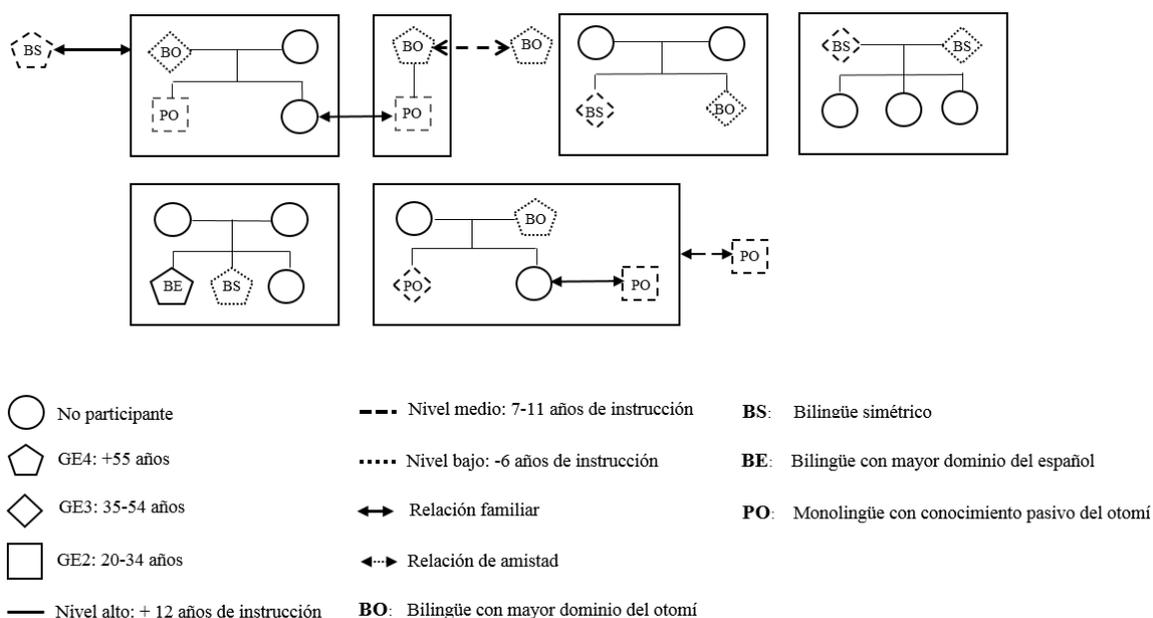
FIGURA 2.12 Clasificación de bilingüismo de los colaboradores de SAC



En San Andrés Cuexcontitlán están representados los cinco grupos lingüísticos. Ocho personas tienen como lengua dominante el otomí (BO), hay seis bilingües simétricos (BS), dos bilingües con mayor dominio del español (BE), seis jóvenes que tienen diferentes grados de conocimiento pasivo del otomí (PO) y seis adolescentes monolingües en español con un conocimiento simbólico del *n̄ható* (ME). En Jiquipilco el Viejo están representados cuatro de los cinco grupos de la tipología. La red consta de cinco personas que tienen como lengua

dominante el otomí (BO), cinco bilingües con un conocimiento alto de las dos lenguas (BS), una sola colaboradora con un conocimiento dominante del español (BE) y cinco colaboradores que no hablan *ñható*, pero tienen la capacidad de entenderlo.

FIGURA 2.13 Clasificación de bilingüismo de los colaboradores de JV



Antes de describir el uso del otomí y el español con diferentes interlocutores y en varios dominios social, es importante hacer algunas anotaciones en torno al uso de las pruebas de habilidades lingüísticas en la caracterización y clasificación de los bilingües. Como he señalado anteriormente (véase §1.2), el bilingüismo es un fenómeno complejo que comprende aspectos psicológicos, lingüísticos, sociales y cognitivos, por esta razón, todavía no se tiene una medida o herramienta que evalúe de manera integral las habilidades de un bilingüe.

Los cuestionarios de esta investigación permitieron medir de forma objetiva las habilidades bilingües de los colaboradores y poder compararlos entre sí a partir de los mismos criterios. No obstante, los instrumentos también tienen limitaciones. En primer lugar, no

contemplé la producción y comprensión en todos los niveles lingüísticos y el número de ítems fue pequeño, por ello, esta herramienta no logra ascir de forma global la complejidad que conlleva el bilingüismo.

Otro de los puntos en contra de las pruebas fue el haber utilizado como índice de ponderación la presencia o ausencia de transferencias cuando la hipótesis central de esta investigación es la existencia de un efecto bidireccional del contacto en un mismo dominio lingüístico (la marcación de número). Sobre esto, es importante decir que la comparación de las habilidades bilingües con las de los monolingües y de las variedades sin contacto con las de contacto ha sido una práctica criticada en la evaluación del bilingüismo, y, al mismo tiempo, considerada como necesaria para comprobar que un fenómeno lingüístico es efecto del contacto y no de la “evolución interna de la lengua”. Práctica que, por cierto, no da cuenta de los cambios indirectos inducidos por contacto, fenómenos multicausales que pueden estar motivados de manera interna y externa (Palacios Alcaine, 2011; Thomason, 2001).

Este problema teórico-metodológico no se ha podido superar hasta el momento y, lamentablemente, no se he podido hallar un mejor método para caracterizar el bilingüismo de los colaboradores de esta investigación. En trabajos futuros sería recomendable hallar una metodología congruente con la idea de bidireccionalidad, que pondere las habilidades de producción y comprensión de los diferentes niveles lingüísticos a partir de un mayor número preguntas o ítems. A continuación, procederé a describir el uso que los bilingües hacen del otomí y el español con diferentes tipos de interlocutores y en varios dominios de uso.

2.2.4 El uso del otomí y el español en las dos comunidades

Los tres grupos de bilingües (BO, BS y BE) presentes en San Andrés Cuexcontitlán y Jiquipilco el Viejo no sólo difieren con respecto al dominio que tienen del otomí y el español,

sino también en el uso que hacen de sus lenguas. Por ello, el objetivo de esta sección será describir el uso que los bilingües hacen del otomí y el español con diferentes interlocutores y dominios sociales, a partir de los datos recabados con la sección “Uso de las lenguas” del cuestionario sociolingüístico (véase Apéndice 2D). Los interlocutores que considera el cuestionario sociolingüístico pueden agruparse en cinco conjuntos, según la edad y el grado de cercanía. El primer grupo está constituido por los abuelos de los entrevistados que representan, al menos, dos generaciones arriba. El segundo conjunto está integrado por la mamá, el papá, los suegros y los tíos del colaborador, que personifican una generación mayor. En la tercera agrupación se encuentran los hermanos, los primos, la pareja, los compadres y amigos de los hablantes. Por último, están los hijos y nietos de los bilingües, que representan una o dos generaciones menores. En cuanto a los dominios sociales, el cuestionario comprende seis espacios: hogar, fiestas, mercados, iglesia, trabajo y ciudad.

La Figura 2.14 muestra tres gráficas con el porcentaje de uso de las lenguas por grupo lingüístico. En la parte superior están los bilingües con mayor dominio del otomí (BO), en medio los bilingües simétricos (BS) y en la parte inferior se muestran los resultados de los bilingües con mayor dominio del español (BE), con los diferentes tipos de interlocutores por grupo lingüístico.⁵² El eje vertical muestra el porcentaje de uso correspondiente al español y al otomí de forma exclusiva y la posibilidad de usar las dos lenguas.⁵³

⁵² La gráfica no contempla los resultados de los monolingües con conocimiento pasivo del otomí (PO) ni de los monolingües en español (ME), debido a que los dos tipos de hablantes usan el español con los diferentes interlocutores considerados en el cuestionario sociolingüístico.

⁵³ Según Grosjean (2012: 14-15), cuando un bilingüe interactúa con un interlocutor o usa la lengua en una situación comunicativa se hace dos preguntas fundamentales: ¿qué lengua hay que usar? y ¿hay posibilidad de utilizar la otra lengua? Al inicio de la interacción, las dos lenguas del bilingüe están inactivas y tiene que elegir entre una lengua A y una lengua B. Si el bilingüe elige la lengua A, esta se activa. Esta primera operación se denomina "elección de lengua" y la lengua elegida se denomina "lengua base". Posteriormente, el bilingüe se plantea la segunda pregunta: "¿hay posibilidad de introducir la lengua B?" Si la respuesta es "no" (imagine que el bilingüe está hablando con alguien que sólo conoce una de sus lenguas), entonces la otra lengua permanece

Los bilingües con mayor dominio del otomí (BO) se comunican con sus abuelos en la lengua otomí. En el grupo de interlocutores de una generación mayor, los BO usan el *ñható* para comunicarse con su madre y padre, con excepción de un hablante. Asimismo, la mayoría de los BO usan exclusivamente el otomí para comunicarse con sus suegros y tíos, pero comienza a haber una mayor presencia del uso bilingüe, incluso uno de los hablantes afirma que sólo habla español con sus tíos.

inactiva y sólo se utilizará una lengua. Esto se denomina modo monolingüe, en el que el bilingüe suele utilizar sólo una lengua y desactivar la otra. Si la respuesta a la segunda pregunta es "sí" (por ejemplo, el bilingüe está hablando con amigos bilingües que comparten sus lenguas), entonces la otra lengua se activa, pero menos que la lengua base, por si el bilingüe la necesita durante la interacción. En este caso, el bilingüe está en modo bilingüe y puede aportar elementos de la otra lengua o incluso cambiar completamente de lengua base.

FIGURA 2.14 Uso de las lenguas con diferentes interlocutores. Grupos lingüísticos de SAC



El uso de la lengua otomí disminuye cuando el interlocutor es de la misma generación, pues la mayoría de los BO utilizan el otomí y el español con sus hermanos y pareja. El uso bilingüe incrementa aún más con los compadres y los amigos. De hecho, con los primos empieza a haber un uso exclusivo del español. Los BO disminuyen el uso bilingüe con los interlocutores una generación menor (hijos, yerno/nuera) e incrementa la exclusividad del español, y con los nietos se vuelve la lengua dominante.

Los bilingües simétricos emplean el *ñható* para comunicarse con sus abuelos. La mayoría de los BS sólo utilizan el otomí con sus padres y suegros, aunque algunos afirman emplear las dos lenguas o únicamente español. El uso exclusivo del *ñható* disminuye cuando se comunican con los tíos e incrementa la alternancia entre el otomí y el español. Este uso bilingüe predomina en la interacción que los BS establecen con individuos de la misma generación (hermanos, amigos, primos y parejas), salvo con sus compadres, con quienes hablan principalmente español. Por último, los BS sólo se comunican en español con sus hijos, yernos, nueras y nietos, con excepción de un colaborador, que usa las dos lenguas con sus hijos.

Los bilingües con mayor dominio del español interactúan con sus abuelos en otomí. En lo que respecta a los hablantes de la siguiente generación, uno de ellos sólo utiliza *ñható* con su mamá, debido a que ella es monolingüe en esa lengua; y el otro hablante alterna en las dos lenguas. Los BE pueden usar el otomí y el español con sus padres y tíos; sin embargo, con sus suegros, sólo se comunican en español. Con los interlocutores de la tercera generación –a la que pertenecen los hermanos, primos, parejas, amigos y compadres–, uno de los BE utiliza las dos lenguas y el otro sólo el español. Finalmente, los BE hablan español con sus hijos, yernos/nueras y nietos.

En resumen, los tres grupos de bilingües (BO, BS, BE) de SAC tienden a utilizar la lengua otomí con los interlocutores que son una o dos generaciones mayores. Predomina un uso bilingüe con las personas de una edad similar, e impera el español con los más jóvenes. Esta diferenciación en el uso de las lenguas según la edad del interlocutor nos permite apreciar un fenómeno de desplazamiento de la lengua otomí en San Andrés Cuexcontitlán.

La Figura 2.15 muestra los resultados de los bilingües de Jiquipilco el Viejo. Los bilingües con mayor dominio del otomí (BO) se comunican con sus abuelos y padres en otomí. En la interacción con los suegros y los tíos, también predomina el uso del *ñható*, pero algunos miembros comienzan a interactuar en las dos lenguas. Con respecto a sus coetáneos, la mayoría de los BO utilizan las dos lenguas con sus hermanos, primos, compadres y amigos; sin embargo, con estos tres últimos el uso del español incrementa. Por último, el español es la lengua imperante en la interacción de los BO con sus hijos, yernos, nueras y nietos, aunque prevalece el uso de las dos lenguas en algunos miembros.

FIGURA 2.15 *Uso de las lenguas con diferentes interlocutores. Grupos lingüísticos de JV*



Los bilingües simétricos (BS) sólo hablan otomí con sus abuelos. El uso de esta lengua también impera cuando interactúan con su mamá, papá, tíos y suegros; no obstante, comienza a ganar espacio el uso de las dos lenguas, sobre todo con los miembros de la familia extensa. La mayoría de los BS utilizan las dos lenguas con interlocutores que les son contemporáneos (hermanos, parejas, primos), y algunos hablan exclusivamente español, principalmente con las personas menos cercanas (compadres y amigos). El español se vuelve la lengua predominante en la comunicación con sujetos de menor edad y exclusiva con sus nietos.

La bilingüe con mayor dominio del español (BE) se comunica en otomí con sus abuelos. Esta hablante emplea las dos lenguas con su mamá; mientras que con su papá interactúa en español debido a que es monolingüe en esa lengua. La BE tiene ese mismo comportamiento con sus suegros y tíos. Por otro lado, esta bilingüe utiliza el otomí y el español con sus hermanos; sin embargo, con los interlocutores de la misma edad y más jóvenes sólo usa el español.

En síntesis, los bilingües con mayor dominio del otomí (BO) y los bilingües simétricos (BS) de JV prefieren comunicarse en *ñható* con las personas de mayor edad, usan el otomí y el español con los interlocutores que tienen la misma edad, y, en menor medida, con los más jóvenes, que hacen un uso exclusivo del español. Por otro lado, la BE usa exclusivamente el español con todos los tipos de interlocutores, con excepción de sus abuelos, mamá y hermanos. Sin embargo, debido a que sólo trabajé con una bilingüe de estas características, no es posible generalizar estos resultados.

Si comparamos el uso del otomí y el español en SAC y JV, observamos una tendencia similar. Los tres tipos de bilingües (BO, BS, BE) se comunican con sus abuelos en otomí porque antes la mayoría de la población era monolingüe en esa lengua. Los colaboradores

utilizan principalmente el *ñható* para interactuar con sus padres; sin embargo, hay algunas diferencias relacionadas con el género: los bilingües hablan otomí con su madre y ambos idiomas con su padre. Este comportamiento está relacionado con los diferentes roles sociales que desempeñaban hombres y mujeres en la comunidad. Las mujeres realizaban la mayoría de sus actividades (trabajo doméstico, agrícola y artesanal) dentro de la comunidad y no necesitaban interactuar con los hispanohablantes. Por otro lado, los hombres eran los que solían trabajar fuera de la comunidad (comercio, albañilería) y debían interactuar con hispanohablantes.

Los bilingües otomí-español también usan el otomí con sus tíos, tías y suegros; sin embargo, el uso exclusivo del español aumenta. Este cambio está relacionado con al menos dos factores: el incremento de matrimonios exogámicos y el grado de intimidad o solidaridad con el interlocutor. Algunos bilingües informaron que sólo se comunican en español con sus suegros porque su pareja es de otra comunidad y no son hablantes de otomí. Otros colaboradores prefieren hablar en español porque se sienten incómodos hablando otomí con personas fuera de su círculo más cercano, aunque hablen esa lengua. Así, el tipo de relación y la cercanía juegan un papel determinante en la elección de una u otra lengua. A pesar de ciertas excepciones, el otomí es la lengua predominante entre las personas mayores.

El uso exclusivo del otomí disminuye con los interlocutores contemporáneos y aumenta sustancialmente la posibilidad de hablar ambas lenguas. Este comportamiento se relaciona con la estructura de la comunidad ya que, según los datos censales, el porcentaje de bilingües incrementó sustancialmente entre las décadas de 1970 y 1980. Los bilingües se comunican en español con sus hermanos y pareja debido a las actitudes lingüísticas y al aumento de los matrimonios exogámicos. Algunos colaboradores decidieron dejar de usar el otomí con estos interlocutores porque experimentaron episodios de discriminación. Otros se

casaron con personas ajenas a la comunidad que son monolingües en español, lo cual restringe las interacciones en otomí. Asimismo, el uso exclusivo del español entre compadres y amigos creció debido a que este tipo de relación no está restringido a los miembros de la comunidad, pues cada vez más personas de SAC y JV buscan establecer relaciones de amistad y compadrazgo fuera de su comunidad.

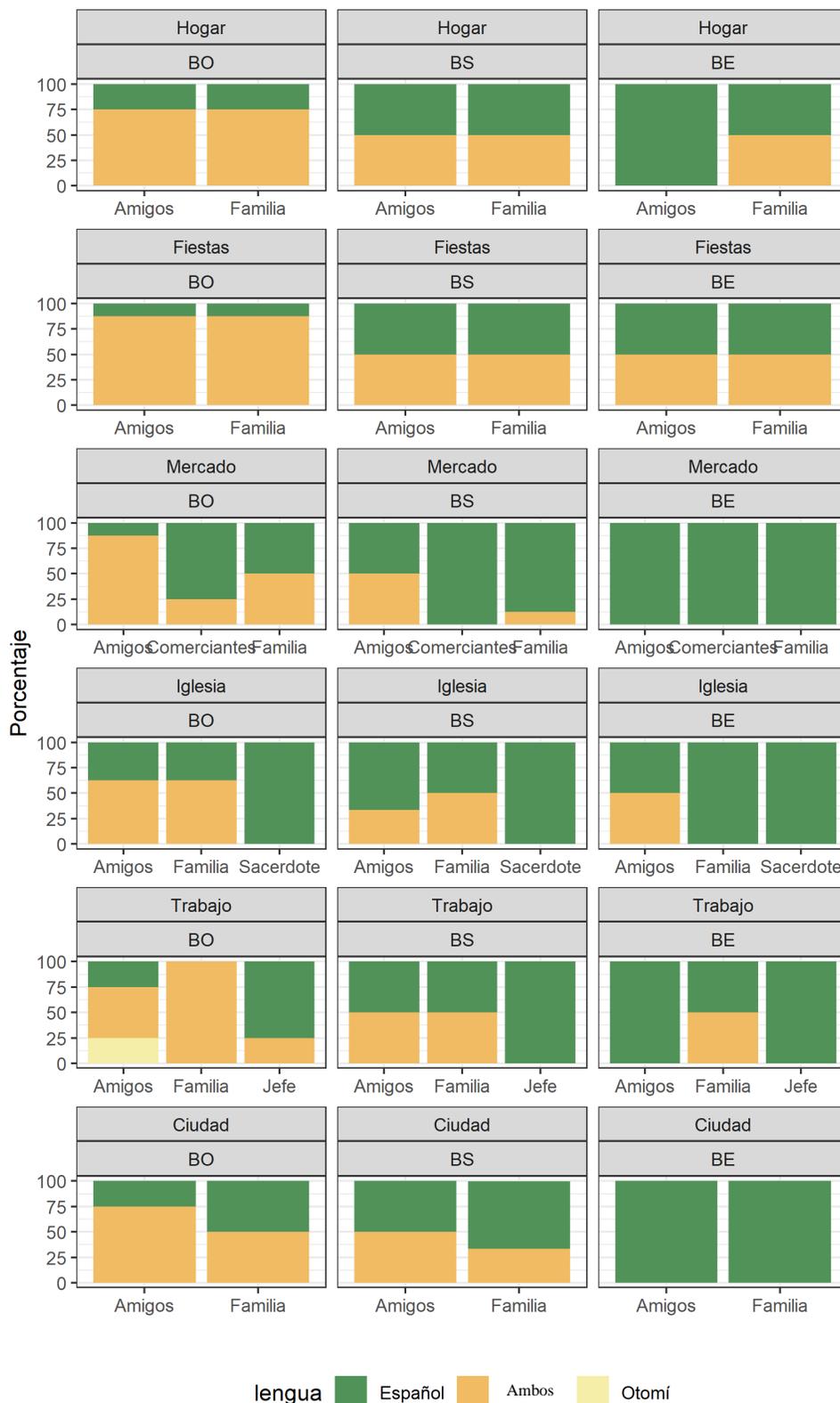
Finalmente, el uso del otomí se reduce con los interlocutores más jóvenes (hijos, yerno/nuera, nietos), y el español se convierte en la lengua dominante. Este hecho es congruente con los cambios en la composición lingüística de los habitantes de SAC y JV, ya que los otomíes interrumpieron la transmisión de su lengua alrededor de la década de 1980, y el porcentaje de hablantes monolingües de español aumentó. Algunos bilingües hablan con sus hijos y nietos en otomí; sin embargo, esto no implica que sean hablantes activos del otomí, sino que tienen un dominio pasivo de esa lengua. Esta diferenciación generacional en el uso de la lengua según la edad del hablante permite vislumbrar un fenómeno de desplazamiento de la lengua otomí en las dos localidades. No obstante, los datos también sugieren que este proceso va más adelantado en SAC, pues en JV todavía se documenta el uso de las dos lenguas con los más jóvenes.

En esta sección también describo el uso del otomí y el español en diferentes dominios sociales con varios interlocutores, a partir de la información recabada en la sección “Uso de las lenguas” del cuestionario sociolingüístico (véase Apéndice 2D). En la Figura 2.16 se muestran los resultados correspondientes a seis dominios sociales (hogar, fiestas, mercado, iglesia, trabajo y ciudad), en los tres grupos de bilingües (BO, BS, BE) de San Andrés Cuexcontitlán.

El primer dominio social es el hogar e incluye dos posibles interlocutores, familiares y amigos. En este espacio, la mayoría de los bilingües dominantes en otomí (BO) usan las

dos lenguas con sus familiares y amigos; mientras que dos de ellos sólo emplean el español. Los bilingües simétricos (BS) también alternan entre sus dos lenguas cuando interactúan con sus familiares y amigos en el hogar, pero el uso exclusivo del español incrementa hasta un 50%. Los bilingües con mayor dominio del español (BE) tienen un comportamiento similar al de los BS en la comunicación con sus familiares; sin embargo, con los amigos, sólo se comunican en español.

FIGURA 2.16 *Uso de las lenguas en diferentes dominios sociales. Grupos lingüísticos de SAC*



En el segundo dominio social son las fiestas del pueblo. En SAC se celebran dos fiestas cívicas importantes, el aniversario de la Independencia de México (15 y 16 de septiembre) y el de la Revolución mexicana (20 de noviembre). En estas fechas se hace honores al lábaro patrio y se realiza un desfile por las principales calles de la localidad donde participan los delegados municipales, los alumnos y los padres de familia de diferentes planteles educativos, también se organizan eventos culturales y contiendas deportivas. La comunidad también tiene dos fiestas religiosas importantes en honor al Santo Patrón Andrés. La primera acaece el 30 de noviembre. En la madrugada del 29 hay fuegos pirotécnicos y grupos musicales, y a la mañana siguiente se realiza una misa y otras actividades religiosas.

El segundo festejo es en la segunda semana de enero. Las actividades inician el primer día del año, cuando los grupos de bailarines (toritos) comienzan a recorrer las casas de los miembros de las cooperativas. Ahí se les ofrece a los danzantes y sus acompañantes diferentes platillos como tamales, taco de plaza, algún guisado, refresco y bebidas. El domingo previo a la fiesta, los fiscales y mayordomos llevan al Santo Patrón a la catedral de Toluca y el día viernes antes de la víspera los grupos de toritos y gran parte de la población van por San Andrés a la catedral y regresan bailando a la comunidad. El día sábado y domingo se queman los toritos en la noche mientras los grupos de bailarines danzan enfrente de la figura del Santo Patrón.

El dominio social “fiestas” contempla dos receptores potenciales: familiares y amigos. La mayoría de los BO utilizan las dos lenguas con familiares y amigos, y sólo uno de ellos utiliza exclusivamente el español. Los BS y los BE también alternan entre el otomí y el español cuando se comunican con familiares y amigos; sin embargo, el uso exclusivo del español en estos bilingües incrementa hasta 50%. Es importante señalar que se observa un mayor uso de la lengua otomí en las fiestas religiosas que en las cívicas.

Los lunes son día de plaza en San Andrés Cuexcontitlán, a ella acuden comerciantes y habitantes de la comunidad, así como mercaderes y personas de pueblos aledaños como San Cristóbal Huichichitlán, San Pablo Autopan y del municipio de Temoaya. En este dominio social se tomaron en cuenta tres posibles interlocutores: familiares, amigos y comerciantes. La mayoría de los BO hacen interactúan en otomí y español con sus amigos y familiares; sin embargo, con los comerciantes, predomina el uso exclusivo del español. Los BS siguen este mismo patrón, pero con una mayor dominancia del español (amigos = 50%, familiares = 87.5%, comerciantes = 100%). Para los BE, el mercado es un dominio de uso exclusivo del español, independientemente del tipo de interlocutor. El español es la lengua predominante en la interacción con los comerciantes, debido a que la mayoría de ellos no son de San Andrés Cuexcontitlán y no son hablantes de otomí. Por último, aunque el mercado es un dominio social que ha demostrado privilegiar el uso de las lenguas indígenas en otras comunidades (Flores Farfán, 1999), el español predomina en este espacio, incluso si el interlocutor tiene alguna relación familiar con el informante.

La iglesia es un ámbito de gran importancia para la comunidad de San Andrés Cuexcontitlán ya que la mayoría de la población es católica. Sin embargo, existen otros grupos religiosos, como los cristianos, los protestantes y los testigos de Jehová. Esta última congregación ha capacitado a personas en la lengua otomí para predicar a las personas que no tienen competencia lingüística en español. También ha habido algunos esfuerzos por incorporar la lengua otomí en actividades dentro de la religión católica. Por ejemplo, en el pueblo vecino de San Cristóbal Huichochitlán, la población ha procurado que las lecturas de la misa sean en español y otomí; sin embargo, el sermón del sacerdote y el resto de la ceremonia es en español.

En el dominio de la iglesia se contemplaron tres tipos de interlocutor: familiares, amigos y sacerdotes. La mayoría de los BO emplea el otomí y el español para comunicarse con sus familiares y amigos, aunque también hay dos bilingües que sólo utilizan español. Este comportamiento se vuelve la regla en la interacción con los sacerdotes. En los BS y los BE predomina el uso exclusivo del español con el sacerdote y los familiares, pero se mantiene un bajo porcentaje de uso bilingüe cuando los colaboradores interactúan con sus amistades. Algunos hablantes indicaron que sólo hablan otomí en voz baja en la iglesia para evitar ser escuchados por personas ajenas a su conversación. El hecho de que el español sea la lengua utilizada por el sacerdote no es sorprendente, ya que esta persona no es originaria de San Andrés y no es hablante de otomí. En resumen, en la iglesia, la lengua dominante es el español, pues en este recinto sólo se predica en esa lengua.

El dominio social correspondiente al trabajo toma en cuenta tres receptores: familiares, amigos y jefes. La mayoría de los BO usan el otomí y el español para comunicarse con sus familiares y amigos, pero algunos prefieren hablar otomí con sus amistades; por otro lado, el español es la lengua predominante en la interacción con sus jefes. La mitad de los BS alterna sus dos lenguas para comunicarse con sus familiares y amigos, y el resto sólo usa español. Esta lengua lidera los intercambios lingüísticos entre los BS y sus jefes. Por último, los BE usan exclusivamente el español para comunicarse con sus amigos y jefes, con excepción de un hablante que usa las dos lenguas con sus familiares.

Esta diferencia está relacionada con tres aspectos: el tipo de trabajo (agricultor, albañil, obrero, comerciante, profesional), el lugar donde trabajan (dentro o fuera de la comunidad) y las características de sus compañeros de trabajo (miembros o no de la comunidad, bilingües o monolingües). Por ejemplo, uno de los bilingües entrevistados trabaja como albañil y comparte este espacio con personas del mismo pueblo, lo que aumenta la

probabilidad de que sus compañeros de trabajo sean bilingües y utilicen las dos lenguas. En cambio, otros colaboradores que trabajan en el comercio han tenido la oportunidad de viajar a diferentes partes de la República Mexicana y convivir con personas ajenas a la comunidad, lo que justifica que el español sea el idioma preferido en la comunicación con sus compañeros de trabajo. Otro informante afirma que se comunica en ambos idiomas con sus diferentes jefes, dependiendo de si es bilingüe o monolingüe, hecho que está muy relacionado con el tipo de trabajo que realiza. Sin embargo, el español es el idioma dominante para dirigirse a sus superiores en el trabajo, ya que la mayoría de las personas se encuentran fuera de San Andrés Cuexcontitlán. En conclusión, aunque el tipo de red de trabajo determina el uso del otomí o del español, no es el único factor ya que existen otros aspectos esenciales como el carácter bilingüe o monolingüe de los miembros de la red y su tipo de relación.

Finalmente, el cuestionario incluye un escenario externo a San Andrés Cuexcontitlán, la ciudad de Toluca, que es el centro urbano más cercano al pueblo. Este espacio considera como interlocutores potenciales a los familiares y los amigos. La mayoría de los BO usan las dos lenguas con sus familiares y amigos, y sólo algunos emplean el español. Los BS también interactúan en español y *n̄ható* con los dos tipos de interlocutores, pero predomina el español. Este comportamiento se maximiza en el grupo de BE, que únicamente hablan español en la ciudad. En conclusión, se puede apreciar que el español ha ganado más fuerza en este ámbito. Este hecho puede estar relacionado con las actitudes de los hablantes hacia el otomí y los episodios de discriminación. Por ejemplo, una informante afirma que le da vergüenza hablar otomí en la ciudad.

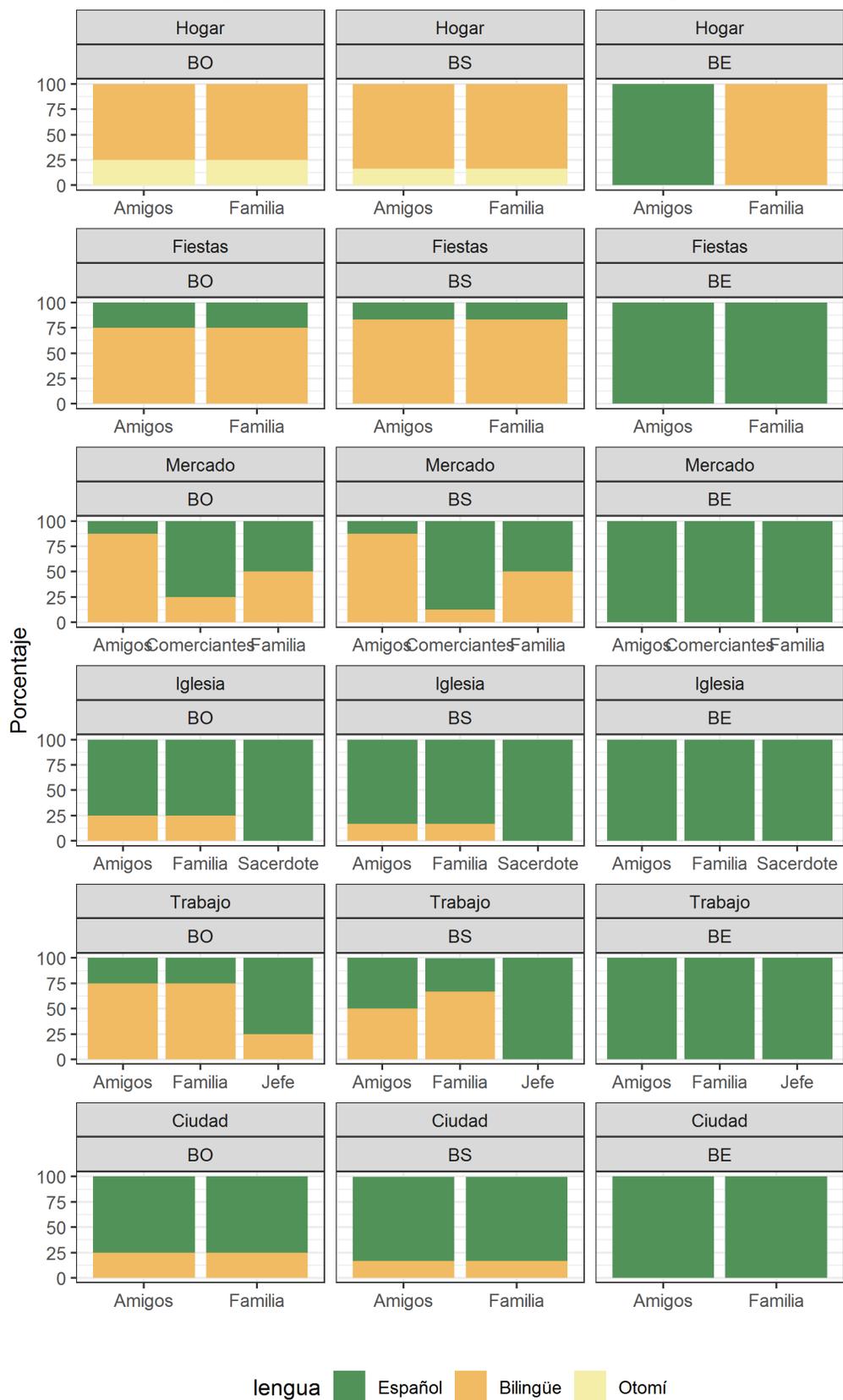
En síntesis, los BO hacen un mayor uso del otomí o alternan entre sus dos lenguas en todos los dominios sociales. Los BS emplean sus dos lenguas en la mayoría de los espacios, pero emplean más el español que los BO. Por último, los BE utilizan mayormente el español

en todos los dominios considerados. Este hecho es congruente con el grado de competencia lingüística de los colaboradores, que tienen un mayor dominio de dicha lengua.

Los datos aquí presentados permiten aproximarse a las funciones y usos de la lengua otomí y del español en seis dominios sociales de SAC. La utilización del otomí y el español dentro de la comunidad está en función de diversos factores, como las características etarias, lingüísticas y sociales de los interlocutores, así como de la situación y lugar en el que se encuentren, aspectos que, en determinado momento, se conjugan de distintas formas. El primer punto que se debe subrayar es que el otomí no es la lengua exclusiva de ninguno de los tres tipos de bilingües y en ninguno de los dominios sociales considerados, incluido el hogar, que se ha caracterizado como un espacio de resistencia en el desplazamiento de las lenguas minoritarias. El uso de las dos lenguas impera en dominios como el hogar, y las fiestas, mientras que el español es la lengua que los bilingües utilizan en el mercado, el trabajo, la iglesia y la ciudad. Asimismo, se observa una constante expansión del español en dominios que anteriormente eran exclusivos del otomí, como el hogar. Estos hechos hacen aún más patente la existencia de un proceso de desplazamiento lingüístico en San Andrés Cuexcontitlán.

En la Figura 2.17 se muestran los resultados correspondientes a seis dominios sociales (hogar, fiestas, mercado, iglesia, trabajo y ciudad), en los tres grupos de bilingües de Jiquipilco el Viejo. En el dominio social correspondiente al hogar, la mayoría de los bilingües dominantes en otomí (BO) y los bilingües simétricos (BS) usan el otomí y el español con sus familiares y amistades, y sólo un bajo porcentaje emplea exclusivamente el otomí. Por otro lado, la bilingüe con mayor dominio del español (BE) se comunica en las dos lenguas con sus familiares, mientras que con sus amigos sólo habla español.

FIGURA 2.17 *Uso de las lenguas en diferentes dominios sociales. Grupos lingüísticos de JV*



En Jiquipilco el Viejo también se celebran fiestas cívicas y religiosas. Al igual que en SAC, se festeja la Independencia de México (15 y 16 de septiembre) y la Revolución mexicana (20 de noviembre) con honores al lábaro patrio y un desfile por las principales calles de la localidad donde participan los delegados municipales, los alumnos y los padres de familia de los diferentes planteles educativos. Asimismo, tienen dos fiestas patronales en honor al señor Santiago. La fiesta principal se realiza el primer martes de agosto y en ella participan miembros de todas las localidades del municipio de Temoaya. Se hace una peregrinación a la iglesia de Jiquipilco el Viejo y ahí conviven y comparten distintos alimentos. Por la noche hay juegos mecánicos, eventos musicales y fuegos pirotécnicos. En las fiestas de la comunidad, la mayoría de los BO y los BS pueden alternar entre sus dos lenguas cuando interactúan con familiares y amigos, y sólo algunos afirman emplear únicamente el español. La BE, por su parte, sólo se comunica en español en este espacio, pues señala que se siente insegura, debido a que no habla *ñható* con mucha fluidez.

El tercer dominio es el mercado que se realiza los días martes en el centro de la localidad. La mayoría de los BO y los BS usan el otomí y el español con sus familiares y amigos, pero algunos de ellos sólo se comunican en español con esos interlocutores. El uso de dicha lengua se incrementa sustancialmente en la interacción con los comerciantes. Por otro lado, la BE hace uso exclusivo del español en el mercado con todos los tipos de interlocutores. En la iglesia, la mayoría de los BO y los BS se comunican en español con familiares, amigos y sacerdotes, y sólo algunos alternan entre sus dos lenguas con sus familiares y amigos. En contraste, la BE hace un uso exclusivo del español en este dominio social.

El cuarto dominio social fue el trabajo. Tanto los BO como los BS pueden alternar en otomí o español cuando se comunican con sus familiares y amigos; sin embargo, en la

interacción con sus jefes, predomina el uso exclusivo del español. En contraste, la BE únicamente utiliza español en este espacio social, debido a su labor como docente. Por último, podemos ver que el español es la lengua imperante en las interacciones lingüísticas que acaecen en la ciudad, salvo por algunos BO y BS, que afirman poder alternar en otomí y en español con sus familiares y amigos.

En resumen, los BO y los BS hacen un uso similar de sus dos lenguas en todos los dominios sociales y con todos los interlocutores considerados. Por otro lado, la BE utiliza la mayor parte del tiempo el español, salvo cuando se encuentra en el hogar y se comunica con sus familiares. En cuanto a los dominios sociales, vemos que el otomí aún se mantiene como lengua de uso exclusivo en el hogar, al menos para algunos hablantes; no obstante, en este espacio predomina la alternancia del otomí y el español. Este uso bilingüe predomina en los dominios sociales fiesta, mercado, iglesia y trabajo, con excepción de ciertos interlocutores externos a la comunidad, como son los comerciantes en el mercado, el sacerdote en la iglesia y el jefe en el trabajo. Por último, en el dominio social de la ciudad, el español es la lengua imperante en todos los grupos de bilingües (BO, BS, BE) y con todos los interlocutores.

Si se compara el uso que los diferentes grupos de bilingües (BO, BS, BE) hacen del otomí y el español en SAC y JV, se puede ver que el otomí ha perdido presencia en la mayoría de los dominios sociales, incluido el hogar; en contraste, el español se está convirtiendo en la lengua imperante en todos los dominios, lo cual sugiere la existencia de un proceso de desplazamiento del *n̄ható* en las dos comunidades, aunque este está más avanzado en SAC.

En la primera parte de este capítulo presenté los principios metodológicos que guían esta investigación, los cuales se basan en la sociolingüística variacionista (Labov, 1996, 2001) y en las propuestas de Poplack y Levey, (2010) y (Torres Cacoullous y Travis, 2018) para el análisis de los cambios inducidos por contacto. Posteriormente describí el tipo de

muestreo utilizado para conformar los grupos de colaboradoras de SAC y JV, así como las herramientas de elicitación de material lingüístico e información sociolingüística. Posteriormente, describí brevemente la historia de SAC y JV, los cambios sociales y económicos que han sufrido estas comunidades en las últimas décadas y cómo esto ha modificado la organización y las características sociales y lingüísticas de sus habitantes. Para ello, caractericé cinco grupos de hablantes (BO, BS, BE, PO, ME) y el uso que hacen del otomí y el español en distintos dominios sociales y con diferentes tipos de interlocutores. Al respecto, pudimos ver que en SAC y JV existe contínuo bilingües que va del monolingüismo en otomí al monolingüismo en español. Además, la descripción del uso de las lenguas nos permitió ver que el otomí está siendo desplazado y se está llevando a cabo un proceso de desplazamiento lingüístico en las dos comunidades. El siguiente capítulo se centrará en el análisis uno de los fenómenos variables asociados al número en el español de contacto, a saber, la concordancia plural.

CAPÍTULO 3. EL ESPAÑOL DE CONTACTO: EL CASO DE LA CONCORDANCIA PLURAL

3.1 INTRODUCCIÓN

Los sistemas de marcación de número de las lenguas del mundo pueden distinguir valores semánticos como el singular, el dual, el trial, el plural y el paucal (Corbett, 2000: 1). A diferencia de la lengua otomí que distingue entre singular, dual y plural (véase § 1.4), el español codifica los valores singular y plural en los sustantivos, los determinantes, los pronombres, los adjetivos y los verbos. El número singular no tiene una marca particular y el plural se expresa con el morfema *-(e)s*, pero también tiene un morfema cero (p. ej. *lunes*, *crisis*) (RAE y ASALE, 2010: 35). A partir de esta propiedad gramatical se establecen relaciones de concordancia entre el nombre y sus modificadores en la frase nominal, entre el sujeto y su predicado verbal o nominal, así como entre los clíticos (acusativo y dativo) con su correferente (Martínez, 1999: 2707).

Sin embargo, en situaciones en las que el español se encuentra en contacto con otras lenguas, la concordancia en ese tipo de relaciones gramaticales puede presentar una mayor variación (Escobar, 2002; Palacios Alcaine, 2011; Torres Sánchez, 2018) que la descrita para el español monolingüe y normativo (Soler Arechalde, 2012). En algunos trabajos sobre el contacto entre el otomí y el español se ha advertido un uso diferente del número gramatical en la relación de concordancia que se establece en la frase nominal (entre el nombre y sus modificadores), entre el verbo y el sujeto de una oración, entre el sujeto y su atributo en oraciones de predicado nominal, y entre los clíticos de acusativo o dativo y sus correferentes (Guerrero Galván & San Giacomo Trinidad, 2014: 1506-1508; Hekking, 2001: 142; Hekking & Bakker, 2007: 134; Lastra, 1987: 241). En el español en contacto con el otomí de San

Andrés Cuexcontitlán y de Jiquipilco el Viejo también se presentan estos fenómenos, como se muestra en algunos ejemplos de nuestro corpus.

- (1) a. **Los** trece **año**
- b. Pues **todos** ya **viene tomados**
- c. O sea/ **las maestra** era muy **malos** entonces
- d. Creyó que echó [**los frijoles**]₁ a la olla/ y el otro creyó pues también **lo**₁ echó
- e. onque sea la secundaria le di nada más mis hijos/ y ahí están/ y ora [**ellos**]₁ **le**₁ digo yo pus que le echen gana

En (1a) la marca de plural sólo aparece gramaticalmente en el artículo definido, el cuantificador expresa pluralidad en el ámbito semántico, pero el núcleo nominal se presenta en su forma singular. En el siguiente ejemplo, el sujeto no concuerda en número con el verbo y el predicado secundario. En la oración de (1c), el núcleo de la frase verbal no concuerda en número con la frase nominal sujeto, *las maestra*, ni con la frase nominal que funciona como atributo, *malos*. En el fragmento de (4d), el clítico de tercera persona con función de objeto directo, *lo*, no concuerda en número con la frase nominal a la que correfiere, *los frijoles*. Por último, en (1e), el clítico de objeto indirecto, *le*, no concuerda en número con el pronombre de 3ª persona, *ellos*.

La PREGUNTA GENERAL de este capítulo es si el contacto con el otomí ha incidido en la concordancia plural del español. De manera particular, se pretende responder las siguientes preguntas:

- 3.1 ¿La variación en la concordancia plural del español tiene su origen o está asociada en algún grado al contacto con la lengua otomí? (PREGUNTA ESPECÍFICA 3.1).
- 3.2 ¿La concordancia plural en el español de contacto tiene el mismo comportamiento lingüístico en el nivel frástico y oracional? (PREGUNTA ESPECÍFICA 3.2).

3.3 ¿El funcionamiento lingüístico y la distribución social de la concordancia de número en el español de contacto, en el nivel frástico y oracional, se comporta de manera similar en las dos comunidades? (PREGUNTA ESPECÍFICA 3.3).

3.4 ¿El funcionamiento lingüístico y la distribución social de la concordancia de número en el español de contacto, en el nivel frástico y oracional, se comporta de la misma manera entre los miembros de las comunidades? (PREGUNTA ESPECÍFICO 3.4).

El OBJETIVO GENERAL de este capítulo es determinar si el contacto con el otomí ha incidido en la concordancia plural del español, mediante el análisis de la variación en la concordancia plural que se establece al interior de la frase nominal, entre el núcleo y sus modificadores (1a), y en el nivel oracional, entre el sujeto y su predicado verbal (1b) o nominal (1c). Los datos provienen de entrevistas en español a hablantes monolingües y bilingües otomí-español de San Andrés Cuexcontitlán y de Jiquipilco el Viejo.

De manera específica, busco determinar si la alternancia en la concordancia de número tiene su origen o está asociada en algún grado con el contacto que el español mantiene con la lengua otomí (OBJETIVO ESPECÍFICO 3.1). En relación con este punto, y dado que la concordancia de número en otomí se organiza de manera diferente en el nivel frástico y oracional (véase §1.4.2 y § 1.4.3), también pretendo discernir si la concordancia plural en el español de contacto tiene el mismo comportamiento lingüístico en el nivel frástico y oracional o si, por el contrario, cada dominio posee características específicas que ameritan un análisis bajo sus propios términos (OBJETIVO ESPECÍFICO 3.2). Por otro lado, pretendo saber si el funcionamiento lingüístico y la distribución social de la concordancia de número en el español de contacto, en el nivel frástico y oracional, se comporta de manera similar en

las dos comunidades (OBJETIVO ESPECÍFICO 3.3) y entre los miembros de estas (OBJETIVO ESPECÍFICO 3.4).

Con respecto al primer objetivo, sobre el origen de la alternancia en la concordancia de número, evaluó la hipótesis de que esta variación es resultado del contacto que el español mantiene con el otomí, de modo que este fenómeno responde a estrategias gramaticales relacionadas con los procesos de la lengua otomí. Por ello, aunque la concordancia plural se expresa con los recursos morfosintácticos del español normativo, la aparición o no de estas marcas estará determinada por alguno(s) de los condicionantes lingüísticos que rigen la expresión gramatical de número en otomí (HIPÓTESIS 3.1). Una variante débil de esta hipótesis es que los recursos del otomí condicionan la forma en que se aprenden o encuentran soluciones al hablar español.

En relación con el segundo objetivo, considero necesario evaluar la idea de que la variación en la concordancia de número en el español de contacto se comporta de manera diferente en el nivel frástico y oracional (HIPÓTESIS 3.2), debido a que en otomí se organiza de forma distinta en los dos niveles (véase § 1.4.2 y § 1.4.3). Por otro lado, con respecto a la distribución de la variación en la concordancia de número en el español de contacto, hipotetizo que tiene el mismo comportamiento lingüístico en San Andrés Cuexcontitlán y Jiquipilco el Viejo, pero difiere en su distribución social en cada comunidad (HIPÓTESIS 3.3).

Aunado a esto, como indiqué en la sección § 2.2.4, en San Andrés Cuexcontitlán y, en menor medida en Jiquipilco el Viejo, se está llevando a cabo un proceso de desplazamiento de la lengua otomí, debido a una ruptura en la transmisión del *n̄ható*, que ha generado una importante diferenciación en el dominio de las lenguas: las personas mayores son las que tienen algún grado de bilingüismo, mientras que los más jóvenes tienden a ser monolingües en español. Con base en estas observaciones y en vista de los objetivos específicos tres y

cuatro, que son determinar si la alternancia en la concordancia de número funciona y se distribuye de la misma manera en las dos comunidades y entre sus miembros, parto de la hipótesis de que en estas comunidades existen diferentes sistemas de concordancia de número en español, relacionados con el tipo de bilingüismo de los hablantes. Estos sistemas no sólo se diferencian en el porcentaje de casos de discordancia, sino en los factores lingüísticos que los organizan y determinan, es decir, que responden a gramáticas diferentes (HIPÓTESIS 3.4).

En la siguiente sección (§3.2) reviso algunas investigaciones sobre la alternancia en la concordancia de número en situaciones monolingües y de contacto con otras lenguas, lo cual nos permitirá contrastar los resultados del otomí con los de otras situaciones de contacto y nos ayudará a identificar algunos factores lingüísticos que podrían ser relevantes para explicar la variación en la concordancia plural del español otomí. En la sección § 3.3 presento algunas precisiones metodológicas en torno al análisis de la concordancia de número y en §3.4 propongo un primer análisis cuantitativo sobre la concordancia de número en el español en contacto con otomí. Con base en los resultados, planteo la cuestión de si la variación en la concordancia de número en el nivel frástico y oracional es un solo fenómeno variable o si, por el contrario, se trata de dos fenómenos que deben ser analizados de manera individual y bajo términos diferentes. Trataré de contestar esta cuestión a lo largo de los apartados § 3.5 y § 3.6, donde propondré un análisis de la concordancia de número en cada uno de los niveles, con el objetivo de entender de forma más precisa su funcionamiento.

Posteriormente, en la sección § 3.7, analizo la cuestión de si la variación en la concordancia de número tiene el mismo funcionamiento lingüístico y la misma distribución en las dos comunidades, San Andrés Cuexcontitlán y Jiquipilco el Viejo. En la sección §3.8, evalúo la HIPÓTESIS 3.4 de que en estas comunidades existen diferentes sistemas de concordancia de número en español y de que estos sistemas no sólo se diferencien en el

porcentaje de casos de discordancia, sino en los factores lingüísticos que los organizan y determinan (es decir, responden a gramáticas diferentes). Por último, en la sección §3.9, presento las conclusiones de este capítulo.

3.2 ANTECEDENTES

En este apartado revisaré algunos trabajos sobre la variación en la concordancia de número en español. En la primera sección muestro un estudio sobre la variación en la concordancia de número en el español de variedades sin contacto (Soler Arechalde, 2012) y, posteriormente, expongo los resultados de algunas investigaciones sobre el español en contacto con sikuni, wayuu y tikuna (Ramírez Cruz, 2009), con quechua en Perú (Escobar, 2000), Argentina (Martínez, 2001, 2012), con mapuzugun (Olate Vinet et al., 2019), con tepehuano del sur (Torres Sánchez, 2018) y con el otomí de Santiago Mexquititlán (Guerrero, 2006).

3.2.1 *La concordancia de número en variedades de español monolingüe*

La investigación de Soler Arechalde (2012: 14) sobre la alternancia en la concordancia de número en variedades de español monolingüe,⁵⁴ analiza cuatro tipos de construcciones en las que la concordancia singular puede alternar con la concordancia plural, debido a que el controlador⁵⁵ de la concordancia es ambiguo con respecto al número, a causa de su semántica

⁵⁴ El corpus de esta investigación está conformado por muestras de español hablado culto de las ciudades de Bogotá, Buenos Aires, Caracas, Madrid, México y Santiago de Chile, que pertenecen al “Proyecto de estudio coordinado de la norma lingüística culta de las principales ciudades de Iberoamérica y de la Península Ibérica” (Soler Arechalde, 2012: 30-31).

⁵⁵ La investigadora concibe el término *concordancia* como una relación asimétrica entre dos elementos. Uno de ellos, el controlador, es el elemento que impone ciertas marcas morfológicas de género, número y persona a otro elemento, el controlado. Asimismo, señala que el controlador es siempre de naturaleza nominal, por lo que puede fungir como núcleo de la frase nominal, del sujeto en la relación sujeto-predicado y del antecedente en la asociación antecedente-pronombre (Soler Arechalde, 2012: 16, 20).

o de su configuración sintáctica. Estas construcciones son: los colectivos (2a), las frases nominales compuestas (2b), las frases con un colectivo más un complemento adnominal plural (2c), y las frases nominales de diferente número en una oración atributiva (2d).

- (2) a. **la familia** de mi padre también **salieron** / [**salió**] del gueto de Varsovia (CA37, 645)
- b. **la práctica y el estudio real viene** / [**vienen**] cuando uno sale de la Universidad (CA3, 45)
- c. **ese tipo de niños necesitan** / [**necesita**] para adaptarse al mundo (ME1, 15)
- d. **la base fundamental de la educación son** / [**es**] los maestros (BO9, 123)

Fuente: adaptado de Soler Arechalde (2012: 14-15)

La autora propone una jerarquía de factores que inciden en la alternancia de la concordancia de número en esos contextos. En primer lugar, las características semánticas de los controladores y otros elementos contextuales que dan lugar a la individuación (p. ej. la delimitación, la homogeneidad o heterogeneidad, la animacidad, la especificidad o generalidad, la concreción y la definición). Acerca de esto, la autora indica que un nombre individuado promueve la concordancia de número, mientras que un nombre no individuado favorece la discordancia (Soler Arechalde, 2012: 230).

El segundo factor es el dominio en el que se establece la relación de concordancia: frástico, oracional y extraoracional. En el nivel frástico, donde se presenta una relación de concordancia entre el núcleo nominal y sus modificadores, existe una mayor cohesión gramatical y, en consecuencia, un menor número de casos de discordancia. En el dominio oracional –en el que se presentan relaciones de concordancia entre el núcleo nominal del sujeto y el núcleo del predicado verbal o nominal– hay una menor cohesión gramatical y surgen más casos de alternancia. Finalmente, en el nivel extraoracional, donde el controlador

y el elemento controlado están en distintas oraciones, tiende a haber más casos de alternancia en la concordancia (Soler Arechalde, 2012: 256).

El tercer factor son las características sintácticas del controlador y el elemento concordante. La autora afirma que cuando el controlador funciona como sujeto tiene un mayor control de la concordancia que cuando funciona como objeto, atributo o adjunto. Con respecto a la función sintáctica del elemento concordante, advierte que este factor está subordinado a la individuación y el dominio. Por ello, propone la siguiente jerarquía: modificador > oración relativa simple > predicado nominal > predicado verbal > pronombre objeto > oración relativa compleja > pronombre oblicuo > pronombre sujeto; y señala que los modificadores son los elementos concordantes que tienen mayor cohesión sintáctica con el nombre y el pronombre sujeto el de menor cohesión (Soler Arechalde, 2012: 257).

Por último, incluye la posición del controlador y del elemento concordante en el contexto (Soler Arechalde, 2012: 15). Al respecto, comenta que en otras investigaciones se ha señalado que el orden controlador + elemento concordante tiende a favorecer la concordancia sintáctica, mientras que el orden elemento concordante + controlador suele desfavorecerla; sin embargo, en su corpus este comportamiento no se sostiene completamente (Soler Arechalde, 2012: 258). Los hallazgos de esta investigación serán un punto de comparación importante en la discusión de nuestros resultados, pues corresponden a variedades de español sin contacto.

La discordancia de número también es un fenómeno recurrente en variedades de español en contacto con otras lenguas, como el náhuatl, el maya, el purépecha o el quechua, por mencionar algunas (Escobar, 2002: 254). La presencia constante de este fenómeno en diferentes situaciones de contacto ha generado controversia en torno a su origen. Al respecto, algunos argumentan que es un fenómeno propio de la adquisición incompleta de una segunda

lengua (Ramírez Cruz, 2009). Y otros sostienen que se trata de una interferencia lingüística que, a pesar de ser común a diferentes situaciones de contacto entre el español y otras lenguas, presenta matices según la lengua en contacto y la situación sociolingüística (Guerrero, 2006; Olate Vinet et al., 2019; Torres Sánchez, 2018).

A continuación, revisaré una investigación, anclada en el primer enfoque, sobre la concordancia de número en el español en contacto con sikuani, wayuu y tikuna. Posteriormente, reseño los principales hallazgos de trabajos, desarrollados desde la segunda perspectiva, en los que se analiza la concordancia de número en el español en contacto con quechua, guaraní, mapuzugun, otomí y tepehuano del sur (Escobar, 2000; Guerrero, 2006; A. Martínez, 2001, 2012, 2017; Olate Vinet et al., 2019; Torres Sánchez, 2018).

3.2.2 La concordancia de número en el español en contacto con sikuani, wayuu y tikuna

Ramírez Cruz (2009: 171) estudia la discordancia de género y número en el español en contacto con sikuani, wayuu y tikuna en la frontera amazónica de Colombia, con el objetivo de determinar si estos fenómenos son un tipo de interferencia, entendida como un fenómeno de variación que surge en el habla de un bilingüe mientras usa la lengua B, como resultado de su conocimiento en la lengua A. El autor señala que para poder catalogar los casos de discordancia como una interferencia es necesario que estos fenómenos tengan estructuras equivalentes en las lenguas sikuani, wayuu y tikuna, las cuales tienen que ir más allá de la ausencia de estas categorías o de un funcionamiento distinto; aunado a esto, no deben existir fenómenos análogos en otras situaciones de contacto (Ramírez Cruz, 2009: 170). Para ello, recopiló un corpus de 60 horas, provenientes de historias de vida de 243 informantes, y localizó los casos de discordancia de número en el dominio nominal y verbal, tal como se ejemplifica en (3).

Discordancia de número en la FN

- (3)
- a. Cuando **vinieron** este, **los batallón rojo** (sikuani)
 - b. Mi mamá tuvo **nueves hijos** (wayuus)
 - c. **Un pepitas**, eso lo pongo un palo, ahí se la, se suena (tikuna)

Discordancia de número entre el verbo y su sujeto

- d. **Su tío, el son guerrilla?** (sikuani)
- e. antes no **conoce** el ron, ni **conoce** manteca antes, **nosotros no conoce** nada (wayuus)
- f. nosotros veníamos trabajando y con mi papá y todos **los abuelos que venía de allá**, de orilla del río hacia acá (tikuna)

Fuente: adaptado de Ramírez Cruz (2009: 173-174)

El segundo paso de la investigación fue revisar el funcionamiento de la concordancia de número en las lenguas sikuani, wayuu y tikuna; y descartó la existencia de correspondencias entre estas lenguas y los fenómenos de discordancia en su español. Posteriormente, el autor inspeccionó la existencia de casos semejantes en una situación de contacto entre el español y una lengua no amerindia, el japonés, y constató la presencia de discordancias de género y número en el español de estos bilingües. Con base en estas observaciones, el autor concluyó que las discordancias de número y género de los bilingües sikuani, wayuu y tikuna no son interferencias, sino procesos de adquisición incompleta de una segunda lengua.

3.2.3 La concordancia de número en el español en contacto con quechua en Perú

En el español en contacto con quechua (Perú) se han documentado casos de discordancia de número en el nivel frástico y oracional (Escobar, 2000). En la frase nominal se presentan

discordancias entre el nombre y sus modificadores, determinantes (4a) y adjetivos (4b). En este dominio también ocurren casos de discordancia entre el nombre y los cuantificadores indefinidos (4c) y numerales (4d). Sobre estos casos, la autora ha señalado que las expresiones cuantificadoras ya tienen en su significado el concepto de pluralidad, razón por la que podrían resultar redundantes las marcas morfológicas de plural en el nombre o en el cuantificador, como en el caso de *bastante* (4c). Además, ha encontrado casos de discordancia entre el verbo y su sujeto, como se aprecia en (4e), y en los adjetivos de predicados atributivos (4f).

- (4) a. le dejan **su hijitas** ahí
b. **niños sucio**
c. se juntan **bastante hombre**
d. **dos hijo** me dejado
e. es necesario que **ellos habla** su quechua
f. pues **los campesinos** estaban **escondido** en el cerro

Fuente: adaptado de Escobar (2000: 30)

La discordancia de número en el español en contacto con quechua ha sido catalogada como una interferencia gramatical o formal –entendida como un elemento lingüístico que se encuentra en la lengua B (receptora), que no corresponde a una característica de la lengua A (donante) ni a un rasgo de la variante monolingüe de la lengua B, pero que está en el habla de los bilingües (Escobar, 2000: 9). Este tipo de interferencia contrasta con aquellas que involucran características funcionales o semánticas del sistema lingüístico y que conllevan un proceso de reorganización o reinterpretación de los contrastes semánticos del sistema

lingüístico del español; por ejemplo, la reduplicación (posesivos), los valores de tiempo (futuro) y el modo (evidencialidad) (Escobar, 2000: 153).

3.2.4 La concordancia de número en el español en contacto con quechua en Argentina

También se han llevado a cabo estudios sobre la concordancia de número en variedades de español en contacto con quechua en Argentina. Uno de ellos fue realizado con la colaboración de hablantes de quechua y de guaraní que cursaban el cuarto año en el Liceo 10 de Buenos Aires (Martínez, 2001). En un corpus de lengua escrita se documentó casos de discordancia de número entre el verbo y los sujetos de tercera persona, que siguen dos patrones: el primero, tiene un sujeto singular y un verbo en plural, como el del ejemplo (5a); y el segundo presenta un sujeto plural y un verbo singular (5b), siendo el primero el más recurrente, con un 68% de la totalidad de casos.

(5) a. ¿Pero **existen una solución** a nuestros problemas?

b. **Y empezó a empañarse los vidrios**

Fuente: adaptado de Martínez (2001)

Martínez (2001) analiza los casos que siguen el primer patrón, como el del ejemplo (5a), con el objetivo de mostrar que las discordancias de número entre el verbo y los sujetos de tercera persona no son un error sino una respuesta a las necesidades comunicativas del hablante, que se basa en principios cognitivos asociados con la conceptualización del número. Con base en la Teoría de prototipos, establece una escala de prototipicidad de sujetos singulares, en la que los elementos nucleares son los pronombres personales y los nombres individuales, y los periféricos son los nombres colectivos y las entidades mixtas (núcleo en singular y complemento plural). Los resultados de esta investigación muestran que la mayoría

de los casos de discordancia de número se presenta en los elementos periféricos y decrecen conforme el sujeto se acerca al prototipo de individualidad. Con base en esto, la autora confirma su hipótesis y muestra que las discordancias de número entre un verbo plural y los sujetos de tercera persona singular tienen su origen en principios cognitivos asociados con la conceptualización del número y no a un error.

En un estudio posterior, Martínez (2012) analiza la concordancia de número en un corpus de textos periodísticos de la comunidad peruana y boliviana en Argentina. En él, documenta casos de discordancia en la frase nominal, entre el nombre y sus modificadores, y en el nivel oracional entre el sujeto y su predicado nominal o verbal, tal como se aprecia en los ejemplos de (6).

- (6) a. [...] el grupo de danzas Amerindia, con todos sus integrantes vestidos con trajes típicos, representó a **los nueve departamento** de Bolivia y al Litoral perdido [...]
- b. [...] el sábado tres de agosto en la localidad de Escobar, como cada año, se realizó el acto cívico con **las autoridades política**, representación diplomática de Bolivia en la Argentina y **las fuerzas vivas** de la zona [...]
- c. Sucede que la **Diablada** no **tenía** número de orden de circulación por ser del interior **tenían** una posición libre ya que además **debían** esperar que la **banda** Sebastián Pagador se desocupe [...]
- d. A las cero horas se abrieron **las puertas del templo** y los fieles comenzaron a entrar para desfilar por unos minutos frente a la imagen del Santo, permaneciendo **abierta** hasta que pase el último peregrino [...]

Fuente: adaptado de Martínez (2012: 143-144)

La autora analiza estos casos de discordancia desde el enfoque funcionalista y plantea la hipótesis de que la alternancia en la concordancia de número surge en contextos portadores de ambigüedad, de la que se valen los hablantes bilingües para desarrollar dos funciones: evaluar entidades respecto al número y valorar el grado de responsabilidad del agente

(Martínez, 2012: 157-158) . Con respecto a la concordancia de número entre verbos activos y su sujeto (6c), la autora encuentra que las discordancias surgen principalmente en oraciones con sujetos individuales no prototípicos (colectivos o frases nominales compuestas), tal como sucedía en los datos de la investigación antes descrita (Martínez, 2001). Por otro lado, en las oraciones pasivas con *se*, advierte que la alternancia en la concordancia de número funge como estrategia de focalización del grado de responsabilidad del agente. Con base en sus resultados, la investigadora señala que el contacto de lenguas permite la aparición de soluciones creativas que parten de las posibilidades que brinda el sistema lingüístico del español (Martínez, 2012: 161).

3.2.5 *La concordancia de número en el español en contacto con mapuzugun*

En una investigación reciente, Olate Vinet et al. (2019) estudian la variación en la concordancia de número en el español de bilingües *mapuzugun*-español, a partir del concepto *copia de código* (véase §1.1.3), definido como un proceso creativo mediante el que los hablantes insertan la copia⁵⁶ de una de sus lenguas (código modelo) a la otra (código básico) con el objetivo de reducir la distancia entre estas y, así, enriquecer y mejorar su funcionalidad (Johanson, 2002 apud Olate Vinet et al., 2019: 218). Con base en este concepto y en la distancia tipológica que existe entre el español y el *mapuzugun* en el dominio de la concordancia de número, los autores proponen la hipótesis de que los hablantes de *mapuzugun* insertan copias selectivas en el español (CB) desde el dominio de número de su lengua materna (CM), tanto en el nivel frástico como oracional.

⁵⁶ Este proceso no es una réplica exacta de su modelo, pues puede diferenciarse en el contenido semántico, en los contextos de aplicación, frecuencia, etcétera (Olate et al. 2019: 224).

Los autores analizaron 3000 oraciones provenientes de 30 entrevistas hechas a bilingües *mapuzugun*-español de la Región de la Araucanía, Chile, y contabilizaron los casos de discordancia de número, en la frase nominal y en la oración, que respondían al esquema proyectado del *mapuzugun* al español,⁵⁷ como se muestra en los ejemplos de (7), donde los hablantes insertan copias selectivas desde la dimensión de número del *mapuzugun* (CM) al español de contacto (CB).

- (7) a. pu papay
 PL abuelita
- Las abuelitas (español estándar)
 Las abuelita (español de contacto)
- b. iladkü - le - n wiño - y pu (+engün)
 enojarse - ESTAT - NR volver - 3agente che (+3.PL)
- Volvió la gente enojada (español estándar)
 Volvieron la gente enojado (español de contacto)

Fuente: adaptado de Olate *et al.* (2019: 219-220).

Los resultados generales mostraron un bajo porcentaje de alternancias en la concordancia de número en el nivel frástico (11%) y en el oracional (7%). En el análisis también consideraron la edad, el bilingüismo y el nivel de instrucción de los colaboradores; al respecto, encontraron un mayor porcentaje de discordancias en los hablantes de entre 45 y 59 años y con un nivel de instrucción básico.

Los investigadores concluyeron que la concordancia de número del español se acomoda al esquema del *mapuzugun*, al menos en la frase nominal, pues los casos de discordancia surgen de la analogía que los bilingües encuentran entre el artículo del español

⁵⁷ En la frase nominal del *mapuzugun* el número se marca con el morfema de plural *pu*, seguido de un nombre animado. Por otro lado, en la oración la concordancia de número singular, dual y plural es opcional en la tercera persona y puede ser influida por el tipo de sustantivo (Olate Vinet et al., 2019: 219-220).

(CB) y la palabra plural del *mapuzugun* (CM), *pu*, que sólo otorga su sentido de número a los determinantes y no al núcleo de la frase nominal, dando como resultado en el español frases nominales con determinantes en plural y núcleos nominales en singular, como en los ejemplos de (7). Sin embargo, los autores señalaron que en el caso de la concordancia de número entre el verbo y los sujetos de tercera persona no tienen evidencia suficiente para afirmar lo mismo.

3.2.6 *La concordancia de número en el español en contacto con tepehuano del sur*

En el español en contacto con el tepehuano del sur, Torres Sánchez (2018) también documenta variación en la concordancia de número en la frase nominal (8a), en la oración, entre el sujeto y su predicado verbal (8b) o nominal (8c), y en los clíticos de acusativo, como se muestra en (8d). Al respecto, la autora subraya la pertinencia de estudiar la alternancia en la concordancia de número en el nivel frástico y oracional (8a)-(8c) como un caso de interferencia –entendido como un elemento foráneo que es pasajero, inestable y ocasional en la lengua receptora (Silva-Corvalán & Enrique-Arias, 2018: 269)– y los fenómenos de número asociados a los clíticos de acusativo como una transferencia, definida como un elemento foráneo que tiene una forma más o menos estable en la lengua que lo recibe (Silva-Corvalán & Enrique-Arias, 2018: 269).

- (8) a. bien **contento-∅ lo-s muchacho-s** (Fabiola, DGO, bilingüe tipo B)
- b. [...] y comen los, **lo-s hongo-s grise-s/** son muy buenos/ pero// otros dicen que no/ pero mi abuelita si ∅ coma (<come)/ cuando crece-∅ como estos meses (Marta, SMO, bilingüe tipo C)
- c. **mi-s padre-s** son tepehuan-∅ (Ofelia, SMO, bilingüe tipo A)
- d. nuestro-∅ padre-s hablaban (Alonso, SMO, bilingüe tipo A)

- e. Ahí andaba cuidando **los duraznos**; pa (<para) que no **lo**; ganen los/ el pájaro (Alondra, SMO, bilingüe tipo B)

Fuente: adaptado de Torres Sánchez (2018: 181, 211)

La autora realizó un análisis cuantitativo de la concordancia de número en el nivel frástico y oracional, en el que consideró tanto factores lingüísticos como extralingüísticos. Dentro de las variables lingüísticas incluyó el tipo de relación de concordancia (frase nominal, sujeto, adjetiva y predicativa) y el tipo de marcación (presencia + presencia, presencia + ausencia y ausencia + presencia). Y como variables extralingüísticas el tipo de bilingüismo,⁵⁸ la edad de aprendizaje del español, el género, la edad, el nivel de instrucción y el lugar de residencia de los hablantes. Los factores que resultaron determinantes en la organización de la concordancia de número fueron el tipo de bilingüismo, la edad del colaborador y el tipo de construcción, siendo las construcciones adjetivas, los hablantes del tipo C, con un conocimiento bajo del español, y con más de 50 años los que promueven la discordancia.

3.2.7 La concordancia de número en el español en contacto con el otomí de Santiago Mexquititlán

En el español en contacto con otomí se han documentado casos de discordancia de número (Guerrero Galván & San Giacomo Trinidad, 2014; Lastra, 1989). Guerrero (2006) realizó un estudio sociolingüístico sobre la variación en la concordancia de número en el español en contacto con el otomí de Santiago Mexquititlán, a partir de la noción de *interferencia*, entendida como un elemento foráneo que es pasajero, inestable y ocasional en la lengua

⁵⁸ A partir de una prueba de competencia en español Torres Sánchez (2018) encontró tres grupos bilingües en su grupo de colaboradores: el tipo A con un conocimiento alto del español; el tipo B, con un conocimiento medio del español; y los del tipo C, con un conocimiento bajo del español. Para mayores detalles véase (Torres Sánchez, 2018, 2019).

receptora (Silva-Corvalán & Enrique-Arias, 2018: 269). En su análisis consideró la concordancia de número entre el nombre y sus modificadores (9a)-(9b), en las relaciones de sujeto (9c), de acusativo (9d), de dativo (9e) y en construcciones predicativas (9f).

- (9) a. debajo de **la camas** están los zapatos
- b. **todo(s) los barrios**
- c. **los personas indígenas** lo **habla(n)** normal
- d. es el único **los** que van a [ha]cer
- e. yo **le(s)** enseñé **a mis hijo(s)**
- f. descués que **estaba(n) un(os) sapos** ahí

Fuente: adaptado de Guerrero (2006: 94-95)

En su análisis cuantitativo consideró como variables explicativas: i) la clase a la que pertenece la concordancia (frástica, sujeto, predicativa, acusativo, dativo, adjetivo), ii) el tipo de marcación u orden (presencia + presencia, presencia + ausencia y ausencia + presencia), iii) la presencia de una palabra o FN que introduzca un referente, iv) la concordancia pragmática asociada al mantenimiento del tópico, v) el tipo de texto (historia de vida, narración, encuesta), vi) el bilingüismo (activo o pasivo), vii) la edad (jóvenes 8-30 años, mayores 31-63 años), viii) la ocupación (estudiante, empleado local, comerciante, empleado migrante, campesino, ama de casa) y ix) la localidad de los colaboradores. De estas, sólo resultaron significativas el tipo de marcación y el tipo de texto, así como la edad y la ocupación de los hablantes. La conclusión general de esta investigación fue que la falta de marcación de número corresponde a una interferencia estructural, cuya aparición es más frecuente en los hablantes mayores y que no era sistemática en una categoría o construcción sintáctica determinada (Guerrero, 2006: 103).

Torres Sánchez (2018) argumenta que la variación en la concordancia de número en el español en contacto con otomí (Guerrero, 2006) y con tepehuano del sur son interferencias estructurales, en las que los hablantes emplean mecanismos que siguen patrones de la gramática de su lengua materna, sin que esto conlleve una copia o exportación exacta de una estructura de la lengua indígena al español (197-199).

Para sustentar esa afirmación explica que los patrones de concordancia de número en estas variedades de español en contacto guardan ciertas correspondencias con las lenguas originarias. En el español tepehuano, la marca de número suele mantenerse en el núcleo y omitirse en el dependiente, de manera similar a lo que sucede en el *o'dam*, donde el objeto de la duplicación es el núcleo nominal, tal como se muestra en (10). En contraste, la marca de número en la frase nominal del español otomí se mantiene en el dependiente y no en el núcleo, tal como sucede en la lengua indígena (11) (Torres Sánchez, 2018: 197-199).

- (10) a. Gu **ma** - mai
 ART **PL** - maguey
 'los magueyes'
- b. el cochinito - **s**
 ART.DEF.SG cochinito - **PL**
 'los cochinitos'

Fuente: adaptado de Torres Sánchez (2018: 183, 198).

- (11) a. **Yá** bätši
3POS.PL hijo
 'sus hijos'
- b. **su** - **s** hijo
3.POS - **PL** hijo
 'sus hijo'

Fuente: adaptado de Guerrero Galván (2006: 103).

Asimismo, Torres Sánchez, (2018: 197) señala que este tipo de interferencia contrasta con la transferencia, en términos de su estabilidad (Silva-Corvalán, 2001: 272), pues argumenta que el hecho de que el tipo de bilingüismo sea la variable con mayor peso

explicativo apoya la idea de que es un fenómeno menos sistemático ya que, aunque se presenta en bilingües con diferentes grados de dominio del español, parece haber una relación con el grado de competencia en español, de modo que a mayor conocimiento de esta lengua, menor número de discordancias (Torres Sánchez, 2018: 197). La autora concluye que la discordancia de número en el español en contacto con tepehuano del sur y con otomí es una interferencia estructural que, a pesar de ser común a diferentes situaciones de contacto entre el español y otras lenguas, presenta matices según la lengua en contacto y la situación sociolingüística (201).

A continuación, presento las características generales del corpus utilizado en este capítulo, hago algunas precisiones conceptuales y describo los factores lingüísticos y sociales que emplearé en el análisis de la concordancia plural, como podremos observar, muchos de estos se basan en los hallazgos de las investigaciones descritas en esta sección. Posteriormente, expongo el análisis general de la concordancia plural del español en contacto.

3.3 DECISIONES METODOLÓGICAS PARA EL ANÁLISIS DE LA CONCORDANCIA PLURAL

3.3.1 Conformación del corpus

Construí el corpus de este capítulo a partir de las narraciones tradicionales y de vida en español de los cuarenta y tres colaboradores. En principio, procuré tener cien ejemplos por colaborador: cincuenta casos de concordancia plural en el nivel frástico y cincuenta en el

oracional,⁵⁹ respetando el principio de exhaustividad;⁶⁰ sin embargo, no fue posible alcanzar esa cifra con todos los hablantes, por lo que al final el corpus quedó conformado por 3 818 casos.

3.3.2 Algunas precisiones conceptuales

En este trabajo entiendo como concordancia la “relación entre al menos dos palabras que se establece con la repetición en cada una de ellas de uno de los morfemas de género, de número o de persona, y que sirve en lo fundamental para relacionar e identificar léxica y sintácticamente las palabras concordantes” (Martínez, 1999: 2697). Asimismo, considero que esta relación es asimétrica, pues uno de los elementos de la relación impone sus marcas morfológicas al resto, ya sea de manera explícita o no⁶¹ (Soler Arechalde, 2012: 16).

En cuanto a la concordancia en la frase nominal, considero que el artículo y los adjetivos calificativos, indefinidos y determinativos (demostrativos, posesivos, numerales cardinales, fraccionarios y multiplicativos, ordinales, cuantificadores definidos, los distributivos, los partitivos), cuya variabilidad morfológica lo permita, concuerdan en género y número con el núcleo de la frase nominal (nombre común, nombre propio, pronombre personal) (Martínez, 1999: 2712-2714, 2722). En el nivel oracional, el verbo concuerda en

⁵⁹ Debido a que la narración de las historias de vida se desarrollaba principalmente en primera persona, singular o plural, decidí considerar 30 ejemplos de la historia de vida y 20 de la historia tradicional, con la finalidad de poder tener representadas las tres personas gramaticales.

⁶⁰ El principio de exhaustividad señala que en el estudio de una variable lingüística es necesario considerar todas sus realizaciones o variantes, en todos los contextos en que es susceptible de aparecer, incluyendo la ausencia (Moreno Fernández, 1998: 310).

⁶¹ Soler Arechalde (2012: 16-17) señala que la concordancia puede enfocarse hacia cuestiones sintácticas o tomar en consideración aspectos semánticos. Esta autora se ancla en este segundo enfoque e indica que “si la coincidencia es tanto formal como semántica, la concordancia sería completa (la más común). Si sólo hay coincidencia formal, se tendrá concordancia sintáctica o formal; y si la coincidencia es únicamente semántica, concordancia *ad sensum* o semántica”.

persona y número con el sujeto, al igual que los nombres, adjetivos y participios con función de predicado nominal (Martínez, 1999: 2754-2755, 2778).

En este trabajo analizo la concordancia plural primordialmente desde una perspectiva sintáctica, por lo que clasifico como concordantes los casos que muestran una coincidencia formal y como discordantes las relaciones en las que, al menos, una marca no está presente. No obstante, también considero algunos casos de concordancia semántica, como las construcciones con un nombre y un numeral: *ocho perros*.

3.3.3 Factores de análisis

En el análisis incluí algunos factores lingüísticos que han sido importantes para explicar la variación en la concordancia de número en variedades de español monolingüe (véase § 3.2.1): el dominio en el que ocurre la relación de concordancia, la función sintáctica del controlador, el tipo de relación de concordancia (Soler Arechalde, 2012) y la categoría del elemento concordante. Para evaluar la HIPÓTESIS 3.1, que sugiere que la variación en la concordancia de plural en español está determinada por los factores que rigen la concordancia plural en otomí, integré al análisis elementos relacionados con la marcación de número en *ñható* (véase §1.4.2 y §1.4.3): la animacidad del nombre, la especificidad de la referencia (Palancar, 2013), el tipo de marcación (Guerrero, 2006; Torres Sánchez, 2018) y la pérdida de segmentos en coda a final de palabra (Guerrero, 2006; Lastra, 1987), con la finalidad de saber si estos factores inciden en la concordancia de número de la variedad de español en contacto. Por último, para determinar la distribución de la variación entre los individuos y en las comunidades, incorporé las variables de localidad, bilingüismo, edad, nivel de instrucción y género. A continuación, presento en más detalle los motivos que me llevaron a seleccionar estos factores, la manera en que los integré al análisis y las predicciones sobre su efecto.

3.3.3.1 Factores lingüísticos generales

El dominio en el que ocurre la relación de concordancia

Consideré pertinente incluir este factor porque ha sido relevante en la comprensión de la alternancia en la concordancia de número en variedades de español monolingüe. En los datos distinguí dos niveles, el frástico y el oracional. Con base en los resultados en dichas variedades,⁶² espero observar una menor frecuencia de casos de discordancia el nivel frástico y una mayor incidencia en el nivel oracional, debido al tipo de relación de constituencia que establecen (Soler Arechalde, 2012: 256).

Función sintáctica del controlador

El siguiente factor fue la función sintáctica del controlador, e integré este rasgo mediante cinco categorías: sujeto, atributo, adjunto, objeto directo e indirecto. Decidí introducir esta variable porque ha sido determinante en la explicación de la alternancia en la concordancia de número en variedades de español monolingüe (Soler Arechalde, 2012: 257). Siguiendo a esta autora, espero encontrar que las funciones sintácticas más nucleares (p. ej. sujeto, objeto directo) favorezcan la presencia de las marcas de concordancia plural, mientras que las funciones periféricas, como la de adjunto, las presentarán en menor proporción.

Tipo de relación de concordante

Decidí considerar este rasgo porque, junto con la función sintáctica, ha demostrado ser importante para explicar la variación en la concordancia de número en variedades de español monolingüe (Soler Arechalde, 2012: 257). En el análisis distingo cuatro tipos de relaciones

⁶² Bogotá, Buenos Aires, Caracas, Madrid, México y Santiago de Chile.

de concordancia, la primera, entre el núcleo nominal y su determinante (D-N) y, la segunda, entre el núcleo nominal y un adjetivo (N-A). En el nivel oracional, la relación de concordancia entre el verbo y su sujeto (N-V), así como la del sujeto y el predicado nominal (Suj-Atri). Decidí implementar este factor porque tengo la expectativa de encontrar diferencias en el uso de la concordancia plural en las diferentes relaciones que se establecen al interior de la frase nominal y en el nivel oracional. Con base en la jerarquía propuesta por Soler Arechalde (2012), espero una mayor concordancia en las relaciones de concordancia en la frase nominal, (D-N) y (N-A), y una mayor presencia de discordancias de número en las relaciones del tipo (N-V) y (Suj-Atri).

Categoría del elemento concordante

Distinguí seis tipos de elementos concordantes: numeral, cuantificador, determinante, adjetivo, sujeto léxico, sujeto flexivo. Decidí incluir este factor porque pienso que algunas categorías podrían favorecer la concordancia y otras desfavorecerla. Por ejemplo, en otomí, los conceptos de propiedad no se expresan mediante una categoría léxica sino a partir de construcciones nominales y verbales (véase §1.4.3.7). Por ello, es posible que las relaciones de concordancia que involucran a los adjetivos presenten una mayor presencia de usos alternos o discordancias de número.

3.3.3.2 Factores lingüísticos relacionados con la marcación de número en otomí

Animacidad

Integré al análisis la animacidad porque tipológicamente ha mostrado ser determinante en la organización de los sistemas de marcación y concordancia de número (Corbett, 2000). Específicamente en la lengua otomí, la concordancia de número en el verbo con sujeto de

tercera persona depende de que el referente sea humano (Palancar, 2013: 118). En el análisis incluyo este factor con tres categorías: inanimado, animado y humano.⁶³ Con base en la HIPÓTESIS 3.1, que señala que la variación en la concordancia de número del español en contacto responde a estrategias gramaticales de la lengua otomí, espero que la aparición de las marcas de concordancia en el español de contacto esté relacionada con la animacidad del controlador, de tal forma que a mayor jerarquía en la escala de animacidad ([+humano]), mayor presencia de marcas de concordancia y, a menor rango, menor concordancia.

Especificidad

Resolví introducir como factor de análisis la especificidad del referente⁶⁴ porque, al igual que la animacidad, condiciona la aparición de las marcas de concordancia no singular en el verbo otomí cuando el sujeto es una tercera persona (Palancar, 2013: 118). En el análisis parto de un enfoque pragmático y considero que una frase nominal tiene referencia específica “cuando el hablante, al emplearlo, dé a entender que se refiere a un objeto o individuo determinado” (Leonetti, 1999: 860) y como inespecífica “cuando el hablante no pretenda referirse a una entidad determinada, bien porque no sea relevante, bien porque el SN sea afectado por un

⁶³ Corbett (2000: 54-87) plantea una escala de animacidad: Hablante (Pronombre 1ªP) > Oyente (Pronombre 2ªP) > 3ª persona > Parentesco > Humano > Animado > Inanimado, en la que cada elemento puede representar un punto de corte para determinar si las oposiciones en los sistemas de marcación son significativas para ciertos elementos y para otros no.

⁶⁴ El término especificidad se ha tratado a partir de diferentes criterios: pragmáticos, lógicos (en términos de alcance o ámbitos) y discursivos (relacionado con la noción de partitividad). En el primer enfoque, se subraya “la intención del hablante de comunicar y hacer manifiesto que pretende referirse a una entidad determinada tanto si el receptor es capaz de identificar el referente como si no” (Leonetti, 1999: 858). El segundo enfoque se ha centrado en analizar la ambigüedad entre lecturas específicas e inespecíficas en contextos en los que una frase nominal indefinida coaparece con un cuantificador y en contextos intensionales. Bajo esos términos, una frase nominal indefinida tendrá una lectura específica si tiene alcance amplio sobre el cuantificador u operador intensional, es decir, si es referencialmente independiente de éstos; en contraposición, un indefinido es inespecífico cuando tiene un alcance estrecho con respecto a otro operador (Leonetti, 1999: 858-859; Pozas Loyo, 2016: 90-91). En cuanto a la especificidad partitiva, una frase nominal indefinida es interpretada como específica si refiere a un elemento de un conjunto previamente introducido en el discurso (Pozas Loyo, 2016: 91).

contexto modal o intensional,⁶⁵ bien porque el conjunto sobre el que se cuantifica no esté previamente establecido en el discurso” (Leonetti, 1999: 860).

En algunos ejemplos retomo algunos indicios gramaticales que son suficientes mas no necesarios para determinar la especificidad o inespecificidad de la referencia. Para la inespecificidad, consideré la presencia de contextos interrogativos (*¿Tienes un libro de matemáticas?*), imperativos (*Lee muchos libros*), con operadores modales (*Puedes comprarte un libro*), oraciones condicionales (*Si ves un niño perdido, avisa a las autoridades*), negaciones (*No tengo muchos amigos*), predicadores a nivel individual (*Es esencial encontrar un amigo*) y predicados intensionales (*Es necesario traer un libro*) (Fábregas, 2018: 50-51). Como indicios de especificidad tomé en cuenta el modo verbal en las oraciones de relativo (*Quiero comprarme un libro **que habla de Italia** vs. Quiero comprarme un libro **que hable de Italia***), la presencia de un objeto directo preposicional (*Busco un amigo vs. Busco **a** un amigo*) y la posición prenominal de los adjetivos (*Quiero comprarme un **estupendo** libro*), entre otros (Fábregas, 2018: 52-53).

En vista de la HIPÓTESIS 3.1 y del comportamiento que este factor tiene en la concordancia en otomí, propongo que las frases nominales cuya referencia sea específica favorecerán la aparición de las marcas de concordancia plural, mientras que cuando hagan referencia inespecífica estas aparecerán menos.

⁶⁵ Según Carnap (1947), “cada expresión posee una intensión y una extensión: la intensión de un predicado es una propiedad y su extensión es la clase de objetos que tienen esa propiedad [...] utilizó el término “intensión” (que había empleado Leibniz en 1701) para llamar una función que relaciona cada término con las propiedades que determinan la aplicabilidad del término, en oposición a “extensión”, entendida como la función que relaciona cada término con la clase de objetos sobre los que puede caer la aplicabilidad del término” (García Fajardo, 2016: 63).

Tipo de marcación

En el análisis tomo en cuenta uno de los factores que Guerrero (2006) empleó en su investigación sobre la concordancia de número en el español en contacto con el otomí de Santiago Mexquititlán, el tipo de marcación de la expresión de número, y distinguí cuatro patrones: tipo 1) presencia de marca en el modificador + presencia de marca en el nominal (*las niñas*); tipo 2) presencia de marca en el modificador + ausencia de marca en el nominal (*las niña*); tipo 3) ausencia de marca en el modificador + presencia de marca en el nominal (*la niñas*); y tipo 4) ausencia de marca en los dos elementos concordantes (*la niña*). Con base en los resultados de Guerrero (2006: 99-100) y en la HIPÓTESIS 3.1, espero que el patrón de tipo 2, asociado a la lengua otomí, sea el que más favorezca la discordancia de número y, en menor medida, el de tipo de marcación 3.

Índice de pérdida de segmentos consonánticos a final de palabra

La pérdida de segmentos consonánticos a final de palabra es un fenómeno recurrente en el español en contacto con otomí. Al respecto, se ha señalado que es uno de los mecanismos⁶⁶ que los bilingües emplean para reducir las diferencias que el español y el otomí tienen en su inventario silábico⁶⁷ y, así, recuperar la estructura básica del otomí (CV), como se ejemplifica en (12). Pienso que este mecanismo podría tener un efecto en la concordancia debido a que,

⁶⁶ Otros de los mecanismos que suelen emplear son la aféresis en sílabas con estructura V o VC (p. ej. *repentí* ‘arrepentí’, *tonse* ‘entonces’, *cabar* ‘acabar’), la inserción de un corte glótico al final de palabras que terminan y/o inician en vocal (p. ej. ‘*abajo*, ‘*ánima*), la inserción de consonantes y la resilabificación (p. ej. *li.gle.sia vi.ri.gen, te.re.se.ro*) (Guerrero Galván & San Giacomo Trinidad, 2014: 1503-1504; Hekking, 2001: 142-143; Hekking & Bakker, 2005: 255; Lastra, 1987: 240).

⁶⁷ Aunque ambas lenguas tienen sílabas de tipo CV, el español tiene una amplia gama de sílabas abiertas y cerradas: CV, CVC, V, CCV, VC, CCVC, VCC, CVCC, CCVCC (Hidalgo & Quilis, 2004: 210-211; Navarro, 1946: 46-47; Quilis, 1997: 67), mientras que el inventario de sílabas en otomí se reduce a CV, CCV, CVCC y CVC, estas últimas con un mayor número de restricciones morfofonológicas (Guerrero, 2015: 243-244; Lastra, 1997: 24).

en algunos contextos, la omisión de los segmentos /s/ y /n/ puede generar estragos en la categoría de número en el nivel sintagmático y oracional (Guerrero, 2006; Guerrero Galván & San Giacomo Trinidad, 2014: 1503; Hekking, 2001: 142-143; Hekking & Bakker, 2005: 255; Lastra, 1987: 240), como se muestra en (13).

- (12) a. *nosotro* ‘nosotros’
b. *lune* ‘lunes’
c. *camió* ‘camión’
- (13) a. *trese año* ‘trece años’
b. *los muchacho* ‘los muchachos’
c. *los hijo* ‘los hijos’

Por esta razón, determiné evaluar el efecto que la pérdida de segmentos en coda al final de palabra tiene en la variación en la marcación de concordancia de número (Guerrero, 2006; Lastra, 1987). Para ello, generé un índice de pérdida de codas por colaborador. El parámetro se obtuvo a partir de un promedio del número de segmentos consonánticos no asociados al número y elididos al final de palabra, a la vista del total de consonantes en coda final de la historia de vida. Incorporé el resultado de este procedimiento en el análisis como una variable continua. Sobre este factor espero que mientras mayor sea el índice de pérdida de consonantes al final de palabra de un colaborador, mayor número de casos no concordantes presentará.

3.3.3.3 Factores sociales

Localidad

El primer factor social que incluí fue el lugar de origen de los hablantes. Aunque en San Andrés Cuexcontitlán y Jiquipilco el Viejo se está llevando a cabo un proceso de desplazamiento de la lengua otomí, en la primera de estas el proceso está más avanzado (véase §2.2). Con base en esta diferencia, espero encontrar un menor índice de discordancia en los miembros de la primera comunidad y un mayor porcentaje en los hablantes de Jiquipilco el Viejo.

Bilingüismo

El bilingüismo es el factor sociolingüístico más importante y lo incorporé a los análisis mediante los grupos de hablantes que identifiqué a partir de las pruebas de habilidades lingüísticas en otomí y español (véase § 2.3.3): bilingüe con mayor dominio del otomí (BO), bilingüe simétrico (BS), bilingüe con mayor dominio del español (BE), monolingües con conocimiento pasivo del otomí (PO) y monolingües en español (ME). Con base en la HIPÓTESIS 3.4 –que indica que existen diferentes sistemas de concordancia de número en español relacionados con el tipo de bilingüismo de los hablantes–, espero que haya diferentes porcentajes de discordancia entre estos grupos, de modo que los bilingües con un alto dominio del otomí (BO, BS) presentarán una mayor presencia de discordancias de número y los hablantes dominantes en español (BE, PO, ME) mostrarán un bajo porcentaje de estos.

Edad

Tomé en cuenta cuatro grupos etarios: adolescentes (14-19 años), jóvenes (20-34 años), adultos (35-54 años) y mayores (más de 55 años) (Martín Butragueño & Lastra, 2011). Dada la HIPÓTESIS 3.4 y la existencia de una importante diferenciación en el dominio de las lenguas, según la edad, en las dos comunidades –las personas de más edad son las que tienen algún grado de bilingüismo y los más jóvenes tienden a ser monolingües en español– espero que los hablantes mayores tengan un sistema con un número considerable de discordancias y que el sistema de los más jóvenes tenga un bajo porcentaje de estas.

Instrucción

Asimismo, incluí el nivel de instrucción, con tres niveles: bajo (menos de 6 años escolares), medio (7-11 años) y alto (más de 12 años escolares). Resolví integrar esta variable porque ha sido importante en el planteamiento de algunas tipologías de bilingüismo (véase §1.2.2), ya que los años de escolaridad pueden afectar el conocimiento que los hablantes tienen de las lenguas, con mucha frecuencia en favor del español y en detrimento de la lengua originaria. De acuerdo con estas observaciones y en vista de la HIPÓTESIS 3.4, espero que los colaboradores con instrucción baja presenten un porcentaje importante de casos de discordancia y, en menor proporción, los hablantes con nivel de instrucción medio y alto.

Género

Como señalé anteriormente, la diferencia entre el habla masculina y femenina no está basada en diferencias biológicas, sino en los roles sociales que el hombre y la mujer tienen en una comunidad de habla determinada (Labov, 2001: 263). En San Andrés Cuexcontitlán y

Jiquipilco el Viejo hay algunas diferencias en los roles sociales que desempeñan mujeres y hombres, principalmente en las generaciones mayores, pues esta disparidad ha disminuido en las nuevas generaciones. Los hombres forman parte de redes laborales dentro y fuera de la comunidad, lo cual ha generado en muchos de ellos un mayor dominio del español. En contraste, las mujeres son amas de casa y suelen moverse en redes familiares dentro de la comunidad, razón por la que suelen presentar un mayor dominio y uso de la lengua otomí. Con base en estas anotaciones, hipotetizo que las mujeres serán el grupo que favorezca la no concordancia de número, mientras que los hombres tendrán una mayor presencia de casos de concordancia.

En la Tabla 3.1 presento un resumen con las variables lingüísticas y extralingüísticas que emplearé en el análisis. En negritas resalto las variantes que podrían presentar una mayor presencia de discordancias de número.

TABLA 3.1 *Variables de análisis*

Factores	Variantes
Lingüísticos generales	
Dominio	*Frástico *Oracional
Función sintáctica	*Sujeto *O. Directo *O. Indirecto *Atributo *Adjunto
Tipo de relación	*D-N *N-A *N-V *Suj-Atr
Categoría	*Numeral *Determinante *Adjetivo *Cuantificador *S. léxico/pron. *S. morfológico
Lingüísticos asociados al contacto	
Animacidad	*Inanimado *Animado *Humano
Especificidad	*Inespecífico *Específico
Tipo de marcación	*Tipo 1 *Tipo 2 *Tipo 3 *Tipo 4
Pérdida de segmentos	
Sociales	
Localidad	*San Andrés Cuexcontitlán *Jiquipilco el Viejo
Bilingüismo	*BO *BS *BE *PO *ME
Edad	*Generación 1 *Generación 2 *Generación 3 *Generación 4
Instrucción	*Bajo *Medio *Alto
Sexo	*Mujeres *Hombres

En el siguiente apartado expongo los resultados del análisis general. En la primera parte ofrezco una descripción sobre la distribución del número de casos de concordancia y discordancia según los factores lingüísticos y extralingüísticos antes descritos. Además, proporciono los resultados de las pruebas de Chi-cuadrada de Pearson (χ^2) y V de Cramer (véase §2.1.3.1 y §2.1.3.2), que indican si los factores seleccionados inciden de manera significativa en la concordancia / discordancia de número, así como la fuerza de esta asociación. Posteriormente, presento los resultados de un análisis multivariado de efectos fijos que utilicé para evaluar la contribución simultánea de las diferentes variables lingüísticas y extralingüísticas en la variación en la concordancia de número (Labov, 1996: 114-115), con el propósito de determinar si los factores asociados al contacto con el otomí

inciden en la concordancia / discordancia de número en el español de contacto y así poder rechazar o aceptar la HIPÓTESIS 3.1.

3.4 MODELO GENERAL DE LA CONCORDANCIA PLURAL EN EL ESPAÑOL EN CONTACTO CON OTOMÍ

3.4.1 Modelo general: análisis descriptivo

En el corpus de 3 818 casos de concordancia plural prevalecen los casos de concordancia, con un 81.22%, y sólo se presenta un 18.78% de discordancias. Este porcentaje contrasta con las predicciones, pues esperaba encontrar un mayor índice de este fenómeno, dada su aparente notoriedad. El porcentaje de discordancias de número es equiparable al documentado por Guerrero (2006: 91) en el español en contacto con el otomí de Santiago Mexquititlán, con un 19%. El comportamiento de las variedades de español otomí en las tres comunidades también es similar al del español en contacto con *mapuzugun*, donde registran un 19% de discordancias de número (Olate Vinet et al., 2019: 224). Sin embargo, contrasta con los resultados de los bilingües tepehuano del sur-español, cuyo porcentaje de discordancias apenas alcanza un 8.29% y el de los monolingües en español 3.89% (Torres Sánchez, 2018: 150).

TABLA 3.2. *Concordancia de número en el corpus*

Concordancia	Frecuencia	Porcentaje
Concordancia	3 101	81.22
No concordancia	717	18.78

Sobre la oposición entre la notoriedad de los casos de discordancia y su bajo porcentaje de aparición, Torres Sánchez (2018: 178) señala que se trata de “un fenómeno

marcado –en el sentido de prominente– para el oído del monolingüe en español”,⁶⁸ cuya relevancia podría atribuirse a los valores que indiza. En el caso particular de este estudio, creo que el porcentaje global bajo enmascara la existencia de diferentes sistemas de concordancia (HIPÓTESIS 3.4), pues como se verá más adelante, este porcentaje puede elevarse arriba del 50% en hablantes cuya lengua dominante es el otomí (BO). No obstante, no descarto la idea de realizar un trabajo más adelante en el que indague acerca del significado social de esta variable y el efecto en su uso y distribución dentro de las comunidades.

A continuación, presento una descripción general de los resultados en una serie de tablas donde se muestra la frecuencia y el porcentaje de casos de concordancia y discordancia de número según los factores de análisis descritos en §3.3.3, así como los resultados de la prueba de Chi-cuadrado de Pearson y de V de Cramer (véase §2.1.3.1 y §2.1.3.2), que indican si existe una relación significativa entre los factores lingüísticos y extralingüísticos, y la fuerza de esta.

3.4.1.1 Factores lingüísticos generales

La Tabla 3.3 muestra la distribución de los casos de concordancia y discordancia de acuerdo con factores lingüísticos generales. La primera gran distinción que hice en los datos fue en torno al dominio sintáctico en el que se presenta la relación de concordancia de número (frástico u oracional). Las pruebas estadísticas señalan que la concordancia plural está asociada de manera significativa con el dominio sintáctico (Fisher, $p < 0.05$), y que esta

⁶⁸ Olate *et al.* (2019) señalan que en la mayoría de los estudios previos sobre el español *mapuzugun* se ha enfatizado la presencia de discordancias de número; sin embargo, estas observaciones contrastan con los índices hallados en su investigación, los cuales muestran un bajo porcentaje de aparición del fenómeno. Los autores explican estas diferencias como un posible proceso de retracción lingüística.

asociación tiene un nivel de fuerza bajo (V de Cramer: 0.140). En el corpus encontré un mayor porcentaje de casos de discordancia (24.5%) en el nivel frástico, en comparación con el nivel oracional (13.5%). Estos resultados contrastan con lo reportado en variedades monolingües de español, en las que la cantidad de discordancias de número en el nivel frástico es inferior a la del nivel oracional, debido a que el núcleo nominal y sus modificadores tienen una relación de constitución más fuerte que la existe entre el núcleo nominal del sujeto y el núcleo del predicado verbal o nominal (Soler Arechalde, 2012: 255-256).

TABLA 3.3 *Distribución de los casos de concordancia y discordancia por variables lingüísticas generales*

Nivel	No concordancia		Concordancia		Indicadores
	F	%	F	%	
Frástico	449	24.5	1385	75.5	Fisher= 0.000 V de Cramer: 0.140 Sign: 0.000
Oracional	268	13.5	1716	86.5	
Relación	No concordancia		Concordancia		Indicadores
	F	%	F	%	
Suj-Atri	34	26.1	96	0.738	$\chi^2= 0.000$ V de Cramer: 0.186 Sign: 0.000
D-N	373	24.5	1170	75.4	
N-A	76	24.1	215	75.8	
N-V	234	12.6	1620	87.3	
Categoría EC	No concordancia		Concordancia		Indicadores
	F	%	F	%	
Numeral	86	31.5	187	68.5	$\chi^2= 0.000$ V de Cramer: 0.189 Sign: 0.000
Cuantificador	32	24.8	97	75.2	
Artículo	306	24.5	942	75.5	
Adjetivo	52	19.9	209	80.1	
Sujeto léxico	126	19.4	524	80.6	
Sujeto flexivo	115	9.1	1142	90.9	
Función	No concordancia		Concordancia		Indicadores
	F	%	F	%	
Objeto indirecto	19	31.1	42	68.9	$\chi^2= 0.015$ V de Cramer: 0.082 Sign: 0.015
Adjunto	190	27.0	513	73.0	
Objeto directo	114	22.4	396	77.6	
Sujeto	392	15.6	2116	84.3	
Atributo	2	5.6	34	94.4	

En (14), muestro algunos ejemplos de concordancia y discordancia de número en el nivel frástico y oracional. En el primer ejemplo, el artículo de la frase nominal *las muchacha*, tiene la marca morfológica de plural *-s* pero el núcleo nominal se encuentra en su forma singular. En este mismo fragmento, el numeral de la frase *treinta grupo* semánticamente indica pluralidad, por lo que el núcleo nominal debería ocurrir en su forma plural, sin embargo, se produce en su forma singular. Posteriormente, está la frase nominal *puras*

mujere, en la que tanto el nombre como el adjetivo presentan la marca morfológica plural. Por último, el adjetivo de la frase *puras pareja* posee la marca de número plural, pero el núcleo nominal no. En (14b), hay un caso de discordancia entre el verbo, *llegó*, que corresponde a una 3ª persona singular, y su sujeto, que refiere a una 3ª persona plural, *los que tejía palma*. En la segunda oración, el verbo *sacar* está conjugado en 3ª persona plural pero su sujeto refiere a una 3ª persona singular. Finalmente, está una oración en la que el verbo copulativo, *ser*, concuerda con el sujeto, *sus bailables*, y con el predicado nominal, *buenos bailables*.

- (14) a. [...] y ora saliero **las muchacha**/ ya hay grupos como que de cuarenta y cinco cuarenta **treinta grupo**/ pero puras muchacha/ **puras mujere** ora/ **puras pareja** (*Bilingüe con mayor dominio del otomí, SAC*).
- b. **llegó** a sacar su torito [**los que tejía palma**] y también <~tambié> este/ el los metatero/ los molcajeteros/ [**todo eso sacaron** sus sus toritos/ pero sus toritos/ su uniforme/ [**sus bailables**] **eran** muy **buenos bailables** (*Bilingüe con mayor dominio del otomí, SAC*).

Con respecto al tipo de relación, los resultados de las pruebas estadísticas muestran que la concordancia / discordancia plural está asociada de forma significativa con dicho factor (χ^2 , $p < 0.05$), pero con un nivel de fuerza bajo (V de Cramer de 0.153). En la Tabla 3.3 (*supra*), la relación Suj-Atri tiene más casos de discordancia (26.1%), como se ejemplifica en (15a). El segundo lugar lo ocupa la relación D-N, con 24.5% de discordancias (15b), le sigue la relación nombre adjetivo (24.1%), como se puede apreciar (15c). En última posición se halla la relación N-V, con sólo 12.6% de discordancias (15d). Estos resultados contrastan con lo documentado en variedades de español monolingüe (Soler Arechalde, 2012: 257), donde hay una mayor incidencia de discordancias entre el sujeto y el predicado verbal (N-

V), seguido del predicado nominal (Suj-Atri) y, en último lugar, entre el nombre y sus modificadores (D-N).

- (15) a. terminamos la/ botella/ pues <~ps> menos de dos horas ya **estábamos borrachos** <~borracho> (*Bilingüe con mayor dominio del otomí, SAC*)
- b. “ [...] pus nomás pura camisa”/ y este/ y sí me quitó **los zapato**/ me puso una mantita/ como cobijita/ amarrado en el pecho (*Bilingüe con mayor dominio del otomí, SAC*).
- c. ya de ahí pus terminando eso pus ya lo van a ir a dejar al panteón con sus flores y ya/ bueno es que como son diferente(s) cada costumbre/ aquí nomás **pura flores**/ sí en otras cosa(s) (*Bilingüe simétrico, JV*).
- d. yo me acuerdo que este había feria **los bailarines/ eran puros señores**/ tenían sus chaleco(s) grande(s) su sombrero/ un sombrero grande (*Bilingüe con mayor dominio del otomí, SAC*).

En lo referente a la categoría gramatical, los estadísticos indican que este factor incide de forma significativa en la alternancia en la concordancia (χ^2 , $p < 0.05$), con un nivel de fuerza bajo (V de Cramer: 0.189). La Tabla 3.3 muestra que el 31.5% de los nombres en contrucciones con un numeral no presenta marca de plural, como se ejemplifica en (16a). Cuando el elemento concordante es un cuantificador (16b) o un artículo (16c) se presenta un porcentaje de discordancia cercano a 25%. Los adjetivos alcanzan un 19.9% de no concordancias, como se muestra en (16d). En el nivel oracional, hay un mayor porcentaje de discordancias cuando el sujeto es léxico (19.4%), que cuando se trata de un sujeto flexivo (9.1%), como se ejemplifica en (16e).

- (16) a. Ah/ sí le decía esa ocasión de/ una vez nos dijo mi papá dice este/ “de aquí de **ocho día** nos vamos a Temoaya” (*Bilingüe con mayor dominio del otomí, SAC*)
- b. le digo antes <~ante> mi/ mi mamá me decía antes <~ante> dice no/ no trabajábamos las señoras pues <~pus> es que luego tenía(n) **muchos hijo** (*Bilingüe con mayor dominio del otomí, JV*)

- c. porque tenía caballo mi papá/ [4:00] tenía puerco/ y no le dimo de comer [a] **los puerco**/ otra vez este o los caballo (*Bilingüe con mayor dominio del otomí, SAC*)
- d. pues <~pus> yo cuando crecí/ pues <~pus> según que había este espanto de donde quiera/ porque había este/ **casas muy retirada** (*Bilingüe con mayor dominio del otomí, JV*)
- e. entonces <~entonce> me dio mucho coraje porque terminé de hacer el aseo y luego **ya iba a entrar los muchachos**/ y como siempre faltaba los maestros/ o **las maestra**/ pues <~pus> este/ **se encajaba** conmigo (*Bilingüe con mayor dominio del otomí, SAC*)

Los estadísticos indican que la función sintáctica del controlador afecta de manera significativa la concordancia o discordancia de número (χ^2 , $p < 0.05$), con un nivel de fuerza bajo (V de Cramer: 0.127). La Tabla 3.3 (*supra*) muestra que la función sintáctica con más casos de discordancia el objeto indirecto (31.1%), como se aprecia en (17a). La función de adjunto presenta un 27% de discordancia (17b). La función objeto directo ocupa la tercera posición con 22.4% de discordancias, sirva de ejemplo la oración (17c). La presencia de discordancias disminuye de manera importante con la función de sujeto (15.6%), tal como se muestra en (17d). En contraste, la función atributo apenas alcanza un 5.6% de discordancias, ejemplo de esto es la oración (17e) que tiene relaciones de concordancia normativas.

- (17) a. porque tenía caballo mi papá/ tenía puerco/ y no le dimo de comer **los puerco** (*Bilingüe con mayor dominio del otomí, SAC*)
- b. [...] ya cuando supe trabajar pus ya me salí/ **las cosecha**/ aunque yo no podía todavía <~todavía>/ no guantaba mi surco pero ahí iba (*Bilingüe con mayor dominio del otomí, SAC*)
- c. porque pus antes no/ tú salías/ yo antes iba ir a cuidar **mis pollo** iba a cuida mis borrego/ nadie te ofendía pero ora sí (*Bilingüe con mayor dominio del otomí, JV*)
- d. entonces <~entonce> pos/ **crecía esos animalitos** entonces/ pues pus lo vendía (*Bilingüe con mayor dominio del otomí, SAC*)

- e. había un hondureño/ y su esposa era de allá y/ la otra era/ bueno habían varias/
las compañeras eran como **puras mujeres**/ bueno las meseras eran puras mujeres (*Monolingüe con conocimiento pasivo del otomí, SAC*)

Estos resultados contrastan ligeramente con las predicciones y con las observaciones sobre la alternancia en la concordancia de número en variedades monolingües, en las que la función sintáctica de sujeto es la que propicia una mayor presencia de concordancia de número canónica (Soler Arechalde, 2012: 257).

3.4.1.2 Factores lingüísticos relacionados con el contacto

La Tabla 3.4 muestra la distribución de los casos de concordancia y discordancia de número de acuerdo con algunos factores lingüísticos asociados al contacto con el otomí: la animacidad, la especificidad, el tipo de marcación y la pérdida de consonantes. Con respecto al primer factor, hay una mayor frecuencia de casos de discordancia en nombres animados (41.2%) e inanimados (24.3%), en comparación con los nombres +humanos, cuyo porcentaje de discordancias es de 15.4%. Los estadísticos muestran que la animacidad sí está asociada con la variación en la concordancia de número (χ^2 , $p < 0.05$), con una fuerza de asociación baja (V de Cramer: 0.137). La distribución de los datos no coincide totalmente con la jerarquía de animacidad (véase nota 63) ni con las predicciones hechas pues, aunque los nombres humanos presentan una menor incidencia de discordancias, son los nombres animados, y no los inanimados, los que muestran un mayor porcentaje de discordancias.

TABLA 3.4 *Distribución de los casos de concordancia y discordancia por variables lingüísticas asociadas al contacto*

Animacidad	No concordancia		Concordancia		Indicadores
	F	%	F	%	
Animado	35	41.2	50	58.8	$\chi^2= 0.000$ V de Cramer: 0.137 Sign: 0.000
Inanimado	293	24.3	912	75.7	
Humano	389	15.4	2139	84.6	
Especificidad	No concordancia		Concordancia		Indicadores
	F	%	F	%	
Inespecífico	462	22.1	1627	77.9	Fisher= 0.000 V de Cramer = 0.094 Sign: 0.000
Específico	255	14.7	1474	85.3	
Pérdida de segmentos en coda	No concordancia		Concordancia		Indicadores
	Promedio= 0.0719 ± 0.0444		Promedio= 0.0243 ± 0.0352		
Tipo de marcación	No concordancia		Concordancia		Indicadores
	F	%	F	%	
Tipo 1 (presencia + presencia)	0	0.00	3101	81.22	
Tipo 2 (presencia + ausencia)	539	14.11	0	0.00	
Tipo 3 (ausencia + presencia)	118	3.09	0	0.00	
Tipo 4 (ausencia + ausencia)	60	1.57	0	0.00	

En (19a), muestro un caso de discordancia de número en el nivel frástico, cuyo núcleo nominal es animado. Más adelante, en el fragmento (19b), hay tres frases nominales que tienen como núcleo un nombre inanimado y no concuerdan en número con sus respectivos determinantes. Por último, en (19c), el núcleo de la frase nominal, *mis papás*, es +humano y concuerda en número con el posesivo de primera persona.

- (19) a. yo cuando yo era chiquita pus este/ yo cuidaba **mis vaca** (*Bilingüe con mayor dominio del otomí, SAC*)
- b. porque este tenía yo que levantar a las cinco/ a las sei(s) de la mañana/ a lavar **los traste** o a ver que/ pero pus tenía que pus levantar/ lavar **las mano** y/ hacer **las tortilla** o una salsa (*Bilingüe con mayor dominio del otomí, SAC*)
- c. cuando iba a la escuela [...] no tenía zapatos/ me iba descalza/ es todo lo que me pedían pues no/ no me podía dar **mis papás** (*Bilingüe simétrico, SAC*)

Otro de los rasgos semánticos considerados en el análisis fue la especificidad de la referencia del controlador. Las pruebas estadísticas indican que la especificidad incide de forma significativa en la alternancia en la concordancia de número (Fisher, $p < 0.05$), con un nivel de fuerza muy bajo (V de Cramer: 0.094). En la Tabla 3.4 se aprecia un mayor porcentaje de discordancias cuando el controlador tiene una referencia inespecífica (22.1%), como se ejemplifica en (20a), y un menor porcentaje cuando la referencia es específica (14.7%), como se aprecia en (20b). Estos resultados coinciden con lo documentado en otras variedades de español monolingüe (Soler Arechalde, 2012: 255) y con las predicciones hechas a partir del comportamiento de este factor en otomí, pues los referentes más específicos favorecen la concordancia de número; no obstante, será necesario realizar otro tipo de pruebas para poder afirmarlo de manera contundente.

- (20) a. porque **muchos se quedan joven viuda**/ y no **sabe** si **tiene** un terrenito no lo saben pus trabajar (*Bilingüe simétrico, JV*)
- b. **mis papás** pus no **querían** que yo bailara en eso/ porque dice(n) que/ hacía daño o hace daño (*Bilingüe simétrico, SAC*)

En el análisis también integré el tipo de marcación, tal como lo realizaron Guerrero (2006) y Torres Sánchez (2018) para el español en contacto con otomí y tepehuano del sur. Las pruebas estadísticas sugieren que el tipo de marcación está asociado significativamente con la alternancia en la concordancia de número (χ^2 , $p < 0.05$). En la Tabla 3.4 (*supra*), el 81.2% de los ítems mostró una concordancia plural canónica con el patrón presencia de marca + presencia de marca (21a). Los casos discordantes (18.8%), por su parte, presentaron tres patrones de marcación. El más recurrente fue presencia de marca + ausencia de marca (21b), con 14.11%. Le siguen los casos del tipo ausencia + presencia (21c), con un 3.09%; por

último, se documentan algunos casos de ausencia + ausencia de la marca (21d), con sólo el 1.57%.

- (21) a. y pues <~pus> mi papá no/ no le interesaba/ si traíamos el hambre o no/ no ya llegaron <~llegaro> y váyanse a dar de comer **los animales**/ y ahí vamos (*Bilingüe con mayor dominio de otomí, SAC*)
- b. eso es lo que traía **los viejito**/ o sea **los viejito** que le llama/ y se hacía que las bartolas (*Bilingüe con mayor dominio del otomí, SAC*)
- c. Eres un mentiroso eres un flojo lo que pasa es que no quieres bajar **la mazorcas** (*Bilingüe simétrico, JV*)
- d. pus ante(s) bueno/ ya **mis hijo ello** ya no fe- bueno esto ya/ no **siguen** por/ ora sí la/ los este/ tradición de nuestros pa-/ bueno/ de nuestros padre(s) (*Bilingüe simétrico, SAC*)

Estos resultados concuerdan con lo hallado por Guerrero (2006) en el español en contacto con el otomí de Santiago Mexquititlán, donde la mayoría de los casos de discordancia siguen el Tipo de marcación 2, que guarda correspondencias con el otomí: *ya yo* ‘los perros’. Con respecto al índice de pérdida de segmentos en coda, una variable continua, usamos una prueba de Anova univariante, cuyo resultado muestra la existencia de una relación entre la pérdida de segmentos y la alternancia en la concordancia ($p=0.001$). De este modo, mientras mayor es el índice de pérdida de codas, mayor número de discordancias y viceversa.

3.4.1.3 Factores sociales

Integré al análisis las variables de género, edad, nivel de instrucción y grado de bilingüismo para determinar la distribución de la variación en la concordancia de número en las comunidades (HIPÓTESIS 3.3) y así poder evaluar la HIPÓTESIS 3.4, que indica que existen al menos dos sistemas de concordancia de número en español en SAC y JV: el de los bilingües

con un buen dominio del otomí, que también son los hablantes de mayor edad y con un nivel de instrucción bajo, y el de los hablantes que tienen como lengua dominante o única el español.

TABLA 3.5 *Distribución de los casos de concordancia y discordancia por variables sociales*

Comunidad	No concordancia		Concordancia		Indicadores
	F	%	F	%	
Jiquipilco el Viejo	205	13.9	1272	86.1	Fisher: 0.000 V de Cramer: 0.100 Sign: 0.000
San Andrés Cuexcontitlán	512	21.9	1829	78.1	

Tipo de bilingüismo	No concordancia		Concordancia		Indicadores
	F	%	F	%	
B. dominantes en otomí (BO)	442	44.7	547	55.3	$\chi^2=0.000$ V de Cramer: 0.434 Sign: 0.000
B. simétrico (BS)	234	19.6	958	80.4	
B. dominantes en español (BE)	5	2.1	233	97.9	
M. conocimiento pasivo del otomí (PO)	25	2.6	938	97.4	
Monolingüe en español (ME)	11	2.5	425	97.5	

Edad	No concordancia		Concordancia		Indicadores
	F	%	F	%	
Grupo 4 (+ 55 años)	389	30.4	892	69.6	$\chi^2=0.000$ V de Cramer: 0.319 Sign: 0.000
Grupo 3 (35-54 años)	292	25.7	846	74.3	
Grupo 2 (20-34 años)	25	2.6	938	97.4	
Grupo 1 (15-19 años)	11	2.5	425	97.5	

Instrucción	No concordancia		Concordancia		Indicadores
	F	%	F	%	
Bajo	654	34.8	1227	65.2	$\chi^2=0.000$ V de Cramer: 0.404 Sign: 0.000
Medio	59	3.6	1567	96.4	
Alto	4	1.3	307	98.7	

Género	No concordancia		Concordancia		Indicadores
	F	%	F	%	
Mujeres	460	18.0	1010	79.7	Fisher: 0.052 V de Cramer: 0.027 Sign: 0.093
Hombres	257	20.3	2091	82.0	

La variable localidad está asociada con la variación en la concordancia de número (χ^2 , $p<0.05$), en una magnitud baja (V de Cramer: 0.100). En San Andrés Cuexcontitlán se presenta un mayor porcentaje de discordancias (21.9%), en comparación con la comunidad

de Jiquipilco el Viejo donde sólo se halla un 13.9% de casos. El fragmento (22a) de la narración de un bilingüe simétrico de San Andrés Cuexcontitlán tiene diversos casos de discordancia en el nivel frástico y oracional. En contraste, el segmento (22b) de la historia de vida de un bilingüe simétrico de Jiquipilco el Viejo presenta una mayor tendencia a la concordancia canónica.

- (22) a. ya se empieza a bailar **todo los grupo** que **participa** y el día sábado en la tarde **se vuelven a reunir** ahí en la iglesia/ para bailar una pieza o **dos pieza** y de ahí **se bajan**. (*Bilingüe simétrico, SAC*).
- b. **Los señores se van de viaje/ se van** por meses pero **vuelven** a regresar/ **se van** y este sí/ **se van** a Durango/ **se van** a Querétaro/ se van a Morelia/ a Querétaro este a Tijuana. (*Bilingüe simétrico, JV*).

Dado que en JV el proceso de desplazamiento de la lengua otomí no está tan avanzado como en SAC, esperaba encontrar el patrón opuesto. Este resultado podría estar relacionado con que los miembros de la comunidad de habla de Jiquipilco el Viejo⁶⁹ parecen mostrar cierto rechazo a “la mezcla” de sus lenguas (préstamos, cambio de código) y a las transferencias del otomí en el español, que etiquetan como “hablar mal español”. No obstante, son cuestiones en las que se debe profundizar más para poder plantear una relación de causalidad.

En cuanto al bilingüismo, los estadísticos indican que este factor está asociado de forma significativa con la alternancia en la concordancia de número (χ^2 , $p < 0.05$), con un nivel de fuerza media (V de Cramer: 0.434). Los resultados en la Tabla 3.5 concuerdan con las

⁶⁹ “Una comunidad de habla está formada por un conjunto de hablantes que comparten efectivamente, al menos, una lengua [o una variedad lingüística], pero que, además, comparten un conjunto de normas y valores de naturaleza sociolingüística: comparten unas mismas actitudes lingüísticas, unas mismas reglas de uso, un mismo criterio a la hora de valorar socialmente los hechos lingüísticos, unos mismos patrones sociolingüísticos [...] una comunidad de habla es básicamente una comunidad de consenso, de sintonía entre grupos e individuos diferentes, donde el conflicto está minimizado” (Moreno Fernández, 1998: 23-24). Para mayores detalles, véase (Gumperz, 1962).

predicciones de §3.3.3, pues hay diferencias porcentuales importantes entre los distintos grupos lingüísticos. Los bilingües con un mayor dominio del otomí (BO) presentan la mayor cantidad de casos de discordancia (44.7%), como se muestra en el ejemplo (23a). Los bilingües simétricos ocupan la segunda posición, con 19.6% (23b). Estos porcentajes disminuyen radicalmente en los grupos que tienen como lengua dominante (BE) o única lengua al español (PO, ME), pues apenas muestran en promedio un 2.4% de discordancias (23c).

- (23) a. **Supongamos y los abuelito** o/ o las **sus mamá sus papá todo ello/ le daban** su comida lo que le gustaba su/ la fruta/ eso **su refrescos**. (*Bilingüe con mayor dominio del otomí, JV*).
- b. Pus me quitó todo o sea que me llevó todo **mis zapatos/ todo eran nuevecito mis zapato/ que eran zapatos dura mil/** de hule/ y pus lugar que mi mamá me consolara me pegó/ porque había perdido **los zapatos**. (*Bilingüe simétrico, SAC*).
- c. **[Las pastoras] son danzas/** sí se por lo regular **se visten** de blanco/ **se visten** de blanco con **sus velos** y todo eso y este/ y hay un a persona adulta que que carga/ es un/ es una especie de poste pero delgado/ este/ este es un poste enorme como **unos seis metros de altura**. (*Monolingüe con conocimiento pasivo del otomí, JV*).

La distribución de nuestros datos parece coincidir con la idea de la existencia de diferentes sistemas de concordancia de número en español (HIPÓTESIS 3.4), al menos porcentualmente: el de los bilingües con un importante dominio del otomí (BO, BS) y el de los hablantes que tienen como lengua dominante o única el español (BE, PO, ME). Sin embargo, estos apenas son indicios, pues es necesario realizar más pruebas para determinar si la diferencia porcentual entre estos grupos es significativa estadísticamente y si estos podrían constituir dos sistemas, no sólo por cuestiones cuantitativas, sino por los factores lingüísticos que los organizan y determinan (véase § 3.8).

El tercer factor sociolingüístico que consideré fue la edad. Los estadísticos señalan que este factor está relacionado significativamente con la variación en la concordancia (χ^2 , $p < 0.05$), con una fuerza de asociación media (V de Cramer: 0.319). Los datos de la Tabla 3.5 indican que las personas con más de 55 años (G4) muestran más casos de discordancia (30.4%), como se ejemplifica en (24a), un fragmento de la historia de vida de un colaborador del grupo etario cuatro, que es bilingüe con mayor dominio del otomí. En él se presentan diferentes casos de alternancia en la concordancia plural, tanto en el nivel frástico como oracional. La segunda posición la ocupan los hablantes de la generación de adultos (G3), con un 25.7%. En el segmento de la narración de un bilingüe de este grupo etario (24b), hay cinco ítems de número plural en el nivel frástico y oracional, de los cuales dos del nivel oracional no concuerdan en número.

- (24) a. pus ya no me dejó mi mamá ni mi papá [estudiar]/ y a trabajar a **las milpa(s)/ a los animales/** así pasaba yo// y pus no me importó/ si terminé mi escuela o no/ pero ahora que yo estoy viendo **toda mis hermana/** todo mi ps/ todo lo que **mis compañerita** lo que yo crecí junto/ **todo está** bien menos yo (*Bilingüe con mayor dominio del otomí, SAC*)
- b. este váyanse allá mientras/ **dejen** el niño es que este **están haciendo** un trabajo **los muchacho/** córrele luego **regresa/** sí/ déjenlo porque si no me **interrumpe** (*Bilingüe simétrico, SAC*)
- c. aquí se le invita a **varias comunidades/** aquí la inicio de la feria es el cinco/ que es la víspera/ el seis el mero día lo que te/ es marte(s) Ndongu que lo **conocemos/** cuando **se queman los castillos** todo eso/ el martes **llegan los invitados** que **son las imágenes/** en compañía de toda su gente también/ **vienen** de **varios pueblos** de alrededor de Temoaya (*Bilingüe con mayor dominio del español, JV*)
- d. Jugaba al avioncito/ que se dibujaba abajo y/ jugaba futbol con **mis tíos/** me llevaban al campo ya de ahí este también/ bueno jugaba a/ al futbol sí/ futbol y/ con **las pistolas de agua compraban todos teníamos** un pus/ ahí **nos estábamos mojando** (*Monolingüe en español con conocimiento pasivo del otomí, JV*)

El número de casos discordantes disminuye drásticamente en los hablantes de la segunda generación (2.6%), tal como se muestra en el ejemplo (24c), en el que hay 10 ítems de número plural en el nivel frástico y oracional con una concordancia canónica. Los hablantes de la primera generación también presentan un porcentaje muy bajo de discordancias (2.5%), como se puede ver en (24d), que presenta una concordancia plural normativa. Estos hallazgos coinciden con las predicciones planteadas en §3.3.3 y sugieren una importante diferenciación entre los hablantes de las generaciones mayores (G4 y G3) y los más jóvenes (G2 y G1), al menos en términos porcentuales, lo cual podría ser un indicio a favor de la HIPÓTESIS 3.4.

El nivel de instrucción de los colaboradores también resultó estar asociado de forma significativa a la concordancia de número (χ^2 , $p < 0.05$), con un nivel de fuerza medio (V de Cramer: 0.404). Las tendencias vistas en el corpus coinciden con lo predicho, pues las personas con un nivel de instrucción bajo son las que obtuvieron el mayor porcentaje de discordancias (34.8%), como se aprecia en (25a). En contraste, los colaboradores de nivel medio y alto apenas alcanzaron un 3.6 y 1.3%, respectivamente (25b)-(25c). Estas observaciones constituyen un segundo indicio a favor de la HIPÓTESIS 3.4, aunque sea en términos numéricos.

- (25) a. sí pues luego yo voy con mi esposo/ o con **mis hijo**/ luego le digo/ pues **vamos/ somos hartito/ vamo**/ luego ya tengo **dos nuera** pus ya/ **vamos**/ ahí **vamos todo**/ sí ajá/ y así a ir a traer leña (*Bilingüe con mayor dominio del otomí, JV*)
- b. plantas que no **tenemos** aquí/ **las compañeras** de **diferentes estados**/ **no** los **intercambiamos**/ los que **ellos no tiene/ nosotros le damos**/ y los que **nosotros no tenemos/ ello** no los **vende**/ en este caso si ni uno ni otro/ pus lo **compramos** en el DF <~defe> (*Bilingüe simétrico, JV*)
- c. Ya que **todos los compañeros** que **solicitaban** plaza pues no **tenían** más que secundaria/ y **eran** bueno/ yo no era de la zona pero como llevaba esa ya preparación de primer año de normal **pudieron** darme la plaza (*Bilingüe con mayor dominio del español, JV*)

Los datos de la Tabla 3.5 indican que los hombres emitieron 20.3% casos de discordancia (26a), en comparación con las mujeres que presentan un 18% de no concordancia (26b). Sin embargo, las pruebas estadísticas muestran que no hay una relación de dependencia entre la concordancia de número y el sexo de los colaboradores (Fisher, $p>0.05$). Las diferencias porcentuales podrían estar relacionados con las diferencias que hallamos respecto a los roles sociales que las mujeres y los hombres tienen en San Andrés Cuexcontitlán y Jiquipilco el Viejo, pues mientras en la primera comunidad las mujeres mayores suelen estar más vinculadas al hogar y a moverse en redes locales, en Jiquipilco el Viejo las mujeres suelen salir constantemente de la comunidad a comerciar productos agrícolas y servilletas a los centros urbanos cercanos, lo cual ha contribuido al desarrollo de un buen dominio tanto del otomí como del español.

- (26) a. Pero antes de eso **los maestro de las escuela eran muy groseros/** ¿por qué **era muy groseros?** (*Bilingüe con mayor dominio del otomí, SAC*).
- b. como **mis hermanos se creciero** más después/ pos no **supieron** qué **cosas** lo que nos **pasaban** antes/ quiero decir **los primeros** entonces **son los que sufrieron** mucho. (*Bilingüe con mayor dominio del otomí, SAC*).

En resumen, los resultados muestran que el comportamiento de la variedad de español en contacto con otomí difiere de las variedades de español monolingüe, con respecto a los tres factores considerados (dominio, tipo de relación de concordancia y función sintáctica del controlador). Contrario a lo esperado, en el nivel frástico se registra el mayor porcentaje de casos de discordancia y las funciones periféricas no son las que más casos de discordancia presentan sino algunas centrales, como el objeto directo.

En lo referente a los factores lingüísticos que rigen la concordancia en otomí, la pérdida de segmentos a final de palabra, la animacidad del nombre y la especificidad de la

referencia están asociados a la alternancia en la concordancia de número. La distribución mostró más o menos el efecto esperado, pues son los nombres animados y la referencia inespecífica los que presentan más porcentaje de casos de discordancia (véase §3.3.3). Estos hallazgos constituyen un primer indicio para la comprobación de la tesis acerca de que la aparición de las marcas de concordancia plural podría estar determinado por los factores que rigen la concordancia de número en otomí. Sin embargo, será necesario realizar otro tipo de análisis para comprobar esta afirmación (véanse §3.5 y §3.6).

Con respecto a los factores sociales, la variación en la concordancia de número está relacionada con la localidad,⁷⁰ el tipo de bilingüismo, la edad y el nivel de instrucción. La disposición de los datos corresponde con las predicciones expuestas en §3.3.3, pues las personas de la generación mayor (G4), con un menor nivel de instrucción y dominantes en la lengua otomí (BO) son quienes obtuvieron un mayor porcentaje de casos de discordancias, en comparación con los más jóvenes, con un nivel de instrucción medio-alto y cuya lengua dominante o única es el español (BE, PO, ME). Estos resultados sugieren la factibilidad de la HIPÓTESIS 3.4, sobre la existencia de dos sistemas de concordancia de número en el español de SAC y JV, al menos en términos cuantitativos. Sin embargo, para poder afirmarlo será necesario hacer otras pruebas y comprobar que, más allá de la diferenciación porcentual, existen discrepancias en cuanto a los factores lingüísticos que los organizan (véase § 3.8).

El análisis estadístico descriptivo (frecuencias, porcentajes, resultados de Chi-cuadrado de Pearson y V de Cramer) proporciona un panorama general de la relación que algunos factores lingüísticos y extralingüísticos guardan con el objeto de análisis. Sin embargo, en la lengua estos factores no actúan individualmente, sino de forma simultánea.

⁷⁰ Es importante advertir que los factores que se encuentran asociados a la concordancia de número en una magnitud mayor son los sociolingüísticos, sobre todo el grado de bilingüismo.

Por esta razón, emplearé una serie de análisis multivariados que evalúen la contribución simultánea de diferentes variables lingüísticas y extralingüísticas en la variación de la concordancia de número, y así establecer la dirección del efecto de estos factores y la jerarquía entre ellos.

3.4.2 Modelo general: análisis inferencial

En el apartado anterior presenté la distribución general de los datos y proporcioné los resultados de la prueba de Chi-cuadrado de Pearson que aprobaban o rechazaba la existencia de relaciones bivariantes, sin considerar la incidencia de otros factores –por ejemplo, la relación entre la concordancia y la animacidad del nombre, o la concordancia y la edad de los colaboradores–; sin embargo, en la lengua, los factores lingüísticos y extralingüísticos no actúan de forma individual sino en conjunto. Por esta razón, los datos proporcionados hasta ahora no son suficientes para afirmar o negar que la variación en la concordancia de número tenga su origen en el contacto con el otomí y, en consecuencia, que la aparición o no de las marcas de concordancia de número en español está determinada por los factores que rigen la concordancia de número en otomí (HIPÓTESIS 3.1).

Para ello, será necesario evaluar el efecto que el conjunto de los factores lingüísticos (generales y asociados al sistema de concordancia en otomí) y extralingüísticos (véase §3.3.3) tienen en la concordancia / discordancia de número, y establecer una jerarquía entre ellos, mediante una prueba de regresión logística escalonada (véase § 2.1.3.4). Con respecto a la HIPÓTESIS 3.1, los resultados de la regresión logística escalonada nos permitirán: i) determinar si alguno(s) de los factores lingüísticos asociados al sistema de concordancia de número en otomí que seleccioné incide(n) en la concordancia / discordancia de número en español; de ser, así también ayudarán a ii) saber si la dirección del efecto de dichos factores

es el mismo en las dos lenguas; y por último, a iii) conocer cuál es la importancia de esos factores en la organización del sistema de concordancia de número en el español de SAC y JV.

Para construir el modelo general tomé como variable dependiente a la concordancia / discordancia de número y como variables explicativas la mayoría de los factores lingüísticos y extralingüísticos que describí en la sección anterior:⁷¹ el dominio de la relación de concordancia, el tipo de relación, la categoría del elemento concordante, la animacidad del nombre, la especificidad de la referencia, la pérdida de segmentos en coda final, la localidad, el tipo de bilingüismo, la edad, el nivel de instrucción y el sexo de los colaboradores.

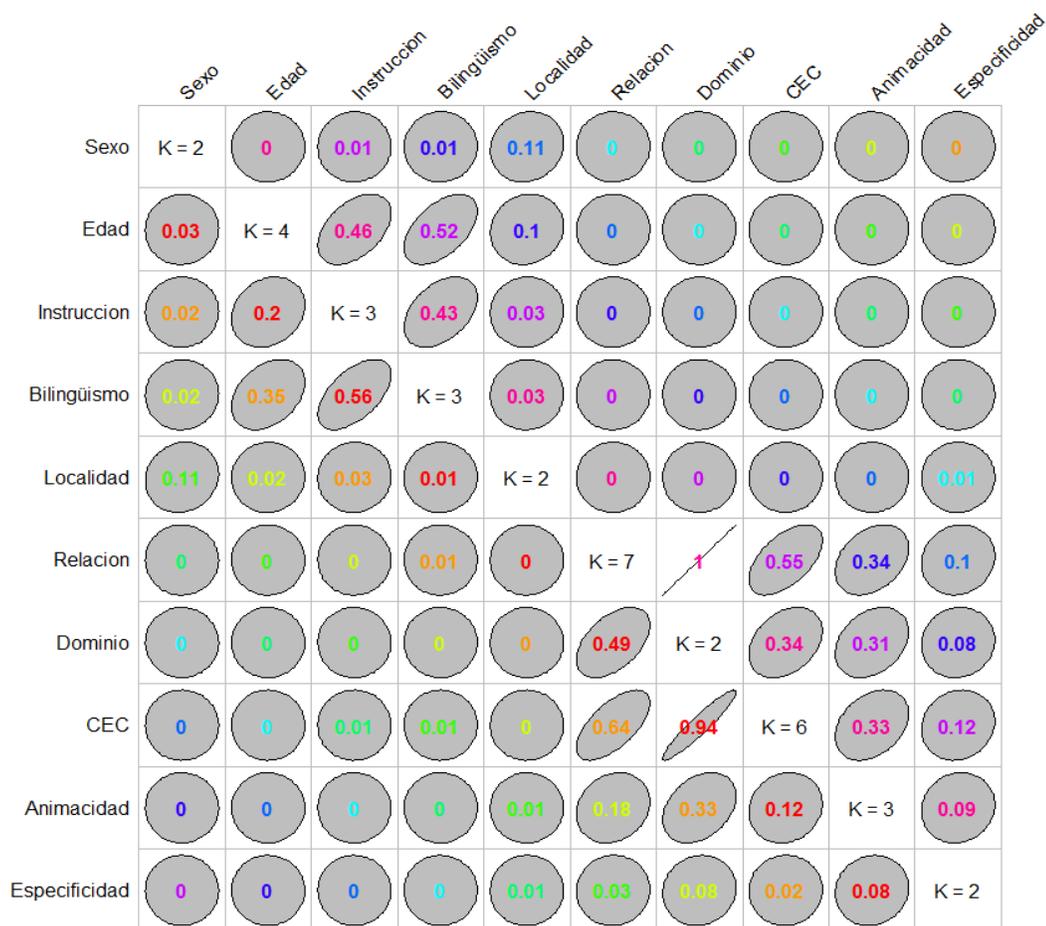
Antes de llevar a cabo la regresión logística escalonada utilicé la medida de Goodman y Kruskal del paquete “The GoodmanKruskal” en la plataforma de R para realizar una prueba de colinealidad entre las variables independientes y evitar errores en el análisis (véase §2.1.3.3). Los resultados se presentan en una escala de 0 a 1, en donde un valor más cercano a 1 indica mayor correlación entre las variables y más probabilidad de problemas de colinealidad.

La Figura 3.1 muestra una correlación entre las variables edad, bilingüismo, e instrucción: **bilingüismo-instrucción (0.56)**, **edad-bilingüismo (0.52)**, **edad-instrucción (0.46)**, instrucción-bilingüismo (0.43), bilingüismo-edad (0.35), instrucción-edad (0.2). La correlación entre estas variables tiene que ver con la constitución de los cinco grupos lingüísticos (BO, BS, BE, PO, ME) que describí en el apartado §2.2.3, donde advertí cierta homogeneidad sociolingüística entre sus miembros. Debido a que los factores edad e

⁷¹ No incluí el factor función sintáctica del controlador porque sólo aplicaba a los casos de concordancia en el nivel frástico. Por último, no integré al análisis el tipo de marcación, debido a que generaba problemas en el análisis por tener variantes con valor categórico.

instrucción pueden subsumirse en la variable bilingüismo y dada la naturaleza de la HIPÓTESIS 3.4, decidí excluir del modelo los factores edad e instrucción y conservar la variable de bilingüismo. Asimismo, se presenta una relación de correlación entre la categoría del elemento concordante (CEC), el tipo de relación y el dominio de la concordancia: relación-dominio (1), relación-CEC (0.55), dominio-relación (0.49), dominio-CEC (0.34), CEC-relación (0.64), CEC-dominio (0.94). Por ello decidí conservar la variable dominio de la concordancia y descartar las otras dos.

FIGURA 3.1 Resultados de la medición Goodman y Kruskal. Modelo completo



Después de ese proceso redefiní el modelo general y realicé la regresión logística por pasos con las siguientes variables: dominio de la concordancia, animacidad del nombre y

especificidad de la referencia, pérdida de codas al final de palabra, localización, bilingüismo⁷² y sexo. A continuación, en la Tabla 3.6 presento los resultados de la regresión logística por pasos.⁷³

TABLA 3.6 *Resultados de la regresión logística escalonada. Modelo completo*

Variables	Estimado	Error Est.	Valor z	Pr(> z)
Intersección	0.05509	0.12339	0.446	0.65527
BE, PO, ME	3.52820	0.17378	20.303	< 2e-16 ***
Bilingües simétricos	1.17045	0.10276	11.390	< 2e-16 ***
Ref. específica	0.52732	0.10520	5.013	5.37e-07 ***
Dominio oracional	0.55691	0.11917	4.673	2.97e-06 ***
Localidad: SAC	-0.43630	0.10901	-4.002	6.27e-05 ***
Sexo masculino	-0.30161	0.10834	-2.784	0.00537 **
Nombre animado	-0.66309	0.27430	-2.417	0.01563 *
Nombre humano	0.10940	0.12617	0.867	0.38589

Códigos de significación: 0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Los resultados muestran que, de las siete variables consideradas, seis resultaron significativas en el siguiente orden de importancia: i) el bilingüismo, ii) la especificidad de la referencia, iii) el dominio de la relación, iv) la localización, v) el sexo y vi) la animación del nombre. En cuanto al grado de bilingüismo, los hablantes que tienen como lengua dominante al español (BE, PO, ME) favorecen la concordancia de número (3.52820) y, en menor proporción (1.17045), los bilingües simétricos (BS). En contraste, los hablantes que tienen

⁷² Para este análisis estadístico fusionamos los cinco grupos lingüísticos en tres, se mantuvo el grupo de bilingües con mayor dominio del otomí (BO) y el de bilingües simétricos (BS), y se fusionó en un solo grupo a los bilingües con mayor dominio del español (BE) y a los dos conjuntos de monolingües en español: PO y ME. Esta reagrupación tiene justificación en los datos de la Tabla 3.12, donde vemos que los BE, PO, ME tienen un comportamiento similar pues todos presentan menos de 2.6% de casos de discordancia.

⁷³ En la regresión logística se tuvo una variable de respuesta dicotómica, concordancia *versus* discordancia. En la primera columna de la tabla se encuentran las variables más relevantes según el modelo aplicado. En la columna con la etiqueta “Estimado” se ofrecen los resultados sobre el sentido de la relación en una escala de log (odds). Los valores en el intervalo de 0 a $+\infty$ indican una relación positiva, es decir, un incremento en la probabilidad de concordancia. Los valores en la escala de 0 a $-\infty$ señalan una relación negativa, es decir, un decrecimiento en la probabilidad de que se presente concordancia de número. En las siguientes dos columnas se proporcionan los valores de error estándar y de Z. En la última columna se presentan los valores de probabilidad, que indican si la variable en cuestión es significativa, de acuerdo con un nivel de significación menor a 0.05.

como lengua dominante el otomí (BO) son los que tienen un mayor índice de discordancias de número.

Asimismo, corroboré la tendencia observada en la Tabla 3.4, las frases nominales específicas favorecen la concordancia de número, mientras que las inespecíficas tienen el efecto contrario. El tercer factor relevante fue el dominio sintáctico en el que tiene lugar la relación de concordancia, los resultados sugieren que en el nivel oracional existe una mayor posibilidad (0.55691) de que se presente una concordancia de número canónica, en comparación con el dominio frástico.

La localidad del colaborador fue el cuarto factor. Como se vio en la Tabla 3.5, la comunidad de San Andrés Cuexcontitlán es la que favorece la aparición de casos de discordancia de número (-0.43630), mientras que en la comunidad de Jiquipilco el Viejo existe una mayor tendencia a la concordancia canónica. Estos resultados se examinarán en mayor detalle en la sección §3.7. El quinto factor en la jerarquía fue el sexo del colaborador, los hablantes de sexo masculino (-0.30161) fomentan la aparición de casos de discordantes, mientras que las mujeres están asociadas a un uso más canónico.

El último factor relevante para la organización de la concordancia fue la animacidad. Los nombres animados potencian la aparición de casos discordantes (-0.66309), mientras que los nombres con el rasgo +humano favorecen la concordancia (0.10940). Estos datos sugieren una organización de la concordancia en función de la animacidad del nombre, en el que el punto de coyuntura se presenta entre el rasgo animado y el humano.

3.4.3 Discusión del modelo general

En vista de los resultados, surge el cuestionamiento acerca de su contribución a las hipótesis de este capítulo. Con respecto a la HIPÓTESIS 3.1 –que sugiere que la aparición o ausencia de

las marcas de concordancia plural está determinada por los factores que rigen la expresión del número en otomí—, encuentro que el factor más importante en la organización de la alternancia en la concordancia de número es el grado de bilingüismo, nuestro principal índice del contacto y la predominancia de las lenguas en los hablantes, el verdadero *locus* del contacto (Weinreich, 1953: 2). De este modo, vemos que a mayor contacto y predominancia del otomí en los hablantes (BO, BS), mayor presencia de discordancias de número; contrario a esto, cuando el contacto y la predominancia del otomí es menor (BE, PO, ME) ocurren menos casos de discordancia.

Asimismo, dos de los factores que organizan la concordancia de número en otomí (véase §1.4.2 y §1.4.3), la especificidad de la referencia (2° factor en la jerarquía) y la animacidad del nombre (6° factor en la jerarquía), resultaron ser determinantes en el funcionamiento del sistema de concordancia de número en el español de contacto. Además, el efecto de estos rasgos es el mismo en las dos lenguas, pues son los nombres con el rasgo +humano y con referencia específica los que promueven la concordancia de número, mientras que los nombres -humanos (animados/inanimados) con referencia inespecífica se asocian a la discordancia. Con base en esto, podría afirmar que el contacto con el otomí incide en la variación en la concordancia de número, pues los factores que organizan la concordancia de número en esa lengua también determinan la variación en la concordancia de número en el español de contacto. Sin embargo, para poder afirmarlo de manera contundente será necesario revisar algunos puntos relacionados con la HIPÓTESIS 3.2 en los párrafos subsecuentes.

Por otro lado, que el factor Dominio resultara decisivo en la alternancia de la concordancia de número abona en favor de la HIPÓTESIS 3.2, la cual sugiere que la variación en la concordancia de número en el español de contacto se comporta de manera diferente en

el nivel frástico y oracional. Además de eso, el efecto de este factor en la variedad de español de contacto dista del documentado en variedades monolingües, pues mientras en el español bilingüe el dominio oracional favorece la concordancia canónica y el dominio frástico la discordancia, en el español monolingüe se presenta el patrón opuesto.

En lo que respecta a la HIPÓTESIS 3.4, relacionada con la distribución de la variación en la concordancia de número y la posible existencia de diferentes sistemas, los resultados de la regresión logística muestran que el bilingüismo es el factor más determinante en la organización de la concordancia de número en el español de SAC y JV. Los hablantes cuya lengua dominante (BE) o única es el español (PO, ME) y los bilingües simétricos (BS) tienden a tener un sistema canónico de concordancia de número, mientras que los bilingües con mayor dominio de la lengua otomí (BO) presentan un sistema de concordancia alterno. Estos hallazgos parecen apoyar la HIPÓTESIS 3.4; sin embargo, estos datos y los relacionados con las variables de localidad y sexo se analizarán y discutirán en mayor detalle en la sección §3.7 y §3.8, donde se revisarán los datos sobre la distribución de la variación en la concordancia de número.

De regreso con la HIPÓTESIS 3.2, resulta llamativa la importancia y el comportamiento que el factor dominio (3° en la jerarquía) tiene en la variación de la concordancia de número, el cual sugiere una distinción importante entre los niveles frástico y oracional, al menos en términos cuantitativos. Aunado a ello, la significación y conducta de la especificidad (2° factor de la jerarquía) y la animacidad (6° factor) nos remite a la cuestión de si estos factores son igual de relevantes en la organización de la concordancia en el nivel frástico y oracional, ya que en la lengua otomí esta relación se organiza de forma distinta en los dos niveles (véase § 1.4.2 y §1.4.3), y la especificidad y la animacidad sólo son relevantes en el nivel oracional. Ante esta cuestión, resulta pertinente recordar la advertencia de algunos autores (Torres

Cacoullos & Travis, 2018; Weinreich, 1953),⁷⁴ que sugieren realizar una revisión exhaustiva del funcionamiento del fenómeno de estudio en las dos lenguas en contacto, con el objetivo de encontrar puntos de divergencia hacia donde se podría dirigir un cambio y, así, poder plantear hipótesis más precisas que permitan entender de mejor manera el funcionamiento de estos fenómenos.

Con base en los primeros resultados, las preguntas que de estos se derivan y las advertencias de otras investigaciones, considero pertinente plantear un análisis de la concordancia de número por nivel (frástico y oracional), que permita aceptar o rechazar la HIPÓTESIS 3.2 y mejorar los resultados de la HIPÓTESIS 3.1 pues, si pretendo determinar la influencia que el otomí tiene en el español, y la concordancia de número en la lengua indígena funciona de manera diferente en el nivel frástico y oracional, resulta pertinente plantear un análisis adecuado en cada nivel y de acuerdo con sus propios términos. En los apartados §3.5 y §3.6 expondré los resultados del análisis de la concordancia por nivel.

3.5 LA CONCORDANCIA DE NÚMERO EN LA FRASE NOMINAL

3.5.1 Introducción

El objetivo principal de este apartado será proponer un análisis pertinente para el estudio de la variación en la concordancia de número en la frase nominal, con la finalidad de determinar si este fenómeno está relacionado de algún modo con el contacto que el español mantiene con el otomí. Para ello, en primer lugar, haré una revisión del comportamiento de la

⁷⁴ Weinreich (1953: 2) señala que en el análisis de una interferencia es necesario describir exhaustivamente las diferencias y similitudes entre las lenguas en contacto con respecto al dominio (fónico, gramatical y léxico) que se esté estudiando. Asimismo, en el método que Torres Cacoullos y Travis (2019) proponen para el estudio de los cambios inducidos por contacto, el segundo paso consiste en revisar el comportamiento del fenómeno de estudio en variedades monolingües de las lenguas en contacto, con el objetivo de encontrar puntos de divergencia estructural hacia donde se podría dirigir un cambio.

concordancia de número en la frase nominal del español y del otomí para localizar los puntos de encuentro y divergencia entre las dos lenguas (Poplack & Levey, 2010) y, así, poder plantear una hipótesis y un análisis apropiado.

A partir de la resolución de ese apartado, realizaré un análisis cuantitativo inferencial (regresión logística escalonada) mediante el que determinaré cuál es el efecto que tienen los factores lingüísticos y extralingüísticos (generales y asociados al otomí) en la concordancia de número en la frase nominal del español otomí.

3.5.2 El número en la frase nominal del español y del otomí

En seguida describo las características principales de la marcación de número plural en la frase nominal del otomí y del español. En la Tabla 3.7 se puede ver que, en ambas lenguas, el núcleo de la frase y el controlador de la relación de concordancia es un nombre, el cual puede ir acompañado de determinantes antepuestos (elementos concordantes), con valores de número, posesión, (in)definitud, deixis, cantidad, etc. No obstante, también difieren con respecto al estatus morfológico de las marcas de número y su distribución. En español, el número plural se codifica mediante el sufijo *-(e)s*, que se adhiere tanto al núcleo nominal (controlador) como a los determinantes (elementos concordantes). Por otro lado, en otomí, el número plural se codifica con el enclítico prenominal *=ya*, que se ancla a los determinantes que lo preceden (elementos concordantes)⁷⁵, pero no al núcleo nominal (controlador). De este modo, el español sigue un patrón de marcación de número del tipo DET-PL + NOMBRE-PL (presencia + presencia) y el otomí uno del tipo DET=PL + NOMBRE (presencia + ausencia).

⁷⁵ La marca de número *ya* puede usarse como enclítico y adherirse a las marcas de deixis e (in)definitud, también puede usarse como morfema libre en construcciones con numerales y cuantificadores (Palancar, 2009: 98).

TABLA 3.7 *Marcación de número en la FN del español y el otomí*

Factores	Frase Nominal plural	
	Español	Otomí
Marca	Sufijo <i>-(e)s</i>	Enclítico = <i>ya</i> / Morfema libre <i>ya</i>
Núcleo (controlador)	Sufijo <i>-(e)s</i>	Sin marca
Dependiente (elemento concordante)	Sufijo <i>-(e)s</i>	Enclítico = <i>ya</i>
Patrón	DET-PL NOMBRE-PL (PRESENCIA + PRESENCIA)	DET=PL NOMBRE (PRESENCIA + AUSENCIA)
Animacidad	Indistinto	Indistinto
Especificidad	Indistinto	Indistinto
Función	Indistinto	Indistinto

Por último, en relación con los factores semánticos y sintácticos que podrían incidir en la concordancia de número, el español y el otomí siguen sus respectivos patrones independientemente de la animacidad del núcleo nominal, la especificidad de su referencia y su función sintáctica.

A continuación, en la Tabla 3.8, presento algunos ejemplos de cómo se codifica el número plural en español y en otomí, con el objetivo de mostrar los dos patrones de marcación en diferentes tipos de frases nominales. En primer lugar, la marca de pluralidad – *s* del español se afija al núcleo nominal, *niña*, a los determinantes (definido, indefinido, posesivo, demostrativo) y el cuantificador en todos los tipos de construcciones, con excepción de los numerales, que tienen un significado de pluralidad, pero no presentan una marca formal. Por otro lado, en otomí, el enclítico plural (=ya) se ancla morfofonológicamente a las marcas de definitud, indefinitud y deixis, así como a los cuantificadores y numerales, en forma morfema libre (*ya*), mas no al núcleo nominal *ngüü*

‘casa’. Las construcciones posesivas, por su parte, forman parte del sistema de número pronominal en otomí, que posee tres valores, singular, dual y plural, los cuales se codifican mediante enclíticos que se adhieren al núcleo nominal, al igual que la marca morfológica de persona gramatical.

TABLA 3.8 *La marcación de número en diferentes estructuras nominal del español y el otomí*

Construcción	Español	Otomí
Definida	DEF-PL + nombre-PL	DEF =PL + nombre
	la-s casa-s	nu= ya ngüü
	DEF-PL casa-PL	DEF =PL casa ‘las casas’
Indefinida	IND-PL + nombre-PL	IND= PL + nombre
	una-s casa-s	na=ya ngüü
	IND-PL casa-PL	IND=PL casa ‘unas casas’
Demostrativa	DEM-PL + nombre-PL	DEM=PL + nombre
	esta-s casa-s	nuyu=ya ngüü
	DEM-PL casa-PL	DEM.PL=PL casa ‘estas casas’
Posesiva	POS-PL + nombre-PL	POS + NOMBRE=PL
	mi-s casa-s	ma ngüü=hu
	1POS-PL casa-PL	1POS casa=PL ‘mis casas’
Cuantificador	CUANT-PL + nombre-PL	CUANT + PL + nombre
	mucha-s casa-s	gatho ya ngüü
	CUANT-PL casa-PL	CUANT PL casa ‘todas las casas’
Numeral	NUM + nombre-PL	NUM + PL + nombre
	ocho casa-s	ñahto ya ngüü
	NUM casa-PL	ocho PL casa ‘ocho casas’

En resumen, la mayor diferencia que existe entre el español y el otomí, con respecto a la marcación de número en la frase nominal, tiene que ver con la ubicación de la marca de

pluralidad, núcleo-controlador *versus* dependiente-elemento concordante, ya que en español la marca de pluralidad se sufija al núcleo nominal (controlador) y al dependiente (elementos concordantes) y, en otomí, sólo se ancla al dependiente.

La importancia de esta diferencia en la concordancia de número del español en contacto con otomí ya había sido advertida en el trabajo de Guerrero (2006) sobre la concordancia plural en la variedad de español otomí de Santiago Mexquititlán (véase §3.2.7). En su análisis cuantitativo integró estas diferencias mediante la variable tipo de marcación, con tres variantes: el tipo 1) presencia de marca + ausencia de marca $\{PL + (EC)\}_1$, el tipo 2) ausencia de marca + presencia de marca $\{(EC) + PL\}_2$, y el tipo 3 que comprende los patrones presencia $\{PL + PL\}_3$ o ausencia $\{(EC) + (EC)\}_3$ de marca en los dos elementos concordantes. Sus resultados mostraron que la mayoría de los casos de discordancia de número en la frase nominal del español en contacto con otomí seguía el primer tipo de marcación, presencia de marca + ausencia de marca $\{PL + (EC)\}_1$, que guarda correspondencia con la marcación de número en otomí, en donde la marca de número se adhiere a los determinantes mas no al núcleo nominal, dando lugar al patrón presencia + ausencia.

Con base en las observaciones de la Tabla 3.8 y en los resultados de Guerrero (2006), pienso que la variación en la concordancia de número de la frase nominal del español en contacto con el otomí de SAC y JV podría estar asociada con la diferencia que existe entre el español y el otomí con respecto a la marcación de número en la frase nominal. De modo que los casos de discordancia de número de esta variedad se acomodan al esquema del otomí, lengua en la que la marca de número sólo se asocia al dependiente (elemento concordante antepuesto) y no al núcleo de la frase nominal (controlador en posición pospuesta), dando como resultado frases nominales con un determinante en plural y un núcleo en singular, es

decir, con un patrón del tipo (presencia de la marca + ausencia de la marca), tal como sucede en el español en contacto con el otomí de Santiago Mexquititlán (Guerrero, 2006).

No obstante, los casos de discordancia de número en la frase nominal del español otomí de SAC y JV también podrían seguir otros dos patrones, en cuyo caso sería más difícil argumentar la influencia del otomí. En el primero de estos, la marca de número sólo se hospedaría en el núcleo y no en el dependiente, generando el patrón (ausencia de la marca + presencia de la marca); mientras que en el otro patrón posible no se presentaría ninguna marca pluralidad (ausencia de la marca + ausencia de la marca).

En la Figura 3.2 muestro los patrones de marcación plural en la frase nominal del español (P-P) y del otomí (P-A), así como los tres posibles patrones de discordancia de número (P-A), (A-P) y (A-A), de los cuales sólo el primero tiene una estructura similar a la del otomí, pues la marca de número sólo se sufixa al dependiente (P-A).

FIGURA 3.2 *Tipos de marcación de número*

Español		Español de contacto				Otomí			
Concordancia		Discordancia							
P-P		A-P		A-A		P-A			
La-s	casa-s	La-∅	casa-s	La-∅	casa-∅	La-s	casa-(s)	nu=ya	ngüü
DEF-PL	casa-PL	DEF-∅	casa-PL	DEF-∅	casa-∅	DEF-PL	casa-∅	DEF=PL	casa

Con base en la HIPÓTESIS I, ahora enfocada en la FN, y en la posible existencia de estos patrones, pienso que, si la estructura de la marcación de número en otomí incide en la variación en la concordancia de número del español, entonces los casos de discordancia presentarán una configuración similar a la del *ñható* (P-A). Para poder determinarlo, en primer lugar, revisaré la proporción de casos de concordancia y discordancia de número en

el corpus, considerando los cuatro patrones (P-P, A-P, A-A, P-A) de la Figura 3.2. Con base en los resultados plantearé un análisis cuantitativo inferencial para determinar qué factores lingüísticos y extralingüísticos inciden en su uso.

3.5.3 Resultados de la concordancia de número en la FN

En el corpus de 1 834 frases nominales predominan los casos de concordancia (75.51%) y las frases discordantes sólo comprenden el 24.48% de los datos. Además, la mayoría de los casos de discordancia sigue el patrón de marcación P-A (22.24%), que tiene la misma configuración que en otomí. El patrón A-A representa el 1.36% de los datos y el patrón A-P apenas alcanza el 0.87% de los casos. Este resultado concuerda con lo hallado por Guerrero (2006) y con la hipótesis que señala que los casos de discordancia de número en el español se adecúan al patrón de la lengua otomí.

TABLA 3.9 Concordancia de número en la frase nominal

Concordancia	Patrón	Frecuencia	Porcentaje
Concordancia	P-P: La-s casa-s	1 385	75.51%
Discordancia	P-A: La-s casa-Ø	408	22.24%
	A-P: La-Ø casa-s	16	0.87%
	A-A: La-Ø casa-Ø	25	1.36%

Los resultados de la tabla muestran que la estructura dominante en los casos discordantes es el tipo de marcación que sigue el esquema del otomí (P-A). Por ello, decidí otorgarle un lugar central al tipo o patrón de marcación, de modo que el objetivo del análisis estadístico será determinar cuáles son los factores lingüísticos y extralingüísticos que están

relacionados con el uso del tipo de marcación concordante (Presencia de la marca + Presencia de la marca) y cuáles con el patrón que sigue el esquema del otomí (Presencia + Ausencia).⁷⁶

Antes de llevar a cabo el análisis estadístico inferencial, realicé tablas de contingencia en las que crucé la variable tipo de marcación con las variables lingüísticas y extralingüísticas planteadas en §3.3.3 y calculé los estadísticos Chi-cuadrada de Pearson y V de Cramer (véase §2.1.3.1 y §2.1.3.2), cuyos resultados se encuentran en el Apéndice 3A. Además, utilicé la medida de Goodman y Kruskal del paquete *The GoodmanKruskal* en la plataforma de R para cerciorarme que no existiera colinealidad entre las variables independientes (véase §2.1.3.3). Con base en los resultados (véase Apéndice 3B), redefiní el modelo para la FN, en el que tomé como variable dependiente el tipo de marcación: concordante (P-P) *versus* discordante con el esquema otomí (P-A) y como variables explicativas la pérdida de codas al final de palabra, la animacidad del nombre, la especificidad de la referencia, la función sintáctica de la frase nominal, el tipo de relación de concordancia, la localidad, el tipo de bilingüismo y el sexo de los colaboradores (véase §3.3.3).

Antes de ver los resultados, es necesario precisar algunos puntos relacionados con la animacidad del nombre y la especificidad de la referencia. Además de la justificación en §3.3.3, decidí integrar estos dos factores al análisis de la concordancia de número en la FN porque en el modelo general resultaron ser importantes para explicar la concordancia de número en el nivel frástico y oracional, a pesar de que sólo son determinantes en la concordancia de número oracional del otomí (véase §1.4.2). En ese sentido, me interesa saber

⁷⁶ Uno de los miembros del Comité Lector indicó que un factor que no tomé en cuenta es la frase nominal contiene un determinante que entraña pluralidad (“más de uno”) o si tiene un marcador lábil o alternante de cuya forma (presencia de -s) depende que se confiera o no un sentido de pluralidad. Este es un comentario muy acertado, que por cuestiones de tiempo no fue posible incorporar a esta investigación, pero que sería sumamente importante incorporar en investigaciones posteriores.

si la animacidad y la especificidad realmente tiene un efecto en el nivel frástico o sólo en el oracional, lo cual abonaría en favor de la hipótesis de que la variación en la concordancia de número en el español de contacto se comporta de manera diferente en el nivel frástico y oracional (HIPÓTESIS 3.2),

Con base en esta hipótesis, y dado que la animacidad y la especificidad no tienen un efecto en la determinación de los patrones de marcación de número en la frase nominal del español y el otomí, espero que no sean factores determinantes en la concordancia de número en el nivel frástico. En caso de que llegasen a tener un impacto, esperarí encontrar una mayor presencia de concordancia (patrón P-P) cuando el núcleo nominal ocupase una posición alta en la jerarquía de animacidad y/o tuviera una referencia específica, y una mayor presencia de discordancias (patrón P-A) cuando el nombre se ubicase en una posición inferior en la escala de animacidad y/ tuviera una referencia inespecífica.

A continuación, en la Tabla 3.10 se exponen los resultados de la regresión logística por pasos (véase §2.1.3.4).⁷⁷ Los resultados muestran que, de las ocho variables consideradas, cuatro resultaron significativas, en el siguiente orden de importancia: i) la pérdida de segmentos consonánticos a final de palabra, ii) el bilingüismo, iii) el sexo del colaborador y la iv) la función sintáctica.

⁷⁷ En la regresión logística se tuvo una variable de respuesta dicotómica: tipo de marcación discordante con el esquema otomí (P-A) *versus* el tipo de marcación concordante (P-P). En la primera columna de la tabla se encuentran las variables más relevantes según el modelo aplicado. En la columna con la etiqueta “Estimado” se ofrecen los resultados sobre el sentido de la relación en una escala de log(odds). Los valores en el intervalo de 0 a $+\infty$ indican una relación positiva, es decir, un incremento en la probabilidad del tipo de marcación concordante. Los valores en la escala de 0 a $-\infty$ señalan una relación negativa, es decir, un decrecimiento en la probabilidad de que se presente el tipo de marcación concordante (P-P). En las siguientes dos columnas se proporcionan los valores de error estándar y de Z. En la última columna se presentan los valores de probabilidad, que indican si la variable en cuestión es significativa, de acuerdo con un nivel de significación menor a 0.05.

TABLA 3.10 *Resultados de la regresión logística escalonada de la variación en el tipo de marcación en la FN*

Variables	Estimado	Error Est	Valor z	Pr(> z)
Intersección	1.5532	0.2532	6.133	8.62e-10 ***
Pérdida de consonantes	-19.6376	2.3306	-8.426	< 2e-16 ***
Español (BE, PO, ME)	3.2543	0.4148	7.846	4.31e-15 ***
Bilingüe simétrico	0.3269	0.1915	1.707	0.087836
Sexo masculino	-0.5064	0.1498	-3.380	0.000724 ***
Función: atributo	1.9787	0.7838	2.524	0.011589 *
Función: objeto indirecto	0.1747	0.3716	0.470	0.638211
Función: sujeto	-0.1157	0.1749	-0.662	0.508168
Función: objeto directo	0.2023	0.1751	1.155	0.248020

Códigos de significación: 0 '***' 0.001 '**' 0.01 '*' 0.05

La pérdida de segmentos consonánticos a final de palabra es el factor que más afecta la alternancia de los tipos de marcación, de modo que, mientras mayor es la pérdida de consonantes a final de palabra menos probabilidad hay de que se presente el tipo de marcación concordante (P-P) y, por tanto, se favorece el patrón discordante (P-A). El siguiente factor más importante es el grado de bilingüismo. Los hablantes que tienen como lengua dominante al español (BE, PO, ME) favorecen el tipo de marcación concordante P-P (3.2543) y, en menor proporción, (0.3269) los bilingües simétricos, como se puede apreciar en los ejemplos (27a) y (27b), respectivamente. En contraste, los hablantes que tienen como lengua dominante el otomí (BO) son los que tienden a emplear el tipo de marcación discordante (P-A) que sigue el esquema otomí, tal como se muestra en (27c) (véase Apéndice 3A).

- (27) a. Yo digo que **los abuelitos** son quien lo hablan mejor (*Monolingüe en español, SAC*)
- b. y así me invitó a salir y así nos conocimos/ y ya tardó y en lo que bailamos en **unos toritos** (*Bilingüe simétrico, SAC*)

- c. I: yo **a los dieciséis año** me junté/ y ya tuve mi primer niño a los diecisiete/ sí
(*Bilingüe con mayor dominio del otomí, JV*)

El tercer factor en la jerarquía es el sexo del colaborador. Los hablantes de sexo masculino (-0.5064) no favorecen el tipo de marcación concordante (P-P) y, por tanto, fomentan el uso del patrón otomí (presencia + ausencia), como se ejemplifica en (28a). En contraste, las mujeres producen un mayor porcentaje de concordancias (28b) (véase Apéndice 3A). El comportamiento de los tipos de marcación según el género de los colaboradores coincide con lo hallado en el análisis general y lo discutiremos con mayor detalle en la sección §3.7.

- (28) a. aquí voy eh ya me hice viejo pero sí todavía este/ sí sí me echo [3:30] **unos viajecito** todavía (*Bilingüe con mayor dominio del otomí, SAC*)
b. pus lugar que mi mamá me consolara me pegó/ porque [4:30] había perdido **los zapatos** (*Bilingüe simétrico, SAC*)

Por último, los datos sugieren que las frases nominales con función de atributo favorecen el tipo de marcación concordante (1.9787), como se muestra en (29). Este comportamiento dista de lo encontrado en variedades monolingües, en las que la función de sujeto es la que suele favorecer la concordancia de número.

- (29) esas son **las dos historias de aquí del pueblo** (*Bilingüe simétrico, JV*)

Por otro lado, los resultados de la regresión logística escalonada confirman que la animacidad del nombre y la especificidad de la referencia no son determinantes en la alternancia de los tipos de marcación de número en la frase nominal, a pesar de que en el primer análisis habían resultado incidir en la organización de la concordancia de número en el nivel frástico y oracional. La irrelevancia de estos factores coincide con el hecho de que ni

en la frase nominal del español y ni en la del otomí son relevantes, y abona en favor de la HIPÓTESIS 3.2 que señala, que la variación en la concordancia de número difiere en el nivel frástico y oracional, sin embargo, sólo podremos saberlo hasta ver los resultados del nivel oracional en la sección §3.6.

3.5.4 Discusión en torno a la concordancia de número en la FN

La HIPÓTESIS 3.1 de este capítulo señala que la variación en la concordancia de número en el español de SAC y JV está asociada con el contacto que mantiene con el otomí, por ello, aunque la concordancia plural se expresa con los recursos morfosintácticos del español, la aparición o no de estas marcas está determinada por alguno(s) de los factores que rigen la expresión gramatical de número en otomí. Para la concordancia de número en la frase nominal, estipulé que la variación podría estar relacionada con la diferencia que existe entre el español y el otomí con respecto a la ubicación de la marca de pluralidad, núcleo-controlador *versus* dependiente-elemento concordante.

Con base en esas observaciones, postulé que si la estructura de la marcación de número en otomí incidía en la variación en la concordancia de número del español, entonces los casos de discordancia de esta variedad se acomodarían al esquema de la lengua otomí, en el que la marca de número sólo se asocia al dependiente (elemento concordante antepuesto) y no al núcleo de la frase nominal (controlador en posición pospuesta), dando como resultado frases nominales con un determinante en plural y un núcleo en singular, es decir, con un patrón del tipo (presencia de la marca + ausencia de la marca).

El primer análisis mostró una prevalencia del tipo de marcación concordante P-P (75.51%), con respecto a los tres tipos de marcación discordante (P-A, A-P, A-A: 24.48%), donde predominó el patrón que sigue el esquema del otomí P-A (22.24%). Estos primeros

resultados sugieren que esta hipótesis se cumple, pues la mayoría de los casos de discordancia de número en la frase nominal del español de contacto sigue el esquema del otomí (P-A).

Posteriormente, realicé una regresión logística escalonada para determinar qué factores lingüísticos y extralingüísticos incidían en la alternancia entre los tipos de marcación concordante (P-P) y discordante (P-A) y, así, establecer la dirección del efecto y la jerarquía de estos factores. Los resultados de la regresión logística escalonada muestran que sólo dos factores lingüísticos inciden en la alternancia de los tipos de marcación: la pérdida de consonantes a final de palabra y la función sintáctica de la frase nominal.

Con respecto al primero, la elisión de consonantes a final de palabra promueve el tipo de marcación discordante con el esquema otomí (P-A). La estrecha relación entre estos dos fenómenos tiene que ver con la naturaleza fonológica del sufijo de número en la frase nominal, *-(e)s*, pues se trata de un segmento consonántico que ocurre al final de palabra, posición donde los bilingües otomí-español tienden a elidir segmentos, como parte de uno de los mecanismos para recuperar la estructura silábica básica CV de su primera lengua, el otomí.

El otro factor lingüístico es la función sintáctica de la frase nominal. Las frases con función de atributo son las que más favorecen el tipo de marcación concordante. Este resultado difiere de lo descrito en variedades de español monolingüe (Soler Arechalde, 2012: 257), donde la concordancia es favorecida por las frases nominales con una función sintáctica nuclear (Sujeto, OD, OI).

En lo referente a la HIPÓTESIS 3.2, que señala que la variación en la concordancia de número difiere en el nivel frástico y oracional, los resultados advierten que la animacidad del nombre y la especificidad de la referencia no son determinantes en la alternancia de los tipos de marcación de número en la frase nominal, a pesar de que en el primer análisis habían

resultado ser determinantes en la organización de la concordancia de número en el nivel frástico y oracional. La irrelevancia de estos factores es consistente con el hecho de que no son relevantes en la marcación de número de la frase nominal del español y del otomí, y abona a favor de la HIPÓTESIS 3.2; sin embargo, sólo podremos saberlo hasta ver los resultados del nivel oracional en la siguiente sección §3.6.

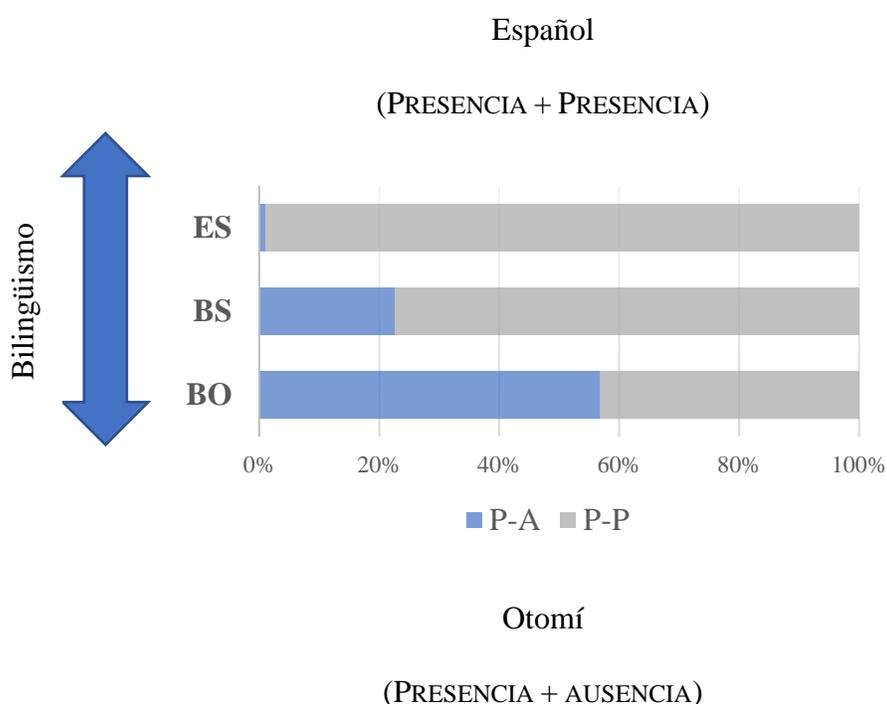
En cuanto a la distribución de la variación en los tipos de marcación y la HIPÓTESIS 3.3, los resultados de la regresión logística escalonada no sugieren un comportamiento diferenciado de la variación en la concordancia de número en la frase nominal en SAC y JV. No obstante, resultaron ser relevantes el tipo de bilingüismo y el sexo de los colaboradores, lo cual podría ser un primer indicio sobre la factibilidad de la HIPÓTESIS 3.4, al menos, en términos cuantitativos. La importancia de la diferenciación por género se discutirá con mayor detalle en la sección §3.8, cuando revisemos la distribución de la variación por comunidad.

En lo referente a la HIPÓTESIS 3.4, los resultados de la regresión confirman lo hallado en el análisis general, pues la alternancia entre los patrones de concordancia y discordancia de número en la frase nominal presenta una distribución diferenciada según el grupo lingüístico. Los hablantes dominantes o monolingües en español (BE, PO, ME) y, en menor proporción, los bilingües simétricos (BS) favorecen el tipo de marcación concordante (P-P); mientras que los bilingües con mayor dominio del otomí hacen un mayor uso del patrón discordante (P-A) (véase Apéndice 3A).

Esto quiere decir que en los hablantes de San Andrés Cuexcontitlán y Jiquipilco el Viejo coexisten los tipos de marcación concordante (P-P) y discordante (P-A), pero en diferentes proporciones. En la Figura 3.3, podemos ver que los bilingües con mayor dominio del otomí (BO) tienen un sistema de concordancia en el que predomina el tipo de marcación que sigue el esquema otomí (P-A), con casi el 60% de los casos, sobre el tipo de marcación

concordante (40%). Mientras que en el resto de los hablantes es mayor el tipo de marcación concordante (P-P), con respecto al tipo discordante (P-A). Este resultado sería un primer indicio sobre la factibilidad de la HIPÓTESIS 3.4, al menos, en términos cuantitativos. Pero será necesario determinar si sólo se trata de diferencias porcentuales o si también hay diferencias en cuanto a los factores lingüísticos que determinan la variación en la concordancia.

FIGURA 3.3 *Distribución de los tipos de marcación según el grado de bilingüismo*



3.5.5 Conclusiones en torno a la concordancia de número en la FN

Los resultados presentados a lo largo de la sección §3.5 permiten afirmar la HIPÓTESIS 3.1 de este capítulo. La variación en la concordancia plural de la frase nominal del español hablado en San Andrés Cuexcontitlán y Jiquipilco el Viejo sí está relacionado con el contacto con la lengua otomí. Para poder explicar cómo el conocimiento del *ñható* condiciona las soluciones que los bilingües emplean para establecer la concordancia plural al hablar español resulta útil

retomar el planteamiento de Johanson (2002) sobre la idea de que una unidad lingüística (simples o complejas, libres o ligadas, léxicas o gramaticales, abiertas o cerradas, frases) posee propiedades materiales (rasgos segmentales, patrones fonotácticos, patrones acentuales), semánticas (contenido denotativo y connotativo), combinacionales (constitución interna, patrones de combinación con otras unidades) y frecuenciales; sin embargo, me abstendré de usar el término *copia de código*, porque parto de la idea de que los hablantes perciben similitudes y diferencias entre sus lenguas y, con base en esto, establecen relaciones de equivalencia entre ellas.

En el caso de la variación en la concordancia plural de la frase nominal, los bilingües encuentran similitudes y diferencias en las propiedades combinacionales de las marcas de número del otomí *ra/ya* y el español *-(e)s* y a partir de eso establecen una equivalencia entre ellas. Esta equivalencia combinacional se puede apreciar en el uso del patrón de marcación alterna (Presencia + Ausencia de la marca: *La-s casa*), que sigue el esquema del *ñható*, en el que la marca de número sólo se hospeda en el dependiente (Presencia + Ausencia de la marca: *Ku=ya yo* ‘Los perros’). Asimismo, se presenta un efecto del factor función sintáctica, que también es relevante en la marcación de número del otomí (véase §4.3). Aunado a esto, la variación en la concordancia plural en la frase nominal está fuertemente relacionada con la pérdida de consonantes a final de palabra, un fenómeno fónico que también tiene su origen en el contacto con el otomí.

La HIPÓTESIS 3.2, que señala que la variación en la concordancia de número en el español de contacto se comporta de manera diferente en el nivel frástico y oracional, será discutida en mayor detalle al final de la sección §3.6, cuando se tengan los resultados de la concordancia plural en el nivel oracional. Por el momento, los datos de la Tabla 3.10 difieren

de los resultados del análisis general (véase §3.6, Tabla 3.6), lo cual es un primer indicio a favor de esta hipótesis.

En cuanto a la HIPÓTESIS 3.3, los resultados de la regresión logística escalonada sugieren que la variación en la concordancia plural de la frase nominal está asociada al género de los colaboradores: los hombres no favorecen la concordancia. Sin embargo, por el momento, no sabemos si este uso diferenciado en función del género se presenta en las dos comunidades o no. Esto lo veremos en mayor detalle en §3.7.

Los resultados de la regresión logística escalonada de la Tabla 3.10 indican que el tipo de bilingüismo es el segundo factor de mayor importancia que incide en la variación en la concordancia de número en la frase nominal del español. Los hablantes cuya lengua dominante es el español (BE, PO, ME) y, en menor medida, los bilingües simétricos favorecen la concordancia plural (Presencia + Presencia), mientras que los bilingües con mayor dominio del otomí promueven el uso del tipo de marcación que sigue el patrón de la lengua otomí (Presencia + Ausencia), como se muestra en la Figura 3.3. Esto es un primer indicio a favor de la HIPÓTESIS 3.4, que supone la existencia de diferentes sistemas de concordancia de número en español, relacionados con el tipo de bilingüismo de los hablantes. En la sección §3.8 discutiré esta hipótesis en mayor detalle y evaluaré la idea de que estos sistemas no sólo se diferencian en el porcentaje de casos de discordancia, sino en los factores lingüísticos que los organizan y determinan, es decir, que responden a gramáticas diferentes.

3.6 LA CONCORDANCIA DE NÚMERO EN EL NIVEL ORACIONAL

3.6.1 Introducción

Los resultados del primer análisis generaron la pregunta de si este era pertinente para evaluar la HIPÓTESIS 3.1, debido a que consideré la variación en la concordancia de número en el nivel frástico y oracional como un solo fenómeno variable, a pesar de que en otomí tiene un comportamiento diferente en cada nivel. Por ello, en esta sección propondré un análisis específico para la variación en la concordancia de número en el nivel oracional, con dos propósitos: el primero es saber si este fenómeno está relacionado con el contacto que el español mantiene con el otomí (HIPÓTESIS 3.1) y, el segundo, si su comportamiento es diferente al de la variación en la concordancia de la frase nominal (HIPÓTESIS 3.2).

En la primera parte de esta sección, realizaré una revisión del funcionamiento de la concordancia de número en español y en otomí en el nivel oracional con la finalidad de determinar cuáles son los puntos de encuentro y divergencia entre las lenguas y, así, poder plantear una hipótesis y un análisis apropiado. A partir de eso, realizaré un análisis cuantitativo inferencial (regresión logística escalonada) para saber qué factores lingüísticos y extralingüísticos inciden en la concordancia plural en el nivel oracional del español en contacto y ver si alguno(s) de estos están relacionados con el otomí.

3.6.2 La concordancia de número en español y en otomí en el nivel oracional

En la Tabla 3.11 se presentan los factores que rigen el funcionamiento de la concordancia plural en el nivel oracional del español y el otomí. Aunque ambas lenguas son *pro-drop* y el sujeto puede ocupar una posición preverbal o posverbal, difieren en algunos aspectos. El primero de estos tiene que ver con la naturaleza de la marca de pluralidad en el verbo, pues

mientras el español tiene una serie sufijos fusionales que codifican los rasgos de tiempo, aspecto, modo, persona y número en el verbo; el otomí posee un conjunto de clíticos verbales especializados en número. Las dos lenguas también discrepan con respecto a la obligatoriedad u opcionalidad de las marcas de concordancia y los factores que la condicionan. En otomí, la concordancia de número no singular es obligatoria cuando el sujeto es una primera y segunda persona, mientras que la aparición de las marcas de concordancia con la tercera persona se restringe a los sujetos con un referente +humano y +específico (Palancar, 2013: 110), para más detalles véase §1.4.2. En contraste, la marcación de la concordancia plural en español es obligatoria, independientemente de la persona, la animacidad y la especificidad del sujeto.

TABLA 3.11 *Características de la concordancia de número en el nivel oracional del español y el otomí*

Lenguas	Español	Otomí
Sujeto	Pro-drop	Pro-drop
Orden	S-V / V-S	S-V / V-S
Marca de número	Sufijo	Clítico
Estructura verbal	raíz verbal-P.TAM.NUM	P.TAM=raíz verbal-OBJ=NUM
Tipo de morfología	Fusional	Aglutinante
Persona gramatical	1ª persona – obligatorio 2ª persona – obligatorio 3ª persona – obligatorio	1ª persona – obligatorio 2ª persona – obligatorio 3ª persona – opcional
Animacidad	Humano – obligatorio Animado – obligatorio Inanimado – obligatorio	Humano – obligatorio Animado – opcional Inanimado – sin concordancia
Especificidad	Específico – obligatorio Inespecífico – obligatorio	Específico – obligatorio Inespecífico – opcional

En síntesis, las principales diferencias entre los sistemas de concordancia de número del español y el otomí en el nivel oracional tienen que ver con la obligatoriedad u opcionalidad de la concordancia de número y los factores que la determinan: persona gramatical, animacidad del nombre y especificidad de la referencia.

Con base en las observaciones de la Tabla 3.11, pienso que la variación en la concordancia de número del español en contacto con el otomí de SAC y JV podría estar asociada con la diferencia que existe entre el español y el otomí con respecto a la marcación de número en el nivel oracional. De modo que, los casos de discordancia de número de esta variedad se acomodan a las restricciones del otomí, lengua en la que la marcación de número es opcional en la 3ª persona gramatical, en función de su animacidad y especificidad. Para poder determinarlo, en primer lugar, revisaré la proporción de casos de concordancia y discordancia de número en el subcorpus oracional y, posteriormente, plantearé un análisis cuantitativo inferencial para determinar qué factores lingüísticos y extralingüísticos inciden en este fenómeno variable.

3.6.3 Resultados de la concordancia plural en el nivel oracional

En el corpus de 1 984 oraciones con sujetos plurales predominan los casos de concordancia (86.49%) sobre los de discordancia (13.50%). Este porcentaje es menor al documentado en el subcorpus de frases nominales, donde se presenta un 24.48% de casos de discordancia. Estos resultados contrastan con lo hallado en variedades monolingües de español, en las que la cantidad de discordancias de número en el nivel frástico era menor que en el oracional, debido a que la fuerza de la relación de constitución entre el núcleo nominal y sus modificadores es mayor a la que existe entre el núcleo nominal del sujeto y el núcleo del predicado verbal o nominal (Soler Arechalde, 2012: 255-256).

Antes de llevar a cabo el análisis estadístico inferencial, realicé algunas tablas de contingencia en las que crucé la variable dependiente (concordancia *versus* discordancia) con las variables lingüísticas y extralingüísticas descritas en §3.3.3. También calculé los estadísticos Chi-cuadrada de Pearson y V de Cramer (véase §2.1.3.1 y §2.1.3.2), cuyos resultados se encuentran en el Apéndice 3C. Asimismo, utilicé la medida de Goodman y Kruskal del paquete *The GoodmanKruskal* en la plataforma de R para cerciorarme que no existiera colinealidad entre las variables independientes (véase §2.1.3.3). Con base en dichos resultados (véase Apéndice 3D), redefiní el modelo de la concordancia oracional y realicé una regresión logística escalonada (véase §2.1.3.4) en la que tomé como variable dependiente la concordancia *versus* discordancia de número y como variables independientes: el tipo de relación, la pérdida de consonantes a final de palabra, la persona gramatical, la animacidad del nombre, la especificidad de la referencia, la localidad, el tipo de bilingüismo⁷⁸ y el sexo de los colaboradores.

Antes de revisar los resultados de la Tabla 3.12, es necesario realizar algunas precisiones en torno a los factores de persona, animacidad y especificidad. Dado que la HIPÓTESIS 3.1 señala que la aparición de las marcas de concordancia en el español de contacto puede estar determinada por algunos de los factores que organizan la concordancia en otomí, consideré pertinente considerar en el análisis la persona gramatical del sujeto, su animacidad y especificidad, debido a que estos factores son determinantes en la marcación de número en el nivel oracional de la lengua otomí, pues la concordancia no singular es obligatoria cuando el sujeto es una primera o segunda persona y opcional en la tercera, cuya marcación está

⁷⁸ Para este análisis estadístico los cinco grupos lingüísticos se fusionaron en tres, se mantuvo el grupo de bilingües con mayor dominio del otomí (BO) y el de bilingües simétricos (BS), y se fusionó en un solo grupo a los bilingües con mayor dominio del español (BE) y a los dos conjuntos de monolingües en español: PO y ME.

subordinada a la animacidad del nombre y la especificidad de la referencia (Palancar, 2013: 118). Al respecto, espero que los sujetos de primera y segunda persona favorezcan la concordancia plural y que los de tercera promuevan los casos de discordancia. Además, supongo que la aparición de las marcas estará asociada a los sujetos humanos y específicos y los casos de discordancia a los nombres de menor animacidad y especificidad.

La Tabla 3.12 muestra los resultados de la regresión logística por pasos⁷⁹ que indican que, de las ocho variables consideradas en el análisis, cinco resultaron significativas, en el siguiente orden de importancia: i) la persona gramatical, ii) el tipo de bilingüismo, iii) el tipo de relación, iv) la animacidad del nombre y v) el sexo del colaborador.

TABLA 3.12 *Resultados de la regresión logística escalonada de la concordancia oracional*

Variab les	Estimado	Error Std.	Valor z	Pr(> z)	
Intersección	3.5000	0.4131	8.473	< 2e-16	***
3ª Persona	-2.5303	0.2697	-9.381	< 2e-16	***
2ª Persona	-0.5617	0.5547	-1.013	0.311280	
Bilingüe simétrico	1.0880	0.1694	6.421	1.35e-10	***
BE, PO, ME	3.0959	0.2288	13.533	< 2e-16	***
Relación: FV-FA	-1.1250	0.2599	-4.329	1.50e-05	***
Relación: FV-FN	-0.7676	0.2176	-3.528	0.000419	***
Nombre animado	-2.2514	0.5732	-3.928	8.56e-05	***
Nombre humano	-0.4964	0.2981	-1.665	0.095855	
Sexo masculino	-0.4935	0.1719	-2.870	0.004105	**

Códigos de significación: 0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05

⁷⁹ En la regresión logística se tuvo una variable de respuesta dicotómica, concordancia *versus* discordancia. En la primera columna de la tabla se encuentran las variables más relevantes según el modelo aplicado. En la columna con la etiqueta “Estimado” se ofrecen los resultados sobre el sentido de la relación en una escala de log (odds). Los valores en el intervalo de 0 a $+\infty$ indican una relación positiva, es decir, un incremento en la probabilidad de concordancia. Los valores en la escala de 0 a $-\infty$ señalan una relación negativa, es decir, un decremento en la probabilidad de que se presente concordancia de número. En las siguientes dos columnas se proporcionan los valores de error estándar y de Z. En la última columna se presentan los valores de probabilidad, que indican si la variable en cuestión es significativa, de acuerdo con un nivel de significación menor a 0.05.

La persona gramatical es el factor lingüístico que más repercute en la alternancia en la concordancia, específicamente, la tercera persona no favorece la concordancia de número (-2.5303), como se ejemplifica en (30a). Los sujetos de segunda persona guardan una relación negativa con la concordancia plural, como se muestra en (30b), pero no resultó ser significativo. En contraste, los sujetos de primera persona favorecen la concordancia plural (30c) (véase Apéndice 3D).

- (30) a. pero yo me mojé mucho/ y **mis hermano estaba** riendo nada más (*Bilingüe con mayor dominio del otomí, SAC*)
- b. como allá donde **vive usted**/ ahí es otra cosa/ ahí le vive/ ahí van a festejar de noche ¿no?/ aquí es de día (*Bilingüe simétrico, JV*)
- c. cuando estaba este/ mis mis güelito este vivía mis agüelito/ y **nosotro no comíamos** nada (*Bilingüe con mayor dominio del otomí, SAC*)

El segundo factor de importancia fue el grado de bilingüismo. Los hablantes que tienen como lengua dominante o única lengua el español (BE, PO, ME) promueven la concordancia de número (3.0959) y, en menor proporción, los bilingües simétricos (1.0880), como se ejemplifica en (31a) y (31b). En contraste, los hablantes que tienen como lengua dominante el otomí (BO) son los que tienen un mayor índice de discordancias de número, como se muestra en (31c) y en el Apéndice 3C.

- (31) a. I: sí/ porque antes **las mamás las abuelas o las tatarabuelas se ponían** chincuetes y orita pus todas son de mezclilla/ de mayones (*Monolingüe con conocimiento pasivo del otomí, JV*)
- b. cuando éramos chicas sí mi mamá luego este a vende/ a vender capulines a cambiar capulines/ que luego nos daban este calabacitas/ y **nosotro les dábamos** capulines/ o **ellos nos daba** maíz y nosotros le dábamos capulines (*Bilingüe simétrico, JV*)
- c. el aseo que íbamos a hacer/ pero nada más yo lo hacía y ellas pus se juntaron <...> más grande y tenían dinero/ **ellos se vestía** bien/ pos/ **tenía** sus pandillas (*Bilingüe con mayor dominio del otomí, SAC*)

El tercer factor en la jerarquía fue el tipo de relación de concordancia. Las construcciones atributivas son las que menos favorecen la concordancia de número (-1.1250), y en menor medida, las oraciones con sujeto en posición posverbal (-0.7676), como se puede observar en (32a) y (32b), respectivamente. En contraste, los predicados verbales con sujeto antepuesto son los que menos casos de discordancia presentan (32c) (véase Apéndice 3C).

- (32) a. I: ah pus sí me puse trabajar para sacar adelante **mis hermanos que estaban chico** porque nos dejaron huérfano (*Bilingüe simétrico, SAC*)
- b. ya es diferente yo veo que sí ya cambió mucho/ ajá sí **orita dice mis niña**/ “no yo el día que yo me quiera/ yo quiera tener un novio/ yo les voy a avisa/ pero no van a pensar que luego me quiero casar” (*Bilingüe simétrico, JV*)
- c. entonces **mis hijo pus se ayudaro** (*Bilingüe simétrico, JV*)

El siguiente factor en la jerarquía fue la animacidad del núcleo nominal del sujeto. Los nombres animados (-2.2514) no promueven la concordancia de número y favorecen la discordancia, como se puede apreciar en el ejemplo (33a). Los sujetos inanimados también muestran un porcentaje importante de discordancias (33b); en contraste, los nombres humanos son los que menos casos de discordancia presentan (33c) (véase Apéndice 3).

- (33) a. cuando estuvimos en Mérida este nos íbamos a/ a Puerto Progreso [9:00] pus tá como a dos horas y media de de Mérida a Puerto Progreso/ era un pueblito muy pequeño pero/ muy hermoso eh **llegaba los barcos** y/ pus a mí me gustaba ver mucho los barco (*Bilingüe con mayor dominio del otomí, SAC*)
- b. I: bueno/ ante sí **aparecía esos pato** que decías/ en ese mismo de/ en ese mismo camino/ la casa del/ de la señora mi Falia que le decían (*Bilingüe con mayor dominio del otomí, SAC*)
- c. I: ah porque **ellos eran/ venían de misiones** aquí/ **se llamaban misioneros/ misioneros de cristo/ entos vinieron aquí**/ nos conocimos y ya de aquí me fui con ellos/ y ya de ahí pus ya me salí (*Monolingüe con conocimiento pasivo del otomí, JV*)

La última variable determinante es el género de los colaboradores. Los hombres no favorecen la concordancia (-0.4935) y promueven la aparición de discordancias de número, como se muestra en (34a). Las mujeres, por su parte, emiten un menor número de discordancias plural, como se observa en el segundo ejemplo.

- (34) a. **Me vine caminando desde allá donde está los bomberos/** porque ahí era un/ una milpa baldío (*Bilingüe con mayor dominio del otomí, SAC*)
- b. I: sí también/ sí/ muchas veces **los jóvenes que se van a otro lado/ regresan con otros conocimientos** y pues lo/ lo practican ¿no? (*Bilingüe simétrico, JV*)

3.6.4 Discusión en torno a la concordancia de número en el nivel oracional

La HIPÓTESIS 3.1 señala que la variación en la concordancia de número es resultado del contacto que el español mantiene con el otomí de modo que, aunque la concordancia plural se exprese con los recursos morfosintácticos del español, la aparición o no de estas marcas está condicionada por alguno(s) de los factores que rigen la expresión gramatical del número en otomí. Para la concordancia de número en el nivel oracional, postulé que, si el otomí influía en la variación en la concordancia de número en el nivel oracional del español de contacto, entonces este fenómeno variable estaría relacionado con la persona gramatical, la animacidad y la especificidad del sujeto. De modo que los sujetos de 1ª y 2ª persona, +humanos y +específicos favorecerían la concordancia de número y los sujetos de 3ª persona, no humanos e inespecíficos promoverían los casos de discordancia.

Los resultados de la regresión logística escalonada coinciden parcialmente con esas predicciones, pues la persona gramatical y la animacidad del sujeto inciden en la variación de la concordancia de número oracional y, además, lo hacen en la misma dirección que en otomí, pues son los sujetos de 3ª persona y no humanos los que promueven la discordancia

de número y los sujetos de 1ª y 2ª persona, +humanos los que favorecen la concordancia. Sin embargo, la especificidad no resultó ser un factor determinante, a pesar de que en el análisis descriptivo se observó una mayor presencia de casos de discordancia de número con sujetos inespecíficos (véase Apéndice 3C). Con base en estas observaciones, la HIPÓTESIS 3.1 se cumple.

La prueba estadística también señala que las construcciones atributivas (FV-FA) y los predicados verbales con sujeto pospuesto (FV-FN), cuyos elementos concordantes están en posición posverbal, no favorecen la concordancia de número. Este comportamiento podría estar relacionado con que en otomí es poco frecuente que los verbos presenten la marca de plural cuando la FN de sujeto está posición posverbal (Hernández Green, 2021).

Sobre la HIPÓTESIS 3.2, los resultados advierten que los factores lingüísticos que determinan la variación en la concordancia de número oracional son distintos a los que actúan en el nivel frástico. En la Tabla 3.13, se puede ver que la pérdida de consonantes a final de palabra y la función sintáctica afectan la marcación de número en la frase nominal; mientras que, en el nivel oracional, actúan la persona gramatical y la animacidad del sujeto, así como el tipo de relación de concordancia. A pesar de estas diferencias, se puede advertir que en los dos niveles inciden las mismas variables extralingüísticas: el sexo y el tipo de bilingüismo de los colaboradores.

TABLA 3.13 *Factores que determinan la concordancia de número en el nivel frástico y oracional*

No.	Concordancia en la FN	Concordancia en la oración
1	Pérdida de consonantes	Persona gramatical
2	Bilingüismo	Bilingüismo
3	Sexo	Relación
4	Función	Animacidad
5		Sexo

La diferencia sobre el efecto de la pérdida de consonantes en los dos niveles sintácticos podría estar relacionada con la naturaleza morfofonológica de las marcas de número en cada nivel. En la frase nominal, la morfología de número tiene un comportamiento aglutinante y el sufijo de número *-(e)s* coincide con el nivel segmental, lo cual explica la fuerte relación entre la elisión de los segmentos consonánticos y los casos de discordancia de número. En contraste, en el nivel oracional, la marcación de número tiene un comportamiento fusional y, en la mayoría de las conjugaciones, las marcas no coinciden con el nivel segmental, por lo que no se ve tan afectado por este fenómeno.

La animacidad no resultó determinante en la alternancia de los tipos de marcación de número en la frase nominal, a pesar de que había sido importante en el modelo general, donde se analizó como un solo fenómeno la variación en la concordancia de número en el nivel frástico y oracional. La irrelevancia de este factor en el nivel frástico es consistente con el hecho de que no es determinante en la marcación de número de la frase nominal del español y del otomí. En contraste, la animacidad sí incide en la variación de la concordancia de número oracional y, además, lo hace en la misma dirección que en otomí, pues son los sujetos +humanos los que favorecen la concordancia. Estos resultados confirman la HIPÓTESIS II, pues la variación en la concordancia de número en el español de contacto se comporta de manera diferente en el nivel frástico y oracional.

Por otro lado, en lo que concierne a la HIPÓTESIS 3.3, los resultados de la regresión no muestran un comportamiento diferenciado de la variación en la concordancia de número oracional por comunidad. No obstante, al igual que en el análisis del nivel frástico, resultó relevante el sexo de los colaboradores, lo cual podría ser un indicio de la factibilidad de esta hipótesis. Por ello, discutiré en mayor detalle la diferenciación por género en la sección §3.7, donde revisaré la distribución de la variación por comunidad.

Con respecto a la HIPÓTESIS 3.4, los resultados confirman lo hallado en el análisis general, pues la variación en la concordancia de número en el nivel oracional también presenta una distribución diferenciada según el grupo lingüístico. Los bilingües con mayor dominio del otomí (BO) tienen un sistema en el que la concordancia de número coexiste con casos de discordancia que se rigen principalmente por los factores que organizan la concordancia de número en el otomí: persona y animacidad. Los bilingües simétricos (BS) muestran un menor porcentaje de casos de discordancia que los BO, pero el tipo de marcación concordante es predominante. Por último, los hablantes dominantes en español muestran un sistema con muy bajo porcentaje de discordancias. Por lo que este resultado sería un primer indicio sobre la factibilidad de la HIPÓTESIS 3.4, al menos, en términos cuantitativos. Pero será necesario determinar si sólo se trata de diferencias porcentuales o si también hay diferencias en cuanto a los factores lingüísticos que determinan la variación en la concordancia (véase §3.8).

3.6.5 Conclusiones en torno a la concordancia de número en el nivel oracional

Los resultados de esta sección §3.5 afirman la HIPÓTESIS 3.1 de este capítulo. La variación en la concordancia plural en el nivel oracional del español hablado en San Andrés Cuexcontitlán y Jiquipilco el Viejo sí está relacionado con el contacto con la lengua otomí. Para poder explicar cómo el conocimiento del *ñható* condiciona las soluciones que los bilingües emplean para establecer la concordancia plural al hablar español, de nueva cuenta, recurro al planteamiento de Johanson (2002) en torno a las propiedades de las unidades lingüísticas (véase §1.1.3).

En el caso de la variación en la concordancia plural en el nivel oracional, los hablantes perciben similitudes y diferencias entre las relaciones de concordancia del otomí y el español,

y, a partir de eso, establecen relaciones de equivalencia entre ellas. Esta equivalencia se puede apreciar en los factores que condicionan la concordancia plural en el nivel oracional de esta variedad de español, a saber: la persona gramatical, la animacidad y el tipo de relación, que son algunos de los factores que determinan la marcación de número del otomí en el nivel oracional. Además, el efecto de estos es el mismo en las dos lenguas: los sujetos de 3ª persona y no humanos promueven la discordancia de número y los sujetos de 1ª y 2ª persona y +humanos son los que favorecen la concordancia.

Los resultados obtenidos en la sección §3.5 y §3.6 confirman la HIPÓTESIS 3.2, que señala que la variación en la concordancia de número en el español de contacto se comporta de manera diferente en el nivel frástico y oracional. Aunque en ambos casos los bilingües toman como modelo las propiedades combinatorias de la concordancia plural del otomí para establecer la relación de concordancia plural en su español, cada nivel tiene especificaciones diferentes, como se resume en la Tabla 3.13, debido a que la concordancia de número en otomí se organiza de forma distinta en los dos niveles (véase §1.4.2 y §1.4.3).

En cuanto a la HIPÓTESIS 3.3, los resultados de la regresión logística escalonada sugieren que la variación en la concordancia plural de la frase nominal está asociada al género de los colaboradores: los hombres no favorecen la concordancia. Sin embargo, por el momento, no sabemos si este uso diferenciado en función del género se presenta en las dos comunidades o no. Esto lo veremos en mayor detalle en la siguiente sección, §3.7.

Los resultados de la regresión logística escalonada de la Tabla 3.12 indican que el tipo de bilingüismo es el segundo factor de mayor importancia que incide en la variación en la concordancia de número en el nivel oracional del español. Los hablantes cuya lengua dominante es el español (BE, PO, ME) y, en menor medida, los bilingües simétricos (BS) favorecen la concordancia plural, mientras que los bilingües con mayor dominio del otomí

(BO) promueven los casos de discordancia. Esto es un primer indicio a favor de la HIPÓTESIS 3.4, que supone la existencia de diferentes sistemas de concordancia de número en español, relacionados con el tipo de bilingüismo de los hablantes. En la sección §3.8 discutiré esta hipótesis en mayor detalle y evaluaré la idea de que estos sistemas no sólo se diferencian en el porcentaje de casos de discordancia, sino en los factores lingüísticos que los organizan y determinan, es decir, que responden a gramáticas diferentes.

3.7 LA CONCORDANCIA DE NÚMERO EN SAN ANDRÉS CUEXCONTITLÁN Y JIQUIPILCO EL VIEJO: ¿UNA O DOS COMUNIDADES?

Los resultados del análisis realizado en §3.4 sugirieron que la variación en la concordancia de número tiene un comportamiento diferente en Jiquipilco el Viejo y San Andrés Cuexcontitlán. Por ello, en esta sección buscaré determinar si el funcionamiento lingüístico y la distribución social de la concordancia de número en el nivel frástico y oracional difiere o no en las dos comunidades (OBJETIVO ESPECÍFICO 3.3). Al respecto, hipotetizo que la variación en la concordancia de número en el español de contacto tiene el mismo comportamiento lingüístico en las dos comunidades, pero difiere en su distribución social en cada comunidad (HIPÓTESIS 3.3). En primer lugar, revisaré el funcionamiento de la concordancia de número en la frase nominal de las dos comunidades a partir del análisis planteado en la sección §3.5. Posteriormente, en la sección §3.7.3, describiré el comportamiento de la concordancia en el nivel oracional con base en el análisis propuesto en §3.6.

3.7.1 *La concordancia plural en la frase nominal del español de San Andrés Cuexcontitlán y Jiquipilco el Viejo*

La Tabla 3.14 muestra la distribución de los tipos de marcación concordante y discordante en las frases nominales de San Andrés Cuexcontitlán y Jiquipilco el Viejo. En las dos comunidades predomina el tipo de marcación concordante (P-P); sin embargo, este porcentaje es mayor en Jiquipilco el Viejo (85.08%) y menor en San Andrés (69.17%). Asimismo, la mayoría de los casos de discordancia siguen el tipo marcación con el esquema del otomí (P-A) en las dos comunidades (JV: 13.67%; SAC: 27.92%) y el resto (A-P, A-A) representa menos del 3%.

TABLA 3.14 *La distribución de los tipos de marcación en SAC y JV*

Comunidad		Jiquipilco		San Andrés	
Concordancia	Patrón	F	f	F	f
Concordancia	P-P: Lo-s niño-s	622	85.08	763	69.17
Discordancia	P-A: Lo-s niño-Ø	100	13.67	308	27.92
	A-P: Lo-Ø niño-s	6	0.82	10	0.90
	A-A: Lo-Ø niño-Ø	3	0.41	22	1.99

Dado que la mayoría de los casos de discordancia siguen el tipo de marcación con el esquema otomí en las dos comunidades, el siguiente paso será realizar un análisis cuantitativo inferencial para saber si los factores lingüísticos y extralingüísticos que determinan la concordancia de número en el nivel frástico son los mismos en San Andrés Cuexcontitlán y Jiquipilco el Viejo.

Antes de llevar a cabo el análisis estadístico inferencial, realicé tablas de contingencia en las que crucé la variable dependiente (concordancia P-P *versus* discordancia P-A) con las variables lingüísticas y extralingüísticas descritas en §3.3.3 para cada comunidad. También

calculé los estadísticos Chi-cuadrada de Pearson y V de Cramer (véase §2.1.3.1 y §2.1.3.2), cuyos resultados se encuentran en el Apéndice 3E. Asimismo, utilicé la medida de Goodman y Kruskal del paquete *The GoodmanKruskal* en la plataforma de R para cerciorarme que no existiera colinealidad entre las variables independientes (véase §2.1.3.3). Con base en los resultados (véase Apéndice 3F), redefiní los modelos para la FN y realicé una regresión logística escalonada (véase §2.1.3.4) en la que tomé como variable dependiente el tipo de marcación: concordante (P-P) *versus* discordante con el esquema otomí (P-A) y como variables explicativas la pérdida de codas al final de palabra, la categoría del elemento concordante, la animacidad, la especificidad, la función sintáctica de la frase nominal, el tipo de bilingüismo y el género de los colaboradores (véase §3.3.3).

En la Tabla 3.15 se puede apreciar que el factor lingüístico que más incide en la alternancia de los tipos de marcación en las dos comunidades es la pérdida de consonantes a final de palabra. El segundo factor que más influye es el tipo de bilingüismo: los hablantes dominantes en español (BE, PO, ME) y, en menor medida, los bilingües simétricos (BS) promueven el uso del tipo de marcación concordante (P-P), mientras que los bilingües dominantes en otomí (BO) son quienes más emplean el tipo de marcación discordante (P-A) en las dos comunidades (véase Apéndice 3E). No obstante, las dos comunidades difieren en la distribución de la alternancia según el género, pues los hombres son quienes favorecen el uso del tipo de marcación (P-A) en San Andrés Cuexcontitlán, pero en Jiquipilco el Viejo este rasgo no resultó significativo.⁸⁰

⁸⁰ En la regresión logística se tuvo una variable de respuesta dicotómica: tipo de marcación discordante con el esquema otomí (P-A) *versus* el tipo de marcación concordante (P-P). En la primera columna de la tabla se encuentran las variables más relevantes según el modelo aplicado. En la columna con la etiqueta “Estimado” se ofrecen los resultados sobre el sentido de la relación en una escala de log(odds). Los valores en el intervalo de 0 a $+\infty$ indican una relación positiva, es decir, un incremento en la probabilidad del tipo de marcación concordante. Los valores en la escala de 0 a $-\infty$ señalan una relación negativa, es decir, un decrecimiento en la

TABLA 3.15 *Resultados de la regresión logística escalonada de la concordancia de número en la FN por comunidad*

Jiquipilco					
No	Variables	Estimado	Error Est.	Valor z	Pr(> z)
	Intersección	1.8287	0.8064	2.268	0.02336 *
1	IPC2	-43.9247	17.5065	-2.509	0.01211 *
2	Dominantes español	2.2909	0.8179	2.801	0.00509 **
3	Bilingües simétricos	1.0906	0.6417	1.700	0.08919 .
San Andrés Cuexcontitlán					
	Variables	Estimado	Error Est.	Valor z	Pr(> z)
	Intersección	3.1376	0.4365	7.188	6.60e-13 ***
1	IPC2	-29.0119	3.6904	-7.861	3.80e-15 ***
2	Dominantes español	2.9968	0.7950	3.770	0.000164***
3	Bilingües simétricos	-1.0098	0.2742	-3.682	0.000231***
4	Hombres	-0.9557	0.1933	-4.945	7.63e-07 ***

Los resultados sugieren que la alternancia en los tipos de marcación de número en la frase nominal tiene el mismo comportamiento lingüístico en San Andrés Cuexcontitlán y Jiquipilco el Viejo pues, en ambos casos, está relacionada con la pérdida de consonantes a final de palabra. Asimismo, las dos comunidades coinciden en la distribución de los tipos de marcación con respecto a la edad, el nivel de instrucción y el tipo de bilingüismo de los colaboradores y sólo difieren con respecto al género (véase Apéndice 3E).

Aunque en las dos comunidades predomina el uso del patrón de concordancia plural canónico en la frase nominal, los hablantes de San Andrés Cuexcontitlán hacen un mayor uso del tipo de marcación que sigue las características combinacionales del otomí (Presencia-Ausencia), en comparación con los bilingües de Jiquipilco el Viejo. Este resultado se opone

probabilidad de que se presente el tipo de marcación concordante (P-P). En las siguientes dos columnas se proporcionan los valores de error estándar y de Z. En la última columna se presentan los valores de probabilidad, que indican si la variable en cuestión es significativa, de acuerdo con un nivel de significación menor a 0.05.

a la idea inicial que teníamos sobre el comportamiento de las comunidades y podría estar relacionado con varios factores.

El primero de ellos tiene que ver con las características de los colaboradores. La mayoría de los bilingües de SAC tienen un mayor dominio del otomí, mientras que en el grupo de JV predominan los bilingües simétricos. Aunado a esto, los datos del Apéndice 3E sugieren algunas diferencias en el comportamiento de los bilingües simétricos de las dos comunidades: los BS de SAC producen un porcentaje importante de discordancias de número, muy parecido al de los BO de ese lugar; en contraste, los BS de JV emplean un bajo porcentaje de éstas, similar al de los hablantes dominantes en español. En relación con este punto, también es importante mencionar que durante la entrevista varios colaboradores de JV mostraron actitudes negativas hacia los fenómenos de contacto, como la discordancia de género en el español o los préstamos en el otomí. Por ello, es posible que estos hablantes cuiden más su habla y eviten producir este tipo de fenómenos.

La HIPÓTESIS 3.3 de este capítulo indica que la variación en la concordancia de número en el español de contacto tiene el mismo comportamiento lingüístico en las dos comunidades, pero difiere en su distribución social. Los resultados de la regresión logística escalonada que se muestran en la Tabla 3.15 confirman esta hipótesis, pues la concordancia plural en la frase nominal del español en SAC y JV está determinado por el mismo factor lingüístico, la pérdida de consonantes en coda a final de palabra, y difieren en la distribución social del fenómeno. En SAC es relevante el género de los colaboradores: los hombres no favorecen la concordancia de número, pero en JV no se presenta este efecto.

Con respecto a la HIPÓTESIS 3.4, que propone que en estas comunidades existen diferentes sistemas de concordancia de número en español relacionados con el tipo de bilingüismo de los hablantes, los datos de la Tabla 3.15 muestran que la distribución de la

variación en la concordancia plural de la FN en función del tipo de bilingüismo se mantiene en las dos comunidades. Evaluaré en mayor profundidad esta hipótesis en la sección § 3.8 donde veremos si estos sistemas no sólo se diferencian en el porcentaje de casos de discordancia, sino en los factores lingüísticos que los organizan y determinan.

3.7.2 La concordancia de número en el nivel oracional del español de San Andrés Cuexcontitlán y Jiquipilco el Viejo

En Jiquipilco el Viejo el porcentaje de discordancias de número oracional es de 12.86% y en San Andrés Cuexcontitlán de 13.90%. La similitud de estos porcentajes contrasta con lo hallado en el nivel frástico, donde los porcentajes y la diferencia entre comunidades son mayores. El siguiente paso será realizar un análisis cuantitativo inferencial para saber si los factores lingüísticos y extralingüísticos que determinan la concordancia de número en el nivel oracional son los mismos en San Andrés Cuexcontitlán y Jiquipilco el Viejo.

Antes de llevar a cabo el análisis estadístico inferencial, realicé tablas de contingencia en las que crucé la variable dependiente (concordancia *versus* discordancia) con las variables lingüísticas y extralingüísticas descritas en §3.3.3 por comunidad. También calculé los estadísticos Chi-cuadrada de Pearson y V de Cramer (véase §2.1.3.1 y §2.1.3.2), cuyos resultados se encuentran en el Apéndice 3G. Asimismo, utilicé la medida de Goodman y Kruskal del paquete *The GoodmanKruskal* en la plataforma de R para cerciorarme que no existiera colinealidad entre las variables independientes (véase §2.1.3.3). Con base en los resultados (véase Apéndice 3H), redefiní los modelos para la concordancia oracional y realicé una regresión logística escalonada (véase §2.1.3.4) en la que tomé como variable dependiente la concordancia *versus* discordancia y como variables explicativas: la persona gramatical, la

animacidad, la especificidad, la categoría del elemento concordante, el tipo de bilingüismo y el sexo de los colaboradores.

En la Tabla 3.16 se presentan los resultados de la regresión logística escalonada. El factor más importante en las dos comunidades es el tipo de bilingüismo: los hablantes dominantes en español (BE, PO, ME) y, en menor proporción, los bilingües simétricos (BS) favorecen la concordancia de número, mientras que los hablantes dominantes en otomí (BO) presentan el mayor porcentaje de casos de discordancia, como se aprecia en el Apéndice 3G. El segundo factor fue la persona gramatical y en las dos comunidades son los sujetos de 3ª persona los que favorecen la discordancia, tal como lo estipulan las predicciones sobre la influencia del otomí (HIPÓTESIS 3.1, véase la Sección §3.6). San Andrés Cuexcontitlán y Jiquipilco el Viejo también coinciden en que los sujetos léxicos y morfológicos promueven los usos concordantes. Asimismo, presentan el mismo patrón de distribución según el sexo de los colaboradores, aunque con diferente nivel de importancia, pues son los hombres quienes favorecen la discordancia de número en las dos comunidades.

TABLA 3.16 *Resultados de la regresión logística escalonada sobre la variación en la concordancia de número en el nivel oracional por comunidad*

Jiquipilco el Viejo					
No	Variables	Estimado	Error Est.	Valor z	Pr(> z)
	Intersección	1.0796	0.7789	1.386	0.165714
1	Dominantes español	2.3422	0.3490	6.711	1.94e-11 ***
	Bilingües simétricos	1.3324	0.2785	4.784	1.72e-06 ***
2	3a Persona	-1.4706	0.4772	-3.082	0.00206 **
3	Hombres	-0.8911	0.3435	-2.594	0.00948 **
4	Sujeto morfológico	1.3835	0.5473	2.528	0.01147 *
	Sujeto léxico	0.6615	0.5436	1.217	0.22362
5	Ref. específica	0.5315	0.3051	1.742	0.08149 .
San Andrés Cuexcontitlán					
	Variables	Estimado	Error Est.	Valor z	Pr(> z)
	Intersección	0.8660	0.4453	1.945	0.0518 .
1	Dominantes español	3.8196	0.3347	11.413	< 2e-16 ***
	Bilingües simétricos	0.8603	0.2203	3.905	9.44e-05 ***
2	3a Persona	-2.7806	0.3107	-8.949	< 2e-16 ***
3	Sujeto morfológico	2.2829	0.4104	5.563	2.66e-08 ***
	Sujeto léxico	1.8688	0.4206	4.444	8.85e-06 ***
4	Hombres	-0.3818	0.2122	-1.799	0.0720 .

En resumen, en las dos comunidades inciden los mismos factores lingüísticos y extralingüísticos, con excepción de la especificidad que sólo fue seleccionada como significativa en Jiquipilco el Viejo. Estos resultados comprueban las predicciones en torno a la HIPÓTESIS 3.1, pues la variación en la concordancia de número en las dos comunidades responde a los mismos factores lingüísticos: la persona gramatical, la categoría del elemento concordante y, en el caso de Jiquipilco el Viejo, la especificidad de la referencia. Además, coinciden en la distribución social de este fenómeno.

En las dos comunidades se presenta un porcentaje similar de casos de discordancia plural en el nivel oracional. Los resultados de la regresión logística escalonada que se

muestran en la Tabla 3.16 comprueban la HIPÓTESIS 3.1 de este capítulo que señala que la variación en la concordancia de número en el español de contacto tiene el mismo comportamiento lingüístico en las dos comunidades, pero difiere en su distribución social.

La concordancia plural en el nivel oracional del español en SAC y JV está determinado por los mismos factores lingüísticos: la persona gramatical y la categoría del elemento concordante. En ambas comunidades, los sujetos de 3ª persona favorecen la discordancia, y los sujetos léxicos y morfológicos promueven los usos concordantes. Por otro lado, SAC y JV difieren en la distribución de la variación en la concordancia plural en función del género: en JV los hombres no favorecen la concordancia y en SAC este factor no es relevante.

Por último, la HIPÓTESIS 3.4 propone que en estas comunidades existen diferentes sistemas de concordancia de número en español relacionados con el tipo de bilingüismo de los hablantes, y los datos de la Tabla 3.16 muestran que la distribución de la variación en la concordancia oracional en función del tipo de bilingüismo se mantiene en las dos comunidades. Evaluaré en mayor profundidad esta hipótesis en la sección §3.8 donde veré si estos sistemas no sólo se diferencian en el porcentaje de casos de discordancia, sino en los factores lingüísticos que los organizan y determinan.

3.8 LA CONCORDANCIA DE NÚMERO EN LOS TRES TIPOS DE BILINGÜISMO: ¿UNA O DOS GRAMÁTICAS?

En los análisis hasta ahora realizados, el tipo de bilingüismo de los colaboradores ha sido un factor constante en la determinación de la concordancia de número en el nivel frástico y oracional. De manera general, he observado importantes diferencias cuantitativas en los usos concordantes y discordantes de los bilingües dominantes en otomí (BO), los bilingües

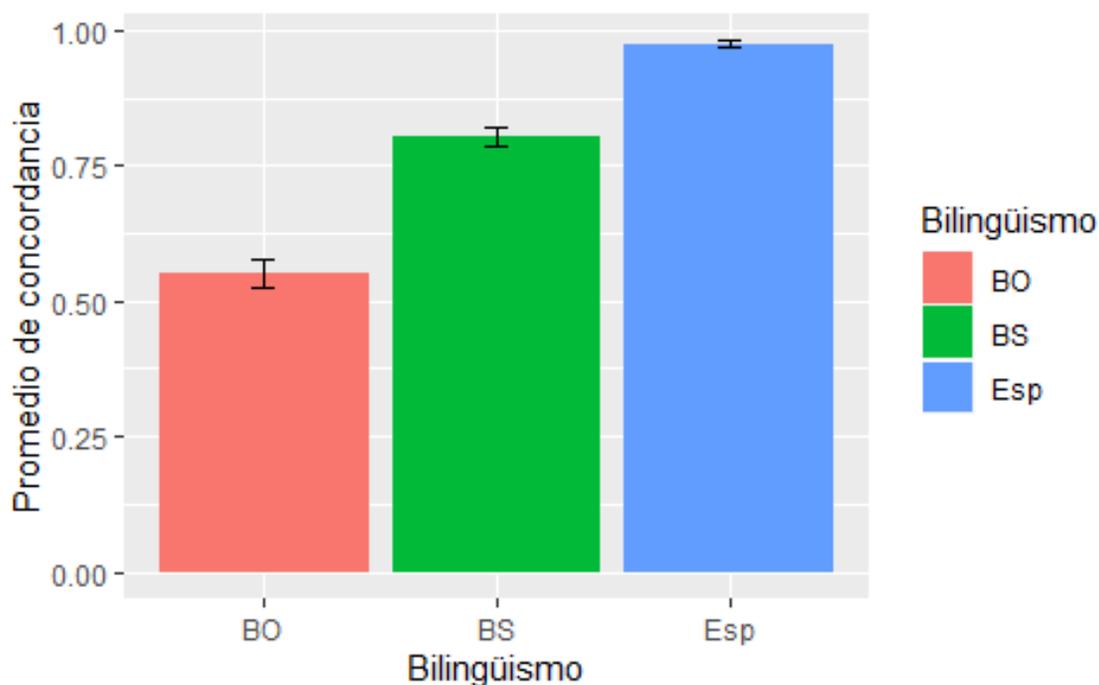
simétricos (BS) y los hablantes dominantes en español (BE, PO, ME). Por ello, en esta sección pondré a prueba la HIPÓTESIS 3.4, la cual postula que en el español de estas comunidades existen diferentes sistemas de concordancia de número relacionados con el tipo de bilingüismo, que no sólo difieren cuantitativamente, sino en los factores lingüísticos que los organizan y determinan, es decir, responden a gramáticas diferentes.

Para este propósito (OBJETIVO ESPECÍFICO 3.4), en primer lugar, determinaré si las diferencias en los porcentajes de usos concordantes / discordantes entre los grupos lingüísticos son significativas. Posteriormente, realizaré una serie de análisis inferenciales mediante los cuales buscaré saber si los factores lingüísticos que determinan la variación en la concordancia de número, en el nivel frástico y oracional, son los mismos o difieren según el grupo lingüístico.

Para determinar si los porcentajes de concordancia/discordancia de número son estadísticamente diferentes en los bilingües con mayor dominio del otomí (BO), los bilingües simétricos (BS) y los hablantes cuya lengua dominante es el español (BE, PO, ME) empleé un modelo lineal generalizado, en el que tomé como variable de respuesta la concordancia de número y como variables independientes el tipo de hablante. El resultado muestra que los tres grupos lingüísticos son estadísticamente diferentes con una significación de 0.001.

La Figura 3.4 muestra gráficamente estas diferencias. En el eje vertical (Y) se encuentra el promedio de la concordancia y en el horizontal (X) los tres grupos lingüísticos: bilingües dominantes en otomí (BO), bilingües simétricos (BS) y hablantes dominantes en español (BE, PO, ME). Además, se muestran los intervalos de confianza para el promedio de concordancia de cada grupo en líneas verticales sobre las barras y, como se puede apreciar, estas no se sobreponen entre sí, lo cual significa que son grupos cuantitativamente diferentes.

FIGURA 3.4 *Promedio de concordancia por grupo lingüístico*



Después de haber probado que las diferencias porcentuales entre los tres grupos lingüísticos son estadísticamente significativas, a continuación, procederé a realizar una serie de pruebas estadísticas para determinar cuáles son los factores lingüísticos que inciden en la alternancia de los tipos de marcación en el nivel frástico y oracional de cada grupo lingüístico.

3.8.1 *La concordancia de número en la frase nominal según el tipo de bilingüismo*

Para el análisis de la alternancia en los tipos de marcación de número en la frase nominal consideré los factores lingüísticos determinantes en el análisis de la Sección §3.5: el tipo de relación de concordancia, la categoría del elemento concordante, la animacidad del nombre, la especificidad de la referencia, la función sintáctica de la frase nominal y la pérdida de consonantes a final de palabra. Antes de realizar el análisis de inferencia multimodelo,

empleé la medida de Goodman y Kruskal del paquete *The GoodmanKruskal* en la plataforma de R para determinar si existían relaciones de colinealidad entre estos factores lingüísticos (§2.1.3.3), cuyos resultados se muestran en el Apéndice 3I.

Posteriormente, usé una aproximación de inferencia multimodelo para identificar el subconjunto de modelos con el mayor soporte empírico, de un total de 50 posibles modelos. Ordené los modelos del mejor al peor, basados en el criterio de información de Akaike. El conjunto de modelos con una diferencia AIC < 2 del mejor modelo fueron considerados con la misma fuerza y plausibilidad. Para evaluar la importancia de cada predictor usé el peso de Akaike (W_i). La importancia relativa de cada predictor se calculó a partir de la suma de los pesos de Akaike ($\sum W_i$) de cada modelo en el que el predictor aparece. Todos los modelos fueron construidos usando el paquete *glmulti* para la versión 3.01 de R (§2.1.3.5).

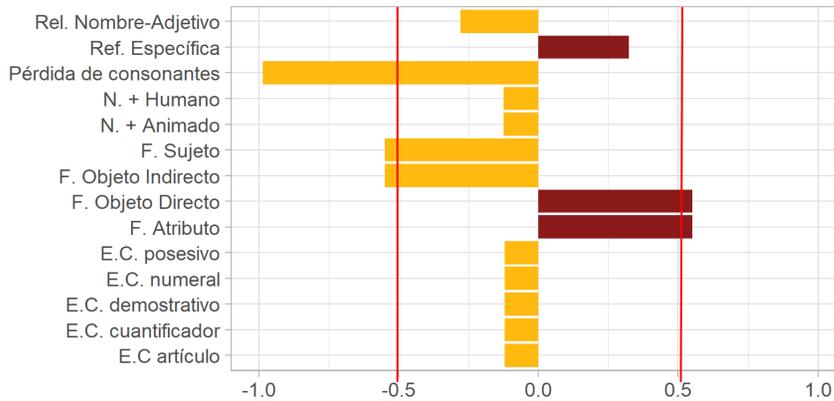
En la Figura 3.11 se puede apreciar que la pérdida de consonantes a final de palabra es el predictor más importante en el mejor modelo de los tres grupos de hablantes y está asociada negativamente con el tipo de marcación concordante (véase Apéndice 3J). En los bilingües con mayor dominio del otomí (BO), la función sintáctica de la frase nominal también actúa como un predictor importante: las frases nominales de sujeto y objeto indirecto favorecen el tipo de marcación discordante (P-A), y las de objeto directo y atributo el tipo de marcación concordante (P-P). Sin embargo, este mismo factor no forma parte de los modelos seleccionados para los bilingües simétricos y los hablantes dominantes en español, debido a que la suma de pesos de Akaike es baja. Por último, la especificidad, la animacidad, la categoría del elemento concordante y el tipo de relación de concordancia no fueron incluidos en los mejores modelos de los tres tipos de hablantes.

Los resultados de las pruebas de inferencia multimodelo sugieren la existencia de dos grupos diferenciados, el de los bilingües dominantes en otomí (BO), y el de los bilingües

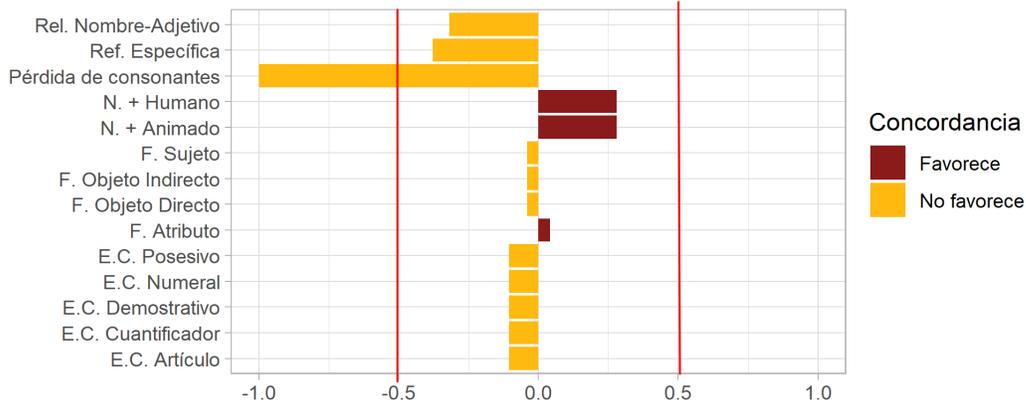
simétricos (BS) y los hablantes dominantes en español (Esp. = BE, PO, ME), cuya principal diferencia radica en la importancia del factor función sintáctica, presente en el modelo de los BO y ausente en los modelos de los BS y los Esp. Este punto es relevante, pues como veremos en el capítulo 4, la función sintáctica de las frases nominales es determinante en la marcación de número en la frase nominal del *n̄ható*.

FIGURA 3.5 *Resultados de la prueba inferencia multimodelo para la alternancia en los tipos de marcación de número en la FN por tipo de hablante*⁸¹

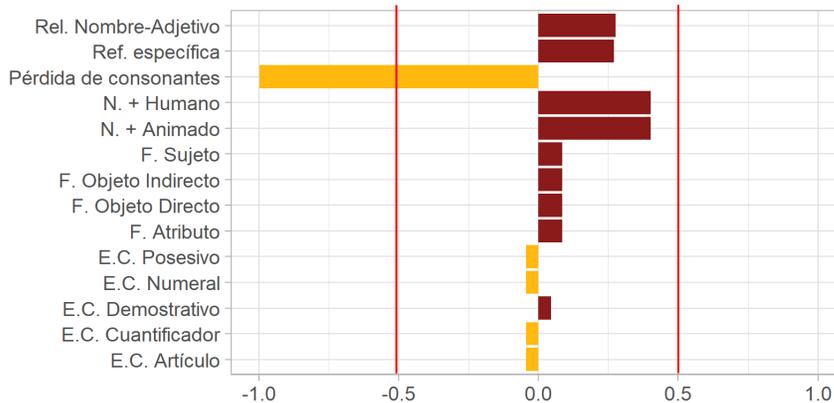
BO



BS



Esp



⁸¹ Las variables predictoras en el conjunto de modelos AICc <2 y el 95% del conjunto de modelos para la concordancia de número en los tres grupos de hablantes. La importancia de cada variable se muestra mediante la suma de los pesos de Akaike. El signo (\pm) de la estimación de los parámetros representa el efecto (positivo o negativo) de cada predictor sobre la variable de respuesta.

3.8.2 La concordancia de número en el nivel oracional según el tipo de bilingüismo

En el análisis de la variación en la concordancia de número oracional según el tipo de bilingüismo, empleé los factores lingüísticos determinantes en el análisis planteado en la Sección §3.6: la pérdida de consonantes a final de palabra, el tipo de relación de concordancia, la categoría del elemento concordante, la persona gramatical, la animacidad y la especificidad del sujeto. Antes de realizar la prueba inferencia multimodelo, utilicé la medida de Goodman y Kruskal del paquete *The GoodmanKruskal* en la plataforma de R para determinar si existen relaciones de colinealidad entre estos factores lingüísticos. Los resultados señalan que no hay colinealidad entre los factores lingüísticos para los tres grupos lingüísticos (véase Apéndice 3K).

Posteriormente, decidí emplear una aproximación de inferencia multimodelo para identificar el subconjunto de modelos con el mayor soporte empírico, de un total de 50 posibles modelos. Ordené los modelos posibles del mejor al peor, basados en el criterio de información de Akaike. El conjunto de modelos con una diferencia en AIC < 2 del mejor modelo fueron considerados con la misma fuerza y plausibilidad. Para evaluar la importancia de cada predictor usé el peso de Akaike (W_i). La importancia relativa de cada predictor se calculó a partir de la suma de los pesos de Akaike (sumatoria W_i) de cada modelo en el que el predictor aparece. Todos los modelos fueron construidos usando el paquete *glmulti* para la versión 3.01 de R.

Los mejores modelos para la variación en la concordancia de número en los bilingües dominantes en otomí (BO) comprende los siguientes predictores en el siguiente orden de importancia: 1) persona gramatical, 2) animacidad y 3) categoría del elemento concordante (véase Apéndice 3L). En primer lugar, hay una ruptura importante en el predictor persona

gramatical, pues los sujetos de 3ª persona favorecen la discordancia de número, mientras que los de 2ª persona favorecen la concordancia. Con respecto a la categoría del elemento concordante, los sujetos léxicos como los morfológicos promueven la concordancia de número. También se halla una importante escisión en el predictor animacidad, pues mientras los sujetos humanos están asociados positivamente con la concordancia, los sujetos animados lo están de manera negativa. Por último, la suma de pesos de Akaike de los factores especificidad, tipo de relación y pérdida de consonantes fue muy baja y no fueron considerados en los mejores modelos.

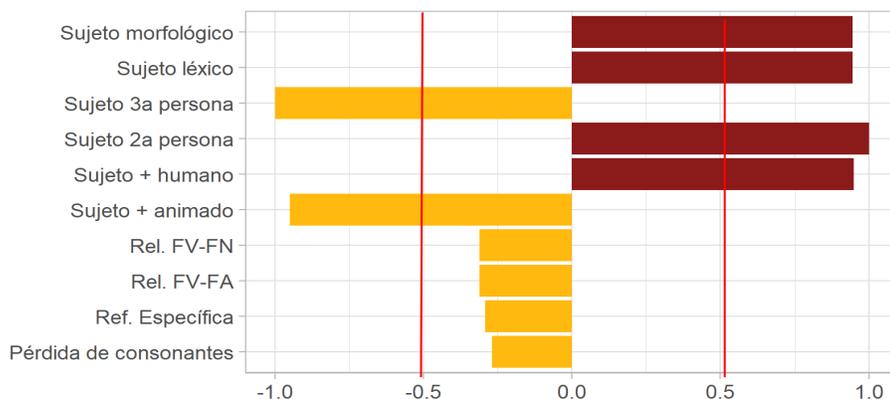
La persona gramatical es el predictor con mayor peso en los mejores modelos del grupo de bilingües simétricos. Sin embargo, no se observa la escisión entre la 3ª y 2ª persona que se encontró en el grupo de bilingües dominantes en otomí, ya que ambos tipos de sujetos guardan una relación negativa con la concordancia de número. El segundo predictor es la pérdida de consonantes y está asociado negativamente con la concordancia de número. El factor categoría del elemento concordante muestra el mismo comportamiento que en los BO, pues tanto los sujetos léxicos como los morfológicos favorecen la concordancia. El cuarto predictor muestra que la relación de concordancia FV-FN no favorece la concordancia, mientras que la FV-FA la promueve.

En el modelo de los bilingües con mayor dominio del español, la pérdida de consonantes es el predictor con mayor peso y está asociada negativamente con la concordancia de número. El segundo factor del modelo es la animacidad: los sujetos animados y humanos favorecen la discordancia de número; mientras que la especificidad favorece la concordancia. Con respecto al cuarto predictor con más peso, se advierte que las relaciones FV-FN no favorecen la concordancia mientras que las relaciones atributivas la promueven. A diferencia de los dos grupos bilingües (BO, BS), en el modelo de los hablantes

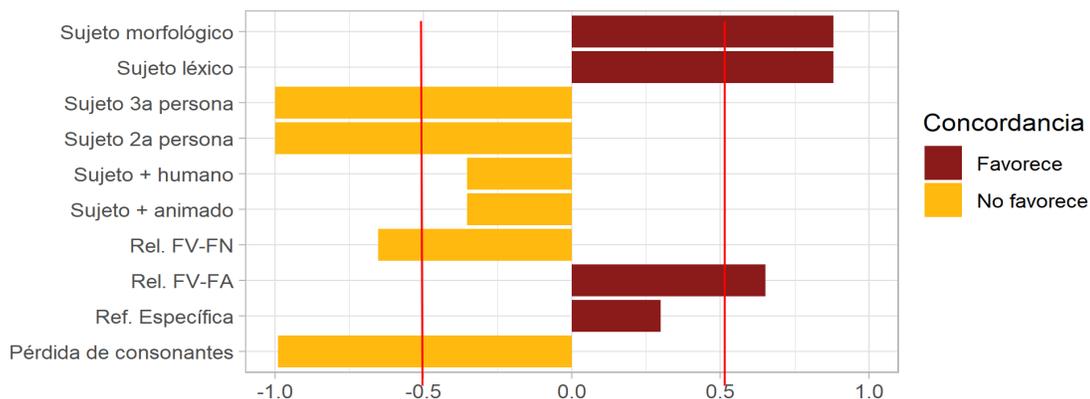
dominantes en español, la persona gramatical es el predictor de menor peso y muestran un comportamiento similar al de los bilingües simétricos.

FIGURA 3.6 *Resultados de la prueba inferencia multimodelo para la alternancia en los tipos de marcación de número en el nivel oracional por tipo de hablante*⁸²

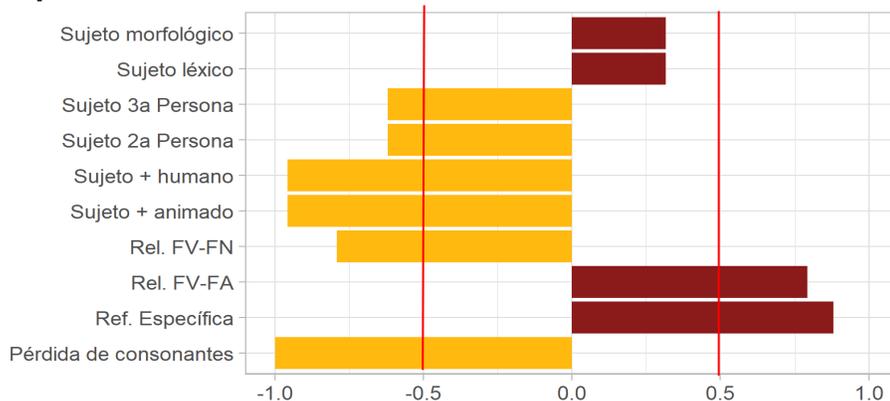
BO



BS



Esp



⁸² Las variables predictoras en el conjunto de modelos AICc <2 y el 95% del conjunto de modelos para la concordancia de número en los tres grupos de hablantes. La importancia de cada variable se muestra mediante la suma de los pesos de Akaike. El signo (\pm) de la estimación de los parámetros representa el efecto (positivo o negativo) de cada predictor sobre la variable de respuesta.

Estos resultados muestran la existencia de dos grupos lingüísticos nítidamente diferenciados, los bilingües dominantes en otomí (BO) y los hablantes dominantes en español (Esp = BE, PO, ME), pues mientras la persona gramatical del sujeto es el predictor más importante en los bilingües dominantes en otomí, es el de menor importancia en los hablantes dominantes en español. La categoría del elemento concordante es el tercer factor en el modelo de los bilingües dominantes en otomí, pero en los hablantes dominantes en español, la suma de pesos de Akaike de este predictor es muy baja y no forma parte del mejor modelo. Aunado a esto, en los hablantes dominantes en español, la pérdida de consonantes, la especificidad y el tipo de relación están entre los cuatro primeros predictores, mientras que en los bilingües dominantes en otomí la suma de pesos de Akaike es muy baja y no figuran en los dos mejores modelos de este grupo. A pesar de que en los dos grupos la animacidad es el segundo mejor predictor, difieren en su efecto.

El modelo de los bilingües simétricos es interesante pues presenta rasgos similares a los bilingües dominantes en otomí y a los hablantes dominantes en español, como una especie de grupo de transición. Los bilingües simétricos se parecen a los bilingües dominantes en otomí en que el predictor de mayor importancia de su modelo, la persona gramatical. Así como en la importancia de la categoría de los elementos concordantes y en la importancia de la especificidad. Sin embargo, difieren con respecto al comportamiento de la animacidad, que es relevante en los BO, pero no en los BS. Por otro lado, los bilingües simétricos muestran el mismo comportamiento que los hablantes dominantes en español con respecto al peso de los predictores tipo de relación y pérdida de consonantes.

Los resultados de los apartados §3.7.1 y §3.7.2 comprueban la HIPÓTESIS 3.4 de este capítulo, que señala que en SAC y JV existen diferentes sistemas de concordancia de número en español, relacionados con el tipo de bilingüismo de los hablantes. Estos sistemas no sólo

se diferencian en el porcentaje de casos de discordancia, sino en los factores lingüísticos que los organizan y determinan, es decir, que responden a gramáticas diferentes.

En el caso de la concordancia en la frase nominal, encuentro, por un lado, el sistema empleado por los bilingües con mayor dominio del otomí (BO), determinado por la pérdida de consonantes en coda a final de palabra y la función sintáctica de la frase nominal, ambos factores relacionados con la marcación de número en la lengua otomí. Por otro lado, está el sistema de los bilingües simétricos (BS) y los hablantes dominantes del español (Esp.= BE, PO, ME), cuyo sistema sólo se encuentra definido por la pérdida de consonantes.

Con respecto a la concordancia en el nivel oracional, hallo dos sistemas nítidamente diferenciados, el de los bilingües dominantes en otomí (BO) y el de los hablantes dominantes en español (Esp. = BE, PO, ME). El sistema del primer grupo está determinado por los factores de persona gramatical, animacidad y categoría del elemento concordante. Por otro lado, el sistema de concordancia de los hablantes dominantes en español (Esp.) está organizado por los factores siguientes: pérdida de consonantes, animacidad y especificidad. El sistema de los bilingües simétricos, por su parte, presenta rasgos similares al de los bilingües dominantes en otomí y al de los hablantes dominantes en español, como una especie de grupo de transición.

3.9 CONCLUSIONES DEL TERCER CAPÍTULO

Al inicio de este capítulo propuse cuatro hipótesis. La primera indica que la variación en la concordancia de número es resultado del contacto que el español mantiene con el otomí, de modo que este fenómeno responde a estrategias gramaticales relacionadas con los procesos de la lengua otomí. Por ello, aunque la concordancia plural se exprese con los recursos morfosintácticos del español, la aparición o no de estas marcas estará determinada por

alguno(s) de los condicionantes lingüísticos que rigen la expresión gramatical del número en otomí (HIPÓTESIS 3.1). La HIPÓTESIS 3.2 señala que la variación en la concordancia de número en el español de contacto se comporta de manera diferente en el nivel frástico y oracional, debido a que la concordancia de número en otomí se organiza de forma distinta en los dos niveles.

A lo largo de los apartados §3.4, §3.5 y §3.6, pudimos comprobar las dos primeras hipótesis. En el apartado §3.4 planteé un análisis general en el que consideré la variación en la concordancia de número en la frase nominal y en el nivel oracional como un solo fenómeno variable. Aunque estos primeros resultados mostraron que este fenómeno está asociado con algunos de los factores que organizan la marcación de número en otomí, la especificidad y la animacidad, también encontré una fuerte diferenciación relacionada con el dominio sintáctico en el que ocurre la relación de concordancia. Por ello, decidí plantear un análisis de la concordancia de número por nivel (frástico y oracional), que permitiera aceptar o rechazar la HIPÓTESIS 3.2 y mejorar los resultados de la HIPÓTESIS 3.1 pues, si pretendía determinar la influencia que el otomí tiene en el español, y la concordancia de número en la lengua indígena funciona de manera diferente en el nivel frástico y oracional, resultaba pertinente plantear un análisis adecuado en cada nivel y de acuerdo con sus propios términos.

En los apartados §3.5 y §3.6 llevé a cabo el análisis de la variación en la concordancia de número en la frase nominal de español. Los resultados afirman la HIPÓTESIS 3.1 de este capítulo. La variación en la concordancia plural de la frase nominal del español hablado en San Andrés Cuexcontitlán y Jiquipilco el Viejo sí está relacionada con el contacto con la lengua otomí. En el caso de la variación en la concordancia plural de la frase nominal, los bilingües encuentran similitudes y diferencias en las propiedades combinatorias de las marcas de número del otomí *ra / ya* y el español *-(e)s* y a partir de eso establecen una

equivalencia entre ellas. Esta equivalencia combinacional se puede apreciar en el uso del patrón de marcación alterna (Presencia + Ausencia de la marca: *La-s casa*), que sigue el esquema del *ñható*, en el que la marca de número sólo se hospeda en el dependiente (Presencia + Ausencia de la marca: *Ku=ya nguu* ‘Las casas’). Asimismo, se presenta un efecto del factor función sintáctica, que también es relevante en la marcación de número del otomí (véase §4.3). Aunado a esto, la variación en la concordancia plural en la frase nominal está fuertemente relacionada con la pérdida de consonantes a final de palabra, un fenómeno fónico que también tiene su origen en el contacto con el otomí.

Los resultados del análisis de la concordancia en el nivel oracional presentados en la sección §3.6 afirman la HIPÓTESIS 3.1 de este capítulo. La variación en la concordancia plural en el nivel oracional del español hablado en San Andrés Cuexcontitlán y Jiquipilco el Viejo sí está relacionada con el contacto con la lengua otomí. En el caso de la variación en la concordancia plural en el nivel oracional, los hablantes perciben similitudes y diferencias entre las relaciones de concordancia del otomí y el español, y, a partir de eso, establecen relaciones de equivalencia entre ellas. Esta equivalencia se puede apreciar en los factores que condicionan la concordancia plural en el nivel oracional de esta variedad de español, a saber: la persona gramatical, la animacidad y el tipo de relación, que son algunos de los factores que determinan la marcación de número del otomí en el nivel oracional. Además, el efecto de estos es el mismo en las dos lenguas: los sujetos de 3ª persona y no humanos promueven la discordancia de número y los sujetos de 1ª y 2ª persona y +humanos son los que favorecen la concordancia.

Los resultados obtenidos en la sección §3.5 y §3.6 confirman la HIPÓTESIS 3.2, pues, aunque en ambos casos los bilingües establecen equivalencias sobre las propiedades combinatorias de las marcas de número y la concordancia plural del otomí y el español para

establecer la relación de concordancia plural en su español, cada nivel tiene especificaciones diferentes, debido a que la concordancia de número en otomí se organiza de forma distinta en los dos niveles.

Los resultados mostrados en el apartado §3.7 comprueban la HIPÓTESIS 3.3 de este capítulo, que señala que la variación en la concordancia de número tiene el mismo comportamiento lingüístico en San Andrés Cuexcontitlán y Jiquipilco el Viejo, pero difiere en su distribución social en cada comunidad. La variación en la concordancia plural de las frases nominales (§véase 3.7.1) está estrechamente relacionada con la pérdida de consonantes en coda a final de palabra, tanto en SAC como en JV. Asimismo, en ambas comunidades el segundo factor de mayor importancia es el tipo de bilingüismo del emisor. La única diferencia radica en un factor social, el género de los hablantes. Este factor sólo resulta relevante en la comunidad de San Andrés Cuexcontitlán, en donde los hombres son quienes promueven el uso del patrón de marcación que sigue el esquema de la lengua otomí (*las casa*). Por otro lado, la concordancia plural en el nivel oracional está determinada por la persona gramatical y el tipo de sujeto en las dos comunidades (véase §3.7.2). Además, en ambos casos, el tipo de bilingüismo del emisor incide en este fenómeno variable. No obstante, de nueva cuenta, encontramos diferencias relacionadas con el género de los colaboradores. Este factor sólo es relevante en JV, donde son los hombres quienes favorecen la discordancia de número.

Por último, los resultados del apartado §3.8 comprueban la HIPÓTESIS 3.4 de este capítulo, que señala que en SAC y JV existen diferentes sistemas de concordancia de número en español, relacionados con el tipo de bilingüismo de los hablantes. Estos sistemas no sólo se diferencian en el porcentaje de casos de discordancia, sino en los factores lingüísticos que los organizan y determinan, es decir, responden a gramáticas diferentes.

En el caso de la concordancia en la frase nominal, encontré, por un lado, el sistema empleado por los bilingües con mayor dominio del otomí (BO), determinado por la pérdida de consonantes en coda a final de palabra y la función sintáctica de la frase nominal, ambos factores relacionados con la marcación de número en la lengua otomí. Por otro lado, está el sistema de los bilingües simétricos (BS) y los hablantes dominantes del español (Esp.), cuyo sistema sólo se encuentra definido por la pérdida de consonantes.

Con respecto a la concordancia en el nivel oracional, hallé dos sistemas nítidamente diferenciados, el de los bilingües dominantes en otomí (BO) y el de los hablantes dominantes en español (Esp). El sistema del primer grupo está determinado por los factores de persona gramatical, animacidad y categoría del elemento concordante. Por otro lado, el sistema de concordancia de los hablantes dominantes en español (Esp.) está organizado por los siguientes factores: pérdida de consonantes, animacidad y especificidad. El sistema de los bilingües simétricos, por su parte, presenta rasgos similares al de los bilingües dominantes en otomí y al de los hablantes dominantes en español, como una especie de grupo de transición.

La concordancia plural (frástica y oracional) del español en contacto con otomí muestra una estratificación asociada al grado de bilingüismo y la edad de los colaboradores, debido al desplazamiento en el que se encuentra la lengua otomí en SAC y JV. Las diferencias cuantitativas y gramaticales entre los dos sistemas de concordancia (BO-Generación 4 *versus* Esp.-Generación 1 y 2) sugieren que la innovación sobre la marcación de número surge en el grupo BO-Generación 4 y se expande ligeramente en los BS-Generación 3; sin embargo, este fenómeno no se extiende a los grupos con mayor dominio o monolingües en español Esp.-Generación 1 y 2, más bien entra en un proceso de retracción.

CAPÍTULO 4. EL OTOMÍ DE CONTACTO: EL CASO DE LA MARCACIÓN DE NÚMERO

4.1 INTRODUCCIÓN

En este capítulo revisaré el efecto que el contacto ha tenido en la lengua otomí a partir de la categoría de número, un área propicia para la manifestación de procesos de variación y cambio en situaciones de contacto (Palacios Alcaine, 2011); específicamente, se centrará en el análisis de la marcación de número en los nombres del otomí. Como se verá más adelante, el efecto del español en el léxico *ñható* ha sido tal, que el porcentaje de préstamos nominales constituye cerca de la mitad de los nombres del otomí. Por esta razón, el análisis de la marcación de número nominal en otomí se planteará desde el origen histórico de los lexemas y se hará una distinción entre el léxico patrimonial y los préstamos del español.

La PREGUNTA GENERAL de este capítulo es si el contacto con el español ha incidido en la marcación de número nominal en otomí. De manera particular, se pretende responder las siguientes preguntas:

4.1 ¿La influencia del español ha sido directa, mediante la importación de la morfología de número del español, o indirecta, a partir de la importación de un patrón o de los condicionamientos lingüísticos del español?

4.2 ¿La influencia del español depende del origen histórico de los nombres (patrimoniales *versus* préstamos)?

4.3 ¿La influencia del español depende del nivel de bilingüismo de los hablantes?

EL OBJETIVO GENERAL de este capítulo será determinar si el español ha influido en la marcación de número del otomí. Para ello, el PRIMER OBJETIVO ESPECÍFICO 4.1 será establecer

cuáles son los recursos morfosintácticos que los bilingües otomí-español emplean para marcar el número en los nombres patrimoniales del otomí y en los sustantivos del español que han ingresado a esta lengua. Esto permitirá saber si ha habido una importación directa de la morfología de número del español. Asimismo, se precisará cuáles son los factores lingüísticos y sociolingüísticos que determinan la marcación de número en los nombres patrimoniales y en los sustantivos del español que han ingresado a esta lengua. Esta información permitirá conocer si ha habido una importación de los condicionamientos lingüísticos del español. El SEGUNDO OBJETIVO ESPECÍFICO 4.2 es determinar si la influencia del español depende del origen histórico de los nombres (patrimoniales *versus* préstamo). Asimismo, precisaré si la influencia del español depende del nivel de bilingüismo de los hablantes (TERCER OBJETIVO ESPECÍFICO 4.3).

La HIPÓTESIS GENERAL de este capítulo es que el español ha influido en la marcación de número nominal del otomí. La PRIMERA HIPÓTESIS ESPECÍFICA 4.1 es que el español no ha influido de manera directa en la marcación de número del otomí. La SEGUNDA HIPÓTESIS ESPECÍFICA 4.2 plantea que el español ha influido de manera indirecta en la marcación de número nominal en otomí. Por ello, aunque la marcación de número se expresa con los recursos morfosintácticos del otomí, la aparición o no de estas marcas estará determinada por los condicionantes lingüísticos que rigen la marcación de número en español.

Además, la HIPÓTESIS 4.2 propone que la fuerza del efecto del español depende del origen de histórico de los nombres. Dicho esto, espero encontrar un mayor efecto de los condicionantes lingüísticos que rigen la marcación de número en español en los préstamos de esta lengua y un menor efecto en los nombres patrimoniales. La HIPÓTESIS 4.3, por su parte, plantea que la fuerza del efecto del español está asociada con el nivel de bilingüismo de los hablantes. Al respecto, hipotetizo que el efecto de los condicionantes lingüísticos que

rigen la marcación de número en español es mayor en los hablantes dominantes de esa lengua y menor en los bilingües con mayor dominio del otomí.

Este capítulo consta de ocho secciones. En §4.2 hago un repaso general de las principales investigaciones sobre el préstamo léxico y su relación con el cambio de código. También reseño algunos trabajos en torno al proceso de relexificación en el náhuatl y a la marcación de número en los préstamos del español en maya yucateco. Por último, reviso los estudios sobre el préstamo léxico y gramatical en la lengua otomí. En la sección §4.3, brindo una descripción general del corpus empleado en este capítulo. Después, en §4.4, describo cuáles son los recursos morfosintácticos que los bilingües emplean para marcar el número en los nombres de las narraciones otomíes. En el apartado §4.5, propongo un análisis de la marcación de número según el origen histórico de los nombres: patrimonial *versus* préstamo. Posteriormente, en §4.6, planteo un análisis de la marcación de número nominal de acuerdo con el tipo de bilingüismo de los hablantes. En el apartado §4.7, proporciono un análisis a partir de un índice léxico de integración. Por último, en §4.8, presento la discusión y las conclusiones de este capítulo.

4.2 ANTECEDENTES

4.2.1 El uso del término préstamo en los estudios de contacto

En este apartado reviso el uso del término *préstamo* en algunas de las obras más importantes sobre el contacto lingüístico. Posteriormente, preciso el concepto de préstamo en el que me baso para el desarrollo de este capítulo. Uno de los primeros autores en proponer una metodología para el estudio del préstamo fue Haugen (1950: 212), quien definió el término

préstamo como “the attempted reproduction in one language of patterns previously found in another”.

El resultado del proceso de préstamo puede tener diferentes grados de similitud con el patrón de origen o modelo. Con base en el grado de semejanza, Haugen (1950: 212) establece dos tipos de reproducción: importación y sustitución. Se trata de una importación cuando el préstamo es muy parecido al modelo y un hablante nativo lo reconoce como propio. Por el contrario, es un proceso de sustitución si el resultado difiere del modelo y es reemplazado por un patrón de la lengua meta.

Con base en los dos tipos de reproducción,⁸³ distingue tres categorías: *loanwords*, *loanblends* y *loanshifts*. (1) *Loanwords* son una importación morfé mica sin sustitución, que pueden tener distintos grados de adaptación fónica (nula, parcial o completa); por ejemplo, *olotl* (náhuatl) > *olote* (esp.) o *vaca* (esp.) > *baga* (otomí). (2) *Loanblends* son un tipo de préstamo que comprende tanto procesos de importación como de sustitución morfé mica. (3) *Loanshift* muestran una sustitución morfé mica sin importación, es decir, implican cambios funcionales en el uso de los morfemas y palabras nativas sin que se lleve a cabo algún cambio fonológico o gramatical; ejemplo de esto es *pocket* ‘bolsillo’ (ing.) > *bocka* ‘bolsillo’ (ing.) + *buch* ‘libro’ (alemán) = *bockabuch* (alemán). Este grupo comprende fenómenos como la traducción de préstamos y el préstamo semántico (Haugen, 1950: 219); por ejemplo, *Mariage de convenance* (fr.) > *Marriage of convenience* (ing.) o *Basketball* (ing.) > *Baloncesto* (esp.).

Posteriormente, en *Languages in Contact*, Weinreich (1953: 1) define el término de *interferencia* como “Those instances of deviation from the norms of either language which occur in the speech of bilinguals as a result of their familiarity with more than one language”

⁸³ Haugen da por sentado que se ha producido una importación semántica en los tres tipos de préstamo. Por otro lado, el término *morfema* no incluye las modificaciones de tipo flexivo (Haugen, 1950: 215).

y distingue tres tipos, de acuerdo con el nivel en el que ocurren: interferencia fónica, gramatical y léxica (véase §1.1.1).

Con respecto a la interferencia léxica, señala que el vocabulario de una lengua puede inferir en el lexicón de diversas formas. La primera posibilidad es que los morfemas de una lengua A se transfieran a una lengua B. La segunda opción es que los morfemas de B comiencen a tener nuevas funciones a partir de las características de los morfemas de A con los que guardan cierta correspondencia (Weinreich, 1953: 47).

Además, las interferencias léxicas pueden ser simples o compuestas. La primera categoría comprende tres fenómenos: i) la transferencia de la secuencia fonémica de una lengua a otra; ii) la extensión del uso de una palabra con base en el modelo de otra lengua; iii) el tercer tipo de interferencia léxica ocurre cuando la forma fónica de un signo cambia a partir de un cognado de la lengua con la que está en contacto, sin afectar su contenido (Weinreich, 1953: 47, 48, 50).

Por otro lado, en *Language Contact, Creolization, and Genetic Linguistic*, Thomason y Kaufman (1988) distinguen dos resultados del contacto lingüístico: el préstamo y la interferencia, a partir de dos criterios fundamentales, la ocurrencia o no de un aprendizaje imperfecto de la segunda lengua y la duración del proceso. El *préstamo* es “the incorporation of foreign features into a group's native language by speakers of that language: the native language is maintained but is changed by the addition of the incorporated features” (37). Este tipo de proceso ocurre a lo largo de cientos de años y suele iniciar con el ingreso de elementos léxicos (41). A diferencia del préstamo, el término *interferencia* refiere:

a subtype of interference that results from imperfect learning during a process of language shift. That is, in this kind of interference a group of speakers shifting to a target language fails to learn the target language (TL) perfectly. The errors made by members of the shifting group in speaking the TL then spread to the TL as a whole when they are imitated by original speakers of that language (Thomason & Kaufman, 1988: 38)

Este tipo de proceso puede tomar sólo una generación y no comienza con el ingreso de vocabulario sino con cambios fónicos, sintácticos e incluso morfológicos (39).

Por último, Heine y Kuteva (2005: 2) indican que la influencia de una lengua a otra se manifiesta a través de la transferencia de una forma, un significado, una unidad de forma-significado, una relación sintáctica o una combinación de éstos. En la Tabla 4.1 se presenta los tipos de transferencia planteados por estos autores y sus correspondencias con la clasificación de Weinreich (1953).

TABLA 4.1 *Tipos de transferencia de acuerdo con Heine y Kuteva (2005) y Weinreich (1953)*

Heine y Kuteva (2005)	Weinreich (1953)
A. Forma: sonidos o una combinación de sonidos.	
B. Significados: gramaticales o funciones.	Interferencia de funciones o significados de formas gramaticales.
C. Unidades de forma-significado o una combinación de éstas.	Interferencia de morfemas.
D. Relaciones sintácticas.	Interferencia de relaciones gramaticales: orden de palabras.
E. Cualquier combinación de (A) a (D)	

Con base en esta distinción, los autores precisan que sólo consideran como *préstamo* aquellos fenómenos que entran dentro de la categoría A y C, es decir, aquellos casos en los que la transferencia implica una forma fonética o una unidad forma-significado (Heine & Kuteva, 2005: 6).

Como se ha podido observar, el término *préstamo* ha sido empleado de diferente manera en los trabajos de contacto aquí descritos. De acuerdo con la clasificación de Weinreich (1953), el fenómeno que es objeto de estudio en este capítulo corresponde a la categoría de interferencia léxica simple, en la que se transfiere una secuencia fonémica de una lengua a otra. Sin embargo, el término interferencia presenta algunos problemas. En principio, hablar de error o desviación de una norma estándar debido a una deficiencia en la adquisición de una segunda lengua, muestra una postura prescriptiva que tiene implicaciones teóricas y prácticas.

Con respecto al primer punto, dar por hecho que todas las interferencias o estructuras alternas son resultado de una adquisición deficiente de una L2, impide ver la complejidad y multicausalidad de estos fenómenos, que pueden tener una motivación cognitiva, semántica, pragmática o, incluso, interna al sistema lingüístico. Por otro lado, en términos prácticos, esta visión prescriptiva ha llevado a la estigmatización de grupos minoritarios, como los indígenas en Hispanoamérica, por considerar su español como una variedad subestándar. Este hecho tiene consecuencias en las actitudes que los hablantes tienen hacia el español, las lenguas indígenas, el bilingüismo y, en consecuencia, el devenir de las lenguas indígenas. Por ello, en este capítulo no emplearé el término *interferencia léxica*.

Por otro lado, la distinción entre préstamo e interferencia que Thomason y Kaufman (1988) plantean se basa en la adquisición (im)perfecta de una segunda lengua y la duración del proceso. Sin embargo, en la situación de contacto entre el otomí y el español de SAC y JV es difícil distinguir entre ambos fenómenos, sobre todo en el ámbito léxico, donde el uso de una palabra del español puede estar asociado tanto a un proceso de préstamo como a uno de interferencia, en términos de estos autores. Por ejemplo, en el vocabulario otomí se encuentra una gran cantidad de nombres del español pertenecientes al ámbito religioso, que

ingresaron alrededor del siglo XVII, y que bien podrían ser considerados como préstamos. En contraste, también se hallan elementos léxicos modernos, cuyas características morfofonológicas podrían catalogarlos como casos de interferencia. Debido a que en este estudio considero todos los elementos léxicos del español sin hacer una distinción con respecto a su temporalidad o al tipo de adquisición de los hablantes (perfecto *versus* imperfecto), decidí no partir de los conceptos de *préstamo* e *interferencia* según Thomason y Kaufman (1988).

En este capítulo parto de un sentido amplio de préstamo, entendido como “the attempted relation in one language of patterns previously found in another” (Haugen, 1950: 212), pues la amplitud de este concepto permite explicar una mayor cantidad de fenómenos lingüísticos que tienen lugar en situaciones de contacto; no obstante, sólo me centraré en el préstamo léxico y no consideraré fenómenos que le son cercanos, como el cambio de código y los calcos semánticos.

4.2.2 *El cambio de código monoléxico y su relación con el préstamo léxico*

Como veremos más adelante, en el corpus se presenta otro fenómeno cercano al préstamo léxico, el cambio de código. Poplack (1980: 583) define este término como “the alternation of two languages within a single discourse, sentence or constituent” y establece una serie de criterios para distinguir entre estos dos fenómenos con base en el nivel de integración.

Esta autora plantea cuatro tipos de construcción. El primero es un constituyente integrado fonológica, morfológica y sintácticamente a la lengua base, razón por lo que no califica como un caso de cambio de código. El segundo tipo sigue la fonología y morfología de la lengua base, pero infringe su sintaxis. El tercero cumple con las reglas fonológicas de la lengua base, pero no su morfología y sintaxis. El tipo cuatro no presenta ningún tipo de

integración a los patrones fonológicos, morfológicos y sintácticos de la lengua base. De este modo, sólo los tres últimos tipos califican como casos de cambio de código (Poplack, 1980: 584).

TABLA 4.2 *Identificación del cambio de código de acuerdo con el tipo de integración a la lengua base (adaptado de Poplack 1980: 584)*

Tipo	Niveles de integración			Cambio de Código
	Fonológico	Morfológico	Sintáctico	
1	Sí	Sí	Sí	No
2	No	No	Sí	Sí
3	Sí	No	No	Sí
4	No	No	No	Sí

En estudios posteriores se mostró que estos criterios son poco efectivos para diferenciar un préstamo de un caso de cambio de código, especialmente cuando se trata de un elemento monoléxico (Myers-Scotton, 1993; Sankoff & Poplack, 1981). Por ello, en un trabajo posterior, Poplack y Meechan (1998) proponen un método comparativo para identificar casos de préstamo y cambio de código que se basa en el análisis cuantitativo de los patrones de comportamiento del habla.⁸⁴

En principio, los autores plantean una serie de supuestos variacionistas aplicados al contacto lingüístico y con base en estos proponen su hipótesis sobre la integración del préstamo (Poplack & Meechan, 1998: 130).

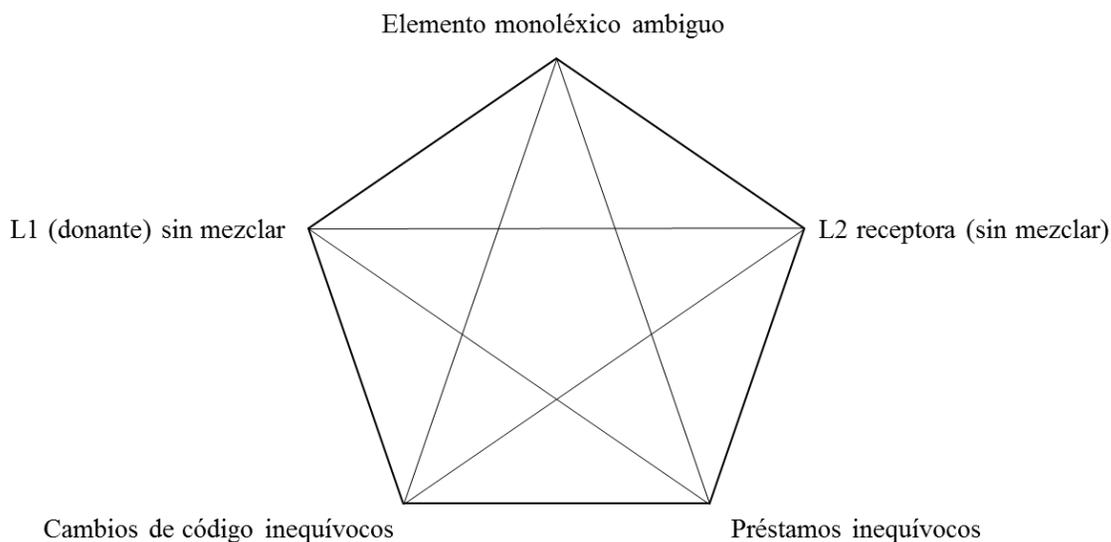
- I) Los condicionamientos (y tasas) de variación lingüística son específicas de cada lengua y pueden ser útiles para determinar si un elemento pertenece a una lengua u otra.

⁸⁴ Poplack y Meechan (1998: 129) definen el término *patrón* como “a series of parallel occurrences (established according to structural and/or functional criteria) occurring at a non-negligible rate in a corpus of language use”.

- II) Si el elemento de la lengua donante (LD) empleado en el discurso de la lengua receptora (LR) tiene las mismas restricciones que sus contrapartes en la LR y difiere de los elementos de la LD en el discurso de la LD. Entonces, la gramática que restringe al elemento de la LD en el discurso de la LR es la LR y por lo tanto se trata de un préstamo.
- III) Si el elemento de la lengua donante (LD) empleado en el discurso de la lengua receptora (LR) presenta la misma jerarquía de restricciones que otros elementos de la LD y al mismo tiempo difiere de sus contrapartes en la LR en el discurso de la LR. Entonces se puede decir que la gramática que restringe al elemento de la LD es la LD y por lo tanto se trata de un caso de cambio de código.

Posteriormente, sugieren una serie de comparaciones entre el presunto préstamo o cambio de código y los patrones o restricciones de las variantes monolingües de la lengua donante y de la lengua receptora, así como con los préstamos establecidos y los casos de cambios de código multiléxico inequívocos en el corpus, tal como se esquematiza a continuación.

FIGURA 4.1 *Comparaciones en el discurso bilingüe* (adaptado de Poplack y Meechan, 1998: 131)



De acuerdo con esta serie de comparaciones y con los supuestos I-III, un elemento monoléxico es un préstamo si tiene patrones como sus contrapartes en el idioma receptor (sin mezclar), mientras que al mismo tiempo difieren de los patrones de la lengua donante y / o de los cambios de código inequívocos. En contraste, se trata de un caso de cambio de código si el elemento mono léxico se comporta como sus contrapartes en la lengua donante (sin mezclar) y de forma simultánea difiere del comportamiento de los miembros de la lengua receptora (sin mezclar) (Poplack & Meechan, 1998: 131).

Por otro lado, Gómez Rendón, (2008: 30) también emplea varios criterios para discernir entre casos de préstamo y cambio de código en su análisis sobre el préstamo léxico y gramatical en el quechua, el guaraní y el otomí:

- a) La integración morfológica y sintáctica en la lengua receptora (inflexión, derivación, orden de palabras).
- b) La adaptación fonológica a la lengua receptora.

c) La resemantización del elemento extranjero en el idioma receptor.

d) La frecuencia de aparición del elemento extranjero en un hablante y entre hablantes.

e) La frecuencia de aparición de clases de palabras en un hablante y entre hablantes.

En la conformación del corpus no hice una distinción entre préstamo y cambio de código monoléxico e integré todos los lexemas nominales del español, independientemente de su nivel de adaptación morfofonológica, incluso aquellos casos que podrían ser candidatos a cambio de código mono léxico. Sin embargo, en la sección §4.7 planteo un análisis de la marcación de número a partir de un índice de integración léxica en el que considero algunos de los criterios empleados para discernir casos de préstamo y de cambio de código, a saber: la adaptación fonológica, la frecuencia de aparición del elemento léxico en un hablante y entre hablantes.

4.2.3 Las escalas de préstamo léxico y gramatical

Dentro de los estudios del préstamo en situaciones de contacto, se han postulado distintas jerarquías y escalas que establecen la (im)posibilidad de que un elemento sea transferido o prestado, en función de algunas restricciones estructurales y gramaticales. A pesar de que ambas gradaciones provienen de la lógica proposicional y dependen de mecanismos deductivos formales, también difieren con respecto a las categorías que comprenden, su aplicabilidad y dimensión temporal (Gómez Rendón, 2008: 59-60).

La jerarquía implicacional se establece a partir de los datos de un gran número de lenguas y comprende diferentes categorías y procesos gramaticales. En este tipo de gradación, la adopción de una categoría es condición previa para el préstamo de otra y puede

tener una dimensión diacrónica, por ello, posee un mayor alcance de aplicabilidad. Por otro lado, la escala de prestabilidad se basa en el estudio cuantitativo de un par de lenguas, generalmente, incluye partes del discurso y tiene una dimensión sincrónica (Gómez Rendón, 2008: 59-64; Matras, 2008: 32). En seguida revisaremos algunas de las escalas / jerarquías de préstamo que han repercutido en los estudios de contacto.

Whitney (1881) fue uno de los primeros en plantear un análisis para el estudio de la mezcla de lenguas y proponer una escala de prestabilidad, a partir de la evolución del inglés. Aunque este autor distingue entre préstamo léxico y gramatical,⁸⁵ considera que ambos fenómenos forman parte de un continuum. La categoría más fácil de transferir de una lengua a otra son los sustantivos, en segundo lugar, se encuentran los adjetivos, seguidos de los verbos y los adverbios. Al respecto, comenta que es fácil transferir adjetivos, verbos y nombres al inglés porque en esta lengua se puede hacer una conversión directa de nombres a adjetivos, y de nombres y adjetivos a verbos sin ninguna modificación formal. Dentro de la categoría de préstamo gramatical, separa las palabras funcionales de los afijos. Con respecto a las primeras, señala que es más fácil transferir preposiciones que conjunciones y pronombres; por último, los afijos derivativos se toman prestados con mayor facilidad que los afijos flexivos (Whitney 1881: 19-20), tal como se muestra a continuación.

⁸⁵ El préstamo léxico comprende la transferencia de nombres, verbos, adjetivos y adverbios; mientras que el préstamo gramatical refiere a la transferencia de palabras funcionales y morfemas (Gómez Rendón 2008: 64).

FIGURA 4.2 *Escala de prestabilidad de Whitney (1881: 19-20)*

Préstamo léxico		Nombres
		Adjetivos
		Verbos
		Adverbios
Préstamo gramatical	Palabra funcional	Preposiciones
		Conjunciones
Afijo		Pronombres
	Derivativos	
		Flexivos

Otra de las escalas de prestabilidad más importante en los estudios de contacto es la de (Haugen, 1950: 224), que se basa en investigaciones con inmigrantes noruegos y suecos en Estados Unidos. Según esta, los elementos que más suelen ser transferidos son los nombres, seguidos de los verbos, los adjetivos, los adverbios y preposiciones y, por último, las interjecciones. Sin embargo, como se verá más adelante, esta jerarquía no siempre se cumple, pues también depende del tipo de contacto y de las características de los sistemas lingüísticos implicados.

FIGURA 4.3 *Escala de prestabilidad de Haugen (1950: 224)*

Nombres	Verbos	Adjetivos	Adverbios/Preposiciones	Interjecciones
>	>	>	>	

Años más tarde, Muysken (1981) estudió la situación de contacto entre el quechua y el español en Ecuador, específicamente los fenómenos de relexificación y translexificación⁸⁶. Con base en sus resultados, propuso otra escala de prestabilidad que considera tanto el préstamo léxico como el gramatical, en el siguiente orden.

FIGURA 4.4 *Escala de prestabilidad de Muysken (1981)*

Préstamo léxico	Nombres
	Adjetivos
	Verbos
Préstamo gramatical	Preposiciones
	Conjunciones coordinantes
	Cuantificadores
	Determinantes
	Pronombres libres
	Pronombres enclíticos
	Conjunciones subordinantes

Según Muysken (1981) para los hablantes de español y quechua es más fácil tomar prestado palabras de contenido que palabras funcionales. Además, considera que las palabras que pertenecen a un paradigma estructurado son más difíciles de prestar que otras palabras que no pertenecen a paradigmas estructurados. Asimismo, señala que las palabras a las que se les asigna un caso son más difíciles de prestarse que aquellas a las que no. Por último, añade que las palabras con mayor complejidad morfológica son más difíciles de prestar que las palabras más simples (Muysken 1981 *apud* Gómez Rendón 2008: 73-74).

⁸⁶ Según la gramática generativa, una entrada léxica comprende diferentes tipos de información: la representación fonológica, los rasgos sintácticos, los rasgos de subcategorización, la representación semántica y los rasgos de selección. Dicho esto, la relexificación es un proceso de sustitución del vocabulario en el que sólo se adapta la representación fonológica de la entrada léxica. Por otro lado, la translexificación es un proceso en el que se adopta la representación fonológica y el resto de la información (Muysken 1981: 60-61).

En un trabajo más reciente, Matras (2008) plantea una jerarquía con base en una muestra de 27 lenguas, entre las que se encuentra el otomí y el yaqui. El autor señala que su jerarquía se basa en la frecuencia en que estas clases de palabras son afectadas por el contacto en las lenguas de la muestra y no es una jerarquía implicacional, pues la adopción de una categoría de menor rango en la escala no necesariamente implica la adopción de las de mayor jerarquía.

FIGURA 4.5 *Escala de prestabilidad de Matras (2008: 61)*

Nombres, conjunciones
Verbos
Marcadores discursivos
Adjetivos
Interjecciones
Adverbios
Otras partículas, adposiciones
Numerales
Pronombres
Afijos derivativos
Afijos flexivos

Las escalas de Whitney (1881), Haugen (1950), Muysken (1981) y Matras (2008) se basan en análisis de la frecuencia con la que una categoría puede verse afectada por cambios inducidos por contacto. Las cuatro escalas coinciden en posicionar a los nombres como la categoría más susceptible al préstamo. Haugen (1950) y Matras (2007) consideran que los verbos son la segunda categoría léxica más prestada, mientras que Whitney (1881) y Muysken (1981) anteponen la categoría adjetival a la verbal.

Gómez Rendón (2008: 69) señala que los verbos son menos prestados que los sustantivos, porque poseen tanto contenido semántico como información estructural. Asimismo, advierte que la falta de consenso sobre la posición de los adjetivos en las escalas de prestabilidad, se debe a que en algunas lenguas los adjetivos son clasificados en la misma categoría que los sustantivos o los verbos, lo cual pone en duda la universalidad de esta categoría desde un punto de vista tipológico. Al respecto, Romaine (1995: 65 *apud* Gómez Rendón 2008: 70) señala que la posición relativa de los verbos y los adjetivos depende más de las categorías de palabras existentes en los sistemas en contacto y menos de la prestabilidad inherente de cada clase.

Aunque no todas las escalas de prestabilidad hacen una distinción entre el préstamo de palabras funcionales y el de morfología derivativa y flexiva, se observa que el vocabulario gramatical no ligado se presta con mayor facilidad que la morfología derivacional y flexiva. Otra diferencia importante entre la escala de Matras (2007) y el resto (Whitney, 1881; Haugen, 1950; Muysken, 1981) es la posición de las conjunciones y los marcadores del discurso. En la mayoría de las escalas estas categorías se encuentran en los rangos más bajos, mientras que en la de Matras (2007) ocupan el primero y tercer lugar, respectivamente. En la sección §4.3 reviso la distribución de los préstamos léxicos y gramaticales en el corpus de esta investigación y la discute a partir de las escalas de prestabilidad antes descritas.

A continuación, reseñaré algunos estudios sobre los préstamos del español en el náhuatl y el maya yucateco. Consideré estas lenguas por dos razones. La primera es la similitud que existe entre el porcentaje de préstamos hallados en el náhuatl y en el corpus de esta investigación. Por otro lado, tomé en cuenta el caso del maya yucateco porque el trabajo de Uth y Gutiérrez Bravo (2018) se centra en la marcación de número en los préstamos nominales, la misma categoría gramatical que es objeto de este capítulo.

4.2.4 Los préstamos léxicos en el náhuatl

Hill y Hill (1977:) estudiaron los procesos de relexificación y muerte lingüística en comunidades nahuas del estado de Puebla y Tlaxcala: San Miguel Canoa, San Pablo del Monte, Acuamanala, San Jerónimo Zacualpan, Tepatlaxco y Cuauhtenco. Según estos autores, dicha zona era de habla náhuatl, pero entre 1870 y 1920 comenzó un proceso de desplazamiento de esta lengua.

Con la colaboración de un nahua hablante realizaron entrevistas que comprendían diferentes actividades, como la narración de historias de vida, tradicionales y de peligro de muerte, la traducción de palabras y enunciados del español al náhuatl, y un cuestionario de actitudes lingüísticas (56). A partir de los datos recolectados, Hill y Hill establecieron tres zonas geográficas: oeste, este y norte, de acuerdo con la ubicación de la autopista Puebla-Tlaxcala y con la situación sociolingüística de las comunidades (58).

La zona oeste comprende localidades como San Jerónimo Zacualpan, en donde no se usa el náhuatl y las personas de mayor edad sólo tienen conocimiento de vocabulario simbólico. El área este está conformado por comunidades más tradicionales que las del oeste y conservan más la lengua náhuatl. La zona norte está dividida en tres partes. La primera incluye poblados como San Miguel Canoa, en donde los niños todavía aprenden náhuatl como primera lengua y los adultos hablan un español rudimentario. En el segundo grupo se encuentra la comunidad de San Pablo del Monte, en ella, las personas mayores de 30 años hablan náhuatl, los más jóvenes la entienden, pero no la hablan, y los niños sólo comprenden algunas palabras. El tercer conjunto, incluye la población de Acuamanala, y se caracteriza porque los habitantes de más edad son semihablantes y las personas menores de 50 años no hablan ni entienden la lengua (58).

En mayor o menor medida en todas las comunidades estudiadas se observa un proceso de pérdida del náhuatl. Hill y Hill explican que el desplazamiento de la lengua náhuatl está relacionado con el crecimiento poblacional, la disminución del trabajo agrícola, la migración a ciudades cercanas, la llegada de la electricidad y los medios de comunicación masivos, las políticas de castellanización y la imposición de la educación primaria universal (59).

Los investigadores observan que el náhuatl de Puebla y Tlaxcala tiene una cantidad de préstamos del español, que ascienden al 40% en habla rápida. Al respecto, señalan que el proceso de relexificación se debe al prestigio que goza el español y advierten que este fenómeno ha tenido un efecto negativo en las actitudes que los hablantes tienen hacia el náhuatl, pues la consideran como una lengua mezclada que ya no funciona como medio de identificación. Esto, a su vez, provoca que los hablantes dejen de usar el náhuatl en diferentes dominios lingüísticos y poco a poco abandonen su lengua, dando lugar a la muerte de esta lengua (60).

Hill y Hill (1981) continuaron analizando el proceso de relexificación en la zona de la Malinche en un trabajo posterior. Esta vez, centraron su investigación en diez comunidades de los estados de Tlaxcala y Puebla: San Miguel Canoa, La Resurrección, San Lorenzo Almecatla. Tlaxcala – San Pablo del Monte, Santiago Ayomnetitla, San Antonio Acuamanala, San Luis Teolocholco, Santa María Acxotla del Monte, Santa Ana Chiautempan, San Rafael Tepatlaxco y San Felipe Cuahutenco.

Mediante un muestreo de redes sociales conformaron un grupo de 75 hablantes de náhuatl, con un rango de edad va de 15 a 100 años y cuyo conocimiento de náhuatl comprendía desde los hablantes dominantes en náhuatl hasta los semihablantes de esta lengua. A cada uno de los colaboradores les realizaron una entrevista, constituida por tareas

de narración, traducción, un cuestionario de actitudes lingüísticas y la lista de Swadesh (1955), de donde provienen sus datos (216).

En el análisis, los investigadores clasificaron los resultados en cuatro tipos: respuestas en náhuatl, respuestas en náhuatl con vacilación, préstamos y omisiones. Con base en esto, ordenaron los 100 ítems de la lista de Swadesh (1955), de acuerdo con la dificultad que representaron para los hablantes, en palabras “difíciles” (42 ítems) y “sencillas” (53 ítems).⁸⁷ Uno de los descubrimientos más importantes de su trabajo fue hallar una relación entre los casos de vacilación, omisión y el uso de préstamos del español: “During the interviews we had the impression that speakers who hesitated, but eventually remembered an item, were doing this often on the same items which other speakers hispanized or missed completely” (Hill & Hill, 1981: 220). Los autores señalan que la relación que existe entre la vacilación, la omisión y el uso de préstamos representa una etapa temprana de la reducción de la prominencia de los elementos léxicos que, posteriormente, llevará a su reemplazo (221).

Hill y Hill también compararon la situación de la Malinche con otras variedades de náhuatl moderno: Mecayapan, Pochutla y Zacapoaxtla (Swadesh 1955) y con el de hablantes de náhuatl en la Ciudad de México (Lastra & Horcasitas, 1976). Con respecto a las primeras, encuentran una correlación alta entre los elementos omitidos o hispanizados en estas variedades y los ítems “difíciles” que omitieron, vacilaron o hispanizaron algunos hablantes de la Malinche (222). Por otro lado, los nahua hablantes de la Ciudad de México mantienen los elementos “fáciles” y apenas pueden recordar los “difíciles”, debido a que estas personas no utilizan el náhuatl y están en etapas extremas de pérdida de esta lengua. Los autores concluyen que el náhuatl de la Malinche está en una etapa temprana de reemplazo léxico,

⁸⁷ Los autores descartaron los primeros siete ítems de la lista de Swadesh, debido a la dificultad que le causaron a la mayoría de los hablantes.

que se manifiesta en la reducción de la prominencia de algunos elementos léxicos, que en otros dialectos del náhuatl son omitidos o hispanizados (225).

Décadas después, San Giacomo Trinidad y Peperkamp (2008) estudiaron los préstamos del español y sus procesos de adaptación fonológica en el náhuatl de Tagcotepec, una población de 500 habitantes, ubicada en la sierra norte de Puebla. En esta comunidad, el náhuatl y el español se encuentran en una relación de diglosia y cada uno cumple funciones distintas: la primera goza de prestigio intracomunitario y la segunda funge como vehículo de comunicación para el comercio, la educación, la salud y las relaciones fuera de su comunidad. Un rasgo geográfico importante es que la localidad está dividida en dos secciones por un río. La primera tiene acceso a la carretera y a otras localidades (rurales / urbanas), y el 81% de su población es bilingüe. Por otro lado, la segunda sección está más alejada, se encuentra en la zona montañosa y sólo el 65% es bilingüe (150).

Los investigadores realizaron 46 entrevistas con 71 hablantes: 31 hombres y 40 mujeres, clasificados en dos grupos etarios: el primero está constituido por 30 hablantes de 12 a 35 años y el segundo por 41 colaboradores de 36 años en adelante. Con el material lingüístico de sus colaboradores, formaron un corpus de 5719 tokens de 840 préstamos y analizaron un total 9630 ocurrencias de (no) adaptación fonológica (151).

San Giacomo y Peperkamp encontraron un 17% de adaptación en sus datos. No obstante, hallaron una gran variación de adaptación individual, que va de 5% a 62%. Mediante un análisis multivariado, analizaron la adaptación fonológica de los préstamos y su asociación con los factores sexo, edad, bilingüismo (individual y comunitario), interlocutor y frecuencia de uso (2008: 151-152).

Los resultados seleccionaron cinco de los factores. El de mayor importancia fue la sección geográfica (bilingüismo comunitario), los hablantes de la sección más alejada y

menos comunicada mostraron un porcentaje de adaptación mayor (21%) al de la primera sección (14%). El segundo factor más relevante fue la frecuencia, los préstamos más recurrentes son los que presentaron un mayor porcentaje de adaptación. El tercer lugar lo ocupa la edad de los colaboradores: los hablantes de entre 36 y 41 años adaptan el 18.4% de los préstamos; mientras que los más jóvenes (12-35 años) sólo adaptan el 13.5% de estos. El último factor fue la relación entre interlocutores, cuando los entrevistados se dirigen a personas de su familia o amigos realizan un mayor porcentaje de adaptación (17%), en comparación con las interacciones con otros habitantes de la comunidad (13.5%) (153-155).

4.2.5 Los préstamos léxicos en el maya yucateco

Pool Balam y Le Guen (2015) describieron las estrategias que los hablantes de maya yucateco emplean para integrar las influencias lingüísticas de otras lenguas, a partir del análisis de datos conversacionales recopilados en las comunidades de Kopchen y Chemax. El corpus de esta investigación estuvo integrado por 13 345 palabras, de las cuales sólo el 9% son préstamos del español (1 258 ítems) (345).

Los autores encuentran que el 46.6% de los préstamos son sustantivos, la siguiente categoría más adoptada son las conjunciones (15.1%), seguidas de las interjecciones (12.8%); mientras que los verbos apenas comprenden el 6.2%. Los numerales del español son más recurrentes (5.7%) que los adverbios (4.5%) y adjetivos (4%). Por último, documentan algunos casos de diminutivos (2%), títulos (1.2%), compuestos (0.9%), el comparativo *más* (0.6%) y algunas preposiciones (0.55%) (372).

Pool Balam y Le Guen (2015) describen algunos procesos de integración fonológica e indican que estos fueron más profundos durante los primeros siglos de contacto, debido a que los mayas tenían un conocimiento limitado del español. Sin embargo, en la actualidad,

han observado una menor recurrencia de los fenómenos de adaptación fonológica a causa de la escolarización y los diversos grados de bilingüismo en las comunidades.

En cuanto a la categoría de número gramatical, Pool Balam y Le Guen (2015) indican que los hablantes de maya usan los números del español con la ayuda del sufijo relacional *-vʌ*.⁸⁸ Al respecto, señalan que “La integración de los números españoles en maya demuestra la sofisticación del sistema gramatical del maya, pues desarrolla una estrategia alterna para hacer uso de los números españoles y, al mismo tiempo, conservar el uso de clasificadores específicos al maya” (Pool Balam y Le Guen, 2015: 355).

Por otro lado, advierten que los préstamos del español pueden ser integrados en su forma plural, a pesar de no tener un significado de pluralidad y no presentar concordancia plural, incluso documentan casos con doble marca plural. Sobre esto advierten que “las clases de palabras del español no son usadas con sus características propias ni tampoco como equivalentes gramaticales en maya, sino que son adaptadas a la estructura gramatical del maya” (Pool Balam y Le Guen 2015: 360).

En un trabajo posterior, Uth y Gutiérrez Bravo (2018) describieron los recursos morfológicos que los hablantes de maya (bilingües y monolingües) emplean para marcar la pluralidad en los préstamos del español, así como los factores lingüísticos y extralingüísticos que determinan el uso de las marcas morfológicas. Para ello, recopilaron un corpus de 31 898 palabras a partir de las narraciones escritas de diez hablantes de maya yucateco, pertenecientes a tres grupos etarios (1: 0-29 años, 2: 30-59 años, 3: 60-89 años) y a dos grupos lingüísticos (monolingües en maya y bilingües otomí-maya) (198).

⁸⁸ El uso de los numerales en maya requiere de un clasificador numeral, según la animacidad del nombre *-p'él* para nombres inanimados, *-túul* para animados o *-kúul* para árbol en pie (Pool Balam y Le Guen, 2015: 355).

En su corpus encontraron un total de 327 préstamos nominales pluralizados. La mayoría estaban pluralizados con los sufijos del español (65%: 214 ocurrencias). La segunda estrategia más recurrente fue el uso del sufijo maya *-o'ob* (22%: 73 ocurrencias).⁸⁹ Por último, sólo el 12% de los casos no presentó ninguna marca de pluralidad (40 ocurrencias) (203).

Posteriormente, los autores determinaron qué factores lingüísticos y extralingüísticos inciden en el uso de los diferentes recursos morfológicos y consideraron: la animacidad, el orden de palabras, las relaciones gramaticales, los determinantes, la edad y el perfil lingüístico de los colaboradores. Con respecto a los factores lingüísticos, encontraron que los 97 préstamos que aparecían con numerales del español sólo marcaban la pluralidad con el morfema del español *-s*. Además, observaron un mayor uso del prefijo maya *-o'ob* con préstamos con referentes humanos (53%) y una menor recurrencia con referentes inanimados (11%). Por último, hallaron un mayor uso de *-o'ob* con préstamos que funcionan como sujeto y una menor recurrencia con préstamos que no son sujeto (206-207).

Uth y Gutiérrez Bravo (2018) también exploraron el efecto de dichos factores lingüísticos y su relación con los factores extralingüísticos de edad y perfil lingüístico. En cuanto a los factores extralingüísticos, encontraron que los bilingües maya-español usan más la marca del español *-s* que la del maya *-o'ob*, además, han generalizado el uso de *-s* con todos los referentes, independientemente de su animacidad. En contraste, los monolingües usan la marca del maya *-o'ob* para pluralizar los préstamos humanos y utilizan el morfema

⁸⁹ Los 73 casos comprenden 41 casos con el patrón *-s + -o'ob*, los autores no consideraron que esta forma corresponda a una marcación duplicada de dos sufijos de plural, sino al hecho de que la *-s* del español se ha reanalizado como parte de la raíz de los préstamos, de modo que el hablante que utiliza tal forma en realidad marca el plural únicamente por medio de *-o'ob* en los contextos correspondientes (Pfeiler, 2014: 2220 *apud* Uth y Gutiérrez Bravo 2018: 192).

del español -s mayormente con los préstamos inanimados. Asimismo, encontraron que los factores de edad y animacidad estaban relacionados. Los hablantes mayores de 30 años hacen un mayor uso de la marca -s con los nombres inanimados que los hablantes más jóvenes. Por otro lado, en los hablantes menores de 30 años, la animacidad del nombre no afecta la marcación de número en los préstamos (208-212).

Con base en los resultados, los autores señalan que el español incide en la marcación de número de los préstamos que han ingresado al maya de forma directa e indirecta. De manera directa, porque el maya yucateco toma prestado los nombres y la marca plural del español. Esto da como resultado un sistema con dos marcas de plural: -s y -o'ob, que compiten de tres formas, dando lugar a un cambio indirecto inducido por contacto. Por un lado, se observa un mantenimiento de la marcación de número del maya yucateco en los préstamos, esto se observa en el uso de -o'ob para pluralizar nombres con referentes humanos y con función de sujeto. Por otro lado, el efecto de la animacidad, propia del sistema de marcación de número del maya yucateco, disminuye con el bilingüismo y con la edad de los hablantes (217-218).

4.2.6 El estudio del préstamo léxico y gramatical en la lengua otomí

Actualmente se cuenta con varios estudios sobre el préstamo léxico en la lengua otomí (Bakker et al., 2008; Bartholomew, 1954; Gómez Rendón, 2008; Hekking, 1995, 2001; Hekking & Bakker, 2007b, 2007a, 2010; Lastra, 1994), en los que se analiza la influencia que el español ha tenido en el léxico de algunas variantes de Querétaro, Hidalgo y el Estado de México.

No obstante, la mayoría de estas investigaciones se ha centrado en describir los préstamos léxicos y gramaticales del español en el otomí de Santiago Mexquititlán y San

Miguel (Querétaro), a partir del Principio de Explicación Funcional (Dik, 1997). Dentro de este enfoque, los cambios lingüísticos son interpretados como una adaptación de los sistemas lingüísticos a los cambios en las circunstancias sociales (factores no lingüísticos), que determinan las necesidades comunicativas de los hablantes y las comunidades de habla (factores lingüísticos). Sin embargo, estas adaptaciones se encuentran restringidas por el sistema lingüístico (factores formales) (Bakker et al., 2008: 174-177; Gómez Rendón, 2008: 31-44; Hekking, 2001). Aunque en los cambios lingüísticos intervienen factores no lingüísticos (sociales, demográficos), lingüísticos (pragmáticos, semánticos) y formales (sintácticas, morfológicas y fonológicas), en este enfoque, se prima el papel de los factores no lingüísticos, ya que son considerados la causa última de los cambios inducidos por contacto.

En el otomí de Querétaro⁹⁰ se documenta un porcentaje de préstamos del español que oscila entre el 10% y el 17%, dependiendo de la metodología y el corpus empleado⁹¹ (Bakker et al., 2008: 215; Gómez Rendón, 2008: 290; Hekking & Bakker, 2007a, 2007b: 119). Los préstamos léxicos comprenden el 51.9% de los elementos del español⁹² (Hekking & Bakker, 2007b: 120).

⁹⁰ La descripción de los préstamos léxicos y gramaticales del otomí de Querétaro se basa en los trabajos de Muysken y Hekking (1995), Hekking (2001), Hekking y Bakker (2007b, 2007c, 2010), Bakker, Gómez Rendón y Hekking (2008) y Gómez Rendón (2008).

⁹¹ En Muysken y Hekking (1995: 103) se documenta un 10% de préstamos del español en un corpus que comprende la elicitación de frases y una entrevista informal con el delegado de la comunidad de Santiago Mexquititlán. Hekking (2001) reporta un 17% de préstamos del español en un corpus de 31 entrevistas. En Hekking y Bakker (2007b: 119, 2007c: 443), se reporta un 14.1% de préstamos del español en un corpus de 115 entrevistas a hablantes de otomí (100,541 significantes). En Hekking y Bakker (2010: 33) se reporta un 19% de préstamos del español en material recopilado a partir de la Lista Básica para la Tipología de los Préstamos Léxicos. Por último, en Gómez Rendón (2008: 290) y Bakker et al. (2008: 215), se reporta un 14.1% de préstamos del español en un corpus de 110,541 tokens recopilados a partir de entrevistas.

⁹² Gómez Rendón (2008: 293) presenta un análisis distinto al de Hekking y Bakker (2007b, 2007c) y encuentra que las categorías léxicas comprenden el 47.5% del total de préstamos.

La categoría léxica más recurrente son los sustantivos (40.7%). Estos elementos suelen tomarse en su forma singular, con su respectiva marca de género gramatical⁹³ (Hekking & Bakker, 2007c: 443). Sin embargo, también se ha documentado la adopción de sustantivos con la marca de plural, como en *kargero-s*, o con el sufijo diminutivo *pastor-sita* (Gómez Rendón, 2008: 289). Al ingresar a la lengua otomí, se acompañan de la morfología nominal de esta lengua (sufijos y proclíticos) y funcionan, principalmente, como núcleo de frases nominales (1a), aunque también pueden fungir como núcleo de frases verbales (1b) o modificadores de frases nominales (1c) (Hekking & Bakker, 2007c: 443; Gómez Rendón, 2008: 309).

- (1) a. *Un'bya ya nguu ya losa - 'bya ya teja*
 Hoy DEF.PL casa DEF.PL concreto - ACT DEF.PL teja
- ya lamina yá njo'mi ya nguu*
 DEF.PL lámina POS.PL techo DEF.PL casa
- ‘Actualmente, las casas son de concreto, y los techos de las casas son de teja y placas metálicas

(Adaptado de Hekking & Bakker, 2007c: 444)

- b. *kwando ar = Alemän mbi = pridente bí = hyo*
 cuando DEF.SG = Alemán 3.PSD = presidente 3.PRS = matanza
- ‘Cuando Alemán era presidente ordenó la matanza’

(Adaptado de Gómez Rendón, 2008: 316)

- c. *ne = r fleko majwi ne da = mät' = ar seda*
 y = DEF.SG fleco manto y 3.FUT = girar = DEF.SG seda

nthu'ye ne gem'bu' - bya ne da = tsut = ya nts'uungu
 pañuelo y después - ACT y 3.FUT = colgar = DEF.PL arete

‘La capa con flecos, y ella hilará el pañuelo de seda y luego se pondrá los aretes’

(Adaptado de Gómez Rendón, 2008: 317)

⁹³ Gómez Rendón (2008: 287) señala que los morfemas de género no son préstamos productivos y no ocurren en las palabras nativas del otomí. Sin embargo, Hekking (2007c: 444) ha encontrado algunos ejemplos donde se introduce los morfemas de género en palabras nativas del otomí: *'beto* ‘nieto’, *'beta* ‘nieta’.

La segunda categoría léxica más adoptada son los verbos, con el 4.8% de la totalidad de los préstamos (Hekking & Bakker, 2007b: 128, 2007c: 444). El proceso de préstamo consiste en aislar la raíz verbal seguida de la vocal temática (/ -a-/, / -e- / o / -i- /),⁹⁴ posteriormente, se inserta de forma directa a la lengua otomí⁹⁵ y se le adhieren los proclíticos y sufijos verbales correspondientes (Gómez Rendón, 2008: 287; Hekking & Bakker, 2007c: 445), tal como se muestra en (2). No obstante, también se documentan algunos casos de verbos adoptados en su forma de participio (*us-ado*) o con marcas de tiempo (*nkambyo*) (Gómez Rendón 2008: 289).

- (2) *Ntonse nu - r txuku tobe*
 Entonces DEM-DEF - SG perro todavía
- mi = molesta - tho nuya kolmenä*
IMPF.3 = molestar - LIM DEM.3PL colmena
 ‘El perro todavía molestaba a las abejas’
 (Adaptado de Hekking & Bakker, 2007c: 445)

Los adverbios son la tercera categoría léxica más adoptada, con el 4.5% del total de los préstamos (Hekking, 2007b: 128, 2007c: 45). Estas formas ingresan a la lengua otomí directamente y se comportan como modificadores de frases verbales, a pesar de que en *ñható* no existe una categoría léxica que cumpla esta función (Gómez Rendón, 2008: 288).

⁹⁴ Gómez Rendón (2008: 287) señala que las vocales / -a- / y / -e- / corresponden a las terminaciones verbales de la tercera persona, mientras que / -i- / es el producto de la elevación de / -e- / en los verbos que cambian su vocal raíz / e / a / i /. Sin embargo, Hekking y Bakker (2007c: 445) señalan que los verbos son tomados en su forma de infinitivo, sin *-r*.

⁹⁵ Gómez Rendón (2008: 287) señala que los verbos del español que ingresan a la lengua otomí no requieren de una forma proverbal o de un afijo de préstamo verbal para funcionar como verbo, tal como sucede en el popoloca o el náhuatl, respectivamente.

- (3) *Ha nu by ga = 'yot'e mälmente ke hinga*
 y cuando 2.PSD = hacer **malamente** ke NEG
 'Señor nuestro, y cuando lo haces mal'
 (Adaptado de Gómez Rendón, 2008: 333)

Los adjetivos son la categoría léxica menos prestada, con sólo 1.9% del total de los préstamos. Estos elementos son tomados en su forma de masculino singular, cuando ingresan a la lengua otomí pueden funcionar como núcleo de una frase verbal y se acompañan de los proclíticos y sufijos verbales del otomí, como se muestra en (4a). También pueden fungir como núcleo de una frase nominal, en cuyo caso se les adhiere la morfología nominal correspondiente (4b). Sin embargo, la mayoría de las veces, los adjetivos funcionan como modificadores de frases nominales (4c), a pesar de que en otomí no existe esa clase léxica (Hekking, 2007c: 444-445; Gómez Rendón, 2008: 287, 318). En esta posición, los adjetivos no se acompañan de morfología verbal y nominal como en los ejemplos anteriores.

- (4) a. *Nä nxutsi to di = ñen - wi*
 DEM niña quien 1.PRS = jugar - DU.INCL

x = ar bïbo
mucho = DEF.SG **inteligente**
 'La chica con la que juego es muy inteligente'
 (Adaptado de Gómez Rendón, 2008: 328)

- b. *¿Ne xä'a ar = bindo tsi mbane?*
 y INTERR DEF.SG = **colorado** DIM padre

Nä'ä mi = thädi hingi 'bui ar = bindo
 PRO.DEM 3.IMPF = responder NEG 3.existir DEF.SG = colorado
 '¿Y el colorado, padre? Él respondió: 'la vaca colorada no está'
 (Adaptado de Gómez Rendón, 2008: 329)

- c. *Hö ya = nei asta 'rato kwadriya ne ya nxutsi*
 Sí DEF.PL = bailarín hasta seis cuadrilla y DEF.PL niña
- ne ya metsi 'rato ya nei prinsipal ge Nxemge*
 y DEF.PL niño seis DEF.PL bailarín principal COP San Miguel
 'Sí, hay hasta seis equipos de bailarines, niños y niñas, y seis bailarines
 principales de San Miguel'

(Adaptada de Gómez Rendón, 2008: 330)

Por otro lado, los elementos gramaticales comprenden el 48.1% de los préstamos.⁹⁶

La categoría gramatical más adoptada son las preposiciones (21.2%) (Hekking & Bakker, 2007b: 122, 2007c: 447), entre las que destaca el uso de *ko* (< con), *pa* (< para), y *de*. En la lengua otomí, la preposición *con* es empleada para marcar la función comitativa, en lugar de los sufijos del otomí *-hú* (PL. INCL.), *-wí* (DUAL INCL.), *-hé* (PL. EXCL.), *-bé* (DUAL EXCL.) (Hekking & Bakker, 2007b: 122-123; Gómez Rendón, 2008: 389); sin embargo, también es posible encontrar oraciones en donde ambas marcas coexisten, como se ilustra en (5a). Además, *con* es utilizada para marcar la función de instrumento (5b), junto o en lugar del sufijo otomí *ir nge* (Hekking & Bakker, 2007c: 448; Gómez Rendón, 2008: 389).

- (5) a. *Mnde ngi = 'ño hú ko hñu ya nxutsi*
 ayer 2PSD = caminar PL.INCL con tres DEF.PL niña
 'Ayer caminaste con tres niñas'

(Adaptada de Hekking y Bakker 2007c: 448)

- b. *Ar jā'i bi = dāk = ar k'eñä kon minge = r*
 DEF.SG hombre 3PSD = atacar = DEF.SG víbora con pico = DEF.SG
 'El hombre atacó a la víbora con el pico'

(Adaptada de Hekking & Bakker, 2007c: 448)

⁹⁶ Gómez Rendón (2008: 293) presenta un análisis distinto al de Hekking y Bakker (2007b, 2007c) y encuentra que las categorías gramaticales comprenden el 52.5% del total de préstamos.

La preposición *para* se emplea para marcar el beneficiario, junto o en lugar del sufijo otomí *-pi* (5c). También se usa para marcar propósito, tal como se muestra en (5d) (Hekking & Bakker, 2007b: 122-123, 2007c: 447; Gómez Rendón, 2008: 389-390).

- (5) c. *nä* - *r* *hyokunguu bi* = *hyok* - *wi*
 DEM - DEF.SG arquitecto 3.PSD = construir - DU.INCL

'nar *nguu pa* - *r* *ts'ut'ubi*
 INDEF.SG casa **para** - DEF.SG gobernador
 'El arquitecto construyó una casa para el gobernador'

(Adaptada de Hekking & Bakker, 2007c: 448)

- d. *Thokü 'nar pont'i zaa pa da t'exü ja* - *r* *'met'e*
 Hacer INDEF.SG cruz madera **para** 3.FUT poner LOC - DEF.SG techo
 'Hacen una cruz de madera para ponerla en el techo'

(Adaptada de Hekking & Bakker, 2007c: 449)

La preposición *de* se usa para marcar la relación de posesión (5e), en vez de yuxtaponer lo poseído y el poseedor, como sucedía en el otomí clásico. Además, sirve para marcar partitivos (5f), materiales y genitivos (Hekking & Bakker, 2007b: 124-125, 2007c: 450; Gómez Rendón, 2008: 390).

- (5) e. *Nixi Independensya nixi Reforma nixi Rebolusyon*
 Ni Independencia ni Reforma ni Revolución

bi = *nkambyo* *yá* *kostumbre de* *ya* *ñãñho*
 3.PSD = cambiar 3.POS *costumbre de* DEF.PL otomí
 'Ni la Independencia, ni la Reforma, ni la Revolución cambiaron las costumbres de los otomíes'

- f. *'Na de ge'u i* = *ndude kaha*
 uno **de** 3.DEM 3.PRES = cargar caja
 'Uno de ellos cargó la caja'

(Adaptados de Hekking & Bakker, 2007c: 450)

Las preposiciones *de*, *desde*, *hasta*, *entre*, *para*, *a* también son utilizadas para marcar una gran diversidad de relaciones locativas, que en el otomí clásico no se marcaban (Hekking

& Bakker, 2007b: 124, 2007c: 449; Gómez Rendón, 2008: 390). Por ejemplo, en (5g), la preposición *de* es empleada para introducir una fuente.

- (5) g. *Ya - dá = pengi de Jalpa*
ya - 1.PSD = regresar de Jalpa
 ‘Ya regresé de Jalpa’

(Adaptada de Hekking & Bakker, 2007c: 449)

La segunda categoría gramatical más prestada son las conjunciones coordinantes, que comprenden el 7.5% del total de los préstamos. Las más recurrentes son: *y, pero, o, ni* y *sino* (Hekking & Bakker, 2007b: 122, 125, 2007c: 447, 450; Gómez Rendón, 2008: 365). Algunas conjunciones sustituyen o coaparecen con las marcas del otomí, como se muestra en (6a). No obstante, también hay casos en los que la conjunción del español no compite con algún elemento equivalente en otomí, tal es el caso de *pero* (6b).

- (6) a. 'bet'o futs'i ne 'ñuni ne ja da = yobi y
 antes barbecho y regar y hacer 3.FUT = doblar y
 después ja da mot'i
 después 3.FUT sembrar
 ‘Primero barbecha y riega y dobla y después siembra’

(Adaptado de Gómez Rendón, 2008: 368)

- b. *Hin = di pä - kq ko = r hñãñho pero*
 NEG = 1.PRS saber - 1.ENFT con = DEF.SG otomí **pero**
ko = r hñãmfo hä di = pä - kq
 con = DEF.SG español sí 1.PRS = saber - 1.ENFT
 ‘No sé en otomí, pero sí sé en español’

(Hekking & Bakker, 2007c: 451)

Los marcadores discursivos son la tercera categoría más prestada (6.5%) (Hekking y Bakker, 2007b: 122, 2007c: 447), los más frecuentes son *entonces, este* y *pues*, según Gómez Rendón (2008: 393). El uso y la frecuencia de *pues* en otomí es muy similar al que tiene en el español mexicano: énfasis, cambio de código, contraste y reanudación, como se muestra

en (7a). El marcador discursivo *este* también tiene funciones parecidas en otomí y español, los otomíes lo emplean para mantener el turno de habla o como marcador de vacilación (7b), tal como sucede en español.

(7) a. *Temu gi = mä - nge*
 Qué 2.PRS = decir - 2.ENFT

Pwes nugä di = mä nga gätho ar zä
Pues yo 1.PRS = decir 1.ENFT todo DEF.SG bien
 ‘¿Qué piensas? Bueno, yo pienso que todo está bien’

(Adaptado de Hekking & Bakker, 2007c: 452)

b. *Ar Xuwa bí este bí = hñuxu 'na - r hē'mi*
 DEF.SG Juan 3.PSD **este** 3.PSD = escribir uno - DEF.SG carta

pa bi = mända = wi ár mpädi Enrike
 para 3.PSD = mandar = DU.INCL 3.POS amigo Enrique
 ‘Juan escribió una carta para mandársela a su amigo Enrique’

(Adaptado de Gómez Rendón, 2008: 398)

Las conjunciones subordinantes ocupan el cuarto lugar, con el 6.1% del total de préstamos. Los elementos más recurrentes son *que, porque, para que, ya que, como, cuando* y *aunque* (Hekking & Bakker, 2007b: 125-126, 2007c: 452; Gómez Rendón, 2008: 372). Las conjunciones del español pueden coaparecer o sustituir a los elementos del otomí, tal como sucede en (7c), donde aparece la conjunción *porque*, en lugar de la partícula otomí *ngetho*.

(7) c. *No = r bätši bi = nzoni porque*
 DEM = DEF niño 3.PSD = llorar **porque**

bi = n - tsät'i na nts'ę di - tho
 PAST.3 = REFL - quemar muy fuerte - LIM
 ‘El niño lloró porque se quemó muy fuerte’

(Adaptado de Hekking & Bakker, 2007c: 453)

En el otomí de Santiago Mexquititlán y San Miguel también se documenta la presencia de pronombres relativos (*que, lo que*), indefinidos (*cada, quien, cualquiera, nada*)

e interrogativos como (*qué, por qué, para qué, cómo, cuánto*) (Hekking & Bakker, 2007b: 126, 2007c: 453-454).

Hekking y Bakker (2001, 2007b: 122-123, 2007c: 447) indican que el préstamo masivo de preposiciones y conjunciones tiene una motivación funcional. Dado que el otomí clásico tenía pocas partículas y sufijos verbales para marcar la relación entre el predicado y otras partes de la oración, la adopción de las categorías gramaticales del español vuelve más explícitas estas relaciones (teoría del hueco funcional y semántico). Asimismo, señalan que otra posible motivación es el olvido o la no adquisición de ciertos morfemas gramaticales del otomí en las nuevas generaciones (Hekking, 2001). Con respecto a los marcadores discursivos, han sugerido que los hablantes adoptan esta categoría para dotar de sabor español su discurso otomí, con la finalidad de obtener un mayor estatus en la comunidad (Hekking & Bakker, 2007b: 122, 131, 2007c: 447-451).

La adaptación fonológica de los préstamos léxicos y gramaticales del español en el otomí

Los préstamos léxicos y gramaticales⁹⁷ pueden estar sujetos a procesos de adaptación fonológica. En los estudios sobre el otomí de Santiago Mexquititlán y San Miguel se han identificado cinco procesos principales (Hekking & Bakker, 2007c: 437-439, 2010: 40; Gómez Rendón, 2008: 285-286). El primero es la nasalización de segmentos vocálicos y consonánticos. La vocal abierta central [a] se nasaliza [ã] ante consonantes nasales [m], [n], [ɲ], como se muestra en los ejemplos (8a-8c). Además, se suelen nasalizar los segmentos sibilantes y oclusivos a inicio de palabra, como ilustra en (8d) y (8e).

⁹⁷ Gómez Rendón (2008: 351-352) señala que la adaptación morfofonológica de los préstamos gramaticales depende de su nivel de gramaticalización. Otros procesos de adaptación morfofonológica que pueden sufrir este tipo de palabras son i) la fusión con elementos contiguos y la afijación de elementos gramaticales de la lengua receptora.

- (8) a. [anima] > [animã] d. [sinko] > [nsinku]
 b. [engaɲa] > [ngaɲã] e. [desde] > [ndezde]
 c. [apenas] > [apenã]

(Adaptados de Hekking & Bakker, 2007c: 438; Gómez Rendón, 2008: 285)

El segundo proceso es la modificación de la estructura silábica con el objetivo de recuperar la estructura silábica CV de la lengua otomí. En algunos préstamos se elide el núcleo vocálico de sílabas átonas en posición inicial (9a) - (9b) y media (9c) - (9d). También se omite los segmentos consonánticos en coda al final de palabra (9e) - (9f). Por último, los grupos consonánticos del español se suelen simplificar en posición inicial o media de palabra (9g).

- (9) a. [alko:l] > [akol] e. [kon] > [ko]
 b. [areglo] > [reglo] f. [kondisyon] > [kondsyo]
 c. [aseyte] > [aste] g. [oktubre] > [otubre]
 d. [bisikleta] > [biskleta]

(Adaptados de Gómez Rendón, 2008: 285)

El tercer proceso comprende cambios en los segmentos vocálicos. El núcleo vocálico de las sílabas átonas puede ser reemplazado por la vocal cerrada central no redondeada [i̠], como se muestra en (10a) y (10b). En algunos préstamos las vocales medias en sílabas átonas se cierran (10c). Además, algunos diptongos se simplifican ante las consonantes [d], [f], [m] y [k], como se muestra en (10d).

- (10) a. [bispera] > [bisp̠ira] d. [kwerpo] > [korpo]
 b. [kaso] > [kasi̠] e. [sikyera] > [sykera]
 c. [besino] > [bisinu]

(Adaptados de Hekking & Bakker, 2007c: 438; Gómez Rendón, 2008: 285)

El último proceso fonológico incluye modificaciones en los segmentos consonánticos. La fricativa velar sorda [x] se glotaliza [h] (11a) - (11b). Otro de los fenómenos es la sonorización de las consonantes oclusivas sordas [p], [t], [k], como se muestra en (11c) y (11d). Por último, la sibilante [s] se palatiza [ʃ] en posición inicial e intervocálica (11e) - (11f). Los dos últimos fenómenos se presentan principalmente en los préstamos más antiguos.

- | | | |
|------|----------------------|--------------------|
| (11) | a. [bruxa] > [bruha] | d. [baka] > [baga] |
| | b. [mexor] > [mehor] | e. [mesa] > [meʃa] |
| | c. [ja ke] > [ja ge] | f. [sebo] > [ʃebo] |

(Adaptados de Gómez Rendón, 2008: 286)

Además de estos procesos de adaptación, (Soustelle, 1993)⁹⁸ documenta otros fenómenos de adaptación que no se documentaron en Santiago Mexquititlán y San Miguel:

- i) la epéntesis del segmento vocálico [ɛ] ante el segmento [l] para formar una nueva sílaba abierta CV, como se ejemplifica en (12a) y (12b); ii) la alternancia entre /d/ y /r/ (12c) - (12d);
- iii) el reemplazo del grupo consonántico /-dr-/ en posición intervocálica por /-dn-/ o /gr/ (12e); iv) la apertura de la vocal [u] en [o] (12f). Por último, registra casos de iv) movimientos

⁹⁸ Este trabajo se basa en una lista de vocabulario recopilado en 33 localidades: Ixmiquilpan, 2. Zimapán, 3. Tolimán, 4. Jiliapan, 5. Santa Catarina, 6. La Cañada (Querétaro), 7 (véase 4), 8. Huichapan, 9. San Miguel, Tolimán, San Pablo, San Antonio, Higuera, 10. Tecozautla, 11. San Salvador, Actopan, con límite en Pachuca, 12. San Ildefonso, 13. Amealco, San Miguel Tlaxcaltepec, Santiago Mezquititlán, San José Itó, San Bartolomé, San Ildefonso, San Pedro y Chintyé. 14. Tetillas, región de Cadereyta, Boyé, Sombrerete, 15. Dolores Hidalgo, Tequisquiapan, La Cruz del Palmar, San Juande la Vega, La Orduña, Santa Cruz, Apaseo, San Bartolo, San Pedro, Obraje de Ixtla, San Miguel de Ixtla y Mayorazgo, 16. Tierra Blanca, Cieneguillas, Corral Viejo, El Picacho, Cerro Colorado, La Cuesta y El Cano, 17. El Cardonal, 18. San Bartolo, Morelos, Sierra de las Cruces, San Marcos, San Gregorio, Malacota y San Felipe, 19. Timilpan, 20. Chapa de Mota (San Juan Tuxtepec, San Felipe Coamango, Dongú, Bodenquí) 21. La región de Jilotepec, noroeste del Estado de México hasta los límites de Querétaro, 22-23. Algunos datos de San Bartolo Oztolotepec, Santa Ana Jilotzingo, 24. Ameyalco, Atarasquillo y Xochicuaatla, 25. Temoaya, Jiquipilco, Capulhuac, 26. Amanalco, 27. Siráhuato, Michoacán (Ciudad Hidalgo, San Matías y San Pedro Jarácuaro, 28. Huixquilucan, 29. Mitad meridional del Valle de Toluca (Acasulco, Atlapulco, Cholula, Tepexoyuca, Coapanoaya, Almaya, Tilapa, Xalatlaco, 30. San José del Sitio, 31. Huayacocotla, Veracruz, 32. Ixtenco, 33. Santa Ana y San Pedro (Soustelle, 1993: 203-207).

acentuales a la última sílaba, como se ilustra en (12g). En contraste, Gómez Rendón (2008: 286) señala que en el otomí de Santiago Mexquititlán el acento primario permanece sin cambios en la mayoría de los préstamos, incluso si se omiten sílabas.

- | | | |
|------|--------------------------|--------------------------------|
| (12) | a. [ánxel] > [ánhele] | e. [pedro] > [bɛdnu] / [pegru] |
| | b. [wakal] > [awakale] | f. [durasno] > [dorasno] |
| | c. [rosario] > [dosario] | g. [asyénda] > [asyendá] |
| | d. [asadón] > [asaro] | |

(Adaptado de Soustelle, 1993: 287-289)

En las comunidades de Santiago Mexquititlán y San Miguel, se ha observado que la adaptación de los préstamos del español a la fonología del otomí está relacionada con la alfabetización en español, el nivel de instrucción, la situación migratoria (Hekking, 1995: 128) y el nivel de bilingüismo de los hablantes (Gómez Rendón, 2008: 284). En otras comunidades otomíes se ha encontrado que el nivel de adaptación fonológica de los préstamos está relacionado con el registro y el tema de conversación. Al respecto, Soustelle (1993: 286-287) señala que cuando los hablantes de otomí conversan sobre temas de su entorno cultural, dentro de los espacios sociales de su comunidad suelen hablar en un otomí con muy pocos préstamos del español, todos integrados fonológicamente al otomí. En cambio, cuando el tópico está vinculado con la religión, los bilingües utilizan una mayor cantidad de préstamos del español sin adaptarlos a la lengua otomí, incluso en casos en que existe un elemento equivalente en su lengua.

El préstamo léxico y gramatical en el otomí desde un enfoque tipológico

El comportamiento de los préstamos léxicos y gramaticales en el otomí de Querétaro se ha contrastado con el de dos lenguas genética y tipológicamente distintas, el quechua y el

guaraní. La categoría léxica más prestada en las tres lenguas son los sustantivos. Según Gómez Rendón (2008: 300), esto se debe a que los sustantivos poseen algunas características que facilitan el proceso de préstamo, por ejemplo, su morfología simple, su independencia sintáctica y semántica, y su capacidad para referir entidades del espacio socio-comunicativo que son fáciles de recuperar.

TABLA 4.3 *Porcentaje de préstamos léxicos y gramaticales en las lenguas otomí, quichua y guaraní (adaptado de Hekking & Bakker, 2007c: 444)*

Categoría léxica	Otomí	Quichua	Guaraní
Nombres	40.7%	54.4%	37.2%
Verbos	4.8%	17.7%	18.3%
Adverbios	4.5%	3.4%	2.3%
Adjetivos	1.9%	8.5%	7.4%
Total	51.9%	83.6%	65.2%

Este mismo autor señala que el menor porcentaje de préstamos verbales se debe a que tienen una morfología compleja y fusional, que inhibe su préstamo en las tres lenguas. Además, advierte que el impacto de este factor depende del tipo de bilingüismo de los hablantes. De este modo, explica que la menor incidencia de verbos del español en otomí se debe a que sus hablantes son bilingües incipientes y no pueden analizar correctamente los morfemas y aislar las raíces verbales para llevar a cabo el préstamo. Contrario a esto, la presencia de una gran cantidad de préstamos verbales en guaraní está relacionado con que la mayoría de sus hablantes son bilingües compuestos y coordinados (Gómez Rendón, 2008: 297).

La principal diferencia entre las tres lenguas se encuentra en el porcentaje de préstamos adjetivales y adverbiales que adoptan, pues mientras en guaraní y quichua los adjetivos son la tercera categoría más adoptada, en otomí apenas comprenden el 1.9% de los préstamos (Hekking & Bakker, 2007b: 128, 2007c: 444). Los autores explican este

comportamiento a partir de las características tipológicas y estructurales de las lenguas. Por un lado, el mayor porcentaje de préstamos adjetivales en quichua y guaraní se debe a que poseen una clase flexible de “no verbos”, que facilita el préstamo de elementos léxicos que pueden integrarse a esta categoría y funcionar como núcleo o modificador de frases nominales, tal como sucede con los adjetivos (Gómez Rendón 2008: 318).

Por otro lado, el bajo porcentaje de préstamos adjetivales en otomí se debe a que esta lengua no tiene una clase léxica que funcione como modificador de una frase nominal⁹⁹ (véase § 1.4.3.1) (Hekking & Bakker, 2007b: 127-129, 2007c: 441-443). Sobre esto, Gómez Rendón (2008) precisa que el efecto y la dirección de este factor estructural dependen del bilingüismo de los hablantes. Con base en esto, plantea que en los bilingües incipientes, hablantes más conservadores, la falta de adjetivos en el otomí inhibe el préstamo de esta clase léxica. Contrario a esto, en los bilingües compuestos y coordinados, hablantes más innovadores, la falta de una clase léxica adjetival en el otomí promueve el uso de este tipo de préstamos.

A partir de estas observaciones, propone que el bajo porcentaje de préstamos adjetivales en el otomí de Querétaro se debe a que la mayoría de los hablantes de otomí son bilingües incipientes, que tienden a conservar las características tipológicas de su lengua y a rechazar los elementos que no se ajusten a esta.¹⁰⁰ Sin embargo, posteriormente, al revisar el

⁹⁹ Existe una discusión en torno a la existencia de la categoría adjetivo en otomí. La mayoría de los investigadores han descrito los conceptos de propiedad como adjetivos, pese a advertir un comportamiento sintáctico verbal (Hekking & Andrés de Jesús, 1984; Lastra, 1997; Voigtlander & Echegoyen, 1985). Por otro lado, Palancar (2006: 327) prima los criterios sintácticos y con base en estos argumenta que no existe la categoría de adjetivo en la lengua otomí, pues los conceptos de propiedad se expresan con compuestos nominales, nombres y verbos.

¹⁰⁰ Las afirmaciones que se realizan en torno al préstamo de adjetivos y su relación con el bilingüismo de los hablantes de otomí no concuerdan con los datos que se presentan en los apéndices de esa investigación (Gómez Rendón 2008: 499-500). Según esos datos, el porcentaje de bilingües incipientes en quichua es de 31.8%, en otomí de 27.2% y en guaraní de 0%. Por otro lado, los bilingües coordinados y compuestos comprenden el

comportamiento sintáctico de los préstamos adjetivales en otomí, encuentra que su función principal es modificar frases referenciales, a pesar de que no existe esa clase léxica (Hekking & Bakker, 2007c: 444-445; Gómez Rendón, 2008: 287, 318). Con base en esto, postula que el adjetivo en otomí se encuentra en un proceso de lexicalización incipiente inducido por contacto, en el que la formación de una nueva clase léxica especializada en la modificación de frases referenciales está reemplazando las estrategias morfosintácticas.

En cuanto a las categorías gramaticales, se puede apreciar que el otomí es la lengua con mayor porcentaje de préstamos gramaticales (48.1%) en comparación con el quichua y el guaraní. Gómez Rendón (2008: 356) señala que las diferencias en la distribución de los datos están relacionadas con la predominancia de las clases gramaticales y la estructura de cada lengua.

TABLA 4.4 *Porcentaje de préstamos gramaticales en las lenguas otomí, quichua y guaraní (adaptado de Hekking & Bakker, 2007c: 447)*

Categorías gramaticales	Otomí	Quichua	Guaraní
Preposiciones	21.2	0.5	0.5
Coordinadores	7.5	6.9	4.4
Marcadores discursivos	6.5	0.6	0.8
Subordinadores	6.1	0.8	4.6
Pronombres	0.6	0.1	0.2
Otros	6.2	7.5	24.3
Total	48.1	16.5	34.8

Gómez Rendón (2008) señala que la teoría del hueco funcional (Hekking, 2001; Hekking & Bakker, 2007b, 2007c) sólo explica parcialmente el préstamo masivo de

68.2% de los hablantes de quichua, el 72.8% de los otomíes y el 100% de los bilingües guaraní-español. Con base en esto, se puede ver que la mayoría de los otomíes no son bilingües incipientes sino bilingües coordinados y compuestos, lo cual no sostendría la idea de que el bajo porcentaje de préstamos adjetivales en el otomí de Querétaro se debe a que la mayoría de los hablantes de otomí son bilingües incipientes. Además, si esto fuera cierto, entonces se observaría un fenómeno similar en quichua, en cuya muestra hay más bilingües incipientes que en otomí.

preposiciones y conjunciones en otomí, ya que la causa última no es lingüística. Al respecto, indica que la situación del contacto otomí-español está determinada, en primer lugar, por las condiciones socio-comunicativas y, en última instancia, por factores estructurales, como la distancia tipológica y estructural entre las lenguas. Con base en esto, señala que la situación diglósica que existe entre los hablantes de otomí y español ha generado el desplazamiento de la lengua originaria y la adopción de una gran cantidad de categorías gramaticales, que les permite a los hablantes de otomí reproducir la estructura comunicativa y discursiva de los hablantes monolingües en español.

En cuanto a los marcadores discursivos, su bajo porcentaje en las tres lenguas se debe a que su uso requiere de una posición específica en el discurso y a que el número de miembros de esta categoría es reducido, ya que esta función puede codificarse mediante recursos sintácticos y prosódicos. No obstante, la lengua otomí introduce una mayor cantidad de marcadores discursivos del español en comparación con el guaraní y el quichua. Según (Gómez Rendón, 2008: 399-440), estas diferencias dependen de la frecuencia y el uso de los marcadores discursivos en las variedades locales de español.

Con base en el comportamiento del préstamo léxico y gramatical del otomí de Querétaro y su contraste con las lenguas guaraní y quichua, los autores señalan que los préstamos del español han tenido un efecto en la estructura de la lengua otomí. Por un lado, la adopción de préstamos adjetivales podría dar lugar a la creación de una categoría adjetival en la lengua otomí. Por otro lado, la adopción de preposiciones y conjunciones coordinantes y subordinantes podría modificar las características morfosintácticas del otomí, que pasaría a ser “una lengua con menos yuxtaposición al nivel de la cláusula y en una lengua más analítica en los niveles sintácticos más bajos” (Hekking & Bakker, 2007b: 121).

Las investigaciones sobre los préstamos léxicos y gramaticales en el otomí de Querétaro aquí descritas se centran en los factores estructurales de las lenguas en contacto y cómo estos modelan los procesos de préstamo. Sus resultados sientan las bases para la formulación de hipótesis sobre la dirección de los cambios inducidos por contacto en situaciones de contacto otomí-español. Sin embargo, aunque estas se basan en el Principio de Explicación Funcional donde se prima el rol de los factores no lingüísticos en los cambios inducidos por contacto, dichos factores juegan un papel secundario en la explicación que los autores proporcionan sobre estos cambios. Se deja de lado las características sociolingüísticas de los hablantes, a pesar de que se presentan afirmaciones e hipótesis relacionadas con el nivel de bilingüismo. Se postulan hipótesis sobre posibles cambios en el otomí, debido a la adopción masiva de préstamos, pero para poder afirmarlos se requieren más estudios.

4.3 DESCRIPCIÓN GENERAL DEL CORPUS

Debido a que este capítulo se centra en el análisis de un fragmento de la gramática del otomí, no consideré a los 17 colaboradores que son monolingües en español (ME) y monolingües con conocimiento pasivo del otomí (PO). El corpus se construyó a partir de las narraciones tradicionales y de vida en otomí de 26 bilingües otomí-español de San Andrés Cuexcontitlán y Jiquipilco el Viejo. Transcribí y traduje diez minutos de las narraciones y recolecté todos los elementos léxicos y funcionales del español, así como los casos de cambio de código. El corpus quedó conformado por 20 902 palabras, de las cuales 4 774 resultaron ser elementos del español (22.83%).

En la Tabla 4.5 se puede ver que la mayoría de los elementos del español son nombres (26.39%), seguidos por las conjunciones (21.82%) y los verbos (10.32%). En cuarta posición se encuentran los pronombres (10.01%) y después las preposiciones (9.06%). La sexta

categoría más prestada son los adverbios (8.12%) y en séptima posición están los adjetivos (4.08%). Los numerales sólo comprenden el 0.75% de los casos y la categoría otros, que comprende artículos y clíticos, el 0.41%. Por último, vemos que los casos de cambio de código representan el 8.98%.

TABLA 4.5 *Categoría de los préstamos del español en SAC y JV*

Categoría	F	%
Nombres	1 260	26.39
Conjunciones	1 042	21.82
Verbos	493	10.32
Pronombres	478	10.01
Preposiciones	433	9.06
Adverbios	388	8.12
Adjetivos	195	4.08
Numeral	36	0.75
Otros	20	0.41
Cambio de código	429	8.98
Elementos del español	4 774	100.00

El hecho de que los nombres sean la categoría léxica más prestada coincide con el orden de todas las escalas de prestabilidad (Haugen, 1950; Matras, 2008; Muysken, 1981; Whitney, 1881) expuestas en el apartado §4.2.3. La posición de las conjunciones y los verbos muestra similitudes con la propuesta de Matras (2008), quien pone en la primera posición de su escala a los nombres y las conjunciones, y en segundo lugar a los verbos. En contraste, difiere de las escalas de Whitney (1881), Haugen (1950) y Muysken (1981), en las que los verbos o los adjetivos son la segunda categoría más prestada y las conjunciones ocupan las últimas posiciones. La disposición de los pronombres y las preposiciones en el corpus de otomí no corresponde con lo propuesto en las escalas de Whitney (1881), Haugen (1950),

Muysken (1981) y Matras (2007), en las que estas categorías son menos prestadas que los adjetivos y los adverbios.

En relación con esto, llama la atención el bajo porcentaje de préstamos adjetivales. Este resultado coincide con lo documentado en otras variedades de otomí, como la de Santiago Mexquititlán (Hekking & Bakker, 2007b; Gómez Rendón, 2008). Al respecto, se ha señalado que el bajo porcentaje de préstamos adjetivales en otomí se debe a que esta lengua no tiene una clase léxica que funcione como modificador de una frase nominal (véase §1.4.3.7 y §4.2.5). Sin embargo, este mismo argumento se ha empleado para explicar la adopción masiva de conjunciones y preposiciones. Si bien, este capítulo no tiene como objetivo plantear una escala de prestabilidad ni determinar cuáles son las motivaciones del préstamo léxico, considero que la hipótesis del “hueco funcional” resulta insuficiente para explicar la motivación del préstamo y creo que no sólo se deben considerar las restricciones estructurales, sino también las motivaciones pragmáticas y comunicativas que inciden en el proceso del préstamo.

En el corpus de este capítulo, los nombres del español son tomados en singular y con su marca de género. Algunos lexemas pueden sufrir modificaciones fónicas y morfológicas para adaptarse a la lengua otomí, de las cuales hablaré en la siguiente sección. El ejemplo (13) muestra cuatro préstamos nominales del español en su forma singular, que mantienen su género gramatical y a los que se les ha añadido las marcas de posesión, indefinitud y número.

(13) *Hindo tót'i ma trabaho tos todo tiempo to pęęfi, to kha na biahe, to paę ya mantel, to paę ya figura*

Hín = do = tót'i **ma** **trábaxo** tós tódo tyémpo
 NEG = 1PSD = encontrar **1POS** **trabajo** entonces todo tiempo

to = pęęfi to = kha **na** Ø **byáxe** to = pę
 1PRFT = trabajar 1.PRFT = hacer **INDEF** Ø **viaje** 1PRFT = vender

ja **mantél** to = pę **ja** **figura**
PL **mantel** 1PRFT = vender **PL** **figura**

‘No encontré trabajo entonces todo el tiempo trabajaba, hacia un viaje, vendía manteles, vendía figuras’

El 37.17% de los préstamos nominales del corpus fueron adaptados a la fonología del otomí y el 49.17% se mantienen sin ninguna adaptación fónica. Además, algunas palabras no sufren ningún proceso de adaptación (13.65%), debido a que su estructura fonológica cumple con las características fonológicas del otomí.¹⁰¹

TABLA 4.6 *Adaptación fonológica de los préstamos nominales en SAC y JV*

Características	Corpus general	
	F	%
Adaptación	471	37.17
Sin adaptación	623	49.17
Sin necesidad de adaptación	173	13.65
Total	1 267	100.00

Con respecto a las estrategias de adaptación, documenté fenómenos similares a los descritos por Hekking y Bakker (2010: 40) y Gómez Rendón (2008: 286). Clasifiqué los mecanismos de adaptación en tres, según la diferencia fonológica entre el español y el otomí

¹⁰¹ Es decir, que el préstamo nominal del español sólo tiene segmentos que forman parte del inventario del otomí del centro / p, b, t, d, k, g, ʔ, m, n, ɲ, s, z, ʃ, ʒ, h, ts, tʃ, y, w, i, i, u, e, o, ε, Λ, ɔ, a, ï, ã, ù / (Lastra, 1989: 21-23, 1992: 16-17). Además, el préstamo del español tiene una estructura silábica abierta CV o patrones del tipo CVC, CVCC, CCV, NCV, NCCV en fronteras morfológicas o en la primera sílaba de un bisílaba, como sucede en el otomí (Guerrero, 2015: 243-244; Lastra, 1997: 24; Palancar, 2009: 11, 30-32). Por último, el préstamo tiene el acento en la primera sílaba, como sucede en otomí (Guerrero, 2015: 236).

que buscan reducir: modificaciones segmentales, silábicas y acentuales. La primera categoría incluye las modificaciones que sufren los préstamos, debido a las diferencias que existen entre el inventario fonológico del otomí y el español. De los 471 préstamos adaptados, 249 sufren alguna modificación en el nivel segmental (52.86%).

Este conjunto comprende fenómenos como: la realización bilabial del fonema labiodental fricativo sordo / ϕ /¹⁰² (14a), la sonorización de oclusivas (14b), la sustitución del fonema fricativo velar sordo /x/ por el fricativo glotal /h/ (14c), o por el fricativo posalveolar /j/ (14d), la palatalización de /s/ (14e); el uso de /r/ en lugar de /l/ (14f). También se documentan casos de simplificación de grupos consonánticos (14g), el cierre de las vocales medias en sílabas átonas (14h); la simplificación de diptongos (14i); y la nasalización de sibilantes al inicio de palabra (14j), siendo los primeros cuatro fenómenos los más frecuentes.

- | | | | | |
|------|----|-------------------|---|---------------|
| (14) | a. | ' ϕ ies.ta | < | 'fies.ta |
| | | ' ϕ a.mi.lia | < | fa.'mi.lia |
| | b. | 'ba.do | < | 'pa.to |
| | | 'go.si.na | < | ko.'si.na |
| | c. | 'pa.re.ha | < | pa.'re.xa |
| | | 'hu.nio | < | 'xu.nio |
| | d. | 'ja.ro | < | 'xa.ro |
| | e. | 'ef.kue.la | < | es.'kue.la |
| | | 'mi.ja | < | 'mi.sa |
| | f. | pe.le.'gri.no | < | pe.re.'gri.no |
| | | 'ra.lo | < | 'ra.ro |

¹⁰² Este tipo de realizaciones prevalecen en el habla de los bilingües de San Andrés Cuexcontitlán. Lastra (1987) documentó realizaciones bilabiales [ϕ] del fonema labiodental fricativo sordo /f/ en el español de los bilingües de esta comunidad y señaló que este fenómeno es resultado de la influencia de la /ph/ del otomí. Sin embargo, Martín Butragueño (2014) afirma que en todo el dominio del español el fonema labial fricativo sordo /f/ tiene principalmente dos alófonos: uno bilabial fricativo sordo [ϕ] y otro labiodental fricativo sordo [f] (334-335). Por otro lado, los datos que Manuel Alvar presenta en *El español en México* (2010, mapa 557), muestran que, en el Estado de México, lugar donde se sitúa San Andrés Cuexcontitlán, el alófono con mayor recurrencia es el labiodental [ϕ]. Con base en esto, resulta cuestionable que esta característica fónica sea, necesariamente, un proceso de adaptación a la fonología del otomí, pues lo que se observa es una variación ya presente en el sistema fónico del español, que es propia de variantes en las que no hay contacto con dicha lengua. No obstante, decidí mantener este fenómeno en el análisis porque sería necesario hacer un análisis más detallado para demostrar que no es un efecto del contacto y descartar completamente la idea de que se trate de un proceso de adaptación.

g.	nim. 'be	<	'nom. bre
	'ngan.de	<	'gran.de
h.	'bis.pi.ra	<	'bis.pe.ra
	'bu.ru	<	'bu.ro
i	'por.ta	<	'puer.ta
j	'nsa.ba.du	<	'sa.ba.do

La categoría de adaptación silábica comprende fenómenos que buscan disminuir las diferencias que existen entre la estructura silábica del otomí y del español,¹⁰³ específicamente aquellas que tienen como objetivo obtener el patrón silábico básico del otomí (CV). Del total de casos de préstamos adaptados (471 ítems), el 64.11% sufren modificaciones silábicas (302 ítems). Cuando el préstamo del español tiene una sílaba inicial con un núcleo vocálico (V), los bilingües suelen elidirlo (15a) o insertar una fricativa glotal (15b). Por otro lado, si el préstamo nominal tiene una sílaba cerrada (CVC), los hablantes eliden la consonante en coda, ya sea en medio (15c) o a final de palabra (15d), o insertan un núcleo vocálico (15e).

- (15) a. 'ma.xen < i.'ma.xen
lu.'sion < i.lu.'sion
- b. 'ʔa.su.ga < a.'su.kar
ʔef.'kue.la < es.'kue.la
- c. 'se.tiem.bre < sep.'tiem.bre
se.le.'sion < se.lek.'sion
- d. pan.ta.'lo < pan.ta.'lon
bai.la.'ri < bai.la.'rin
'mar.te < 'mar.tes
'hue.be < 'xue.bes
lu.'ga < lu.'gar
'a.pos.to < a.'pos.tol
- e. 'kos.ko.ro.ne < kos.ko.'ro.ne

¹⁰³ Aunque ambas lenguas tienen sílabas de tipo CV, el español tiene una amplia gama de sílabas abiertas y cerradas: CV, CVC, V, CCV, VC, CCVC, VCC, CVCC, CCVCC (Quilis, 1997: 67), mientras que el inventario de sílabas en otomí se reduce a CV, CCV, CVCC y CVC, estas últimas con un mayor número de restricciones morfológicas (Guerrero, 2015: 243-244; Lastra, 1997: 24).

Por último, los procesos de adaptación acentual corresponden a movimientos de la prominencia al extremo derecho o izquierdo. De los 471 préstamos adaptados, el 39.06% (184 ítems) sufren alguna modificación acentual. En la mayoría de los casos, los bilingües desplazan el acento a la primera sílaba (80.43%: 148 ítems). En (16a), se muestran tres palabras bisilábicas agudas en las que los bilingües otomí-español desplazaron el acento a la primera sílaba, donde se alberga el acento en otomí.¹⁰⁴ En (16b), se observa el mismo fenómeno, con la diferencia de que el acento fue adelantado dos sílabas. En (16c), hay tres palabras trisilábicas graves, en las que el acento se desplazó una sílaba y, en (16d), dos sílabas. Por otro lado, el 19.57% consiste en un movimiento a la última sílaba, como se muestra en (16e).

- | | | | | |
|------|----|-----------------|---|-----------------|
| (16) | a. | 'ma.ma | < | ma.'ma |
| | | 'pa.pa | < | pa.'pa |
| | | 'ta.rea | < | ta.'rea |
| | b. | 'es.pa.jol | < | es.pa.'jol |
| | | 'al.ba.nil | < | al.ba.'nil |
| | | 'tah.ke.ria | < | tah.ke.'ria |
| | c. | 'tra.ba.xo | < | tra.'ba.xo |
| | | 'La.re.do | < | La.'re.do |
| | | 'ka.ba.jo | < | ka.'ba.jo |
| | d | 'bo.ro.na.so | < | bo.ro.'na.so |
| | | 'kuer.na.bah.ka | < | kuer.na.'bah.ka |
| | | 'tʃín.ga.da.so | < | tʃín.ga.'da.so |
| | e. | ʔo.'jo | < | 'o.jo |
| | | gru.'po | < | 'gru.po |
| | | por.'ta | < | 'pwer.ta |

¹⁰⁴ El acento en español tiene tres funciones: contrastiva, distintiva y culminativa. Existen cuatro tipos de acento, según la sílaba acentuada: el agudo cae en la última sílaba, el grave en la penúltima, el esdrújulo en la antepenúltima y el sobreesdrújulo (Hualde, 2005: 220-22). En contraste, “el acento otomí tiene una función demarcativa que permite al hablante reconocer el inicio del radical y es en esta sílaba en la que se localizan los tres contrastes tonales” (Guerrero, 2015: 236), que son uno alto (A), uno bajo (B) y otro ascendente (AB) (Guerrero, 2015: 236; Hekking & Andrés de Jesús, 1984: 19-20; Lastra, 1989: 25, 1992: 17).

Los préstamos nominales también pueden ir acompañados de diferentes determinantes, por ejemplo, el indefinido *ná* (17a), las marcas de posesivo (17b) y los demostrativos (17c). Asimismo, se les pueden afijar el morfema diminutivo *tsí* (17d) y el citativo *ná* (17e), tal como se muestra en los siguientes ejemplos.

- (17) a. *Ora sí do má na ra fiesta a sentro a San Andrés*
 ora sí do = má **na** ra **fyésta** a séntro a san andrés
 ora sí 1.PSD = ir **INDEF SG fiesta** LOC centro LOC San Andrés
 ‘Ahora sí, fui a una fiesta al centro de San Andrés’
- b. *Hímbi alkansa ka ra dómi ma pápa mi tá3*
 hímbi = bi = alkansá ka ra dómi **ma** **páhpa** mi tá
 NEG = 3.PSD = alcanzar DEM.SG SG dinero **1.POS papá** 1POS papá
 ‘No le alcanzaba el dinero a mi papá’
- c. *Bueno ro pá yo tsíanima i dura ñú pá*
 bwéno ro pá **jo tsí - ánima** i = dúra jú Ø pá
 bueno SG día **PL DIM - ánima** 1.PRS = durar tres Ø día
 ‘Bueno el día de muertos dura tres días’
- d. *Ka ra kuento bi kuentagí ma máma ke este u mi lehkuná bi k’ati hú ya baduná ke mbra nada a déhéná*
 ka ra kwénto bi = kwénta - gí ma máhma késte
 DEM.SG SG cuento 3.PSD = contar - 1.OBJ.PR 1POS mamá ke este
- u **mi** = **lehku** - **ná** bi = kʔati = hí
3.IMPF = **pequeño** - **CIT** 3.PSD = ver.de.lejos = PL.INCL
- (j)a **'ba.du - ná** ke mbra = 'na.da a **déhé - ná**
PL pato - CIT que 3IMPF.CONT = nadar LOC **agua - CIT**
- ‘El cuento que mi mamá me contó, que cuando era niña vieron de lejos unos patos que nadaban en el agua’

En el ejemplo (18), se muestra cuatro verbos del español, cuya forma comprende la raíz verbal + la vocal temática, y a los que se les ha añadido la morfología verbal del otomí.

(18) *Dimá hardinero, dimá rahpa ya pasto, dimá rega ya duuní ya planta, dimá kuida gaathó, dimá rega ya maseton*

da-má = hárdinero **ndi-mí** = **ráhpa** ja pásto **di-mí** = **regá**
 1.IMPF = jardinero 1.IMPF = **rapar** PL pasto **1.IMPF** = **regar**

ja dñí ja plánta **di-mí** = **kuída** gǒthó **di-mí** = **réga**
 PL flor PL planta **1.IMPF** = **cuidar** todo **1.IMPF** = **regar**

ja máseton

PL macetón

‘Era jardinero, rapaba el pasto, regaba las flores, las plantas, cuidaba todo, regaba los macetones’

La tercera categoría que más adoptan los bilingües otomí-español son los adverbios.

Generalmente, los hablantes insertan estas palabras de forma directa en sus narraciones, pero también pueden adaptarlas al sistema fonológico del otomí. En (19) se aprecia el segmento de una narración donde el colaborador emplea el adverbio temporal *siempre*.

(19) *Desde ndo nasegó siempre do ñható dimá ñható híndimá päädi ñampho*
 désde ndo = nasé - gó **syémpre** do = pathó
 desde 1.SIM = nacer - 1.ENFT **siempre** 1.PSD = otomí

ndi-má = pathó hín = di-má = pädi ñampho

1.IMPF = otomí NEG = 1.IMPF = saber español

‘Desde que nació, siempre he sido otomí, no sabía hablar otomí’

Los adjetivos fueron la categoría léxica menos prestada. Este comportamiento también se ha documentado en otras variedades de otomí (Gómez Rendón, 2008; Hekking & Bakker, 2010) y lo han relacionado con que el otomí posee un número reducido de adjetivos léxicos, y que la mayoría de los conceptos de propiedad en esta lengua se expresan mediante compuestos, nombres y verbos, tal como se describe en la sección §1.4.3.7.

- (20) *Na ro kámpana bi ñógí mbro grande ka ro kámpana*
 na ro kámpana bi = jó - gí
 INDEF SG campana 3.PSD = despertar - 1.OBJ.PR
- mbro grande ka ro kámpana**
3.IMPF.CONT grande DEM.SG SG campana
 ‘Una campana me despertó, una campana grande’

El préstamo de palabras funcionales del español en el otomí

En la Tabla 4.5 se observó que los bilingües otomí-español adoptan más palabras funcionales que cualquier categoría léxica. Entre los más recurrentes están la conjunción coordinativa, *y*, la disyuntiva, *o*, y la adversativa, *pero*; aunque también se documentan varios casos con la conjunción negativa, *ni*, y contrastiva, *sino*. En (21a), se presentan dos oraciones coordinadas por la conjunción adversativa, *pero* y, en (21b), una enumeración de tres frases nominales, cuyo último elemento está ligado mediante la conjunción *o*.

- (21) a. *Di négó ma hníní a la mehor buu ma ya nchala hníní pe da négó ma hníní*
 ndi = né - gó ma hníní a la mexór bñi ma
 1.PRES = querer - 1ENFT 1.POS pueblo a lo mejor 3.haber más
- ja ntʃála hníní **pe** da = né - gó ma hníní
 PL bonito pueblo **pero** 1.PRES = querer - 1.ENFT 1.POS pueblo
 ‘Quiero a mi pueblo, a lo mejor hay pueblos más bonitos, pero yo quiero a mi pueblo’
- b. *I ofrese na ra refresko na ra serbesa o ra baso ra déhé*
 i = óφrese na ra réfresko na ra sérbesa **o**
 3PRES = ofrecer INDEF SG refresco INDEF SG cerveza **o**
- ra básó ra déhé
 SG vaso SG agua
 ‘(El casero) ofrece un refresco, una cerveza o un vaso de agua’

Otra de las palabras funcionales que más emplean los bilingües otomí-español son las preposiciones, entre las que destacan por su frecuencia: *con*, *de*, *en*, *para* y *sin*. En (22a), se observa la inserción de la preposición *de* en una construcción de genitivo que, anteriormente, sólo se codificaba mediante yuxtaposición (22b). En el siguiente ejemplo, uno de los

colaboradores empleó la preposición del español, *con*, para introducir un participante con función comitativa; en lugar de la marca de número dual, que tradicionalmente era empleada para marcar construcciones aplicativas comitativas (Hekking & Bakker, 2007a: 121; Hernández Green, 2015: 229; Palancar, 2013), tal como se muestra en (22d).

- (22) a. *Ka ro bétó marte de ka ro agosto go khá ka ro fiesta*
 ka ro bétó 'mar.te **de** ka ro agosto
 DEM.SG SG primero martes **PREP** DEM.SG SG agosto

 go = khá ka ro fyésta
 3.FUT = hacer DEM.SG SG fiesta
 'El primer martes de agosto se hará la fiesta'
- b. *Go 'ini ra pelota ra futbol*
 go = 'íni **ra** **pélota** **ra** **fútbol**
 1.FUT = jugar SG **pelota** SG **futbol**
 'Iré a jugar con la pelota de futbol'
- c. *Dimá 'inihé ga pelta ko ma kompañero y pus dimá dibertihé*
 di-mé = 'íni = hé ga pélta **ko** ma kómpañero i pus
 1.IMPF = jugar = PL.EXCL de pelota **con** 1.POS compañero y pus

 di-má = dibertí = hé
 1.IMPF = divertir = PL.EXCL
 'Jugaba con mi compañero a la pelota y pus nos divertíamos'
- d. *Ma 'behñá do konose sesentaiicho do nthätibé sesentaiuno*
 ma 'béhñá do = konosé como seséntayótʃo
 1POS esposa 1PSD = conocer como sesentaiicho

 do = nthāti = **bé** seténtayúno
 1PSD = casarse = **DU.EXCL** setentaiuno

A mi esposa, la conocí como en el sesentaiicho y nos casamos en el setentaiuno

Por último, se presenta un caso de cambio de código en la narración de un bilingüe simétrico, en la que cuenta los hechos relacionados con la muerte de algunos niños del pueblo, a causa de los ataques de una bruja.

- (23) *Antes se moría muchos niños, antes xo anteriormente*
ántes se moría mút̃fos nípos, ántes ʃo **ánteríorménte**
 antes se morían muchos niños, antes pues anteriormente

mi tú góthé ya bäätsí gāathó ya bäätsí mi tú
 mi = thũ gǔthó ja bǎhtsĩ gótho ja bǎhtsĩ mi = thũ
 3.IMPF = morir todo PL niño todo PL niño 3.IMPF = morir

y este pus dúnthí ya bäätsí mi tú dúnthí
i éste pus dúnthí ja bǎhtsĩ mi = thũ dúnthí
 y este pus mucho PL niño 3.IMPF = morir mucho

pero imahinate pues existi xi k'u ya bruha
pero imahínate pues eksistí ʃi = kʔí ja brúxa
 pero imagínate existían 3.PRFT = haber PL bruja
 ‘Antes se morían muchos niños, antes pues anteriormente morían todos los niños,
 todos los niños morían y este pues muchos niños morían, muchos, pero imagínate,
 pues existían, había brujas’

En el siguiente apartado describo la distribución de los nombres patrimoniales y de origen español en el corpus de acuerdo con las variables sociolingüísticas. Posteriormente, expongo cuáles son los recursos morfosintácticos que los bilingües emplean para marcar el número en los nombres de las narraciones otomíes y cuáles son los factores lingüísticos y sociolingüísticos que los determinan.

4.4 LA MARCACIÓN DE NÚMERO EN LOS NOMBRES DE LAS NARRACIONES OTOMÍES

En la primera sección de este apartado (§4.4.1) describo la distribución de los nombres patrimoniales y del español de acuerdo con las variables sociolingüísticas. Posteriormente, en §4.4.2, muestro cuáles son los recursos morfosintácticos que los bilingües otomí-español emplean para marcar el número en los nombres de las narraciones otomíes (patrimoniales y préstamos), con la finalidad de saber si ha habido una importación directa de la morfología de número del español. En la tercera sección (§4.4.3), preciso cuáles son los factores lingüísticos y sociolingüísticos que determinan la marcación de número en los nombres de

las narraciones otomíes (patrimoniales y préstamos), para diagnosticar si ha habido una importación de los condicionamientos lingüísticos que rigen la marcación de número en español al otomí (al hilo de la HIPÓTESIS 4.1).

4.4.1 La distribución de los nombres en las narraciones otomíes

Como mencioné en la introducción de este capítulo, el español ha tenido un gran impacto en el léxico del otomí. En el corpus encontré 3 091 ocurrencias de nombres, de los cuales el 51.5% son patrimoniales y el 48.5% son de origen español. Si se revisa la distribución de los nombres por tipos, el impacto del español se vuelve aún más evidente, pues contribuye con el 63.1%, mientras que el otomí apenas comprende el 36.9%.

TABLA 4.7 *Porcentaje de palabras de origen otomí y español*

Origen	Ocurrencias		Tipos	
	F	f	F	f
Otomí	1591	51.5	252	36.9
Español	1501	48.5	431	63.1
TOTAL	3091	100.0	683	100.0

La Tabla 4.8 muestra la distribución de los nombres patrimoniales y de origen español en el corpus, de acuerdo con las variables sociolingüísticas de localidad, edad, nivel de instrucción, género y bilingüismo. En primer lugar, se puede ver que en las narraciones de los bilingües de San Andrés Cuexcontitlán predomina el uso de nombres de origen español (55.8%) sobre los patrimoniales (44.2%). Contrario a esto, en Jiquipilco el Viejo, se presenta un mayor porcentaje de nombres patrimoniales (63.1%) que de lexemas del español (36.9%). Este hecho coincide con la idea de que la comunidad de Jiquipilco es más conservadora que la de San Andrés Cuexcontitlán en cuanto a la adopción de préstamos. No obstante, en ambas comunidades se observa una fuerte presencia de nombres del español en el léxico otomí.

Los hablantes de la cuarta generación (+55 años) muestran un mayor uso de los nombres patrimoniales (54.9%), en comparación con los miembros de la tercera generación (47.4%), en quienes predomina la utilización de los nombres del español (52.6%). Esto sugiere que existe un mayor influjo del español en los hablantes de otomí más jóvenes, lo cual coincide o podría estar relacionado con el incremento en la intensidad del contacto otomí-español en las últimas décadas.

En la Tabla 4.8, también se observa una relación negativa entre el nivel de instrucción y el uso de préstamos nominales del español. Los hablantes con un nivel alto emplean menos nombres del español (31.5%) y un mayor porcentaje de nombres patrimoniales (68.5%). En contraste, los hablantes con un nivel de instrucción bajo son los que presentan un mayor uso de préstamos nominales (49.5%). Estos resultados se oponen a lo documentado en otras comunidades otomíes, en donde los hablantes con un mayor nivel de instrucción hacen un mayor uso de préstamos del español (Hekking & Bakker, 2007b: 115).

TABLA 4.8 *Porcentaje de nombres de origen otomí y español de acuerdo con las variables sociolingüísticas*

Localidad		Otomí		Español	
$\chi^2=105.073a$, gl = 1, p = 0.000					
V de Cramer= 0.184, Sig. Aproximada: 0.000					
San Andrés Cuexcontitlán	Ocurrencias	839	44.2	1061	55.8
	Tipos	153	29.8	360	70.2
Jiquipilco el Viejo	Ocurrencias	752	63.1	440	36.9
	Tipos	181	49.6	184	50.4
Edad		Otomí		Español	
$\chi^2=17.031^a$, gl = 1, p = 0.000					
V de Cramer = 0.074, Sig. Aproximada: 0.000					
Generación 3	Ocurrencias	673	47.4	746	52.6
	Tipos	148	35.5	269	64.5
Generación 4	Ocurrencias	918	54.9	755	45.1
	Tipos	196	40.4	289	59.6
Instrucción		Otomí		Español	
$\chi^2=11.866^a$, gl = 2, p = 0.003					
V de Cramer = 0.062, Sig. Aproximada: 0.003					
Bajo	Ocurrencias	1361	50.5	1335	49.5
	Tipos	230	35.9	411	64.1
Medio	Ocurrencias	180	55.7	143	44.3
	Tipos	73	47.7	80	52.3
Alto	Ocurrencias	50	68.5	23	31.5
	Tipos	18	52.9	16	47.1
Género		Otomí		Español	
$\chi^2= 57.651a$, gl = 1, p = 0.000					
V de Cramer = 0.137, Sig. Aproximada: 0.000					
Hombres	Ocurrencias	464	42.3	634	57.7
	Tipos	101	30.6	229	69.4
Mujeres	Ocurrencias	1127	56.5	867	43.5
	Tipos	211	41.8	294	58.2
Bilingüismo		Otomí		Español	
$\chi^2= 57.651^a$, gl = 1, p = 0.000					
V de Cramer = 0.137, Sig. Aproximada: 0.000					
Bil. + Otomí	Ocurrencias	694	49.5	709	50.5
	Tipos	169	38.5	270	61.5
Bil. Simétricos	Ocurrencias	786	52.4	713	47.6
	Tipos	172	38.9	270	61.1
Bil. + Español	Ocurrencias	111	58.4	79	41.6
	Tipos	31	44.9	38	55.1

Con respecto al género, los hombres emplean un mayor porcentaje de nombres del español (57.7%) y un menor número de nombres patrimoniales (42.3%). Contrario a esto, las

mujeres son las que hacen un mayor uso de los nombres patrimoniales (56.5%) que de lexemas foráneos (43.5%).

Por último, se aprecia que la mitad de los nombres empleados por los bilingües con mayor dominio del otomí son nombres del español, 50.5%, y el 49.5% son nombres patrimoniales. La presencia de préstamos del español disminuye ligeramente en el grupo de bilingües simétricos (47.6%) y desciende aún más en los bilingües con mayor dominio del español (41.6%), quienes hacen un mayor uso de nombres otomíes (58.4%). Estos datos se contraponen a los descritos en otras comunidades otomíes (SM), en donde se ha observado que los hablantes que tienen un mayor dominio del español emplean un mayor número de préstamos del español (Hekking & Bakker, 2007b: 115).

4.4.2 *Los patrones de marcación de número en los nombres de las narraciones otomíes*

Como se revisó en el primer capítulo, en las investigaciones que se han realizado sobre el otomí del centro (Lastra, 1989, 1992) se ha señalado que las marcas prenominales *ra* y *ya* codifican la categoría de número en la frase nominal. La marca *ra* aparece cuando el referente es singular y *ya* cuando es plural, como se ve en (24).

- (24) a. *Ra ngüü*
ra ngü^ǰ
SG casa
'La casa'
- b. *Ya ngüü*
ja ngü^ǰ
PL casa
'Las casas'

(Otomí de SAC, adaptado de Lastra, 1992: 18)

Sin embargo, hasta el momento no se ha realizado una investigación que estudie de manera específica la categoría de número en la variedad de otomí considerada en esta investigación. El único trabajo del que se tiene noticia es el de Acosta Aguilera (2020), quien estudia de forma detallada la sintaxis y la semántica de las marcas *ra* y *ya* en el otomí del Valle del Mezquital. En esta investigación, la autora señala que las frases nominales marcadas con *ra* ‘SG’ y *ya* ‘PL’ pueden ser argumentos verbales y funcionar como sujeto, objeto tema, objeto receptor y complemento locativo (91-94). El otomí del centro (SAC y JV) parece tener un comportamiento similar al del Valle del Mezquital, pues las frases nominales con las marcas *ra* y *ya* también pueden funcionar como sujeto (25a), objeto primario (25b), objeto secundario (25c) y complemento locativo (25e).

- (25) a. *Bi ntsääti ra hmī / ya hmī*
 bi = ndzāti **ra hmī** / **ja hmī**
 3PSD = quemar **SG tortilla** / **PL tortilla**
 ‘La(s) tortilla(s) se quemó / quemaron’
- b. *Ra t’iixú bi únibí ya hmī ra bäätsí / ya bäätsí*
 ra tʔiʃú bi = úʔni - bí ja hmī **ra bäätsí** / **ja bäätsí**
 SG señora 3PSD = dar - 3OBJ.PR PL tortilla **SG niño** / **PL niño**
 ‘La señora le(s) dio tortillas al niño/ a los niños’
- c. *Ra bäätsí bo tú ra fuui / ya fuui*
 ra bäätsí bo = tú **ra phi** / **ja phi**
 SG niño 3PSD = traer **SG sombrero** / **PL sombrero**
 ‘El niño trajo un sombrero / los sombreros’
- d. *Ya xaro bi ‘buui a ra mesa*
 ja fāro bi = ʔbhi a **ra ‘me.sa**
 PL jarro 3PSD = estar P.LOC **SG mesa**
 ‘Los jarros estaban en la mesa’

(Otomí de SAC y JV)

Acosta Aguilera (2020: 103-120) también mostró que la aparición de las marcas *ra* y *ya* es opcional en frases nominales donde otro determinante codifica la información de

número. Esto se observa en construcciones con demostrativos (proximales / no proximales) (26a) y con la marca de indefinitud 'na (26b),¹⁰⁵ pero no con los numerales, donde es obligatoria la presencia de las marcas de número, como se muestra en (26c).

(26) a. *núya ya / ø nxutsi tsits'i 'na ra t'ũ'yo*
núya (ya) nxutsi ø = tsits'i 'na ra t'ũ'yo
 DEM.PROX.PL (PL) niña 3.SUJ = llevar INDEF.SG SG perrito
 'Estas niñas llevan un perrito'
 (Otomí del Valle del Mezquital, adaptado de Acosta Aguilera, 2020: 107-108)

b. *'na ra / ø déti bi yut'i mbo ha ra ngũ*
 'na (ra) déti bi = yut'i mbo ha ra ngũ
 INDEF.SG (SG) borrego 3PSD = entrar dentro.de PREP SG casa
 'Un borrego entró a la casa'
 (Otomí del Valle del Mezquital, adaptado de Acosta Aguilera, 2020: 110)

c. *Jose bi pá góho ya / ø fáni*
 jose bi = pá **góho** ya / Ø* **fáni**
 José 3PSD = vender **cuatro** PL / Ø* **caballo**
 'José vendió cuatro caballos / caballo'
 (Otomí del Valle del Mezquital, adaptado de Acosta Aguilera, 2020: 118, 119)

Asimismo, Acosta Aguilera (2020: 134-147) indagó sobre la interpretación exclusiva e inclusiva de las marcas *ra* y *ya*. Los resultados muestran que, en contextos ascendentes (afirmaciones perfectivas),¹⁰⁶ la marca *ra* se emplea para codificar referentes singulares y la marca *ya* plural. Sin embargo, en contextos descendentes (construcciones condicionales, negativas e interrogativas), la marca *ra* también puede tener una interpretación inclusiva y referir tanto singulares como plurales, mientras que la forma *ya* sólo refiere plurales.

¹⁰⁵ En nuestro corpus también encontramos frases nominales con un núcleo nominal otomí en las que no aparecen las marcas de número *ra* y *ya* cuando está presente un demostrativo o la marca de indefinitud; sin embargo, sería necesario realizar otro tipo de pruebas para comprobar que, en efecto, la presencia de las marcas de número en presencia de otros determinantes es opcional en el otomí de SAC y JV, tal como lo mostró Acosta Aguilera (2020) en el otomí del Valle del Mezquital.

¹⁰⁶ Para mayores detalles, véase Acosta Aguilera (2020).

Por último, la autora señala que la aparición de las marcas de número no está determinada por la semántica léxica del sustantivo, ya que pueden combinarse con nombres animados e inanimados, de masa¹⁰⁷ e, incluso, con nombres propios (Acosta Aguilera, 2020: 147-161). En el otomí de SAC y JV, se halla un comportamiento similar. Las marcas de número singular y plural pueden aparecer con nombres humanos (27a), animados (27b) e inanimados (27c), de masa (27d) - (27e)¹⁰⁸ y propios (27f), tal como se muestra a continuación.¹⁰⁹

(27) a. *Ra ntsooya ra ndi noohó*

ra ndzōja ra ndi nõhó
SG señor 3PRS.CONT muy gordo
 ‘El señor es muy gordo’

(Otomí de SAC, adaptado de Lastra, 1992: 188)

b. *Bi rági ra kjuá*

bi = rági **ra khuá**
 3PSD = escapar **SG conejo**
 ‘El conejo se escapó’

(Otomí de SAC, adaptado de Lastra, 1992: 94)

c. *I tsuungí ni ra míní*

i = ts̃ni - gí ni **ra míní**
 3PRS = picar - 1OBJ.PR DEM.SG **SG espina**
 ‘La espina me pica’

(Otomí de SAC, adaptado de Lastra, 1992: 119)

¹⁰⁷ Sobre esto, Acosta Aguilera (2020: 150-157) señala que cuando los nombres de masa del otomí van acompañados de la marca plural, la frase nominal puede tener una interpretación de tipos, porciones o intensificación.

¹⁰⁸ La frase nominal de este ejemplo parece tener la interpretación de intensificación que Acosta Aguilera (2020: 150-157) documenta en el otomí del Valle del Mezquital, sin embargo, sería necesario realizar otro tipo de pruebas para poder afirmarlo.

¹⁰⁹ A pesar de las semejanzas que muestran el otomí del Valle del Mezquital (Acosta Aguilera, 2020) y el otomí de SAC y JV, no asumo que el sistema de marcación de número en las dos variedades sea necesariamente el mismo, pues para poder demostrarlo sería necesario hacer un estudio específico. No obstante, los hallazgos de Acosta Aguilera (2020) representan una base de partida importante para el análisis y la comprensión de nuestro objeto de estudio.

d. *Di né ra déhé*

di = né ra déhé
 1PRES = querer SG agua
 ‘Quiero agua’

(Otomí de SAC, adaptado de Lastra, 1992: 64)

e. *Mindikhá k'ú nuya nuya ya déhé*

mi = ndi = khá kʔi nuja nuja ja déhé
 3IMPF = INTENS = haber DEM.PL DEM.PL DEM.PL PL agua
 ‘Había mucha agua’

(Otomí de SAC, adaptado de Lastra, 1992: 304)

f. *Ná ndi ndra huntéhé gége ra Maria, ra Andres, ra Pala, ra Ino do máhmé ko mi Boni*

Ná ndi ndra = hunte = hé gége ra mária
 uno vez 1IMPF.CONT = juntar = PL.EXCL 3PRO SG María

ra ándrea ra pála ra íno do = má = hmé
 SG Andrea SG Paola SG Inocencio 1PSD = ir = PL.EXCL

ko mi'bo.ni
 con la mamá de Boni

‘Una vez nos juntamos, ellos, María, Andrea, Paola, Inocencio, fuimos a casa de la mamá de Boni’

Sin embargo, se ha observado que las marcas de número se distribuyen de manera diferente cuando el núcleo de la frase nominal es un préstamo léxico del español. En (28), se muestra algunos ejemplos de frases nominales que tienen como núcleo un lexema del español y que no presentan ninguna marca de número, a pesar de ser argumentales y funcionar como sujeto (28a), objeto secundario (28b), objeto primario (28c) y complemento adverbial (28d).

(28) a. *Ka ma mama bi tsó ke ispektor da zúhu*

ka ma máma bi = tsó ke Ø íspektor da = zíhi
 DEM.SG 1POS mamá 3PSD = asustar que Ø inspector 3S.FUT = llegar
 ‘A mi mamá le asustó que llegara el inspector’

- b. *K'ú ya maestra mi únigí clase bi tramagíhu*
 kʔí ja maestra mi = úni - gí kláse
 DEM.PL PL maestra 3IMPF = dar - 1OBJ.PR clase
 bi = 'tra(u)ma - gí = hí
 3PSD = traumar - 1OBJ.PR = PL.EXCL
 'Los maestros que me dieron clase, nos traumaron'
- c. *Híndo né go má eskuela porke min golpe alumno*
 Hín = do = né go = má Ø éskwela porke
 NEG = 1PSD = querer 3FUT = ir Ø escuela porque
 min = 'gol.pe Ø alumno
 3IMPF = golpear Ø alumno
 'No quise ir a la escuela porque (el maestro) golpeaba a los alumnos'
- d. *Pe ka ra bes do ntíhé menos de yóho ora do théhmé ka ra b'otia*
 pe ka ra bés do = ntí = hé menos de
 pero DEM.SG SG vez 1PSD = emborracharse = PL.EXCL menos
 jóhó Ø ora do = thé - Ø = hmé ka ra bʔotia
 dos Ø hora 1PSD = terminar - 3OBJ = PL.EXCL DEM.SG SG botella
 'Pero una vez nos emborrachamos, en menos de dos horas terminamos la botella'

(Otomí de SAC y JV)

La pregunta general de este capítulo es si el español ha incidido en la marcación de número nominal en otomí y si esta influencia ha sido mediante la importación directa de la morfología de número del español. Si se piensa en los recursos morfosintácticos que el otomí y el español poseen para la codificación del número en la frase nominal se podría encontrar, al menos, cuatro posibilidades. La primera es que los bilingües codifiquen el número gramatical mediante los recursos morfosintácticos del otomí (Patrón 1), tal como se muestra en los siguientes ejemplos.

- (29) a. *Nunka da pétsi ka ra oportunidad*
 núnka da = pétse ka ra oportunidad
 nunca 1PRES = tener DEM.SG SG oportunidad
 'Nunca tuve la oportunidad (de estudiar)'

- b. *K'ŷ ya bruha i buela a ma ñá*
kʔi ja brúxa i = 'bue.la a ma ñá
DEM.PL PL bruja 3PRS = volar LOC 1POS cabeza
 'Las brujas vuelan a mi cabeza'
- c. *Hímbi alkansa ka ra dómi ma pahpa mi thá*
 hím = bi = alkansa **ka ra dómi** ma páhpa mi thá
 NEG = 3.PSD = alcanzar **DEM.SG SG dinero** 1.POS papá 1.POS papá
 'No le alcanzaba el dinero a mi papá'
- d. *Túxí ya dethá*
 Ø = túʃi **ja dethá**
 2.IMP = llevar **PL maíz**
 'Lleva los maíces'

La segunda posibilidad es que los bilingües otomí-español no presenten las marcas de número del otomí: *ra* y *ya*, ni la marca plural del español *-(e)s* (Patrón 2), como se puede ver en el ejemplo (30).¹¹⁰

- (30) a. *Dimá 'yotéhé ya figura ga resina*
 Di-má = ʔyoté = hé ja 'figura ga **résina**
 1IMPF = hacer = PL.EXCL PL figura de **resina**
 'Hacia figuras de resina'
- b. *Go má xahé páxi*
 go = má = ja = hé **páʃi**
 1FUT = ir = cortar = PL.EXCL **hierba**
 'Vamos a ir a cortar hierba'

La tercera posibilidad es que el número sea codificado mediante los recursos morfológicos del español (Patrón 3), *-(e)s*, tal como se muestra en el ejemplo (31a). Por último, la codificación de número se podría realizar con los recursos morfosintácticos de las dos lenguas (Patrón 4), como se observa en (31b).

¹¹⁰ Según un miembro del comité lector, los ejemplos en (30) siguen un patrón netamente del español, pues en esta lengua los términos de masa, genéricos y los términos de preposición suelen no llevar marcas de número. Es decir, el patrón no es meramente la ausencia de marca, sino la ausencia de marcas con ciertos tipos de nominales o nominales en determinadas funciones sintácticas.

- (31) a. *Na ra muchacho ya bi xi damá ke mbri nobio pero hím bri nobios*
 na ra mutʃatʃo já bi = ʃi da-má ke mbri nóbio
 INDEF SG muchacho ya 3PSD = decir 1IMPF que 1.PRS.CONT novio
- pero **hím** = **bri** **nóbio** - **s**
 pero **NEG** = **1PRS.CONT** **novio** - **PL**
 ‘Un muchacho dijo que éramos novios, pero no éramos novios’
- b. *Xo minda ‘uuníhkí ma koskorroneshé*
 ʃo mi nda = ʔní - hkí **ma** **kóskoron** - **(e)s** = **hé**
 Pues 3IMPF INTENS = dar - 1OBJ **1POS** **coscorrón** - **PL** = **PL.EXCL**
 ‘El maestro me daba mis coscorrones’

Después de haber establecido los cuatro posibles patrones de marcación de número en los nombres del español, revisé su distribución en el corpus. En la Tabla 4.9, se puede ver que la mayoría de los nombres (patrimoniales y préstamos) reciben la morfología de número del otomí (Patrón 1). El segundo patrón más frecuente es la ausencia de marcas morfológicas de número (Patrón 2). En contraste, sólo se documenta un nombre con el morfema plural del español (Patrón 3) y otro con las marcas de número plural de las dos lenguas (Patrón 4).

TABLA 4.9 *Frecuencia de los patrones de marcación de número en los nombres de las narraciones otomíes*

Tipo de marcación	F	f
Patrón 1: <i>ra / ya + N</i>	2184	70.63
Patrón 2: N	906	28.23
Patrón 3: N + -s	1	1.09
Patrón 4: <i>ra / ya + N + -s</i>	1	0.03
TOTAL	3092	100.00

Debido a que sólo encontré dos nombres con los Patrones 3 y 4, pensé que era posible que los bilingües hayan reanalizado el sufijo plural *-(e)s* como parte de la raíz del nombre, de modo que podría reasignarlos dentro de los primeros dos patrones. Los nombres con el Patrón 3 (N-s) podrían reasignarse al segundo (N) y el único nombre con el Patrón 4 (*ra / ya + N + -s*) al primero (*ra / ya + N*).

Con base en los resultados de la Tabla 4.9, se advierte que no hay una importación directa de la morfología de número del español en el otomí. En la siguiente sección determinaré cuáles son los factores lingüísticos y sociolingüísticos que inciden en la marcación de número en los nombres de las narraciones otomíes (patrimoniales y préstamos).

4.4.3 La distribución de los patrones de marcación de número en los nombres de las narraciones otomíes

Después de haber definido cuáles son los patrones morfosintácticos que los bilingües otomí-español emplean para marcar el número en los nombres patrimoniales y los préstamos del español, procederé a determinar cuáles son los contextos, los factores lingüísticos y sociolingüísticos que permiten la aparición de las marcas de número (Patrón 1: *ra / ya + N*) y cuáles su omisión (Patrón 2 *N*). Indagaré si se emplea el Patrón 1 y 2 en los mismos contextos que se han documentado para el otomí de otras variedades de otomí (Acosta Aguilera, 2020) o si inciden otros factores relacionados con la marcación de número en español o si influyen factores de ambas lenguas (HIPÓTESIS 4.1).

Para ello, en el análisis incluí un factor relacionado con la marcación de número en la frase nominal del otomí: la presencia de otros determinantes que codifican la categoría de número (Acosta Aguilera, 2020). También tomé en cuenta algunos rasgos semánticos de los nombres que han demostrado ser relevantes en la marcación de número en español y otras lenguas: la distinción entre nombre propio y común, la concreción, la animacidad, la (dis)continuidad del nombre, la función sintáctica de la FN, la referencialidad y el tipo de constituyente.

Por último, para determinar la distribución del uso de la presencia (*ra / ya + N*) y ausencia de las marcas de número (*N*) entre los individuos incorporé las variables localidad,

edad, nivel de instrucción, género y tipo de bilingüismo. En seguida, describo en más detalle los motivos para seleccionar estos factores, la manera en que los integro al análisis y las predicciones sobre su efecto.

La presencia de determinantes que codifican la categoría de número

Quiero saber si los factores que rigen la marcación de número en los nombres del otomí de SAC y JV (patrimoniales y préstamos) son los mismos que se han documentado para el otomí de otras variedades. Para ello, en el análisis considero si los nombres son o no modificados por otros determinantes que codifican la información de número en la FN, debido a que Acosta Aguilera (2020: 103-120) encontró que las marcas de número *ra* y *ya* pueden ser opcionales cuando en la frase nominal están presentes los demostrativos o la marca de indefinitud. Dentro de este factor consideré cuatro categorías: nombres sin determinante que codifique la categoría de número, nombres sin determinante con numeral, nombres con determinante indefinido y nombres con determinantes definidos.

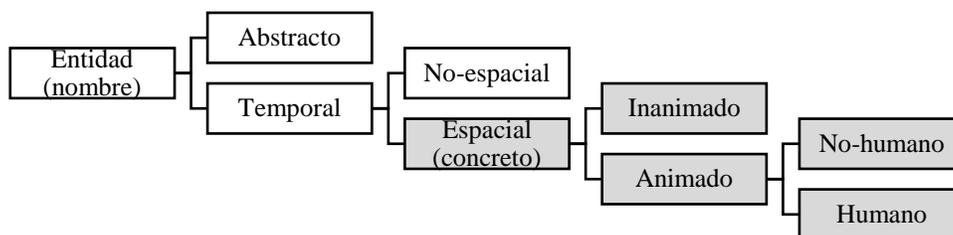
Nombres propios y comunes

La segunda distinción que planteé fue entre nombres propios y comunes. Consideré este factor porque muestra un comportamiento diferenciado en las dos lenguas. Los nombres propios del español son empleados mayormente sin determinantes ni marcas de número (Patrón 2 N) (Laca, 1999: 924-925). En contraste, los nombres propios en otomí suelen ir acompañados de la marca de número singular *ra* (Patrón 1) (Acosta Aguilera, 2020: 147-161). Tomé como nombre común a los lexemas “que expresa[n] la pertenencia de la[s] cosa[s] a alguna clase” y como nombre propio a aquellos elementos que “identifica[n] una cosa entre los demás elementos de su misma clase” (Bosque, 1999: 5).

La concreción del nombre

Para determinar si un nombre es abstracto o concreto, tomé como referencia la jerarquía de rasgos semánticos de la Figura 4.6 (Givón, 2001: 56), específicamente, las nociones de temporalidad y espacialidad. Consideraré como abstractas, aquellas entidades que no existen en el tiempo y el espacio (recuadros con fondo blanco), y como nombres concretos las entidades que existen en ambas dimensiones (recuadros en fondo gris). A pesar de que no se ha observado que la concreción del nombre determine la marcación de número en español y en otomí (Acosta Aguilera, 2020: 147-161), decidí incluir este factor porque pienso que podría tener algún efecto.

FIGURA 4.6 *Jerarquía de los principales rasgos semánticos de los nombres*



La animacidad del nombre

A pesar de que se ha observado que la animacidad del nombre no determina la marcación de número en español y en otomí (Acosta Aguilera, 2020: 147-161), consideré este factor en el análisis porque tipológicamente ha mostrado ser determinante en la organización de los sistemas de marcación de número. Corbett (2000: 54-87) plantea una escala de animacidad y señala que, en la mayoría de las lenguas, es más frecuente que se marque la categoría de

número en los elementos de mayor rango en la jerarquía: Hablante (Pronombre 1ªP) > Oyente (Pronombre 2ªP) > 3ª persona > Parentesco > Humano > Animado > Inanimado. Integro este factor al análisis mediante tres categorías: inanimado, animado y humano. Con base en las observaciones de Corbett (2000), espero que, si este factor es determinante, la aparición de las marcas de número (Patrón 1) esté asociada con los nombres de mayor animacidad y su ausencia (Patrón 2) con los de menor animacidad.

La dis(continuidad) del nombre

Resolví incluir el rasgo de \pm continuidad porque ha mostrado tener un efecto en la morfosintaxis y semántica de la categoría de número. De manera general, los nombres discontinuos pueden tener una forma singular y otra plural, mientras que los nombres continuos se presentan principalmente en singular y su pluralización implica un proceso de recategorización (Corbett, 2000: 83). Además, la marcación de número muestra un comportamiento diferenciado según la (dis)continuidad del nombre en las dos lenguas. En el español, por ejemplo, los nombres continuos tienen menos restricciones que los discontinuos para aparecer sin ningún determinante (Patrón 2). Por otro lado, en otomí, los nombres continuos parecen tener menos restricciones, pues pueden aparecer con las marcas de singular y plural (Acosta Aguilera, 2020).

En el análisis distingo entre nombres continuos y discontinuos, a partir de las nociones de “limitación” y “estructura interna”. Considero como discontinuos los nombres que son limitados y tienen estructura interna, y como nombres continuos aquellos que no son limitados y no tienen estructura interna, de modo que la principal diferencia entre estas

categorías es el rasgo \pm limitado (Jackendoff, 1991 apud Corbett, 2000: 80).¹¹¹ Por ejemplo, no se puede dividir el referente del nombre libro (discontinuo) y obtener otra entidad que pueda denominarse con la misma expresión lingüística; en cambio, sí se puede dividir el referente del nombre agua y seguir llamándolo del mismo modo.

TABLA 4.10 *Categorías semánticas de una FN (adaptado de Jackendoff, 1991 apud Corbett, 2001: 80)*

Valor de las características	Categoría
+limitado, -estructura interna	Individuos
+limitado, +estructura interna	Grupos
-limitado, -estructura interna	Sustancias
-limitado, +estructura interna	Agregados

Función sintáctica

En el análisis integré la función sintáctica de la frase nominal porque este factor es determinante en la distribución de las frases determinadas y escuetas (Patrón 2 N). En español, las frases nominales con determinantes pueden desempeñar más funciones sintácticas (argumentos y adjuntos) que las frases nominales escuetas, cuya aparición es frecuente en los complementos predicativos, pero tienen restricciones para funcionar como sujetos u objetos indirectos y directos (Laca, 1999: 894). Incluí cinco categorías: sujeto, objeto primario, objeto secundario, atributo, adjunto.

¹¹¹ Algunos autores señalan que la distinción entre nombres continuos y discontinuos en el nivel léxico es inadecuada y que esta diferenciación se da en el nivel de la frase nominal, es decir, que se trata más de una cuestión sintáctica (Allan, 1980: 566 apud Corbett, 2000: 82). Sin embargo, en este análisis se considera que el rasgo \pm continuo sí es propio a los nombres.

Referencialidad

Seleccioné este factor porque es determinante en la distribución de las frases determinadas y escuetas. En el análisis hice la distinción entre frases nominales referenciales y atributivas. Consideré que un nombre era empleado de forma referencial cuando “el hablante usa la frase nominal con el fin de que su oyente pueda saber de qué o quién se habla” (Donellan, 1966 apud García Fajardo, 2016: 56-57). Valoré que el nombre era empleado atributivamente “si el hablante usa la frase nominal describiendo las propiedades de quien quiera que las cumpla” (Donellan, 1966 apud García Fajardo, 2016: 57).

Términos de preposición

Otro de los factores que tomé en cuenta fue si el nombre era el núcleo de una frase nominal o el núcleo del término de una preposición, debido a que uno de los contextos en donde más ocurren los nombres escuetos en español es en los términos de preposición (Laca, 1999: 911).

Localidad

La primera variable sociolingüística que integré al análisis es la comunidad de origen de los bilingües otomí-español. Dada las diferencias sociolingüísticas entre San Andrés Cuexcontitlán y Jiquipilco el Viejo, descritas en §2.2, supongo que los bilingües otomí-español de JV emplearán la morfología del otomí para codificar el número en los nombres con mayor recurrencia que los bilingües de SAC, quienes tenderán a omitir estas marcas.

Edad

Los bilingües otomí-español pertenecen a dos grupos etarios: adultos (35-54 años) y mayores (más de 55 años). Como revisamos en §2.2, los hablantes de mayor edad en SAC y JV son

quienes tienen un mayor dominio del otomí. En vista de la HIPÓTESIS 4.3, espero que los bilingües mayores empleen con mayor frecuencia las marcas *ra* y *ya* (Patrón 1) para codificar el número; mientras que los más jóvenes (adultos) tenderán a omitirlas (Patrón 2).

Instrucción

Asimismo, incluí el nivel de instrucción, con tres niveles: bajo (menos de 6 años escolares), medio (7-11 años) y alto (más de 12 años escolares). Resolví integrar este factor porque el nivel de instrucción ha sido un factor importante en el planteamiento de algunas tipologías de bilingüismo (Palacios, 2005), debido a que los años de escolaridad pueden afectar el conocimiento que los hablantes tienen de las lenguas, con mucha frecuencia en favor del español y en detrimento de la lengua originaria. Al respecto, espero que los hablantes con mayor nivel de instrucción y dominantes en español harán un menor uso de las marcas de número *ra* y *ya*; en contraste, los hablantes que tengan como lengua dominante en otomí y menor nivel de instrucción tenderán a hacer un mayor uso de las marcas de número.

Género

Por último, consideré el género del colaborador. Como señalé anteriormente, la diferencia entre el habla masculina y femenina no está basada en diferencias biológicas, sino en los roles sociales que el hombre y la mujer tienen en una comunidad de habla determinada (Labov, 2001: 263). En San Andrés Cuexcontitlán y Jiquipilco el Viejo hay algunas diferencias en los roles sociales que desempeñan mujeres y hombres, principalmente en las generaciones mayores, pues esta disparidad ha disminuido en las nuevas generaciones. Los hombres forman parte de redes laborales dentro y fuera de la comunidad, lo cual ha generado en ellos un mayor dominio del español. En contraste, las mujeres son amas de casa y suelen

moverse en redes familiares dentro de la comunidad, razón por la que suelen presentar un mayor dominio y uso de la lengua otomí.

Bilingüismo

El factor sociolingüístico más importante del análisis es el grado de bilingüismo. Integré este factor al análisis mediante tres de los grupos lingüísticos que identifiqué a partir de las pruebas de habilidades bilingües (véase §2.2.3): bilingües con mayor dominio del otomí (BO), bilingües simétricos (BS) y bilingües con mayor dominio del español (BE). Con base en la HIPÓTESIS 4.3, espero que los BO hagan un mayor uso de las marcas de número y los BE usen una menor proporción de estos.

En la Tabla 4.11 presento un resumen con las variables lingüísticas y extralingüísticas que emplearé en el análisis y sus variantes.

TABLA 4.11 *Factores de análisis y predicciones*

Factor	Variantes		
(i) Determinantes	FN. s/determinante FN. s/determinante + numeral FN. determinante indefinido FN. determinante definido	(viii) Constituyente	Núcleo Término de prep.
(ii) Propios	Comunes Propios	(ix) Localidad	San Andrés Jiquipilco el Viejo
(iii) Concreción	Concreto Abstracto	(x) Edad	Generación 3 Generación 4
(iv) Animacidad	Inanimado Animado Humano	(xi) Instrucción	Alto Medio Bajo
(v) (Dis)continuidad	Continuo Discontinuo	(xii) Género	Masculino Femenino
(vi) Función	Sujeto Objeto primario Objeto secundario C. adverbial C. nominal Predicativo	(xiii) Bilingüismo	BE BS BO
(vii) Referencialidad	F. referencial F. atributiva		

En el siguiente apartado §4.4.4 presento los resultados de un análisis general. En la primera parte describo la distribución de los patrones de marcación de número según los factores lingüísticos y extralingüísticos antes descritos. Además, proporciono los resultados de la prueba de Chi-cuadrado de Pearson (χ^2) y V de Cramer (véase §2.1.3.1 - §2.1.3.2), que indican si los factores lingüísticos y extralingüísticos seleccionados inciden de manera significativa en la marcación de número, así como la fuerza de esta asociación. Posteriormente, proporciono los resultados del análisis multivariado de efectos fijos, que utilicé para evaluar la contribución simultánea de las diferentes variables lingüísticas y extralingüísticas en la variación de la marcación de número (Labov, 1996: 114-155), con el propósito de determinar si los factores asociados al contacto con el español inciden en la marcación de número en el otomí de contacto y así poder rechazar o aceptar la HIPÓTESIS 4.1.

4.4.4 Análisis descriptivo de la marcación de número en los nombres de las narraciones otomíes

En la Tabla 4.12 se presenta la distribución de los patrones de marcación de número en los nombres de las narraciones otomíes de acuerdo con los factores lingüísticos: la presencia de determinantes que codifican el número, la distinción entre nombre propio y común, la concreción, la animacidad, la (dis)continuidad del nombre, la función sintáctica, la referencialidad de la frase nominal y el tipo de constituyente.

TABLA 4.12 *Distribución de los patrones de marcación de acuerdo con los factores lingüísticos*

Determinantes con rasgos de número $\chi^2=346.922^a$, gl = 3, p = 0.000 V de Cramer = 0.335, Sig. aproximada: 0.000	Patrón 1 (ra / ya + N)		Patrón 2 (N)	
	F	%	F	%
FN c / Determinante definido	674	92.5	55	7.5
FN c / Determinante indefinido	231	86.5	36	13.5
FN s / Determinante	1271	62.3	770	37.7
FN s / Determinante + Numeral	9	16.4	46	83.6
Nombres propios y comunes $\chi^2= 77.939$, gl = 1, p = 0.000 V de Cramer = 0.157, Sig. aproximada: 0.000	Patrón 1 (ra / ya + N)		Patrón 2 (N)	
	F	%	F	%
Propio	91	43.8	117	56.2
Común	2094	72.6	790	27.4
Concreción $\chi^2= 241.360^a$, gl = 1, p = 0.000 V de Cramer = 0.279, Sig. aproximada: 0.000	Patrón 1 (ra / ya + N)		Patrón 2 (N)	
	F	f	F	f
Concreto	1767	78.4	486	21.6
Abstracto	418	49.8	421	50.2
Animacidad $\chi^2= 81.203^a$, gl = 2, p = 0.000 V de Cramer = 0.162, Sig. aproximada: 0.000	Patrón 1 (ra / ya + N)		Patrón 2 (N)	
	F	f	F	f
Animado	97	87.4	14	12.6
Humano	768	79.8	194	20.2
Inanimado	1320	65.4	699	34.6
Contable $\chi^2= 80.959^a$, gl = 1, p = 0.000 V de Cramer = 0.162, Sig. aproximada: 0.000	Patrón 1 (ra / ya + N)		Patrón 2 (N)	
	F	%	F	%
Contable	1678	75.2	552	24.8
No contable	507	58.8	355	41.2
Función $\chi^2= 378.693^a$, gl = 5, p = 0.000 V de Cramer = 0.502, Sig. aproximada: 0.000	Patrón 1 (ra / ya + N)		Patrón 2 (N)	
	F	%	F	%
Sujeto	682	93.6	47	6.4
Objeto Secundario	60	89.6	7	10.4
Objeto Primario	924	83.5	182	16.5
Compl. Nominal	82	51.6	77	48.4
Adverbial	356	45.7	423	54.3
Predicativo	81	32.1	171	67.9
Referencialidad $\chi^2= 254.45$, gl = 1, p = 0.000 V de Cramer = 0.286, Sig. Aproximada: 0.000	Patrón 1 (ra / ya + N)		Patrón 2 (N)	
	F	%	F	%
Frases nominales referenciales	1597	80.41	389	19.58
Frases nominales atributivas	588	53.16	518	46.83

Constituyente $\chi^2=238.745^a$, gl = 1, p = 0.000 V de Cramer = 0.278, Sig. Aproximada: 0.000	Patrón 1 (<i>ra / ya</i> + N)		Patrón 2 (N)	
	F	%	F	%
Núcleo nominal	1981	76.2	620	23.8
Término de preposición	204	41.5	287	58.5

Número $\chi^2=245.469^a$, gl = 1, p = 0.000 V de Cramer = 0.282, Sig. Aproximada: 0.000	Patrón 1 (<i>ra / ya</i> + N)		Patrón 2 (N)	
	F	%	F	%
Pronominal	515	99.2	4	0.8
Nominal	1670	64	903	35.1

Origen del nombre $\chi^2=16.056^a$, gl = 1, p = 0.000 V de Cramer = 0.072, Sig. Aproximada: 0.000	Patrón 1 (<i>ra / ya</i> + N)		Patrón 2	
	F	%	F	%
Patrimonial	1175	73.9	416	26.1
Préstamos	1010	67.3	491	32.7

El primer factor lingüístico que incluí en el análisis fue la presencia/ausencia de otros determinantes que codifican el número gramatical. En las frases nominales con determinantes definidos predomina la presencia de las marcas de número (92.5%), como se muestra en (32a), y sólo en 7.5% de los casos están ausentes. Las frases nominales con la marca de indefinitud *'na* muestran un comportamiento similar, pero el porcentaje de ausencias aumenta a 13.5% (32b). En las frases nominales sin determinante, la ausencia de las marcas *ra* y *ya* alcanza el 37.7% de los nombres (32c). Este porcentaje incrementa radicalmente en los sintagmas nominales con numerales, en los que sólo el 16.4% de los nombres van acompañados de las marcas de número (32d).

- (32) a. *Bi gustagé ka ra hníní*
bi = gusta - gé **ka ra hníní**
3.PSD = gustar - 1OBJ **DEM.SG SG pueblo**
‘Me gustó el pueblo’
- b. *Bi náhkí na ra dūyúní*
bi = ná - hkí **na ra dñí**
3.PSD = dar - 1OBJ.PR **INDEF SG flor**
‘Me dio una flor’

- c. *Di pá di pá biahe*
 di = pá di = pá **byáxe**
 1.PRS = ir 1.PRS = ir **viaje**
 ‘Voy, voy de viaje’
- d. *I ‘badi yóho semana para ka ro fiesta*
 i = ʔbodi **jóho** Ø **sémana** para ka ro fiesta
 3PRS = faltar **dos** Ø **semana** para DEM.SG SG nombre
 ‘Faltan dos semanas para la fiesta (del pueblo)’

La prueba de Chi-cuadrada de Pearson indica que existe una relación estadísticamente significativa entre la presencia/ausencia de las marcas de número y la presencia/ausencia de determinantes que codifican el número (χ^2 , $p < 0.05$), con un nivel de fuerza moderado (V de Cramer de 0.335). Sin embargo, la distribución de los datos se opone a lo documentado en otras variedades de otomí. Mientras en el OVM las marcas de número son obligatorias con los numerales y opcionales con la marca de indefinitud y los demostrativos, en SAC y JV la ausencia de las marcas de número *ra* y *ya* (Patrón 2: N) no está relacionada con la presencia de los determinantes definidos e indefinidos, sino con las frases nominal sin determinantes o con numerales. Estos resultados tampoco coinciden con el patrón de marcación del español del centro de México, en donde la presencia del sufijo plural *-s* es obligatoria con los numerales. No obstante, concuerda con lo descrito para el español de los bilingües simétricos y dominantes en otomí (véase § 3.5).

Las pruebas estadísticas muestran que la marcación de número depende de la distinción entre nombre propio y común (Fisher, $p < 0.05$), con un nivel de fuerza bajo (V de Cramer: 0.157). La presencia de las marcas de número (Patrón 1: *ra* / *ya*) es más recurrente

en los nombres comunes (72.6%), como se ejemplifica en (33a), y la mayoría de los nombres propios no presenta marcas de número (52.6%),¹¹² como se muestra en (33b).

- (33) a. *Da 'yot'e ya mádá*
 da = ʔyoté **ja mádá**
 1.PRS = hacer **PL molcajete**
 ‘Hago molcajetes’
- b. *Thó xi 'ké Ros*
 thó ʃi - ʔké **'ros**
 INTERR decir - 1.OBJ **Ros**
 ‘¿Qué me dices Ros?’

Con respecto a la concreción del nombre, se aprecia una diferencia importante entre los nombres concretos y abstractos. La presencia de las marcas de número (Patrón 1: *ra / ya*) es mayor en los nombres concretos (78.4%), como se muestra en (34a); mientras que en los nombres abstractos la distribución de los patrones es similar (Patrón 1: 49.8%, Patrón 2: 50.2%) (34b).¹¹³ Las pruebas estadísticas indican que la concreción del nombre afecta de manera significativa la marcación de número (Fisher, $p < 0.05$), con un nivel de fuerza bajo (V de Cramer: 0.279). Los resultados coinciden con las predicciones en torno a este factor.

- (34) a. *Bi k'ati ka ma ntsooya na ra 'behñáná*
 bi = kʔátsi ka ma ntsʔya **na ra ʔbéhñá ná**
 3.PSD = ver DEM.SG 1POS señor **INDEF SG señora CIT**
 ‘Dice mi esposo que vio una señora’

¹¹² Uno de los miembros del comité lector menciona que los nombres vocativos suelen presentar patrones divergentes a los argumentales (p. ej. en español o catalán, donde suelen llevar artículo definido, este nunca se presenta en el vocativo). Por ello, es posible que en otomí la función vocativa no esté favoreciendo la presencia de las marcas.

¹¹³ Uno de los miembros del comité lector señala que suele presentarse un traslape entre los factores concreción y (dis)continuidad, pues los nombres abstractos suelen ser continuos en una proporción mayor a los nombres concretos. En el análisis no realicé una clasificación que conjugara de forma simultánea ambos factores con clasificaciones del tipo “nombre abstracto continuo” *versus* “nombre abstracto discontinuo”. No obstante, sí cuide el posible empalme de estos factores mediante el uso de pruebas de colinealidad.

- b. *Luego nda kháhké daño*
 luego nda = khá - hké **dáño**
 luego 1.IMPF = hacer - 1.OBJ **daño**
 ‘Luego me hacía daño’

Aunque la animacidad no es un factor determinante para la marcación de número ni en español y ni en otomí, las pruebas estadísticas señalan que existe una relación significativa entre este factor y el tipo de patrón (χ^2 , $p < 0.05$), con un nivel de fuerza bajo (V de Cramer: 0.162). La presencia de las marcas de número (Patrón 1: *ra / ya*) es más recurrente en los nombres animados (87.4%), seguidos de los nombres humanos (79.8%) y, en último lugar, los inanimados (65.4%), en los que la ausencia de marcas de número comprende el 34.6% de los casos, tal como se ejemplifica en (35). Los resultados concuerdan parcialmente con la escala de animacidad (Corbett, 2000), pues en los nombres inanimados hay una mayor ausencia de las marcas de número (Patrón 2: N) y una mayor presencia de estas en los nombres animados y humanos, aunque el orden entre estas dos categorías no es el esperado.

- (35) a. *Do ekstraña nge do nú ka ra charro*
 do = ekstrañá nge do = nú **ka ra t̃járo**
 3PSD = extrañar porque 3PSD = ver **DEM.SG SG charro**
 ‘Se extrañó porque vio el charro’
- b. *Nda yega nda foodí nara dóró*
 nda = jéga nda = ðödí **na ra dóró**
 1IMPF = llegar 1IMPFV = cuidar **INDEF SG toro**
 ‘Llegaba, cuidaba un toro’
- c. *Do tági ta hqi*
 do = tógi ta **hñi**
 1.PSD = caer hasta **tierra**
 ‘Caí hasta el piso’

Las pruebas estadísticas sugieren que existe una asociación significativa entre el tipo de marcación de número y la (dis)continuidad de los nombres, con un nivel de fuerza bajo

(V de Cramer: 0.162). Los resultados concuerdan con lo predicho, la mayoría de los nombres discontinuos presenta las marcas de número (Patrón 1 *ra / ya*: 75.2%), como se observa en (36a); y en los nombres continuos la distribución de los patrones es similar (Patrón 1: 58.8%, Patrón 2: 41.2%) (36b).

- (36) a. *Dimá k'átsi ya mbagho*
 di-má = kʔátsi ja mbõho
 1.IMPF = ver PL **señor de ciudad**
 'Veía a los señores'
- b. *Dimá 'té 'yot'ehé ya figura ga resina*
 di-má = ʔyoté = hé ja figura ga **resina**
 1.IMPF = hacer = PL.EXCL PL figura de **resina**
 'Hacíamos figuras de resina'

La prueba de Chi-cuadrada de Pearson muestra que la función sintáctica incide de forma significativa en la marcación de número (χ^2 , $p < 0.05$), con un nivel de fuerza moderado (V de Cramer: 0.502). Se observa una mayor presencia de las marcas de número (Patrón 1: *ra / ya*) en las frases nominales que funcionan como sujeto (93.6%), objeto primario (89.6%) y secundario (83.5%), tal como se ilustra en (37a) - (37c). Contrario a esto, la ausencia de marcas de número (Patrón 2: N) es más recurrente en las frases nominales que fungen como complemento nominal (48.4%), complemento adverbial (54.3%) y complemento predicativo (67.9%), como se ejemplifica en (37d) - (37f).

- (37) a. *Ka ra maestro ya mi ña español*
ka ra maestro yá mi = ɲa español
DEM.SG SG maestro ya 3.IMPF = hablar español
 'El maestro hablaba español'
- b. *Gi kúhki ya thá*
 gi = kúhki **ja thá**
 2.PRS = bajar PL **mazorca**
 'Baja las mazorcas'

- c. *Do rá tú rá hmĩ ka ra tʔĩfú*
do = tú rá hmĩ **ka** **ra** **tʔĩfú**
1.PSD = traer 3.POS tortilla **DEM.SG** **SG** **señora**
‘Le traje su taco a la señora’
- d. *Bundo fũudí do mágó a eskuela*
bu ndo = fũdí do = má - gó **a** **éjkwela**
cuando 1.SIM = empezar 1.PSD = ir - 1.ENFT **LOC** **escuela**
‘Cuando empecé a ir a la escuela’
- e. *Mágé gi má ñäägí ya bote ga aguardiente*
má - gé gi = má ñǎgí ja bóte **ga** **aguardiente**
ir - 2.ENFT 1.PRS = ir esconder PL bote **de** **aguardiente**
‘Vete, ve a esconder los botes de aguardiente’
- f. *Porke damá pęęfĩ porke damá hardinero*
porke da-má pęęfĩ porke **da-má** **hárdinero**
porque 1.IMPF trabajar porque **1.IMPF** **jardinero**
‘Porque trabajaba, porque era jardinero’

Otro de los factores lingüísticos que se tomó en cuenta fue el tipo de constituyente. Las pruebas estadísticas sugieren que este factor está asociado significativamente con el uso de los patrones de marcación (Fisher, $p < 0.05$), con un nivel de fuerza bajo (V de Cramer: 0.278). Los bilingües emplean con más frecuencia las marcas de número (Patrón 1: *ra / ya*) en los nombres que son núcleo de frases nominales (76.2%), como se ejemplifica en (38a), y tienden a omitirlas (Patrón 2: N) en los nombres que son término de preposición (58.5%), como se ve en (38b).

- (38) a. *Dimá paa ya losa ya xaro ya ts'ũĩ*
di-má = pǎ **já** **lósa** **ja** **járó** **ja** **ts'ũĩ**
1.IMPF = vender **PL** **plato** **PL** **taza** **PL** **olla**
‘Vendía platos, tazas, ollas’
- b. *Dimá 'inihé ga pelta*
di-mé = ʔini = hé ga **pélta**
1.IMPF = jugar = PL.EXCL de **pelota**
‘Jugábamos a la pelota’

Las pruebas estadísticas señalan que el origen de los nombres está asociado con los patrones de marcación de número (Fisher, $p < 0.05$), en una magnitud baja (V de Cramer: 0.072) (HIPÓTESIS 4.2). El Patrón 1: *ra / ya* es más recurrente en los nombres patrimoniales (73.9%), como se ejemplifica en (39a), y el porcentaje del Patrón 2: N es más frecuente (32.7%) en los préstamos nominales del español, tal como se ilustra en (39b).

- (39) a. *Dimá túxí ra tuhmĩ*
 di-má = tújí **ra tuhmĩ**
 1.IMPF = llevar **SG pan**
 ‘Llevaba el pan’
- b. *Ndra mágó chimixa*
 ndra = má - gó **t̃jí - míja**
 1.IMPF.CONT = ir - 1.ENFT **DIM - misa**
 ‘Estaba yendo a misita’

Por último, en los nombres se hizo una distinción entre los nombres con marcas de número nominal y pronominal. Los nombres con número pronominal tienen un comportamiento casi categórico, pues los nombres carecen de marcas de número (Patrón 2: N) sólo en cuatro ocasiones. Por ello, en los siguientes análisis sólo consideraré los nombres con número nominal.

- (40) a. *Híndo t̃at’í ka ma trabaxo*
 hín - do = t̃t’í ka ma trábaxo
 NEG - 1PSD = encontrar DEM.SG 1.POS trabajo
 ‘No encontré trabajo’
- b. *Híngi má eskuela ya*
 Hín - gi = má éskwela já
 NEG - 2.PRS = ir escuela ya
 ‘No vayas a la escuela ya’

En la Tabla 4.13 se presenta la distribución de los patrones de marcación de acuerdo con las variables localidad, edad, instrucción, género y grado de bilingüismo. Las pruebas

estadísticas muestran que hay una relación de dependencia entre la marcación de número y la comunidad de los colaboradores (Fisher, $p < 0.05$), con un nivel de fuerza bajo (V de Cramer: 0.036). Este resultado coincide con las expectativas planteadas, pues los hablantes de San Andrés Cuexcontitlán son los que utilizan un mayor porcentaje de nombres escuetos (Patrón: N) (30.6%), en comparación con los hablantes de Jiquipilco el Viejo (27.3%).

TABLA 4.13 *Distribución de los patrones de marcación de acuerdo con los factores sociolingüísticos*

Localidad	Patrón 1 (ra / ya + N)		Patrón 2 (N)	
	F	%	F	%
$\chi^2 = 4.005^a$, gl = 1, p = 0.047 V de Cramer = 0.036, Sig. Aproximada: 0.045				
San Andrés Cuexcontitlán	1318	69.4	582	30.6
Jiquipilco el Viejo	867	72.7	325	27.3
Edad	Patrón 1 (ra / ya + N)		Patrón 2 (N)	
	F	%	F	%
$\chi^2 = 7.369^a$, gl = 1, p = 0.003 V de Cramer = 0.077, Sig. Aproximada: 0.003				
Generación 3	1037	73.1	382	26.9
Generación 4	1148	68.6	525	31.4
Instrucción	Patrón 1 (ra / ya + N)		Patrón 2 (N)	
	F	%	F	%
$\chi^2 = 5.014^a$, gl = 2, p = 0.082 V de Cramer = 0.040, Sig. Aproximada: 0.082				
Alto	60	82.2	13	17.8
Medio	231	71.5	92	28.5
Bajo	1894	70.3	802	29.7
Género	Patrón 1 (ra / ya + N)		Patrón 2 (N)	
	F	%	F	%
$\chi^2 = 0.000^a$, gl = 1, p = 1.000 V de Cramer = 0.000, Sig. Aproximada: 0.994				
Hombres	776	70.7	322	29.3
Mujeres	1409	70.7	585	29.3
Bilingüismo	Patrón 1 (ra / ya + N)		Patrón 2 (N)	
	F	%	F	%
$\chi^2 = 15.407^a$, gl = 2, p = 0.000 V de Cramer = 0.071, Sig. Aproximada: 0.000				
Bil. +Esp	158	83.2	32	16.8
Bil. +Otomí	985	70.2	418	29.3
Bil. Simétricos	1042	69.5	457	30.5

Las pruebas estadísticas muestran que hay una relación de dependencia entre la marcación de número y la edad de los colaboradores (Fisher, $p < 0.05$), con una magnitud baja (V de Cramer: 0.077). Sin embargo, la distribución de los datos no concuerda con las expectativas, pues son los hablantes de la tercera generación los que más emplean las marcas de número (Patrón 1: *ra / ya*), con un 73.1% de los casos, mientras que los hablantes de la cuarta generación son quienes más nombres escuetos (Patrón 2: N) usan (31.4%).

El comportamiento del factor nivel de instrucción no corresponde con las expectativas planteadas, pues los nombres escuetos (Patrón 2: N) son más recurrentes en los hablantes con nivel de instrucción bajo (29.7%) y medio (28.5%), y menor en los hablantes con un nivel de instrucción alto (17.8%). La prueba de Chi-cuadrada de Pearson indica que el nivel de instrucción de los colaboradores no está asociado de forma significativa con la marcación de número (χ^2 , $p > 0.05$).

La distribución de los nombres escuetos (Patrón 2: N) es idéntica en hombres y mujeres (70.7%) y la prueba estadística señala que no hay una relación de dependencia entre el género de los colaboradores y la marcación de número (Fisher, $p > 0.05$). Por último, la alternancia en la marcación de número mantiene una relación de dependencia con el tipo de bilingüismo (χ^2 , $p < 0.05$), con un nivel de fuerza bajo (V de Cramer: 0.071). Sin embargo, los resultados se oponen a las expectativas planteadas, pues los bilingües simétricos (BS) y dominantes en otomí (BO) hacen un mayor uso de nombres sin marcas de número (Patrón 2: N), con un 30.5% y 29.3% de los casos, mientras que en los bilingües con mayor dominio del español (BE) este porcentaje disminuye hasta 16.8%.

Los resultados parecen coincidir con la idea de la existencia de diferentes sistemas de marcación de número en otomí, al menos porcentualmente: el de los bilingües con un importante dominio del otomí (BO, BS) y el de los hablantes que tienen como lengua

dominante el español (BE). Sin embargo, estos apenas son indicios, pues es necesario realizar más pruebas para determinar si la diferencia porcentual entre estos grupos es significativa estadísticamente y si estos podrían constituir dos sistemas, no sólo por cuestiones cuantitativas, sino por los factores lingüísticos que los organizan y determinan (HIPÓTESIS ESPECÍFICA 4.3).

En resumen, los factores lingüísticos que rigen la marcación de número en los nombres de las narraciones otomíes son: la presencia de determinantes que codifican el número, la distinción entre nombre propio y común, la concreción, la animacidad y la (dis)continuidad del nombre, la función sintáctica y la referencialidad de la frase nominal, el tipo de constituyente y el origen histórico del lexema. La distribución mostró que los nombres comunes, concretos, animados, discontinuos y de origen otomí, así como las frases nominales con función nuclear y referencial favorecen la aparición de las marcas de número.

Por otro lado, los resultados muestran que el comportamiento de la variedad de otomí en contacto con español (SAC y JV) difiere de otras variedades de otomí (OVM), con respecto al efecto que tiene la presencia de determinantes que codifican el número, la distinción nombre propio y común, la animacidad y el tipo de constituyente. Estas diferencias pueden ser propias de la variación geográfica o estar relacionadas con el contacto que el otomí de SAC y JV han mantenido con el español. Debido a que no se cuenta con una descripción previa sobre la marcación de número en el otomí de SAC y JV, es difícil poder afirmar en qué medida las discrepancias encontradas entre el otomí del centro y el otomí del Valle del Mezquital se deben o no al contacto. Por ello, será necesario revisar el comportamiento de la marcación de número en los nombres según su origen histórico. Ello permitirá saber si hay una diferencia en el comportamiento de la marcación del número en

los nombres patrimoniales y de origen español (HIPÓTESIS 4.2) y así discernir si existe un efecto del contacto con el español (HIPÓTESIS 4.1).

En cuanto a la distribución de los patrones de marcación de número, se encontró que esta presenta relaciones de dependencia con las variables de comunidad, edad y tipo de bilingüismo. Las personas de la generación mayor (G4) y dominantes en la lengua otomí (BO, BS) son los que obtuvieron un mayor porcentaje de nombres sin marcas de número (Patrón 2: N), mientras que los de la tercera generación y cuya lengua dominante es el español (BE) mostraron un mayor uso de las marcas de número (Patrón 1: *ra/ya*). Aun así, la diferenciación entre los bilingües con mayor dominio del español (BE) y los dominantes en otomí (BO, BS) sugiere la factibilidad de la HIPÓTESIS 4.3, que predice que la fuerza del efecto del español está asociada con el nivel de bilingüismo de los hablantes. Sin embargo, para poder afirmarlo será necesario hacer otras pruebas y comprobar que, más allá de la diferenciación porcentual, existen discrepancias en cuanto a los factores lingüísticos que los organizan.

Estos primeros resultados (frecuencias, porcentajes, resultados de Chi-cuadrado de Pearson y V de Cramer) proporcionan un panorama general de la relación que algunos factores lingüísticos y extralingüísticos guardan con nuestro objeto de análisis. A continuación, realizo un análisis regresión logística escalonada (§2.1.3.4) para evaluar la contribución simultánea de diferentes variables lingüísticas y extralingüísticas en la marcación de número, y así establecer la dirección del efecto y la jerarquía entre estos.

4.4.5 Análisis inferencial de la marcación de número en los nombres de las narraciones

otomíes

En el apartado anterior presenté la distribución general de los datos y proporcioné los resultados de la prueba de Chi-cuadrado de Pearson que aprobaba o rechazaba la existencia de relaciones bivariantes. Sin embargo, en la lengua, los factores lingüísticos y extralingüísticos no actúan de forma individual sino en conjunto. Por ello, será necesario evaluar el efecto que el conjunto de los factores lingüísticos y extralingüísticos tienen en la marcación de número, y establecer una jerarquía entre ellos, mediante una prueba de regresión logística escalonada (véase §2.1.3.4).

Para construir este modelo consideré como variable dependiente la presencia (*ra / ya* + N) y ausencia (N) de las marcas de número y como variables explicativas los factores lingüísticos y extralingüísticos que se describieron en la sección anterior (véase § 4.4.3): la presencia de determinantes que codifican el número, la distinción entre nombre propio y común, el tipo de nombre, la animacidad, la (dis)continuidad, la función sintáctica, la referencialidad, el tipo de constituyente, el origen histórico del nombre, la localidad, la edad, la instrucción, el género y el bilingüismo.

Antes de llevar a cabo la regresión logística escalonada realicé una prueba para cerciorarme de que no existiera colinealidad entre las variables independientes (i-xiv), pues esto podría generar errores en el análisis. Para ello, se empleó la medida de Goodman y Kruskal del paquete “The GoodmanKruskal” en la plataforma de R (véase §2.1.3.3). Los resultados se presentan en una escala de 0 a 1, en donde un valor más cercano a 1 indica mayor correlación entre las variables y más probabilidad de problemas de colinealidad.

En la Figura 4.7 se puede apreciar que la mayoría de los factores explicativos no guardan relación entre sí. Sólo se aprecia cierta asociación entre el nivel de bilingüismo y el nivel de instrucción, pues el 0.44 de la variabilidad en el nivel de bilingüismo es explicable por la variabilidad en el nivel de instrucción; sin embargo, sólo el 0.08 de la variabilidad en los niveles de instrucción es explicable por la variabilidad en el nivel de bilingüismo. Por ello, decidí mantener ambos factores.

FIGURA 4.7 Resultados de la medición Goodman y Kruskal. Nombres en las narraciones

otomíes

	Sexo	Edad	Instrucción	Bilingüismo	Localidad	Lengua	Función	Constituyente	Tipo	Propio	Animación	Contable	Determinantes	Referencialidad
Sexo	K = 3 0.01	0.04	0.03	0.23	0.02	0	0.01	0	0	0	0	0	0	0
Edad	0.01	K = 2	0.01	0.19	0	0	0	0	0	0	0	0	0	0
Instrucción	0.04	0.01	K = 2	0.08	0.03	0	0	0	0	0	0	0	0	0
Bilingüismo	0.07	0.23	0.44	K = 3	0.05	0	0	0	0	0	0	0	0	0
Localidad	0.22	0	0.03	0.04	K = 2	0.04	0.01	0.01	0	0	0	0.01	0	0.01
Lengua	0.02	0	0	0	0.04	K = 2	0	0.01	0.01	0.01	0	0	0	0
Función	0.01	0.01	0	0	0.02	0.03	K = 6	0.31	0.07	0.05	0.2	0.03	0.06	0.13
Constituyente	0.01	0	0	0	0.01	0.01	0.11	K = 2	0	0.01	0.02	0	0.02	0
Tipo	0	0	0	0	0	0.01	0.02	0	K = 2	0.03	0.02	0.05	0	0.09
Propio	0	0	0	0	0	0.01	0.01	0.01	0.03	K = 2	0.03	0.01	0	0.06
Animación	0	0.01	0	0	0	0	0.06	0.03	0.04	0.04	K = 3	0.06	0	0.02
Contable	0	0	0	0	0.01	0	0	0	0.05	0.01	0.04	K = 2	0.02	0.02
Determinantes	0	0	0	0	0.01	0	0.02	0.03	0.03	0.01	0	0.03	K = 4	0.1
Referencialidad	0	0	0	0	0.01	0	0.02	0	0.09	0.06	0.02	0.02	0.06	K = 2

Después de este proceso realicé la regresión logística por pasos con las siguientes variables: la presencia de determinantes que codifican el número, la distinción entre nombre propio y común, el tipo de nombre, la animacidad, la (dis)continuidad, la función sintáctica, la referencialidad, el tipo de constituyente, el origen histórico del nombre, la localidad, la edad, la instrucción, el género y el bilingüismo. A continuación, se muestran los resultados.¹¹⁴

TABLA 4.14 *Resultados de la regresión logística escalonada. Nombres en las narraciones otomíes*

Variab les	Estimado	Error Est.	Valor z	Pr(> z)
Intersección	3.49247	1.68926	2.067	0.038691 *
FN s / Determinante + Numeral	-3.78699	0.46228	-8.192	2.57e-16 ***
FN s / Determinante	-1.89743	0.17148	-11.065	< 2e-16 ***
FN c / Determinante indefinido	-0.33866	0.25446	-1.331	0.183229
Nombre concreto	0.96935	0.12573	7.709	1.26e-14 ***
F. Sujeto	2.13802	0.20179	10.595	< 2e-16 ***
F. Objeto Primario	1.70059	0.15481	10.985	< 2e-16 ***
F. Objeto Secundario	1.68829	0.47985	3.518	0.000434 ***
F. Compl. Nominal	0.40103	0.22628	1.772	0.076350 .
F. Predicativa	-0.38922	0.21885	-1.779	0.075320 .
Nombre propio	-1.25991	0.21949	-5.740	9.46e-09 ***
Término de preposición	-0.84208	0.16311	-5.163	2.43e-07 ***
Instrucción medio-alto	-0.96761	0.24683	-3.920	8.85e-05 ***
Frasas referenciales	0.34612	0.12524	2.764	0.005715 **
San Andrés Cuexcontlá	-0.40950	0.13986	-2.928	0.003413 **
Bilingües + otomí	-1.20695	0.36858	-3.275	0.001058 **
Bilingües simétricos	-1.08452	0.35420	-3.062	0.002200 **
Género: femenino	-1.79646	1.63783	-1.097	0.272705
Género: masculino	-1.40858	1.64268	-0.857	0.391175
Nombre discontinuo	0.24930	0.12330	2.022	0.043178 *
Nombre patrimonial	0.05471	0.11401	0.480	0.631313

Códigos de significación: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

¹¹⁴ En la regresión logística se tuvo una variable de respuesta dicotómica, presencia (Patrón 1: *ra / ya*) versus ausencia (Patrón 2: N) de las marcas de número. En la primera columna de la tabla se encuentran las variables más relevantes según el modelo aplicado. En la columna con la etiqueta “Estimado” se ofrecen los resultados sobre el sentido de la relación en una escala de log(odds). Los valores en el intervalo de 0 a +∞ indican una relación positiva, es decir, un incremento en el uso de las marcas de número *ra / ya*. Los valores en la escala de 0 a -∞ señalan una relación negativa, es decir, un decrecimiento en la probabilidad de que se emplee las marcas de número *ra / ya*. En las siguientes dos columnas se proporcionan los valores de error estándar y de Z. En la última columna se presentan los valores de probabilidad, que indican si la variable en cuestión es significativa, de acuerdo con un nivel de significación menor a 0.05.

Los resultados muestran que, de las catorce variables consideradas, diez resultaron significativas, en el siguiente orden de importancia: (i) la presencia de determinantes que codifican el número, (ii) la función sintáctica, (iii) la distinción entre nombre propio y común, (iv) el tipo de nombre, (v) el nivel de instrucción, (vi) el tipo de constituyente, (vii) la referencialidad, (viii) el tipo de bilingüismo, (ix) la comunidad y (x) la (dis)continuidad.

En cuanto a la presencia de determinantes que codifican el número, las FNs sin determinante + numeral (-3.78699) y, en menor proporción, las FNs sin determinantes (-1.89743) inhiben el uso de las marcas de número *ra* y *ya*. El segundo factor relevante fue la función sintáctica de la frase nominal. Los resultados sugieren que las funciones de sujeto (2.13802), objeto primario (1.70059) y secundario (1.68829) favorecen el uso de las marcas de número (Patrón 1: *ra* / *ya*). El tercer factor en la jerarquía fue la distinción nombre propio y común, y son los nombres propios (-1.25991) los que inhiben la utilización de las marcas de número (Patrón 1: *ra* / *ya*) y promueven la ausencia de las marcas (Patrón 2: N). Por otro lado, los nombres concretos (0.96935) favorecen la aparición de las marcas de número (Patrón 1: *ra* / *ya*) y retraen el segundo patrón.

El nivel de instrucción es el quinto factor en la jerarquía. Los hablantes con un nivel de instrucción medio-alto no favorecen el uso de las marcas de número (-0.96761) y son quienes emiten más nombres sin marcas. El sexto factor fue el tipo de constituyente, los nombres que son término de preposición inhiben la aparición de las marcas de número (-0.84208) y promueve el uso de nombres sin marcas (Patrón 2: N). En contraste, el uso de las marcas de número (Patrón 1: *ra* / *ya*) se ve favorecido en las frases referenciales (0.34612).

La siguiente variable fue el tipo de bilingüismo de los hablantes. Al respecto, se observa que los hablantes con mayor dominio del otomí (-1.20695) y los bilingües simétricos (-1.08452) inhiben el uso de las marcas de número (Patrón 1: *ra* / *ya*), ya que son los que

emiten más nombres sin codificar el número (Patrón 2: N). El penúltimo factor en la jerarquía es la comunidad de los hablantes, los hablantes de San Andrés Cuexcontitlán son quienes menos uso hacen utilizan las marcas *ra* y *ya* (-0.40950). El último factor relevante para la marcación de número fue la (dis)continuidad del nombre, los nombres discontinuos promueven (0.24930) la aparición de las marcas de número (Patrón 1: *ra / ya*).

En vista de los resultados, surge el cuestionamiento acerca de su contribución a las hipótesis de este capítulo. Con respecto a la HIPÓTESIS 4.1 –que sugiere que la aparición o ausencia de las marcas de número en otomí está determinada por los factores que rigen la expresión del número en español–, se observa que el factor más importante en la marcación de número del otomí es la presencia/ausencia de determinantes que codifican el número, tal como sucede en otras variedades del otomí (OVM); sin embargo, el efecto de los determinantes es el opuesto. No obstante, como mencioné en la sección anterior, las diferencias pueden ser propias de la variación geográfica o estar relacionadas con el contacto con el español, por lo que será necesario realizar otro tipo de análisis.

En cuanto a la HIPÓTESIS 4.2, que señala que la fuerza del efecto del español depende del origen histórico de los nombres. Los resultados de la regresión logística escalonada no indican que el origen histórico de los nombres determine la presencia o ausencia de las marcas de número *ra* y *ya*. Sin embargo, para saber si hay una diferencia en el comportamiento de la marcación del número en los nombres patrimoniales y de origen español (HIPÓTESIS 4.2) y así discernir si existe un efecto del contacto con el español (HIPÓTESIS 4.1), será necesario revisar el comportamiento de la marcación de número en los nombres según su origen histórico.

Por otro lado, que el factor nivel de bilingüismo resultara decisivo en la alternancia de la marcación de número abona en favor de la HIPÓTESIS 4.3, que sugiere que la fuerza del

efecto del español está asociada con el nivel de bilingüismo de los hablantes. Sin embargo, para poder afirmarlo será necesario hacer otras pruebas y comprobar que, más allá de la diferenciación cuantitativa, existen discrepancias en cuanto a los factores lingüísticos que organizan

Con base en los primeros resultados, consideré pertinente plantear un análisis de la marcación de número de acuerdo con el origen histórico de los nombres (patrimonial *versus* préstamo), que permita aceptar o rechazar las HIPÓTESIS 4.1 y 4.2. En el siguiente apartado se exponen los resultados del análisis de la marcación de número según el origen histórico de los nombres.

4.5 LA MARCACIÓN DE NÚMERO SEGÚN EL ORIGEN HISTÓRICO DE LOS LEXEMAS

4.5.1 La marcación de número en los nombres patrimoniales del otomí

Los nombres del otomí son emitidos con las marcas de número singular y plural (Patrón 1: *ra / ya*) el 68.14% de las veces y aparecen sin las marcas (Patrón 2: N) en un 31.85%.

TABLA 4.15 *Distribución de los patrones de marcación en los nombres patrimoniales*

Tipo de marcación	F	f
Presencia de las marcas <i>ra / ya</i> + N	890	68.14
Ausencia de las marcas N	416	31.85
TOTAL	1306	100.00

Posteriormente, utilicé una regresión logística escalonada (véase §2.1.3.4) para determinar qué factores lingüísticos y extralingüísticos inciden en la presencia o ausencia de las marcas de número y cuál es la jerarquía entre ellos. Para construir el modelo consideré como variable dependiente la presencia (Patrón 1: *ra / ya*) y ausencia de las marcas de número

(Patrón 2: N) y como variables explicativas los factores lingüísticos y extralingüísticos descritos en §4.4.3): la presencia de determinantes que codifican el número, la distinción entre nombre propio y común, la concreción, la animacidad y la (dis)continuidad del nombre, la función sintáctica y la referencialidad de la frase nominal, el tipo de constituyente, la localidad, la edad, el nivel de instrucción, el género y el tipo bilingüismo de los colaboradores.¹¹⁵ A continuación, en la Tabla 4.16 se muestran los resultados.¹¹⁶

¹¹⁵ Antes de llevar a cabo el análisis estadístico inferencial, realicé tablas de contingencia en las que crucé la variable dependiente (Presencia *versus* ausencia de las marcas de número) con las variables lingüísticas y extralingüísticas descritas en §4.3.1.3. También calculé los estadísticos Chi-cuadrada de Pearson y V de Cramer (véase §2.1.3.1 y §2.1.3.2), cuyos resultados se encuentran en el Apéndice 4A. Asimismo, utilicé la medida de Goodman y Kruskal del paquete The GoodmanKruskal en la plataforma de R para cerciorarme que no existiera colinealidad entre las variables independientes (véase §2.1.3.3). Los resultados se encuentran en el Apéndice 4B, debido a que no se observaron relaciones de colinealidad entre las variables independientes, incluí todas en la regresión logística escalonada.

¹¹⁶ En la regresión logística se tuvo una variable de respuesta dicotómica, presencia (Patrón 1: *ra / ya*) *versus* ausencia (Patrón 2: N) de las marcas de número. En la primera columna de la tabla se encuentran las variables más relevantes según el modelo aplicado. En la columna con la etiqueta “Estimado” se ofrecen los resultados sobre el sentido de la relación en una escala de log(odds). Los valores en el intervalo de 0 a $+\infty$ indican una relación positiva, es decir, un incremento en el uso de las marcas de número *ra / ya*. Los valores en la escala de 0 a $-\infty$ señalan una relación negativa, es decir, un decrecimiento en la probabilidad de que se emplee las marcas de número *ra / ya*. En las siguientes dos columnas se proporcionan los valores de error estándar y de Z. En la última columna se presentan los valores de probabilidad, que indican si la variable en cuestión es significativa, de acuerdo con un nivel de significación menor a 0.05.

TABLA 4.16 *Resultados de la regresión logística escalonada. Nombres patrimoniales*

Variables	Estimado	Error Est.	Valor z	Pr(> z)
Intersección	1.2000	0.5734	2.093	0.036356 *
FN s / Determinante + Numeral	-3.0120	0.5830	-5.166	2.39e-07 ***
FN s / Determinante	-1.3969	0.2395	-5.832	5.48e-09 ***
FN c / Determinante indefinido	-0.5096	0.3445	-1.479	0.139100
F. Sujeto	1.6629	0.2667	6.234	4.54e-10 ***
F. Objeto primario	1.4958	0.2209	6.771	1.28e-11 ***
F. Objeto secundario	2.1006	0.8060	2.606	0.009155 **
F. Predicativa	-1.1516	0.3073	-3.748	0.000178 ***
F. Compl. nominal	-0.4836	0.4081	-1.185	0.236007
Nombre concreto	1.1790	0.1788	6.594	4.27e-11 ***
Nombre discontinuo	0.6792	0.1726	3.935	8.32e-05 ***
Nombre propio	-1.2161	0.3479	-3.495	0.000474 ***
Término de preposición	-0.7707	0.2479	-3.109	0.001875 **
Instrucción medio-alto	-0.6750	0.3157	-2.138	0.032531 *
Bilingües + Otomí	-1.0945	0.4876	-2.245	0.024793 *
Bilingües simétricos	-0.8732	0.4663	-1.873	0.061134 .
Localidad: SAC	-0.2792	0.1683	-1.659	0.097111 .

Códigos de significación: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 '.' 1

Los resultados de la regresión logística escalonada muestran que, de las trece variables consideradas, ocho resultaron significativas, en el siguiente orden de importancia: (i) la presencia o ausencia de determinantes que codifican el número, (ii) la función sintáctica de la frase nominal, (iii) la concreción, (iv) la (dis)continuidad, (v) la distinción común/propio, (vi) el tipo de constituyente, (vii) el nivel de instrucción y (viii) el tipo de bilingüismo.

El primer factor en la jerarquía fue la presencia o ausencia de determinantes que codifican el número. Las frases nominales sin determinantes con un numeral (-3.0120), las frases nominales sin determinantes (-1.3969) y, en menor proporción, las frases nominales con un determinante indefinido (-0.5096) inhiben el uso de las marcas de número *ra* y *ya*, como se ejemplifica en (41a), (41b) y (41c), respectivamente. En contraste, la mayoría de las

frases nominales que tienen un determinante definido (91.3%) van acompañadas de las marcas de número (Patrón 1: *ra / ya*), como se puede apreciar en (41d).

- (41) a. *Pida dura ñú pá ka ro fíesta*
 pi - da = 'dura **ñú pá** ka ro fíesta
 otra vez - 3.FUT = durar **tres día** DEM.SG PL fíesta
 'Otra vez durará tres días la fiesta'
- b. *Do pęfí mi hää dómi*
 do = pęfí mi = hä **dómi**
 1.PSD = trabajar 1.IMPF = tener **dinero**
 'Trabajé, tenía dinero'
- c. *Dimá päänti ko na ra tsühkwa k'ü ya tako*
 di-má = pänti ko **na ra** tsühkwa kʔi ja táko
 1.IMPF = envolver con **INDEF SG papel** DEM.PL PL taco
 'Envolvía los tacos con un papel'
- d. *Bi má yootí ka ra déhé*
 bi = má = yotí **ka ra déhé**
 3.PSD = ir = sacar **DEM.SG SG agua**
 'Fue a sacar el agua'

Como se había observado en el análisis general, donde se incluyó tanto los nombres patrimoniales como los préstamos del español, el uso de las marcas de número en el otomí de SAC y JV difiere de lo documentado en el otomí del Valle del Mezquital, en donde el uso de las marcas de número es obligatorio con los numerales y opcional con determinantes y la marca de indefinición.

El segundo factor en la jerarquía es la función gramatical de la frase nominal. Las frases nominales que funcionan como sujeto, objeto primario y secundario favorecen el uso de las marcas de número (Patrón 1: *ra / ya*), como se puede observar en los ejemplos de (42a) - (42c). Contrario a esto, la función predicativa inhibe su aparición y promueve el empleo de nombres sin marcas (Patrón 2: N), como se aprecia en (42d).

- (42) a. *Gaathó k'ú yo khääní ba íhí*
gǎthó kʔi jo khǎní ba = ʔíhí
todo DEM.PL PL persona 3.PSD = venir
 ‘Todas las personas vinieron’
- b. *Mi yódí ka ra tsühkwa ma certificado*
 mi = yódí **ka ra tsühkwa** ma sértifikado
 3.IMPF = pedir **DEM.SG SG papel** 1.POS certificado
 ‘Me pedía mi papel, mi certificado’
- c. *Nda koohmé a Múndaq*
 nda = kǒ = hmé a ʔmúndǒ
 1.IMPF = regresar = PL.EXCL LOC **México**
 ‘Cuando regresamos de México’
- d. *Ambi pétsi ya arsina ra nduhté*
 ʔam bi = pétsi ja ársina ra **ndihté**
 INTERR 3.PSD = tener PL arcina 3.PRS.CONT **grande**
 ‘Dónde tenía las arsinas grandes’

La concreción del nombre es el tercer factor en la jerarquía. Los nombres concretos fomentan el uso de las marcas de número (Patrón 1: *ra / ya*), como se ejemplifica en (43a). En contraste, los nombres abstractos ocurren mayormente sin marcas de número (Patrón 1: *ra / ya* = 39.6%; **Patrón 2: 60.4%**), tal como se muestra en (43b).

- (43) a. *To paq ya 'oni*
 to = pǒ **ya ʔáni**
 1.PERF = vender **SG pollo**
 ‘He vendido pollos’
- b. *Da tsúhu mbra ndasinto*
 da = tsíhi **mbra** = **nda** - **sinto**
 3.FUT = llegar **3.IMPF.CONT** = **INTENS** - **malo**
 ‘Llegará muy mal’

La siguiente variable es la (dis)continuidad del nombre. Los nombres discontinuos favorecen el empleo de las marcas *ra* y *ya* (Patrón 1), como se ejemplifica en (44a). Los nombres continuos, por su parte, se presentan sin marcas de número el 46.2% de las veces, tal como se muestra en (44b).

- (44) a. *Ya riko mbgho porque mi háhtéhé ya traje*
 ja ríko mbgho porke mi = háhté = hé **ja tráxe**
 PL rico señor.de.ciudad porke 3.IMPF = vestir = PL.EXCL **PL traje**
 ‘Los señores de ciudad ricos porque vestían trajes’
- b. *Bundo naségó siempre do ñható*
 bu ndo = nasé - gó siempre do = **ñathó**
 cuando 1.PSD = nacer - 1.ENF siempre 1.PSD = **otomí**
 ‘Cuando nací yo siempre fui otomí’

El quinto factor en la jerarquía es la distinción entre nombre común y propio. Los nombres comunes favorecen el uso de *ra / ya + N* (Patrón 1), como se puede apreciar en (45a). Por otro lado, en los nombres propios, se presenta un menor porcentaje de marcas de número (57.7%) e incrementa la ausencia de estas (42.3%), como se ejemplifica en (45b).

- (45) a. *Pida huutse ra ‘oni*
 pi - da = hĩtse **ra ?áni**
 otra vez - 3.FUT = cocinar **SG pollo**
 ‘Otra vez se cocina el pollo’
- b. *Híndi má pqaahmé a Ntsuuhni*
 hín - di = má = pǎ = hmé a **ntsĩhni**
 NEG - 1.PRES = ir = vender = PL.EXCL LOC **Toluca**
 ‘No vamos a vender en Toluca’

El tipo de constituyente es la siguiente variable, los nombres que son término de preposición retraen el uso de las marcas de número y favorecen su ausencia (Patrón 2: N), como se ejemplifica en (46a). En contraste, se observa un mayor uso de las marcas de número (Patrón 1: *ra / ya*) en los nombres que son núcleo de una frase (72.5%), como se ejemplifica en (46b).

- (46) a. *Dimá ‘yótho ga wá’ó’*
 di-má = ?yótho **ga wá-?ó?**
 1.IMPF = andar **de pata de perro**
 ‘Andaba de pata de perro’

- b. *Bi náhkí na ra dóhni*
 bi = ná - hkí **na** **ra** **dóhni**
 3.PSD = dar - 1.OBJ **INDEF** **SG** **flojera**
 ‘Me daba flojera’

El nivel de instrucción es la séptima variable en la jerarquía. Los hablantes con nivel medio-alto no favorecen el uso de *ra* y *ya* (Patrón 1) y, por ende, emiten más nombres sin marcas de número (Patrón 2: N). El último factor en la jerarquía es el tipo de bilingüismo, y se observa que los bilingües con mayor dominio del otomí -en comparación con los otros tipos de bilingües- no favorecen el uso del Patrón 1: *ra* / *ya* y promueven la aparición de nombres sin especificación de número. En la siguiente sección revisaremos qué factores lingüísticos y sociales promueven o inhiben el uso de *ra* y *ya* en los préstamos del español de las narraciones otomíes.

4.5.2 La marcación de número en los nombres del español que han ingresado al otomí

En la Tabla 4.17 se muestra la distribución de los patrones de marcación en los préstamos nominales del español. El 61.56% de los lexemas del español están acompañados de las marcas de número singular (*ra*) y plural (*ya*) y el 38.43% no codifican el número.

TABLA 4.17 *Los patrones de marcación de número en los préstamos nominales del español*

Tipo de marcación	F	f
Presencia de las marcas <i>ra</i> / <i>ya</i> + N	780	61.56
Ausencia de las marcas N	487	38.43
TOTAL	1267	100.00

Posteriormente, utilicé una regresión logística escalonada (véase §2.1.3.4) para determinar qué factores lingüísticos y extralingüísticos inciden en la presencia o ausencia de las marcas de número en los préstamos y cuál es la jerarquía entre ellos. Para construir el modelo consideré como variable dependiente la presencia (Patrón 1: *ra* / *ya*) y ausencia de

las marcas de número (Patrón 2: N) y como variables explicativas los factores lingüísticos y extralingüísticos descritos en §4.4.3: la presencia de determinantes que codifican el número, la distinción entre nombre propio y común, la concreción, la animacidad y la (dis)continuidad del nombre, la función sintáctica y la referencialidad de la frase nominal, el tipo de constituyente, la localidad, la edad, el nivel de instrucción, el género y el tipo bilingüismo de los colaboradores.¹¹⁷ A continuación, se muestran los resultados.¹¹⁸

¹¹⁷ Antes de llevar a cabo el análisis estadístico inferencial, realicé tablas de contingencia en las que crucé la variable dependiente (Presencia versus ausencia de las marcas de número) con las variables lingüísticas y extralingüísticas descritas en §4.3.1.3. También calculé los estadísticos Chi-cuadrada de Pearson y V de Cramer (véase §2.1.3.1 y §2.1.3.2), cuyos resultados se encuentran en el Apéndice 4C. Asimismo, utilicé la medida de Goodman y Kruskal del paquete The GoodmanKruskal en la plataforma de R para cerciorarme que no existiera colinealidad entre las variables independientes (véase §2.1.3.3). Los resultados se encuentran en el Apéndice 4D, debido a que no se observaron relaciones de colinealidad entre las variables independientes, incluí todas en la regresión logística escalonada.

¹¹⁸ En la regresión logística se tuvo una variable de respuesta dicotómica, presencia (Patrón 1: *ra / ya*) versus ausencia (Patrón 2: N) de las marcas de número. En la primera columna de la tabla se encuentran las variables más relevantes según el modelo aplicado. En la columna con la etiqueta “Estimado” se ofrecen los resultados sobre el sentido de la relación en una escala de log(odds). Los valores en el intervalo de 0 a $+\infty$ indican una relación positiva, es decir, un incremento en el uso de las marcas de número *ra / ya*. Los valores en la escala de 0 a $-\infty$ señalan una relación negativa, es decir, un decrecimiento en la probabilidad de que se emplee las marcas de número *ra / ya*. En las siguientes dos columnas se proporcionan los valores de error estándar y de Z. En la última columna se presentan los valores de probabilidad, que indican si la variable en cuestión es significativa, de acuerdo con un nivel de significación menor a 0.05.

TABLA 4.18 Resultados de la regresión logística escalonada. Préstamos

VARIABLES	ESTIMADO	ERROR EST.	VALOR Z	PR(> Z)
Intersección	-7.10654	535.41233	-0.013	0.989410
FN s / Determinante + Numeral	-4.90927	0.89852	-5.464	4.66e-08 ***
FN s / Determinante	-2.37954	0.24791	-9.599	< 2e-16 ***
FN c / Determinante indefinido	-0.03082	0.39544	-0.078	0.937869
F. Sujeto	2.87314	0.34739	8.271	< 2e-16 ***
F. Objeto primario	1.92916	0.22741	8.483	< 2e-16 ***
F. Compl. Nominal	0.81266	0.28811	2.821	0.004792 **
F. Objeto secundario	1.14803	0.68199	1.683	0.092306 .
F. Predicativa	0.37018	0.31635	1.170	0.241933
Nombre propio	-1.40361	0.29695	-4.727	2.28e-06 ***
Nombre concreto	0.83005	0.17622	4.710	2.47e-06 ***
Término de preposición	-0.96784	0.22922	-4.222	2.42e-05 ***
Frasas referenciales	0.62709	0.18370	3.414	0.000641 ***
Instrucción medio-alto	-1.24775	0.42416	-2.942	0.003264 **
Bilingües simétrico	-1.53679	0.58908	-2.609	0.009087 **
Bilingües + otomí	-1.55026	0.61209	-2.533	0.011318 *
San Andrés Cuexcontitlán	-0.36558	0.20990	-1.742	0.081562 .
Género: masculino	9.99581	535.41197	0.019	0.985105

Códigos de significación: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

El primer factor en la jerarquía es la presencia o ausencia de determinantes que codifican la categoría de número. Las frases nominales sin determinantes con numerales (-4.90927) y las frases nominales sin determinantes (-2.37954) inhiben el uso de las marcas de número (Patrón 1: *ra / ya + N*) y favorecen el uso de nombres sin marcas (Patrón 2: *N*), como se ejemplifica en (47a) y (47b). En contraste, las frases nominales con determinantes indefinidos y definidos llevan las marcas de número más del 90% de las veces (véase Apéndice 4C), como se puede apreciar en la segunda dupla de ejemplos.

- (47) a. *Go 'ópi da regalagí kũuta béxo*
 go = ?ó - pi da = regalá - gí **kĩtá** **béfo**
 1.FUT = pedir - BEN 3.FUT = regalar - 1.OBJ **cinco** **peso**
 'Voy a ir a pedir que me regale cinco pesos'

- b. *Do khapi kaso ka ma kú*
do = kha - pi **káso** ka ma kú
1.PSD = hacer - 3.OBJ **caso** DEM.SG 1.POS hermano
‘Le hice caso a mi hermano’
- c. *Entos buyi na ra tiempo en ke da rebela*
entos **bïi na ra tyémpo** en ke da = rébela
entonces haber **INDEF SG tiempo** en que 1.PSD = rebelar
‘Entonces llegó el tiempo en el que me rebelé’
- d. *Híndimá entiendehé ka ra español*
hín - di-má entiende = hé **ka ra español**
NEG - 1.IMPF entender = PL.EXCL **DEM.SG SG español**
‘No entendía el español’

La siguiente variable es la función sintáctica de la frase nominal, la función de sujeto (2.87314), objeto primario (1.92916) y complemento nominal (0.81266) promueven el uso del primer patrón: *ra* y *ya*, y disminuye la posibilidad de que los nombres sean emitidos sin marcas, como se puede ver en los ejemplos (48a) y (48b). Por otro lado, el uso de las marcas de número disminuye en las frases nominales con función adverbial o predicativa, tal como se advierte en (48c) y (48d).

- (48) a. *Mi zagní na ra bebe*
mí = zǎní **na ra bebé**
3.IMPF = llorar **INDEF SG bebé**
‘Lloraba un bebé’
- b. *Rara ka ra ‘behña ba tú ra refresko*
‘rara ka ra ʔbehñá ba = tú **ra réfresko**
rara DEM.SG SG señora 3.PRS = traer **SG refresco**
‘Rara la señora que trae un refresco’
- c. *Mi ‘yó ka kaye*
mí = ʔyó **ka káje**
3.IMPF = andar **DEM.SG calle**
‘Anda en esa calle’

- d. *Ka ra persona mbra charróná*
 ka ra pèrsona **mbra** = **tjáro** - ná
 DEM.SG SG persona **3.IMPF.CONT** = **charro** - CIT
 ‘La persona era un charro, dice’

El tercer factor en la jerarquía es la distinción entre nombre común y propio. Los nombres propios retraen el uso de las marcas de número (-1.40361), como se muestra en (49a). Los nombres comunes, por su parte, llevan las marcas de número el 64.61% de las veces (véase Apéndice 4C), como se aprecia en (49b).

- (49) a. *Do máhmé Akapulko*
 Do = má = hmé **Akapúlko**
 1.PSD = ir = PL.EXCL **Acapulco**
 ‘Fuimos a Acapulco’
- b. *Bi yíxkí dimá pēfī ka ra obra*
 bi = jíj - kí di-má pēfī **ka ra óbra**
 1.PSD = llevar - 1.OBJ 1IMPF trabajar **DEM.SG SG obra**
 ‘Me llevó, trabajaba en la obra’

La concreción del nombre es la cuarta variable seleccionada. Los nombres concretos favorecen (0.83005) el uso de las marcas de número (Patrón 1: *ra / ya*) e inhiben la ausencia de estas (Patrón 2), como se puede apreciar en (50a). Por otro lado, en los nombres abstractos, los patrones de marcación tienen una distribución cercana al 50% (véase Apéndice 4C), como se ilustra en (50b).

- (50) a. *Después da pone ra fruta*
 después da = poné **ra frúta**
 después 3.FUT = poner **SG fruta**
 ‘Después se pone la fruta’
- b. *Di päädígó ñahé de gaathó kosa*
 di = pãdi - gó = ja = hé de **gǎthó kósa**
 1.PSD = saber - 1.ENFT = hablar = PL.EXCL de **todo cosa**
 ‘Sabíamos hablar de todas las cosas’

El tipo de constituyente es el quinto factor más importante. Los nombres que son término de preposición disminuyen la posibilidad de que se usen las marcas de número (-0.96784) y favorecen su ausencia, como se ejemplifica (51a). En contraste, la aparición de las marcas de número (Patrón 1: *ra / ya*) es más recurrente en nombres que son núcleo de una frase nominal (70.3%) (véase Apéndice 4C), como se aprecia en (51b).

- (51) a. *Ma pahpa mi tói ya aso ga marketa*
 ma páhpa mi = tói ja aso **ga markéta**
 1.POS papá 3.IMPF = comprar PL azúcar **de marketa**
 ‘Mi papá compraba el azúcar por marqueta’
- b. *Mbra hñé ka ra pantalón ka ra chamarra*
 mbrá = hñé ka ra pántalon **ka ra t̃ámara**
 3.IMPF.CONT = vestir DEM.SG SG pantalón **DEM.SG SG chamarra**
 ‘Vestía el pantalón y la chamarra’

La referencialidad de la frase nominal es el sexto factor más importante. Las frases nominales referenciales promueven el empleo del primer patrón: *ra / ya + N*, como se advierte en (52a). En contraste, la ausencia de marcas (Patrón 2) es más recurrente en las frases nominales con una interpretación atributiva (52.69%), como se ilustra en (52b).

- (52) a. *Dimá atende ka ra jardín*
 di-má = atendé **ka ra xardín**
 1.IMPF = atender **DEM.SG SG jardín**
 ‘Atendía el jardín’
- b. *Bi gustagí do má a baile*
 bi = gusta - gí do = má a **báile**
 3.PSD = gustar - 1.OBJ 1.PSD = ir LOC **baile**
 ‘Me gustaba ir a los bailes’

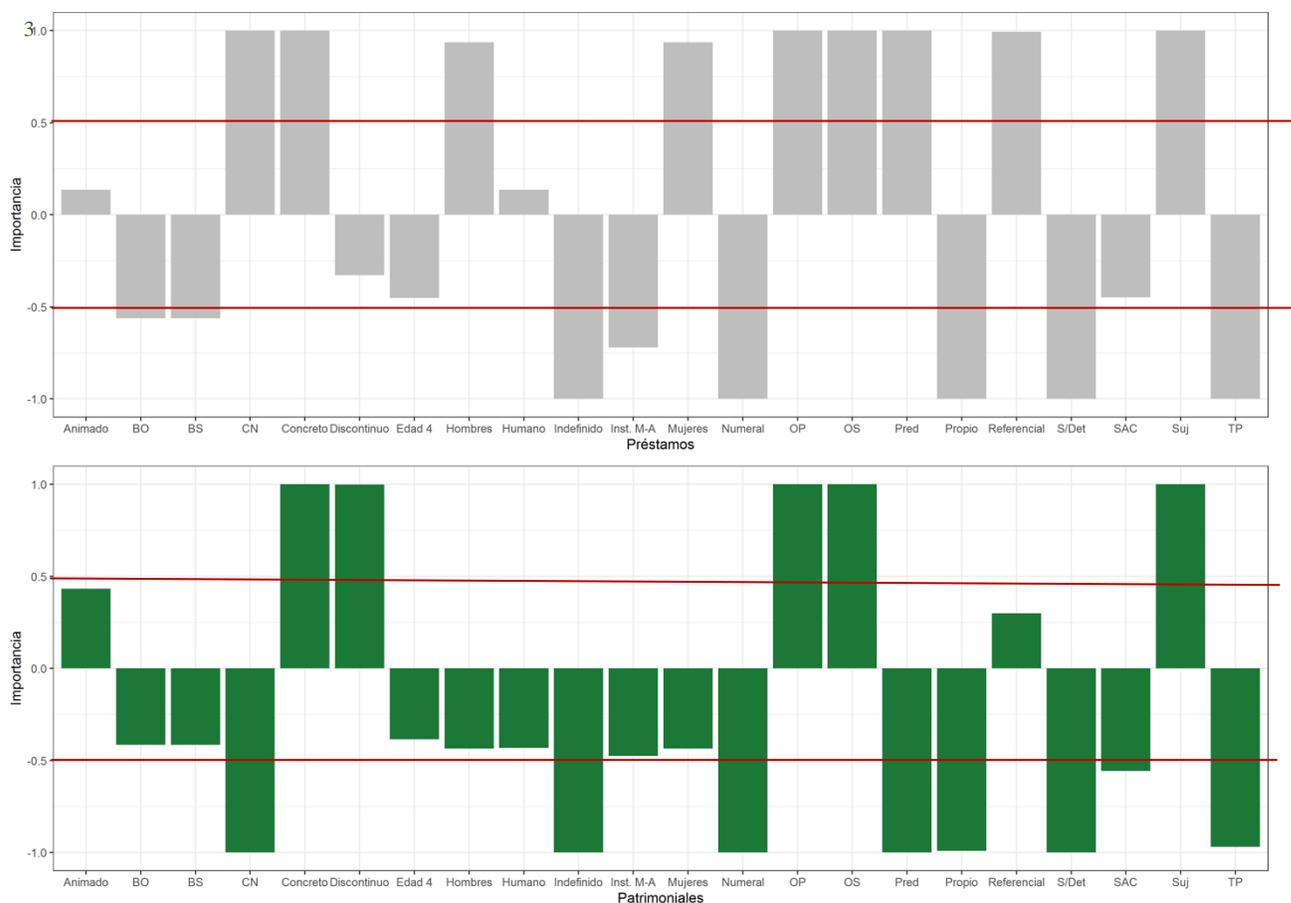
La siguiente variable en la jerarquía es el nivel de instrucción. Los hablantes de nivel medio-alto inhiben el uso de las marcas de número (Patrón 1: *ra / ya + N*) y favorecen su ausencia. El último factor en la jerarquía es el tipo de bilingüismo. Tanto los bilingües con

mayor dominio del otomí (-1.536) como los simétricos (-1.55026) no favorecen el uso de las marcas de número (Patrón 1: *ra / ya*), pues son los que presentan más nombres sin marcas (véase Apéndice 4C).

4.5.3 Conclusiones en torno a la marcación de número según el origen histórico de los lexemas

Antes de concluir esta sección, comparo los factores lingüísticos y sociales que determinan la presencia (Patrón 1: *ra / ya*) y ausencia (Patrón 2: N) de marcas de número en los nombres patrimoniales y los préstamos del español. Para este propósito, empleé una prueba de inferencia multimodelo (§2.1.3.5) con la finalidad de identificar el subconjunto de modelos con el mayor soporte empírico, de un total de 50 posibles modelos. Ordené los modelos posibles del mejor al peor, basados en el criterio de información de Akaike. El conjunto de modelos con una diferencia en $AIC < 2$ del mejor modelo fueron considerados con la misma fuerza y plausibilidad. Para evaluar la importancia de cada predictor usé el peso de Akaike (W_i). La importancia relativa de cada predictor se calculó a partir de la suma de los pesos de Akaike ($\sum W_i$) de cada modelo en el que el predictor aparece. Todos los modelos fueron construidos usando el paquete `glmulti` para la versión 3.01 de R. En la siguiente figura se ilustran los resultados.

FIGURA 4.8 *Resultados de la prueba inferencia multimodelo para la alternancia en los tipos de marcación de número por origen histórico*¹¹⁹



Las variables predictoras en el conjunto de modelos AICc <2 y el 95% del conjunto de modelos para la concordancia de número en los tres grupos de hablantes. La importancia de cada variable se muestra mediante la suma de los pesos de Akaike. El signo (\pm) de la estimación de los parámetros representa el efecto (positivo o negativo) de cada predictor sobre la variable de respuesta.

El predictor de mayor importancia en los dos grupos es la ausencia o presencia de determinantes de número. Las frases nominales sin determinantes, con numerales y con indefinidos inhiben el uso de las marcas *ra* y *ya*. El segundo predictor de mayor importancia

¹¹⁹ BO = bilingües con mayor dominio del otomí, BS = bilingües simétricos, CN=complemento nominal, Edad 4 = hablantes de +55 años, Inst. M-A = nivel de instrucción medio – alto, OP= objeto primario, OS= objeto secundario, Pred. = función predicativa, S/Det. = Frases nominales sin determinantes que codifican el número, SAC = San Andrés Cuexcontitlán, Suj = sujeto, TP = término de preposición.

en los nombres patrimoniales y los préstamos es la función sintáctica de la frase nominal. Sin embargo, en los nombres *ñható* se observa una escisión que en los préstamos no: las funciones nucleares (sujeto, objeto primario, objeto secundario) favorecen la marcación de número y las funciones periféricas (complemento nominal, predicativo, adverbial) la inhiben.

El tipo de constituyente es el tercer predictor de mayor importancia en los nombres patrimoniales y el cuarto en los préstamos. En ambos casos, los nombres que son término de preposición no favorecen el uso de las marcas *ra/ya*. La concreción del nombre es el cuarto predictor en los nombres *ñható* y el tercero en los préstamos, y los nombres concretos son los que promueven el uso del Patrón 1: *ra / ya*. Los dos grupos de nombres difieren en el quinto predictor de mayor importancia. En los nombres *ñható* es la (dis)continuidad, y los nombres discontinuos favorecen el uso de las marcas de número. En los préstamos, es la referencialidad de la frase nominal, y son las frases referenciales las que promueven la aparición de *ra* y *ya*.

Por otro lado, en cuanto a los predictores sociolingüísticos, se observan ciertas diferencias. En el grupo de nombres patrimoniales sólo es relevante la comunidad y son los hablantes de SAC, quienes favorecen el empleo del Patrón 1: *ra/ya*. En contraste, en los préstamos del español, los predictores sociales más importantes son el género, el nivel de instrucción y el grado de bilingüismo. Estos resultados sugieren que la marcación de número tiene un comportamiento lingüístico homogéneo en los dos grupos léxicos, pero una mayor variación sociolingüística en el uso de las marcas de número en los préstamos del español.

En esta sección determiné cuáles son los factores lingüísticos y sociales que inciden en la presencia / ausencia de *ra* y *ya* en los nombres patrimoniales y los préstamos. Los resultados muestran un comportamiento lingüístico similar en los dos grupos léxicos, lo cual, de momento, no nos permite negar ni afirmar las HIPÓTESIS 4.1 y 4.2. En la siguiente sección

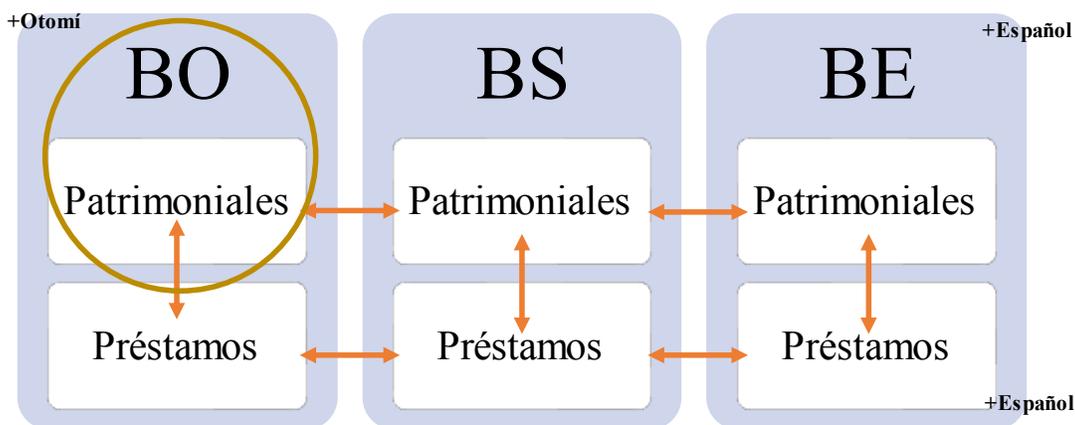
realizaré un análisis en tiempo aparente a partir del tipo de bilingüismo, con la finalidad de encontrar diferencias en la marcación de número en los nombres patrimoniales y en los préstamos.

4.6 ANÁLISIS EN TIEMPO APARENTE: EL TIPO DE BILINGÜISMO

Los estadísticos empleados para determinar qué factores (socio)lingüísticos inciden en la marcación de número de los nombres patrimoniales del otomí y de los préstamos del español seleccionaron al tipo de bilingüismo como el último factor en la jerarquía. En ambos conjuntos léxicos, los bilingües con mayor dominio del otomí y los bilingües simétricos no favorecen el uso del Patrón 1: *ra / ya* (véase §4.4 y §4.5).

A pesar de que el tipo de bilingüismo no fue el factor más relevante, consideré pertinente revisar el uso de las marcas de número en los nombres patrimoniales y los préstamos en los tres grupos de bilingües porque cada uno representa un grado diferente de contacto e influencia del español. La marcación de número en los nombres patrimoniales empleados por los bilingües con mayor dominio del otomí será el punto de referencia a partir del cual estableceré las similitudes y diferencias con el resto de los grupos lingüísticos, debido a que este conjunto representa un menor grado de contacto e influencia del español.

FIGURA 4.9 Pauta para el análisis por grupo lingüístico



Para ello, en primer lugar, determinaré cuáles son los factores lingüísticos que inciden en la marcación de número de los nombres patrimoniales en el grupo de bilingües con mayor dominio del otomí, los bilingües simétricos y los bilingües con mayor dominio del español, mediante un análisis multimodelo. Posteriormente, estableceré las similitudes y diferencias entre los grupos a partir del comportamiento de los bilingües con mayor dominio del otomí.

4.6.1 El uso de las marcas de número en los nombres patrimoniales según el tipo de bilingüismo

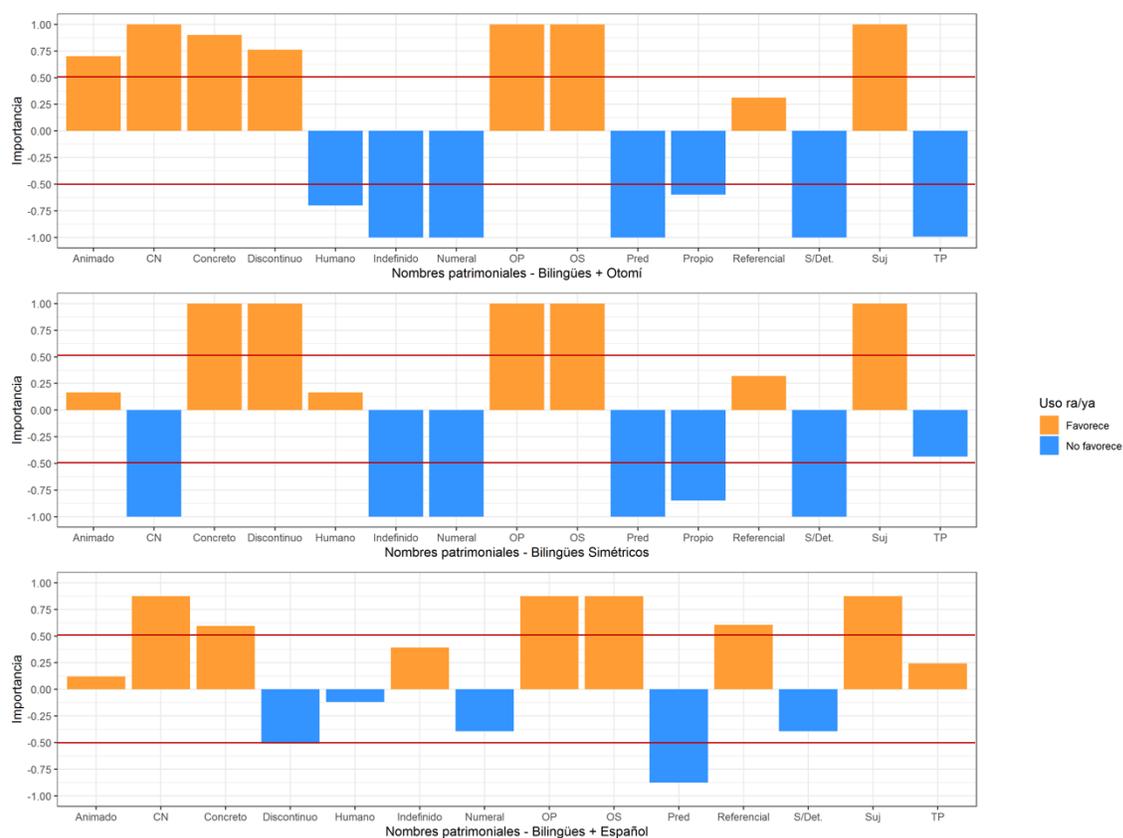
En la Tabla 4.19 se presenta el uso de las marcas de número con los nombres patrimoniales según el tipo de bilingüismo. Los bilingües con mayor dominio del español emiten un mayor porcentaje de *ra* y *ya*, en comparación con los bilingües simétricos y los bilingües con mayor dominio del otomí que emitieron un menor porcentaje de las marcas de número en los nombres patrimoniales.

TABLA 4.19 *Uso de las marcas de número en los nombres patrimoniales. Tipo de bilingüismo*

Tipo de bilingüismo	Patrón 1 <i>ra / ya + N</i>	Patrón 2 N
Bilingües con mayor dominio del otomí	66%	34%
Bilingüe simétrico	69%	31%
Bilingüe con mayor dominio del español	73%	27%

Los resultados de la tabla anterior se contraponen a la idea inicial de que los hablantes dominantes en otomí hacen un mayor uso de *ra* y *ya* que los bilingües más dominantes en español. Por ello, será necesario indagar acerca de los factores lingüísticos que inciden en la codificación de número de los tres grupos de bilingües para determinar qué contextos están favoreciendo la extensión de estas formas. Para este propósito, utilicé una prueba de inferencia multimodelo (§2.1.3.5), en la que consideré los siguientes factores: la presencia de determinantes que codifican el número, la distinción entre nombre propio y común, la concreción, la animacidad y la (dis)continuidad del nombre, la función sintáctica y la referencialidad de la frase nominal y el tipo de constituyente. A continuación, se presentan los resultados para cada grupo de bilingües.

FIGURA 4.10 Resultados de la prueba de inferencia multimodelo del uso de las marcas de número en los nombres patrimoniales de los tres grupos lingüísticos¹²⁰



Las variables predictoras en el conjunto de modelos AICc <2 y el 95% del conjunto de modelos para la concordancia de número en los tres grupos de hablantes. La importancia de cada variable se muestra mediante la suma de los pesos de Akaike. El signo (\pm) de la estimación de los parámetros representa el efecto (positivo o negativo) de cada predictor sobre la variable de respuesta.

En primer lugar, se observa que 6 de los 8 factores lingüísticos empleados en el análisis son relevantes para alguno de los grupos lingüísticos. La función sintáctica y la concreción del nombre inciden en la marcación de número de los nombres patrimoniales en los tres grupos lingüísticos. Por otro lado, la presencia de determinantes que codifican información de número y la (dis)continuidad del nombre sólo son importantes en la

¹²⁰ CN=complemento nominal, OP= objeto primario, OS= objeto secundario, Pred. = función predicativa, S/Det. = Frases nominales sin determinantes, Suj = sujeto, TP = término de preposición.

marcación de número de los BO y los BS. Por último, el tipo de constituyente y la distinción nombre común/propio sólo son importantes en los bilingües con mayor dominio del otomí y los bilingües simétricos, respectivamente.

En primer lugar, revisaremos el comportamiento de los factores que son determinantes para la marcación de número en los BO, BS, BE para saber si tienen el mismo efecto o difieren. El factor de mayor importancia en los tres grupos de bilingües es la función sintáctica. Las frases nominales que funcionan como sujeto, objeto primario y secundario favorecen el uso de las marcas de número. Este efecto se observa en la distribución de los datos, pues en las frases nominales con función de sujeto (**Patrón 1 ra / ya: BO = 87%; BS = 90%; BE = 100%**), objeto primario (**Patrón 1 ra / ya: BO = 87%; BS = 100%; BE = 100%**) y secundario (**Patrón 1 ra / ya: BO = 82%; BS = 85%; BE = 77%**) predomina el uso de las marcas de número, como se muestra en los siguientes ejemplos.

- (53) a. *Hímbi alkansá ka ra dómi ma páhpa*
 hímbi = bi = alkansá **ka ra dómi** ma páhpa
 NEG = 3.PRFT = alcanzar **DEM SG dinero** 1.POS papá
 ‘No le alcanzaba el dinero a mi papá’
- b. *Dimí ‘aaté ya máda*
 di-mí = ʔǎté **ya máda**
 1.IMPF = hacer **PL molcajete**
 ‘Hacia molcajetes’

Los sujetos que no llevan las marcas de número no aparecen de forma escueta, sino que están modificados por el diminutivo *tsí-*, el cuantificador *bastante* (54a), los adjetivos *cada* y *puro* (54b), el adverbio *nada más* (54c), o alguna preposición.

- (54) a. *Ya khääní mi ‘b̥u̥i ra z̥a mi ‘yó k ara káye bastante khääní mi ‘b̥u̥i*
 ya khääní mi = ʔb̥i ra z̥a mi = ʔyó
 PL persona 3.IMPF = estar SG bien 3.IMPF = caminar
- ka ra káje **bastante khääní** mi = ʔb̥i
 DEM PL calle **bastante persona** 3.IMPF = estar
 ‘Las personas están bien, caminaban en la calle, hay bastantes personas’

b. *Na ndía ná bi nkháwa bi nkháwa bi buuní púrya ngéro*
 na ndí na bi = nkháwa bi = bñi **púrya engéro**
 INDEF vez CIT 3.PFT = suceder 3.PFT = salir **puro encuerada**
 ‘Una vez sucedió que salieron puras encueradas’

c. *Todabýa do partisipáhé porque do nú ndi má o sea namá ntsǎya*
 todabýa do = partisipá = hé porque do = nú
 todavía 1.PRFT = participar = PL.EXCL porque 1.PRFT = ver

ndi = má o sea **namá ntsǎya**
 3.SUBJ = ir o sea **nada más señor**
 ‘Todavía participamos porque vi que iban o sea nada más señores’

Por otro lado, las frases nominales con función de objeto que no llevan las marcas de número tienen como núcleo un nombre continuo, como se muestra en (55a) y (55b), o se limitan a cierto tipo de construcciones como ‘prender vela’ (55c) o ‘vender pollo’ (55d).

(55) a. *Go má xahkíhé paxí*

go = má = fahkí = hé **pají**
 1.FUT = ir = segar = PL.EXCL **hierba**
 ‘Vamos a cortar hierba’

b. *Ma mé ma thá xá’tho mi háwí dómi*

ma mé ma thá fǎ’tho mi = há = wí **dómi**
 1.POS mamá 1.POS papá no.haber 3.IMPF = tener = DU.INCL **dinero**
 ‘Mi mamá, mi papá no tenían dinero’

c. *Da má thiiti ‘yǒ da má kā ra tsíduuní*

da = má = thiti **ʔyǒ** da = má kā ra tsí - dñí
 1.PRFT = ir = prender **vela** 1.PRFT = ir traer SG DIM - flor
 ‘Voy a prender las velas, voy a traer flor’

d. *Go má pǎghmé áni*

go = má pǎ = hmé **áni**
 1.FUT = ir vender = PL.EXCL **pollo**
 ‘Vamos a vender pollo’

En contraste, las frases nominales con función predicativa no promueven su aparición. Este efecto se puede observar en la distribución de los datos, donde la mayoría de

las frases con función predicativa no lleva las marcas de número (**Patrón 2 N: BO** = 80%; **BS** = 83%; **BE** = 100%), como se muestra a continuación.

(59) *Desde ndra ma leḗhkugó da má a 'Mundaa*
 desde **ndra** **ma** **lḗhku** - gó da = má a ʔmúndǎ
 desde **1.CONT.IMPF** **más** **chico** - 1.ENF 1.PRF = ir PREP México
 'Desde que era chico fui a México'

b. *Do hant'ihé na ra persona mbra bot'i*
 do = hantʔí = hé na ra persona **mbra** **botʔi**
 1.PRFT = ver.de.lejos = PL.EXCL INDEF SG persona **3.CONT.IMPFT** **negro**
 'Vimos una persona de negro'

La principal diferencia entre los grupos de bilingües radica en el efecto de los complementos nominales. En el grupo de BO y BE promueven la aparición de las marcas de número (**Patrón 1 ra / ya: BO** = 56%; **BE** = 100%), pero en menor medida que las funciones sintácticas nucleares. En contraste, la función de complemento nominal no favorece las marcas de número en los BS, como se muestra a continuación.

(60) *Mi hant'i ro gooxthí nikha*
 mi = hantʔí **ro** **gǒthí** **nikha**
 3.IMPF = ver.de.lejos **SG** **puerta** **iglesia**
 'Veía la puerta de la iglesia'

La concreción del nombre también es un factor determinante para la codificación de las marcas de número en los tres grupos de bilingües, aunque con un rango de importancia diferente (**BO** = 4° lugar; **BS** = 3° lugar; **BE** = 3° lugar). No obstante, en los tres conjuntos, los nombres concretos favorecen el uso de las marcas de número. Este comportamiento se constata en la distribución de los datos, que muestran una mayor presencia de las marcas de número en los nombres concretos (**Patrón 1 ra / ya: BO** = 75%; **BS** = 81%; **BE** = 85%), en

comparación con los nombres abstractos, en los que predomina la ausencia de estas (**Patrón N: BO = 56%; BS = 64%; BE = 61%**), como se muestra en los siguientes ejemplos.

(61) a. *Yxá'tho ma hmĩ ka ra biithĩ*
 i jáʔtho ma hmĩ **ka** **ra** **bĩthĩ**
 y no.haber más tortilla **DEM.SG** **SG** **chiquihuite**
 'Y no había más tortillas en el chiquihuite'

b. *Do hant'ihé na ra persona mbra bot'i*
 do = hantʔí = hé na ra pèrsona mbra **botʔi**
 1.PRFT = ver = PL.EXCL INDEF SG persona 3.CONT.IMPF **negro**
 Vimos una persona de negro

Por otro lado, la presencia de determinantes con rasgos de número y la (dis)continuidad del nombre sólo son importantes en la marcación de número de los bilingües con mayor dominio del otomí y los simétricos. Ante esto aparecen dos cuestiones que intentaré responder enseguida: la primera es si dichas variables tienen el mismo efecto en los BO y los BS; y la segunda es ¿por qué estos factores son determinantes para los BO y los BS, pero no para los BE?

El tipo de determinante incide en la marcación de número de los bilingües con mayor dominio del otomí y los bilingües simétricos, pero con un rango de importancia diferente (**BO = 2º lugar; BS = 3º lugar**). Las frases nominales con determinantes definidos favorecen el uso de *ra* y *ya*, como se puede ver en la distribución de los datos (**Patrón 1 ra / ya: BO = 91%; BS = 91%**) y en el ejemplo (62a). Las frases nominales con determinantes indefinidos también promueven la aparición de *ra* y *ya*, como se muestra en los porcentajes (**Patrón 1 ra / ya: BO = 79%; BS = 81%**) y en el ejemplo (62b).

(62) a. *Bi má yo'ti k ara déhéú*
 bi = má yoʔti **ka** **ra** **déhé** - ú
 3.PRFT = ir sacar **DEM.SG** **SG** **agua** - entonces
 'Fue a sacar el agua, entonces'

- b. *Ndra ma léhku do nú na ra t'ixú*
 ndra ma léhku do = nú na ra tʔixú
 1.CONT.IMPF más chico 1.PFT = ver INDEF SG señora
 ‘Cuando era niña vi a una señora’

Las frases con demostrativos en donde no se marca el número, corresponden a usos adverbiales de estos, como se ejemplifica en (63a). Por otro lado, las frases con la marca de indefinitud que no llevan las marcas de número están relacionadas con la construcción *na ndí* ‘una vez’ (63b). La ausencia de la marca singular podría deberse a que se trata de un recurso narrativo fijo, pero sería necesario realizar pruebas adicionales.

- (63) a. *Do nú na ra t'ixú ka ñú*
 do = nú na ra tʔixú ka ñú
 1.PFT = ver INDEF SG señora DEM camino
 ‘Vi una señora en el camino’

- b. *Na ndí bi inbihtági na ra fyésta*
 na ndí bi = inbihtá - (gí) na ra fyésta
 INDEF vez 3.PRFT = invitar - 1.OBJ INDEF SG fiesta
 ‘Una vez me invitaron a una fiesta’

En ambos grupos, las frases nominales sin determinantes y las frases nominales con un numeral inhiben la aparición de las marcas *ra* y *ya*. Este comportamiento se refleja en la distribución de los datos, donde un porcentaje importante de las frases nominales sin determinantes no lleva las marcas de número (**Patrón 2 N: BO = 42%; BS = 40%**), como se ejemplifica en (64a). Asimismo, casi la totalidad de las frases nominales con numerales no presentan las marcas de número (**Patrón 2 N: BO = 86%; BS = 73%**), como se ve en (64b).

- (64) a. *Xo dimá hóni o hína damá zo a hwäähĩ*
 ʃo di-má = hóni o hína da-má = zo a huáhĩ
 pues 1.IMPF = buscar o NEG 1.IMPF = dejar PREP milpa
 ‘Pues me iba a buscar o si no iba a dejarlo en la milpa’

- b. *I dura ñú pá ka ro fyésta*
 i = dúra **nú** **pá** ka ro **fyésta**
 1.PRS = durar **tres** **día** DEM.SG SG fiesta
 ‘Dura tres días la fiesta’

Aunque los dos grupos de hablantes tienen un comportamiento muy parecido y las frases nominales sin determinantes y con numerales no favorecen la marcación de número, también se advierten pequeños matices que podrían ayudar a contestar la segunda pregunta. En los BO, sólo documenté un caso de frase nominal con numeral *ʔyóhó* ‘dos’ y marca de plural *ya*, cuyo núcleo es +humano (65a). Estos casos aumentan en los BS, donde registro cuatro frases nominales con numeral y marca plural, cuyos núcleos son +humano o +animado, como se muestra en (65b). A pesar de que este factor no fue relevante en el grupo de bilingües con mayor dominio del español, estos hablantes hacen un uso importante de *ra* y *ya* en frases nominales con un numeral (**Patrón 1** *ra / ya* = **50%**; Patrón 2 N = 50%), independientemente de su animadidad (65c), en comparación con los BO y BS en los que su uso era excepcional y estaba vinculado a nombres +humanos o +animados.

- (65) a. *Y cuando b̄u mi p̄etsi ‘yóhó ya b̄éhña na ra ntsaaya*
 y kuán.do b̄i mi = p̄etsi **ʔyóhó** **ya** **ʔbéhñΛ**
 y cuando cuando 3.IMPF = tener **dos** **PL** **señora**
 na ra n̄ts̄ya
 INDEF SG señor
 ‘Y cuando tenía dos mujeres un señor [...]’

- b. *Komo a la una de la mañana ko ma kú do nthébé na ra ‘yó*
 komo a la úna de la mañána ko ma kú
 como a la una de la mañana con 1.POS hermano
 do = nthé = bé **na** **ra** **yo?**
 1.PFT = encontrar = DU.EXCL **INDEF** **SG** **perro**
 ‘como a la una de la mañana con mi hermano encontramos un perro’

- c. *Do taqt’i ñú ya x̄āt’á*
 do = t̄t’í **nú** **ya** **ʃāt’á**
 1.PRFT = encontrar **tres** **PL** **nopal**
 ‘Encontramos tres nopales’

La (dis)continuidad del nombre es un factor relevante para la marcación de número en los bilingües con mayor dominio del otomí y en los bilingües simétricos, aunque con un rango de importancia diferente (**BO** = 5° lugar; **BS** = 4° lugar). No obstante, en los dos conjuntos, los nombres discontinuos favorecen el uso de las marcas de número, como se puede apreciar en la distribución de los datos (**BO** = 71%; **BS** = 75%) y como se muestra a en los siguientes ejemplos.

- (66) a. *Da ze ka ra pá damá tsǎtǎ*
 da = ze ka ra pá da-má tsǎtǎ
 3.FUT = llegar DEM.SG SG día 1.IMPF quemar

ka ra tsí handri
DEM.SG SG dim torito
 ‘Cuando llegue el día, quemar el torito’

- b. *Do ínihé ma móhihé ga hǎi*
 do = íni = hé ma móhi = hé ga hǎi
 1.PFT = jugar = PL.EXCL 1.POS plato = PL.EXCL de **tierra**
 ‘Jugamos con nuestros platos de tierra’

Por último, el tipo de constituyente, la distinción nombre común/propio y la referencialidad sólo son importantes en los bilingües con mayor dominio del otomí y los bilingües simétricos, respectivamente. De esto se desprenden la interrogante ¿por qué estos factores sólo son importantes para un grupo lingüístico y no para el resto?

El tipo de constituyente sintáctico sólo fue relevante en la organización del número en los bilingües con mayor dominio del otomí. En los nombres que son núcleo de una frase nominal prevalece el uso de las marcas de número (**Patrón 1 ra / ya = 73%**; Patrón N = 27%), como se muestra en (67a). En contraste, los nombres que son término de preposición no favorecen el uso de las marcas de número como se aprecia en la distribución de los datos (67b), donde la mayoría de los términos de preposición no especifica el número gramatical

(Patrón 1 *ra / ya* = 21%; **Patrón N = 79%**). De hecho, sólo se presentan 14 términos de preposición con las marcas de número y la mayoría de estos esta asociada con el uso de la preposición del español *ko(n)*, como se muestra en el tercer ejemplo.

- (67) a. *Entóse b̄u ndo tsóhni da pénti ka ra ñú*
 entóse bi ndo = tsóhni da = pénti **ka ra** **ñú**
 entonces cuando 1.SIM = llegar 1.PRFT = agarrar DEM.SG SG **camino**
 'Entonces cuando llegué agarré el camino'
- b. *Mi péhki ga zá ma mamá*
 mi = pé - hkí **ga zá** ma mamá
 3.IMPF = pegar - 1.OBJ **PREP palo** 1.POS mamá
 'Me pegaba con el palo mi mamá'
- c. *Pero dimá paanti ko na ra tsühkwá k'ü ya táko*
 péro di-má = pãnti **ko na ra tsühkwá** kʔi ya táko
 pero 1.IMPF = envolver **con INDEF SG papel** DEM.PL PL taco
 'Pero envolvía con un papel los tacos'

A partir de algunas observaciones preliminares, consideré importante revisar el origen de las preposiciones en este tipo de constituyentes, pues pensé que las preposiciones del español podrían estar asociadas con el uso de *ra* y *ya*. Al respecto, encontré que la mayoría de las preposiciones que usan los BO son de origen otomí (84% = 56) y sólo un bajo porcentaje proviene del español (16% = 11). Aunado a esto, vi que los términos de preposición que llevan las marcas de número están asociados con el uso de la preposición del español *ko(n)*.

Sin embargo, como vimos en la Figura 4.10, este factor deja de tener un efecto en la marcación de número de los nombres patrimoniales usados por los bilingües simétricos y los bilingües con mayor dominio del español. Aunque la mayoría de los términos de preposición de los BS no especifican el número gramatical (Patrón 1 *ra / ya* = 48%; **Patrón 2 N = 52%**), este porcentaje es mayor en los bilingües con mayor dominio del otomí (Patrón 1 *ra / ya* = 21%; **Patrón N = 79%**). Debido a que el uso de las marcas de número en los términos de

preposición del grupo BO estaba asociado con el uso de las preposiciones del español, consideré oportuno revisar la proporción de preposiciones según su origen en el grupo de BS. Sobre esto, encontré que los BS adoptan un mayor porcentaje de preposiciones del español (51%) que los bilingües con mayor dominio del otomí (16%). Asimismo, los nombres que son término de una preposición del otomí tienden a no codificar el número (Patrón 1 *ra / ya* = 14%; **Patrón N = 86%**), mientras que los nombres que son término de una preposición del español suelen llevar las marcas de número (**Patrón 1 *ra / ya* = 56%**; Patrón 2 N = 44%), como se muestra a continuación.

- (68) a. *Ya khääní do konosé mi ña ga ñható*
 ya khäñí do = konosé mi = ja ga pathó
 PL gente 1.PFT = conocer 3.PFT = hablar PREP otomí
 ‘La gente que conocí hablaba otomí’
- b. *Do nthähtíbé po ra tsínikhá al año*
 do = nthähtí = bé po ra tsí - níkhá al 'año
 1.PRFT = casarse = DU.EXCL por SG DIM - iglesia al año
 ‘Nos casamos por la iglesia al año’

En los bilingües con mayor dominio del español, hallé un comportamiento similar al de los simétricos. Los BE han adaptado un importante porcentaje de preposiciones del español (50%). Asimismo, la mayoría de los nombres que son término de una preposición otomí no especifican el número (Patrón 1 *ra / ya* = 14%; **Patrón N = 86%**), mientras que los nombres que son término de una preposición del español tienden a llevar las marcas de número (**Patrón 1 *ra / ya* = 56%**; Patrón 2 N = 44%), como se muestra enseguida.

- (69) a. *Dimá paahmé a hwáhi*
 di-má = pǎ = hmé a hwáhi
 1.IMPF = vender = PL.EXCL PREP milpa
 ‘Vendíamos en la milpa’

- b. *Komo ra ócho de ra nxüüí dimá 'áde ya ruido*
 komo ra ótʃo **de ra njũí** di-má = ʔáde ya ruído
 como SG ocho **de SG noche** 1.IMPF = escuchar PL ruido
 'Como a las ocho de la noche escuchaba ruidos'

En síntesis, la marcación de número en los tres grupos lingüísticos está determinada por la función sintáctica y la concreción del nombre, no obstante, hay ciertas diferencias. En primer lugar, hay un incremento en el uso de las marcas de número del otomí conforme incrementa el conocimiento del español de los bilingües y decrece el dominio de la lengua otomí. Este resultado parece contraintuitivo, sin embargo, podría responder al debilitamiento de algunas restricciones que determinan la presencia / ausencia de las marcas de número en la lengua otomí. Sin embargo, esta es sólo una hipótesis que tendría que ponerse a prueba en otro tipo de investigación y con un corpus más amplio.

La primera restricción debilitada corresponde al uso de las marcas de número en frases nominales con numerales. En el otomí de los bilingües con mayor dominio del otomí, y en menor medida, en el de los bilingües simétricos, el empleo de las marcas de número con frases nominales con numerales es excepcional y los pocos casos que se documentan se restringen a nombres +humanos o +animados. En contraste, en los bilingües con mayor dominio del español, la aparición de las marcas *ra* y *ya* con numerales asciende al 50%.

Encontré un efecto similar con el tipo de constituyente sintáctico. En el grupo de los BO se presenta una fuerte distinción en el uso de las marcas de número según sea el nombre núcleo de una FN o término de preposición. En este último caso, el uso de *ra* / *ya* no se ve favorecido y los pocos casos que se registran están asociados con el empleo de preposiciones del español, como *ko(n)*. Esta restricción se debilita en el grupo de bilingües simétricos debido a la adopción de un porcentaje importante de preposiciones del español, que, a su vez,

conllevar al uso de las marcas de número en un contexto antes restringido. En el grupo de BE, este comportamiento se agudiza.

4.6.2 El uso de las marcas de número en los préstamos según el tipo de bilingüismo

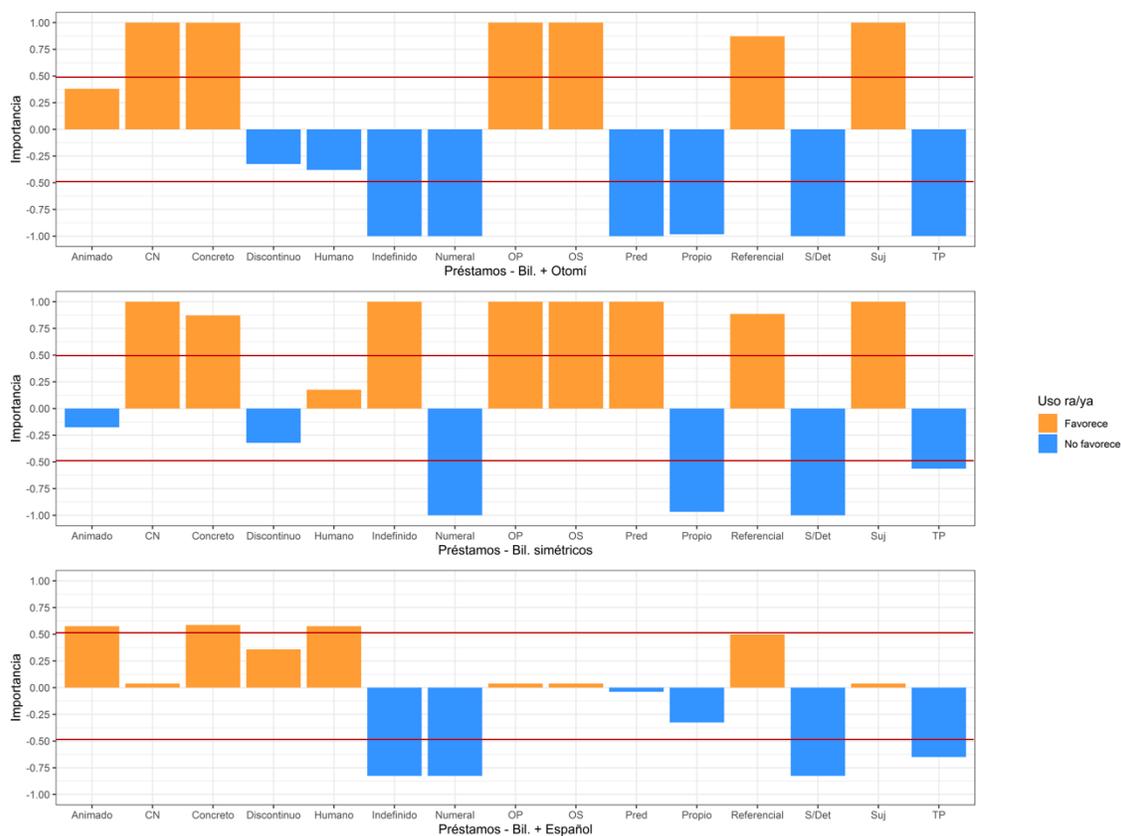
En la Tabla 4.20 se presenta el uso de las marcas de número con los nombres patrimoniales según el tipo de bilingüismo. Los bilingües con mayor dominio del español emiten un mayor porcentaje de *ra* y *ya* en los préstamos del español, en comparación con los bilingües con mayor dominio del otomí y los bilingües simétricos.

TABLA 4.20 *Uso de las marcas de número en los préstamos. Tipo de bilingüismo*

Tipo de bilingüismo	Patrón 1 <i>ra</i> / <i>ya</i> + N	Patrón 2 N
Bilingües con mayor dominio del otomí	64%	36%
Bilingüe simétrico	58%	42%
Bilingüe con mayor dominio del español	75%	25%

Los resultados de la tabla anterior se contraponen a la idea inicial de que los hablantes dominantes en otomí hacen un mayor uso de las marcas de número que los bilingües más dominantes en español. Por ello, será necesario indagar acerca de los factores lingüísticos que inciden en la codificación de número de los tres grupos de bilingües para determinar qué contextos están favoreciendo la extensión de estas formas. Para este propósito, utilicé una prueba de inferencia multimodelo (§2.1.3.5), en la que consideré los factores: la presencia de determinantes que codifican el número, la distinción entre nombre propio y común, la concreción, la animacidad y la (dis)continuidad del nombre, la función sintáctica, la referencialidad de la frase nominal y el tipo de constituyente. A continuación, se presentan los resultados para cada grupo de bilingües.

FIGURA 4.11 *Resultados de la prueba inferencia multimodelo del uso de las marcas de número en los préstamos de los tres grupos lingüísticos*



En primer lugar, se observa que 7 de los 8 factores lingüísticos empleados en el análisis fueron relevantes para alguno de los grupos lingüísticos. La presencia de determinantes que codifiquen la información de número, el tipo de constituyente y la concreción del nombre son determinantes en la marcación de número de los préstamos de los tres grupos lingüísticos. Por otro lado, la función sintáctica, la distinción nombre común/propio y la referencialidad de la frase nominal son importantes en la marcación de número de los bilingües con mayor dominio del otomí y los simétricos. Por último, la animacidad sólo es importante en los bilingües con mayor dominio del español.

En primer lugar, revisaré el comportamiento de los factores que son determinantes para la marcación de número en los BO, BS, BE: el tipo de determinante, el tipo de constituyente y la concreción del nombre, para saber si estos tienen el mismo efecto o difieren en los tres grupos.

El tipo de determinante es el factor más importante para la marcación de número de los préstamos en los tres grupos lingüísticos. En las frases nominales con determinantes definidos predomina el uso de las marcas de número (**Patrón 1 ra / ya: BO = 93%; BS = 89%; BE = 100%**), como se muestra en (70a). En las frases nominales con determinantes indefinidos se halla un comportamiento similar (**Patrón 1 ra / ya: BO = 91%; BS = 91%; BE = 83%**), pues en ellos predomina el uso de las marcas de número, como se muestra en el ejemplo b.

- (70) a. *Péro ka ra kwénto bi kwéntagí ma mahmá*
 péro **ka** **ra** **kuénto** bi = kuénta - gí ma mahmá
 pero **DEM.SG** **SG** **cuento** 3.PFT = contar - 1.OBJ 1.POS mamá
 ‘Pero el cuento nos lo contó mi mamá’
- b. *Bi kámbihé na ra chamárra*
 bi = kámbi = hé **na** **ra** **t̂jamára**
 3.PFT = cambiar = PL.EXCL **INDEF** **SG** **chamarra**
 ‘Cambiaron una chamarra’

Las frases con demostrativos en donde no se marca el número, corresponden a usos adverbiales de estos, como se ejemplifica en (71a). Por otro lado, las frases con la marca de indefinitud que no llevan las marcas de número están relacionados con el nombre *na ndí* ‘una vez’ (71b).

- (71) a. *Mi gāath'i ka ra káye mi 'yó ka káye bastante khǎáni mi b̥uui*
 mi = gǔthʔí ka ra káje mi = ʔyó ka káje
 3.IMPF = andar DEM.SG SG calle 3.IMPF = caminar DEM calle
 bastante khǎáni mi = b̥i
 bastante persona 3.IMPF = estar
 'Andaban en la calle, caminaban en esa calle bastantes personas habían
- b. *Por eso xo do ntsúhé na ndí*
 por eso ʃo do = n̄tsú = hé na ndí ú
 por eso pues 1.PRFT = miedo = PL.EXCL INDEF vez entonces
 'Por eso nos espantamos una vez entonces

Las frases nominales sin determinante inhiben la aparición de las marcas *ra* y *ya*. Este efecto se puede ver en la distribución de los datos, en donde las frases nominales sin determinantes (**Patrón 2 N: BO = 53%; BS = 56%; BE = 33%**) prescinden de las marcas de número, como se muestra en (72a). Las frases nominales con un numeral tampoco favorecen la aparición de las marcas *ra* y *ya* (**Patrón 2 N: BO = 100%; BS = 75%; BE = 0%**), como se puede ver en el segundo inciso.

- (72) a. *'Bu ndo ʃuudí do mágó a éxkwela*
 ʔbi ndo = ʃídí do = má - gó a éʃkwela
 cuando 1.SIM = empezar 1.PFT = ir - 1.ENF LOC escuela
 'Cuando yo empecé a ir a la escuela'
- b. *Xo go 'ópí da regaláǵi k̥uutá béxo*
 ʃo go = ʔó = pí da = regalá - gí k̥ítá b̥éfo
 pues 1.FUT = pedir = 3.OBJ 3.FUT = regalar - 1.OBJ cinco peso
 'Pues les pedía que me regalarán 5 pesos'

Aunque los BO y los BS tienen un comportamiento muy parecido y las frases nominales sin determinantes y con numerales no favorecen la marcación de número, también se advierten pequeños matices. En los BO, no se presenta ni una sola frase nominal con numeral y marca plural. Estos casos aumentan en los BS, donde el 25% de las frases nominales con numeral lleva la marca plural. Este comportamiento se agudiza en el grupo de

bilingües con mayor dominio del español, en los que la totalidad de las FNs con numeral lleva la marca *ya* (**Patrón 1 ra / ya**: **BO** = 0%; **BS** = 25%; **BE** = 100%).

La concreción del nombre es determinante para la marcación de número en los tres tipos de bilingües, aunque con un rango de importancia diferente (**BO** = 4° lugar; **BS** = 5° lugar; **BE** = 3° lugar). Los nombres concretos promueven el uso de las marcas *ra* y *ya*. Este efecto se constata en la distribución de los datos, donde la mayoría de los nombres concretos llevan las marcas de número (**Patrón 1 ra / ya**: **BO** = 71%; **BS** = 62%; **BE** = 84%), como se muestra en los siguientes ejemplos.

- (73) a. *Este níhí mi zǎǎní na ra bebé*
 este níhí mi zǎǎní **na ra bebé**
 este también 3.IMPF llorar **INDEF SG bebé**
 ‘Este también lloraba un bebé’
- b. *Di päädigó ñahé gaǎthó ra ohtomí di päädigó ñahé de gaǎthó kósa*
 di = pädi - gó ja = hé gǎthó ra ohtomí
 1.PRS = saber - 1.ENF hablar = PL.EXCL todo SG otomí
- di = pädi - gó ja = hé **de gǎthó kósa**
 1.PRS = saber - 1.ENF hablar = PL.EXCL **de todo cosa**
 ‘Sabemos hablar todo el otomí, de todas las cosas’

El tipo de constituyente sintáctico también es determinante para la marcación de número. Los nombres que son término de preposición inhiben el uso de las marcas de número en los tres tipos de bilingües. Este comportamiento se constata en la distribución de los datos, pues en los nombres que son término de preposición prevalece la ausencia de estas (**Patrón 2 N**: **BO** = 78%, **BS** = 71%; **BE** = 50%), como se ejemplifica en (74a). Al igual que con los nombres patrimoniales, el uso de las marcas de número con nombres que son término de preposición tiene que ver con la adopción de preposiciones del español (74b), el cual incrementa con el dominio del español.

- (74) a. *Xo do 'yóthóhú ta Ntsuuhni*
 fo do = ʔyó - thó = hí ta nts̃hni
 pues 1.PRFT = caminar - COMPL = PL.INCL hasta Toluca
 'Pues caminábamos hasta Toluca'
- b. *Désde ka ra bés ka ya núnka da bwélbe da puuts'é a hítsi*
désde ka ra bés ka ja núnka da = buélbe
desde DEM.SG SG vez DEM.SG ya nunca 1.PFT = volver
- da = p̃ts̃ʔé a hítsi
 1.PFT = subir LOC tapanco
 'Desde esa vez ya nunca volvía subir al tapanco'

La mayoría de las preposiciones que usan los BO son de origen otomí (58%), en comparación con el 42% de preposiciones del español. Aunado a esto, se observa un uso diferenciado de las marcas de número según el origen de la preposición: los términos de preposiciones otomíes se caracterizan por no llevar las marcas de número (Patrón 1 *ra/ya*: 7%; **Patrón 2 N: 93%**), mientras que cuando la preposición es del español se presenta una distribución de **50% (Patrón 1 *ra/ya*) 50% (Patrón 2 N)**. Los bilingües simétricos adoptan un porcentaje de preposiciones del español (40%) similar al de los BO y también presentan un uso diferenciado de las marcas de número según el origen de la preposición: los términos de preposiciones otomíes no suelen llevar las marcas de número (Patrón 1 *ra/ya*: 22%; **Patrón 2 N: 78%**) y los de preposiciones del español sí (**Patrón 1 *ra/ya*: 63%**; Patrón 2 N: 37%). Aunque en los bilingües con mayor dominio del español no se presenta un incremento en la adopción de preposiciones del español (29%), sí se acentúa el uso de las marcas de número tanto si son preposiciones otomíes como españolas.

La función sintáctica, la distinción nombre común/propio y la referencialidad de la frase nominal determina el uso de las marcas de número en los BO y los BS. La función sintáctica es el segundo factor más importante en los dos grupos de bilingües. Las frases nominales que funcionan como sujeto, objeto primario, objeto secundario y complemento

nominal favorecen el uso de las marcas de número. En los datos, la mayoría de los sujetos (**Patrón 1 ra / ya: BO = 96%; BS = 92%; BE = 92%**), objetos (**Patrón 1 ra / ya: BO = 79%; BS = 79%; BE = 73%**) y, en menor proporción, los complementos nominales llevan las marcas de número, como se ejemplifica en (75a) - (75c). Las frases nominales con función de objeto sin marcas de número se limitan a cierto tipo de verbos o casos en el que el núcleo nominal es continuo, como se presenta en (75d).

- (75) a. *Do k'átsíhé ka ra persona bi thoogi*
do = kʔátsí = hé **ka ra p**érs^ona bi = thǒgi
1.PRFT = mirar = PL.EXCL **DEM.SG SG persona** 3.PFT = pasar
'Vimos que la persona pasó'
- b. *Dimá häxí ya máncá i hún^ota ya basúra*
di-má häjí ya mántj^o i hún^ota **ya basúra**
1.IMPF cortar PL elote y juntar **PL basura**
'Cortaba elotes y juntaba basura'
- c. *Ma hníníhé bi zqá' ka ro tsí khá' ro señor Santiágo*
ma hníní = hé bi zǒʔ **ka ro tsí khá'**
1.POS pueblo = PL.EXCL 3.PRFT bien **DEM.SG SG DIM Dios**

ro sepór santyágo
SG señor Santiágo
'Nuestro pueblo está bien, Dios, el señor Santiago'
- d. *Xo dimá 'ó táko*
fo di-má = ʔó **táko**
pues 1.IMPF = pedir **taco**
'Pues pedía tacos'

En contraste, las frases nominales con función adverbial y predicativa no promueven el uso de las marcas de número en los dos grupos de bilingües. En la distribución de los datos, la mayoría de las frases con función adverbial no lleva las marcas de número (**Patrón 2 N: BO = 63%; BS = 72%**), al igual que las predicativas (**Patrón 2 N = 70%; BS = 69%**), como se ejemplifica enseguida.

- (76) a. *Xo do k'átsí tákerya*
 fo do kʔátsí a **tákerya**
 pues 1.PFT mirar LOC **taquería**
 ‘Pues miré en la taquería’
- b. *Bi k'átsí ka ma ntsaaya na ra 'behña ná mbra nda largo ka ra kabéyo ná*
 bi kʔátsí ka ma ntsɔyΛ na ra ʔbehɲa ná
 3.PRFT mirar DEM.SG 1.POS señor INDEF SG mujer CIT
- mbra nda **largo** ka ra kabéyo ná
 3.CONT.IMPF muy **largo** DEM.SG SG cabello CIT
 ‘Mi esposo vio una mujer, dice, con cabello muy largo, dice’

La distinción entre propio/común es determinante para la marcación de número en los bilingües con mayor dominio del otomí y los bilingües simétricos. En los dos grupos, los nombres propios inhiben el uso de *ra* y *ya*, tal como se advierte en la distribución de los datos (**Patrón 2 N: BO = 58%; BS = 72%**) y en el ejemplo (77a). En contraste, los nombres comunes presentan un mayor porcentaje de marcas de número (**Patrón 1 *ra/ya*: BO = 66%; BS = 62%**).

- (77) a. *Ma pahpá damá paabé a Méhiko*
 ma pahpá da-má pǎ = bé a **méhiko**
 1.POS papá 1.IMPF vender = DU.EXCL LOC **México**
 (Mi papá y yo) íbamos a vender a México
- b. *Xo go má xahé pasí thónde tóka ra trábaho*
 fo go = má ʃa = hé paʃí thónde tóka ra **trábaho**
 pues 1.FUT = ir cortar = PL.EXCL hierba donde toca **sg trabajo**
 ‘Pues nos íbamos a cortar hierba donde tocaba el trabajo’

Aunque los nombres propios no favorecen el uso de las marcas de número, en el corpus hay una presencia importante de *ra* y *ya* (42%). Con base en algunas observaciones preliminares, consideré pertinente revisar si la aparición y la ausencia de las marcas de número en este grupo léxico podría estar relacionado con algún subconjunto, específicamente con la distinción entre antropónimos y topónimos. Para ello, clasifiqué los nombres propios en estas dos categorías y revisé su comportamiento con respecto a las marcas de número. Los

resultados mostraron que todos los nombres propios que llevaban la marca de número *ra* eran antropónimos y los nombres propios que prescindían de las marcas eran topónimos, como se muestra en los siguientes ejemplos.

- (78) a. *Do má a Monterrey do má a Láredo*
do = má a monteréi do = má a **Láredo**
1.PRFT = ir LOC Monterrey 1.PFT = ir LOC **Laredo**
‘Fui a Monterrey, fui a Laredo’
- b. *Ka na ndí ndra húntehé gége ra María ra Andrea ra Pála ra Ino do máhmé*
ka na ndí ndra = húnte = hé gége **ra márya**
DEM INDEF vez 1.CONT.IMPF = juntar = PL.EXCL 3.PRON **SG María**
- ra ándrea ra pála ra íno** do = má = hmé
SG Andrea SG Paula SG Inocencio 1.PFT = ir = PL.EXCL
‘Esa vez nos juntamos con ellos María, Andrea, Paula, Inocencio, fuimos’

La referencialidad de la frase nominal determina el uso de las marcas de número en los BO y los BS. Las frases nominales referenciales promueven el uso de *ra* y *ya*, como se advierte en la distribución de los datos (**Patrón 1 ra / ya: BO = 76%; BS = 65%**) y en (79a). Mientras que en las frases nominales con interpretación atributiva prevalece la ausencia de las marcas de número (**Patrón 2 N: BO = 55%; BS = 51%**), como se ejemplifica en el segundo inciso.

- (79) a. *Por eso tanto ka ma ‘iidá ra Salba bi súfre ni tanto gi gó*
por éso tánto ka ma ʔidá **ra Sálba**
por eso tanto DEM.SG 1.POS hermano.de.mujer **SG Salvador**
- púro ga gólpe mi = gólpe ga ʔalúmno
puro PREP golpe 3.IMPF = golpear PREP alumno
‘Ya no quise ir a la escuela porque puros golpes, les pegaban a los alumnos’

- b. *Ya híndo né da má éskwela porke púro ga gólpe mi golpe ga alúmno*
 ja hín = do = né Da = má **éskwela** porke
 ya NEG = 1.PFT = querer 1.PFT = ir **escuela** porque
- bi = súφre ní tánto gi gó
 3.PFT = sufrir ní tanto que 1.ENF
 'Por eso mi hermano Salvador no sufrió tanto como yo'

En síntesis, la marcación de número en los préstamos del español de los BO, BS y BE está determinada por la presencia de determinantes que codifiquen la información de número, el tipo de constituyente y la concreción del nombre; sin embargo, hay ciertas diferencias. Al igual que con los nombres patrimoniales, hay un incremento en el uso de las marcas de número del otomí conforme incrementa el conocimiento del español de los bilingües y decrece el dominio de la lengua otomí. Este resultado podría responder al debilitamiento de algunas restricciones que determinan la presencia / ausencia de las marcas de número en la lengua otomí. Sin embargo, esta es sólo una hipótesis que tendría que ponerse a prueba en otro tipo de investigación y con un corpus más amplio.

La primera restricción debilitada corresponde al uso de las marcas de número en frases nominales con numerales. En el otomí de los bilingües con mayor dominio del otomí, y en menor medida, en el de los bilingües simétricos, el empleo de las marcas de número con frases nominales con numerales es excepcional; mientras que, en los bilingües con mayor dominio del español, la aparición de las marcas *ra* y *ya* con numerales asciende al 100%.

Se presenta un efecto similar con el tipo de constituyente sintáctico. En el grupo de los BO se presenta una fuerte distinción en el uso de las marcas de número según sea el nombre núcleo de una FN o término de preposición. En este último caso, el uso de *ra* / *ya* no se ve favorecido y los pocos casos que se registran están asociados con el empleo de preposiciones del español. Esta restricción se debilita en el grupo de bilingües simétricos, ya que utilizan

las marcas de número en un contexto donde antes no se usaba. El debilitamiento de esta restricción se agudiza en el grupo de los bilingües con mayor dominio del español.

Conclusiones preliminares

En la sección §4.5 determiné qué factores lingüísticos y sociales inciden en la marcación de número de los nombres patrimoniales y los préstamos del español. Al contrastar los resultados, no hallé grandes diferencias en el comportamiento lingüístico que permitieran distinguir alguna influencia del español. Por ello, no pude negar o afirmar las HIPÓTESIS 4.1 y 4.2. En este apartado §4.6 presenté un análisis de la marcación de número en los nombres patrimoniales y los préstamos a partir del tipo de bilingüismo, con el propósito de identificar alguna influencia del español en los distintos grupos léxicos y en los diferentes tipos de hablantes.

El análisis de los nombres patrimoniales y los préstamos según el tipo de bilingüismo (véase §4.6.1 y §4.6.2) muestra un aumento en el uso de las marcas de número del otomí conforme incrementa el conocimiento del español de los bilingües y decrece el dominio de la lengua otomí. Este comportamiento está relacionado con el debilitamiento de algunas restricciones que determinan la presencia o ausencia de las marcas de número en dos contextos: las frases nominales con numerales y los términos de preposición.

Con respecto a las frases nominales con numeral, parece haber un fenómeno variable inducido por contacto, en el que el otomí podría pasar de ser una lengua en la que los nombres no establecen una relación de concordancia semántica con el numeral a una que sí, tal como sucede en español y el otomí del Valle del Mezquital (Acosta Aguilar 2020). Esta coincidencia dificulta argumentar de forma clara una influencia del español. Sin embargo, el

hecho de que este proceso esté más avanzado en el grupo de los BE y esté menos latente en los BO abona en favor de las HIPÓTESIS 4.1 y 4.3.

Los resultados afirman la HIPÓTESIS 4.1 —que señala que el español ha influido de manera indirecta en la marcación de número de modo que la aparición de las marcas estará determinada por alguno(s) de lo(s) condicionamientos lingüísticos del español— porque los bilingües encuentran similitudes y diferencias en las propiedades combinatorias de las marcas de número del otomí *ra / ya* y el español *-(e)s* en las FNs con numerales y, a partir de eso, establecen una equivalencia entre ellas. Esta equivalencia combinatorial se puede apreciar en el incremento de la concordancia semántica en las FNs con numeral en otomí (*Kuutá ya nguu* ‘Cinco casas’), como sucede en español (*Cinco casas*). Asimismo, afirman la HIPÓTESIS 4.3 —que indica que la fuerza del efecto de los condicionamientos lingüísticos que rigen la marcación de número en el español es mayor en los hablantes dominantes de esa lengua y menor en los bilingües con más dominio del otomí—, pues la presencia de concordancia semántica en las FNs con numerales es mayor en los BE que en los BS y BO.

En cuanto a la presencia de las marcas de número en los términos de preposición, también parece haber un fenómeno variable inducido por contacto, en el que el otomí podría pasar de ser una lengua en la que los nombres que son término de preposición no llevan las marcas de número a una que sí, por influencia del español. Además, el hecho de que este proceso esté más avanzado en el grupo de los BE y esté menos latente en los BO abona en favor de las HIPÓTESIS 4.2 y 4.3.

El influjo del español comprende al menos dos pasos (HIPÓTESIS 4.1). El primero es la adopción de preposiciones del español que pueden o no llevar las marcas de número, dependiendo de la preposición. La segunda parte ocurre cuando los bilingües encuentran similitudes y diferencias en las propiedades combinatorias de las marcas de número del

otomí (*ra / ya*) y el español (*-(e)s*) con las preposiciones *y*, a partir de eso, establecen una equivalencia entre ellas. Esta equivalencia combinacional se puede apreciar en una mayor presencia de las marcas de número en frases con preposiciones del español que se expande a las del otomí. Asimismo, prueban la HIPÓTESIS 4.3 pues la presencia de concordancia semántica en las FNs con numerales es mayor con los BE que con los BS y BO.

En cuanto a la HIPÓTESIS 4.2 —que señala que la fuerza del efecto del español depende del origen histórico de los nombres—. Los resultados muestran un comportamiento lingüístico similar en los dos grupos léxicos; sin embargo, los préstamos parecen ir a la cabecera en la expansión de las marcas de número, lo cual abona en favor de dicha hipótesis. A continuación, seguiré evaluando esta hipótesis a partir de un índice de integración de las piezas léxicas y el efecto de la frecuencia léxica.

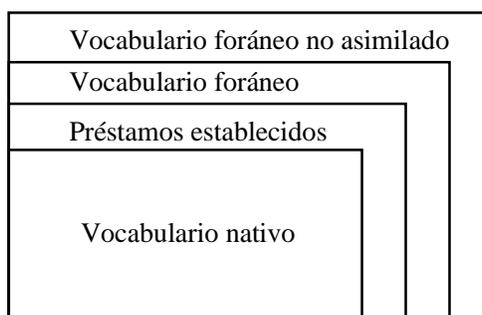
4.7 ANÁLISIS EN TIEMPO APARENTE: ÍNDICE DE INTEGRACIÓN LÉXICA

En este apartado realizo un análisis en tiempo aparente del uso de las marcas de número *ra* y *ya* mediante el uso de un índice de integración léxica (IIL) y una clasificación categórica del léxico que, espero, permita establecer un eje temporal dentro de la organización del lexicón otomí. Para la construcción de este índice me basé en dos fuentes principales. En primera instancia, partí de la idea de estratificación de léxico de Ito y Mester (1995) y retomé algunos de los parámetros que se han propuesto para diferenciar entre casos de préstamo y cambio de código (Gómez Rendón, 2008; Poplack, 1980; Poplack & Meechan, 1998) véase § 4.2.2.

Con respecto a la propuesta de la estratificación del léxico, Ito y Mester (1995) señalan que el léxico de las lenguas está organizado en una estructura que comprende un núcleo y una periferia. En el núcleo se encuentra el vocabulario nativo, que cumple con todas las restricciones léxicas. En el otro extremo se halla la periferia, conformada por ítems

foráneos no asimilados que sólo cumplen con un subconjunto de restricciones (Ito & Mester, 1995: 182), tal como se esquematiza a continuación.

FIGURA 4.12 *Estratificación del léxico. Adaptado de (Ito & Mester, 1995: 182)*



Más adelante clasificaré los nombres con base en esta propuesta, sin embargo, debido al tamaño limitado del corpus, sólo retomaré tres de las cuatro categorías, a saber: vocabulario nativo, préstamos establecidos y préstamos no establecidos. Asimismo, utilizaré algunos de los parámetros empleados para diferenciar entre cambio de código monoléxico y préstamo: adaptación fonológica (Gómez Rendón, 2008; Poplack, 1980; Poplack & Meechan, 1998), frecuencia de uso (Gómez Rendón, 2008; Poplack & Sankoff, 1984), dispersión (Gómez Rendón, 2008; Poplack & Sankoff, 1984), equivalente semántico (Poplack & Sankoff, 1984) y documentación del ítem léxico (Poplack & Sankoff, 1984).

4.7.1 *La construcción del índice de integración léxica*

Con el objetivo de tener un indicador que diera cuenta del grado de integración de los préstamos al léxico otomí, implementé un índice de integración léxica (IIL), en forma de gradiente, mediante una aproximación de índice ponderado. Construí el IIL a partir de cinco criterios: 1) la adaptación fonológica (adap), 2) la frecuencia de uso del ítem léxico en el corpus (freq), 3) la dispersión del ítem léxico entre los hablantes (disp.), 4) la existencia de

un equivalente semántico en otomí (equiv) y 5) la documentación del préstamo (doc) en los compendios léxicos y gramaticales del otomí del centro (Lastra, 1989, 1992).

La adaptación fonológica de los préstamos es una variable categórica con cuatro valores: préstamos léxicos sin adaptación, préstamos que no necesitan ser adaptados porque su estructura sigue las pautas fonológicas del otomí, préstamos adaptados y nombres patrimoniales de la lengua otomí.

La frecuencia del ítem léxico es una variable continua que se calculó del siguiente modo. Primero, determiné la frecuencia absoluta de cada lexema y la dividí entre el número total de nombres en el corpus. Debido a que el resultado es una cifra pequeña, lo multipliqué por diez. Por ejemplo, el préstamo *escuela* aparece 31 veces en el corpus, dividí esta cantidad entre el número total de nombres $31 / 2573$, y obtuve como resultado 0.012048. Multipliqué esta cifra por diez y así conseguí la frecuencia relativa de 0.1204. La dispersión del ítem léxico también es una variable continua, la determiné calculando el número de hablantes que emplearon cada ítem léxico y dividí esta cantidad entre el número total de hablantes. Retomando el ejemplo de *escuela*, este préstamo fue empleado por 11 de los 26 bilingües, lo cual dio como resultado la cantidad de 0.423076.

El factor equivalencia del préstamo es una variable categórica, que comprende tres categorías: las palabras originarias (PO), como es el caso de *dethá* ‘maíz’; los préstamos sin equivalente en otomí (PresSEO), por ejemplo, la palabra *cemento*; y los préstamos con equivalente en otomí (PresCEO), como sucede con la palabra *camino*. Por último, la documentación del préstamo en los compendios léxicos y gramaticales del otomí del centro (Lastra, 1989, 1992) es un factor categórico, que comprende tres grupos: las palabras originarias (PO), como es el caso de *ϕuui* ‘sombrero’; los préstamos documentados

(PresDoc), por ejemplo, *dómi* ‘dinero’; y los préstamos no documentados (PresNoDoc), tal es el caso de *escuela*.

Aunado a esto, establecí una jerarquía entre los cinco factores. A pesar de que la adaptación fónica no es un factor infalible para diferenciar un préstamo de un caso de cambio de código o para determinar la antigüedad de un préstamo, la consideré el factor de mayor relevancia en la construcción del IIL porque es el rasgo menos dependiente de las características de nuestro corpus, en comparación con el resto de las variables.

La frecuencia y la dispersión del ítem léxico ocupan la segunda y la tercera posición en la jerarquía, debido a que su valor puede estar determinado por el tópico de las narraciones. La equivalencia y la documentación del ítem léxico son el cuarto y quinto factor, debido a que la metodología empleada en la recolección de los datos, así como los temas tratados inciden en los valores calculados y no necesariamente son representativos de la realidad.

Para estimar el IIL es necesario que los indicadores utilizados sean numéricos. Por ello, transformé los indicadores categóricos (adaptación fonológica, equivalencia y documentación) a valores numéricos ordinales, como se muestra en la Tabla 4.21. Dentro de estos indicadores, asigné los valores más altos a los ítems asociados con un mayor nivel de integración, mientras que los valores más bajos corresponden a palabras asociadas un menor grado de integración. Por ejemplo, con respecto a la adaptación fonológica, le asigné el valor 0 a las palabras que no presentan ningún tipo de adaptación y el valor 3 a las palabras patrimoniales del otomí. Debido a que la frecuencia y dispersión son indicadores cuya naturaleza es numérica no fue necesario aplicar ninguna transformación. Cada uno de los factores y sus valores se resumen enseguida.

TABLA 4.21 Factores para construir el índice de integración léxica

JERARQUÍA	FACTOR	TIPO	VALORES
1	Adaptación fonológica	Categoría	0 = Sin adaptación 1 = Sin necesidad de adaptación 2 = Adaptación 3 = Patrimonial
2	Frecuencia	Continua	0 1
3	Dispersión	Continua	0 1
4	Equivalencia	Categoría	0 = PresCEO 1 = PresSEO 2 = PO (palabra originaria)
5	Documentación (Lastra 1989, 1992)	Categoría	0 = PresNoDoc 1 = PresDoc 2 = PD (palabra originaria)

La fórmula para calcular el IL fue la siguiente:

Índice ponderado

$$\text{Índice de integración} = \sqrt{v_1 w_1 + v_2 w_2 + v_3 w_3 + v_4 w_4 + v_5 w_5}$$

$$IL = \sqrt{\underset{\substack{\downarrow \\ \text{Normalización 0-1}}}{adap} * w_1 + \underset{\substack{\downarrow \\ \text{Normalización 0-1}}}{freq} * w_2 + disp * w_3 + equiv * w_4 + doc * w_5}$$

Escala 0 - 100

Valores de ponderación usados

$$W_1 = 1$$

$$W_2 = 3$$

$$W_3 = 3$$

$$W_4 = 1.5$$

$$W_5 = 1.5$$

Esto es, la raíz cuadrada de la sumatoria de los cinco indicadores antes descritos multiplicado por un valor de ponderación (w). En este caso, el valor de ponderación permite

que cada elemento en la formula aporte por igual al cálculo del IIL ya que, como lo evidencia la Tabla 4.30, cada indicador difiere en su valor máximo. Para conseguir esto, los valores de ponderación aplicados fueron 1, 3, 3, 1.5 y 1.5 para w_1 , w_2 , w_3 , w_4 y w_5 respectivamente, de modo que el resultado del producto de cada indicador por su valor de ponderación tuviera como valor máximo 3 en cada caso. Finalmente, apliqué la fórmula 1. Para facilitar su interpretación, el IL fue reescalado en un rango de 0 a 100, a través de la función *rescale* de la paquetería *scales* de R. Es así que los valores más cercanos a 100 denotan una mayor integración mientras que los valores más cercanos a 0 una menor integración.

4.7.2 Clasificación categórica del léxico

Además del índice de integración léxica, realicé una clasificación categórica del léxico con base en tres de las categorías propuestas por Ito y Mester (1995) para la estratificación del léxico: vocabulario nativo, préstamos establecidos y préstamos no establecidos. Dentro de la categoría vocabulario nativo consideré los nombres de origen otomí, analizados en § 4.5.1.

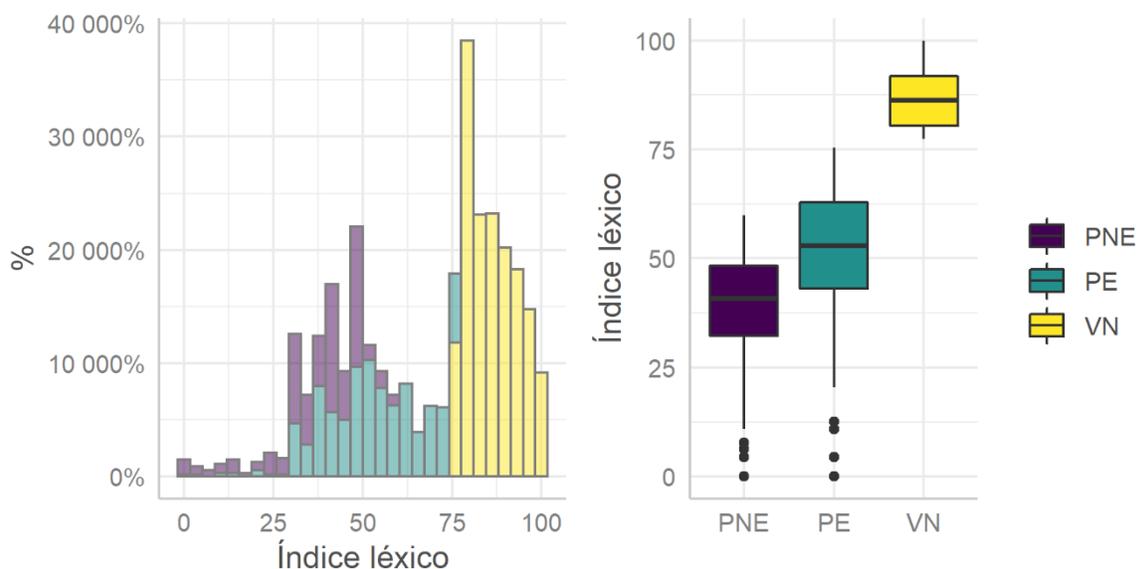
La categoría préstamo establecido comprende aquellos elementos léxicos que 1) son considerados como préstamos “antiguos” en la lista propuesta por Hekking y Bakker (2010) 2) han sido previamente documentados en los compendios léxicos y gramaticales de esta variedad de otomí (Lastra, 1989, 1992), 3) son frecuentes en el corpus, 4) fueron empleados por más de un colaborador y 5) muestran una adaptación fónica completa.

En la categoría de préstamos no establecidos comprende aquellos elementos léxicos que 1) forman parte de la lista denominada “préstamos recientes” de Hekking y Bakker (2010) pueden o no estar documentados en los compendios léxicos y gramaticales del otomí del centro (Lastra 1989, 1992); 3) no son tan frecuentes como los préstamos establecidos; 4)

no las usa más de un hablante y 5) no tienen una adaptación completa a la fonología del otomí.

Con el objetivo de hacer una evaluación del índice léxico en términos de estadística descriptiva analicé la correspondencia que existe entre el índice léxico y las tres categorías antes descritas. De esta forma se puede observar en la siguiente figura que el índice léxico, se corresponde bastante bien con las categorías léxicas, pues los valores más bajos del índice se asocian a las palabras que son identificadas en la categoría de préstamos no establecidos, los valores intermedios se asocian a los préstamos establecidos y finalmente los valores más altos al vocabulario nativo. Esto sugiere entonces que el índice propuesto es adecuado para describir un gradiente de cambio entre préstamos no establecidos y vocabulario nativo.

FIGURA 4.13 *Evaluación del índice léxico*



4.7.3 *El uso de las marcas de número ra y ya a través del índice de integración léxica*

Después de construir el índice léxico, realicé una regresión logística escalonada para determinar el efecto que este factor en conjunto con otros factores lingüísticos y

extralingüísticos tienen en la marcación de número (Patrón 1: *ra/ya*) y (Patrón 2: N), a saber: la presencia de determinantes que codifican el número, la referencialidad, el tipo de constituyente, el índice de integración, la localidad, la edad, el género y el bilingüismo.¹²¹ La Tabla 4.22 indica que, de los ocho factores considerados, sólo cuatro resultaron significativos, en el siguiente orden de importancia: (i) la presencia o ausencia de determinantes que codifican el número, (ii) el tipo de constituyente, (iii) la referencialidad y (iv) el índice de integración.

TABLA 4.22 *Resultados de la regresión logística escalonada. Índice léxico*¹²²

Variables	Estimado	Error Est.	Valor z	Pr(> z)
Intersección	1.939737	0.225773	8.592	< 2e-16 ***
Frases sin determinantes	-1.975304	0.156899	-12.590	< 2e-16 ***
Frases con numeral	-3.982035	0.419853	-9.484	< 2e-16 ***
Frases con indefinido	-0.600118	0.237430	-2.528	0.0115 *
Término de preposición	-1.582267	0.127721	-12.388	< 2e-16 ***
Referencialidad	0.677915	0.096708	7.010	2.39e-12 ***
Índice léxico	0.004240	0.002068	2.050	0.0404 *
Localidad: SAC	-0.197666	0.111697	-1.770	0.0768 .
Edad: mayores	-0.165727	0.096876	-1.711	0.0871 .
Sexo: hombres	0.224079	0.115255	1.944	0.0519 .

Como habíamos observado en análisis anteriores, el tipo de determinante es el factor más importante en el uso de las marcas de número. Las frases nominales sin determinantes

¹²¹ Antes de llevar a cabo la regresión logística escalonada se empleó la medida de Goodman y Kruskal del paquete “The GoodmanKruskal” en la plataforma de R para determinar si existe colinealidad entre las variables independientes (i-xii). En el Apéndice 4F, se presenta los resultados en una escala de 0 a 1, en donde un valor más cercano a 1 indica mayor correlación entre las variables y más probabilidad de problemas de colinealidad.

¹²² En la regresión logística se tuvo una variable de respuesta dicotómica, (Patrón 1: *ra/ya*) versus (Patrón 2: N). En la primera columna de la tabla se encuentran las variables más relevantes según el modelo aplicado. En la columna con la etiqueta “Estimado” se ofrecen los resultados sobre el sentido de la relación en una escala de log(odds). Los valores en el intervalo de 0 a $+\infty$ indican una relación positiva, es decir, un incremento en el uso de las marcas de número *ra/ya*. Los valores en la escala de 0 a $-\infty$ señalan una relación negativa, es decir, un decrecimiento en la probabilidad de que se emplee las marcas de número *ra/ya*. En las siguientes dos columnas se proporcionan los valores de error estándar y de Z. En la última columna se presentan los valores de probabilidad, que indican si la variable en cuestión es significativa, de acuerdo con un nivel de significación menor a 0.05.

(-1.975304), con un numeral (-3.982035) o con la marca de indefinitud (-0.600118) inhiben la aparición de *ra* y *ya*. El segundo factor de mayor importancia fue el tipo de constituyente sintáctico. Los nombres que son término de preposición también disminuyen la probabilidad de que los bilingües empleen las marcas de número (-1.582267). En tercera posición está la referencialidad de la frase nominal y son las frases referenciales las que favorecen la aparición de la morfología de número (0.677915). El último factor de importancia es el índice de integración léxica. Al respecto, vemos que a mayor integración de un lexema mayor es la posibilidad de que los bilingües utilicen las marcas de número *ra* y *ya* (0.004240).

4.7.4 El uso de las marcas de número ra y ya a través de la clasificación categórica del léxico

En este apartado exploro el efecto que tiene la estratificación del léxico en la marcación de número, mediante la clasificación categórica que hice de los nombres en el corpus: vocabulario nativo (VN), préstamos establecidos (PE) y préstamos no establecidos (PNE). Para ello, también realicé una regresión logística escalonada con una variable de respuesta dicotómica, (Patrón 1: *ra* / *ya*) *versus* (Patrón 2: N) y como variables explicativas: la función sintáctica de la FN, el tipo de constituyente, la concreción del nombre, la distinción entre nombre propio y común, la animacidad, la continuidad del nombre, la presencia de determinantes que codifican el número, la referencialidad, la categoría léxica, la localidad, la edad, el género y el bilingüismo.¹²³

¹²³ Antes de llevar a cabo la regresión logística escalonada se empleó la medida de Goodman y Kruskal del paquete “The GoodmanKruskal” en la plataforma de R para determinar si existe colinealidad entre las variables independientes (i-xii). En el Apéndice 4G, se presenta los resultados en una escala de 0 a 1, en donde un valor más cercano a 1 indica mayor correlación entre las variables y más probabilidad de problemas de colinealidad.

Los resultados muestran que, de las trece variables consideradas, cuatro resultaron significativas, en el siguiente orden de importancia: (i) la presencia o ausencia de determinantes que codifican el número, (ii) el tipo de constituyente, (iii) la referencialidad y (iv) la categoría léxica.

TABLA 4.23 *Resultados de la regresión logística escalonada. Grupos léxicos: clasificación categórica*¹²⁴

Variables	Estimado	Error Est.	Valor z	Pr(> z)
Intersección	2.07901	0.19257	10.796	< 2e-16 ***
Frases sin determinantes	-1.99412	0.15742	-12.668	< 2e-16 ***
Frases con numeral	-3.99728	0.42136	-9.487	< 2e-16 ***
Frases con indefinido	-0.60583	0.23780	-2.548	0.01084 *
Término de preposición	-1.57135	0.12786	-12.290	< 2e-16 ***
Referencialidad	0.68846	0.09700	7.097	1.27e-12 ***
ClasificaciónPNE	-0.03038	0.13826	-0.220	0.82607
ClasificaciónVN	0.29556	0.11304	2.615	0.00893 **
Edad: mayores	-0.17942	0.09741	-1.842	0.06550 .
Sexo: hombres	0.22333	0.11521	1.939	0.05256
Localidad: SAC	-0.17686	0.11203	-1.579	0.11442

Al igual que en la prueba anterior, el tipo de determinante fue el factor de mayor importancia. Las frases nominales sin determinantes (-1.99412), con numerales (-3.99728) o con la marca de indefinitud *na* (-0.60583) inhiben la aparición de las marcas de número. De igual forma, los nombres que son término de preposición no favorecen la presencia de *ra* y *ya* (-1.57135). En tercer lugar, se encuentra la referencialidad de la frase nominal y son las frases referenciales las que favorecen el uso de la morfología de número (0.68846). El último

¹²⁴ En la regresión logística se tuvo una variable de respuesta dicotómica, (Patrón 1: *ra* / *ya*) versus (Patrón 2: N). En la primera columna de la tabla se encuentran las variables más relevantes según el modelo aplicado. En la columna con la etiqueta “Estimado” se ofrecen los resultados sobre el sentido de la relación en una escala de log(odds). Los valores en el intervalo de 0 a $+\infty$ indican una relación positiva, es decir, un incremento en el uso de las marcas de número *ra* / *ya*. Los valores en la escala de 0 a $-\infty$ señalan una relación negativa, es decir, un decrecimiento en la probabilidad de que se emplee las marcas de número *ra* / *ya*. En las siguientes dos columnas se proporcionan los valores de error estándar y de Z. En la última columna se presentan los valores de probabilidad, que indican si la variable en cuestión es significativa, de acuerdo con un nivel de significación menor a 0.05.

factor en la jerarquía es la categoría léxica. Al respecto, vemos que los nombres pertenecientes al grupo vocabulario nativo promueve el uso de las marcas de número (0.29556), mientras que los préstamos establecidos parecen inhibirlo.

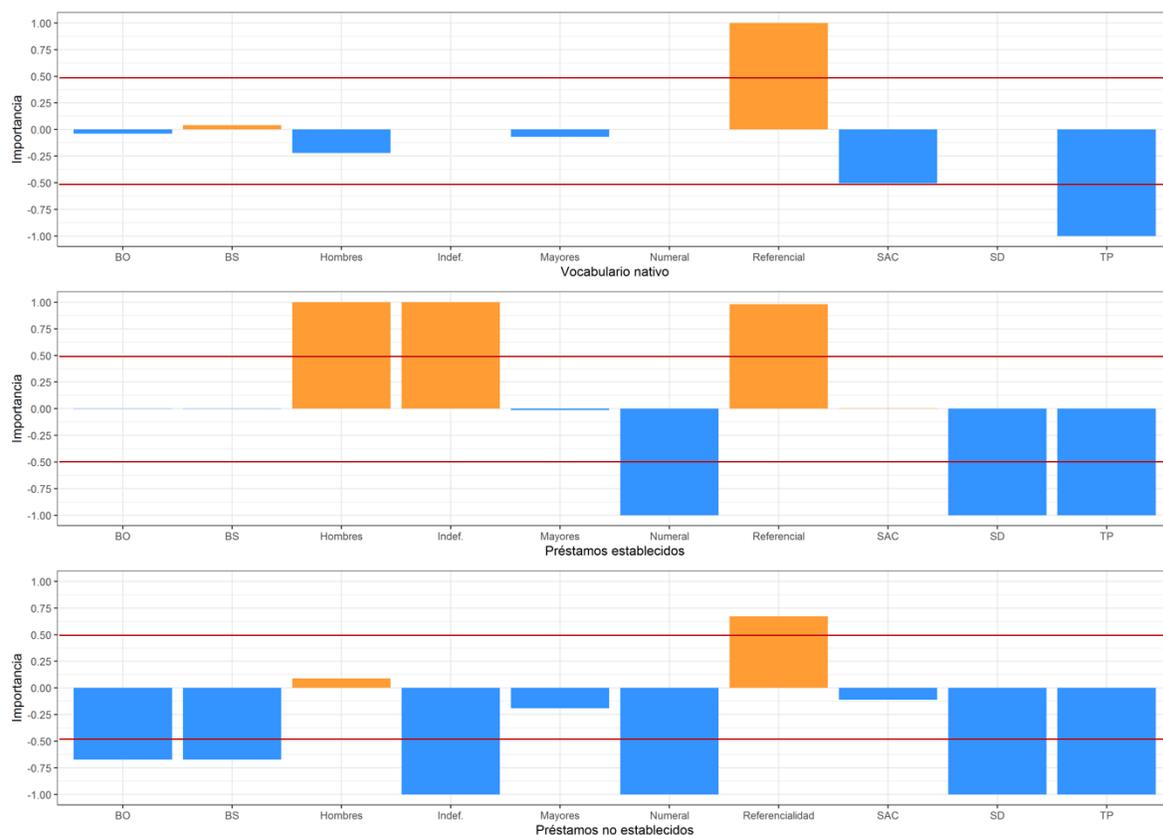
Estos resultados sugieren un uso diferenciado en la marcación de número en los grupos léxicos. Por ello, consideré necesario indagar acerca de la marcación de número en cada uno de los conjuntos léxicos para determinar si los factores que determinan el uso de la morfología de número son los mismos en los tres grupos o difieren, mediante el uso de una prueba de inferencia multimodelo (véase §2.1.3.5).

En la Figura 4.14 se resumen los resultados de las palabras nativas, los préstamos establecidos y los préstamos no establecidos. La marcación de número está determinada por los mismos factores lingüísticos en los tres grupos léxicos y, además, estos inciden en la misma dirección salvo en un contexto. Las frases nominales sin determinantes, con numerales y con la marca de indefinitud no favorecen el uso de las marcas de número, excepto en el conjunto de préstamos no establecidos, en donde la marca de indefinitud favorece el uso de *ra* y *ya*. Los nombres que son término de preposición inhiben la aparición de las marcas de número y las frases referenciales la favorecen.

La principal diferencia entre los tres grupos léxicos se halla en los factores sociolingüísticos. En el vocabulario nativo, sólo incide la comunidad de origen de los hablantes, y son los habitantes de San Andrés Cuexcontitlán quienes inhiben el uso de las marcas de número. En los préstamos establecidos se presenta un uso diferenciado de la morfología de número según la edad y el género de los hablantes. Los hablantes de más de 55 años inhiben el uso de las marcas *ra* y *ya* y los hombres lo favorecen. Por último, en la marcación de número de los préstamos no establecidos interviene el tipo de bilingüismo de

los hablantes: los bilingües simétricos y los dominantes en otomí no promueven la utilización de la morfología de número.

FIGURA 4.14 *La marcación de número en los tres grupos léxicos*



El efecto de la variable comunidad en la marcación de número del vocabulario nativo del otomí ha sido constante en los análisis antes realizados (véase §4.5.3) y en el presente análisis. Este comportamiento concuerda con lo previsto, pues es esperable la existencia de una norma de uso distinta en diferentes comunidades de habla, como ya lo habíamos observado en la sección §4.5.1, cuando advertimos algunas diferencias en el uso de las marcas de número entre el otomí del Valle del Mezquital (Acosta Aguilera, 2020) y el otomí del centro.

La diferenciación en el uso de las marcas de número en el vocabulario nativo de los bilingües de SAC y JV, genera la pregunta de si estas discrepancias sólo son cuantitativas o si detrás de ellas existe una diferencia de tipo gramatical. Por el momento, no daré respuesta a esta pregunta; sin embargo, es posible que las diferencias sólo sean cuantitativas y no estén relacionadas con la existencia de gramáticas distintas, tal como sucede con el fenómeno de concordancia plural en el español de estas comunidades (véase §3.6).

Por otro lado, hay una mayor variación sociolingüística en los préstamos establecidos y los no establecidos. Resulta interesante el efecto que la edad de los colaboradores tiene en la marcación de número de los préstamos establecidos pues, a pesar de que gran parte de estos ítems llevan un tiempo considerable en el lexicón otomí (p. ej. el léxico asociado a prácticas religiosas), hay un uso diferenciado según la edad de los hablantes: los bilingües con más de 55 años inhiben el uso de las marcas de número y los más jóvenes (20-34 años) lo promueven.

Este hecho concuerda con uno de los planteamientos de la sección §4.6, donde advertía un incremento en el uso de las marcas de número del otomí conforme incrementaba el conocimiento del español de los bilingües y decrecía el dominio de la lengua otomí, a causa del debilitamiento de algunas restricciones que determinan la presencia / ausencia de las marcas de número en la lengua otomí. En relación con este punto, es relevante el efecto que el tipo de bilingüismo tiene en la marcación del número de los préstamos no establecidos, pues el efecto que advertíamos en la sección §4.5.2 parece no aplicar a todo el conjunto de préstamos del español, sino sólo a aquellos que se encuentran en la periferia del lexicón otomí. Por otro lado, la homogeneidad en el comportamiento lingüístico de la marcación de número en los tres grupos lingüísticos señala que, a pesar de los cambios cuantitativos en el uso de la morfología de número, detrás de estos actúan los mismos factores lingüísticos.

4.7.5 El efecto de la frecuencia léxica en el uso de ra y ya

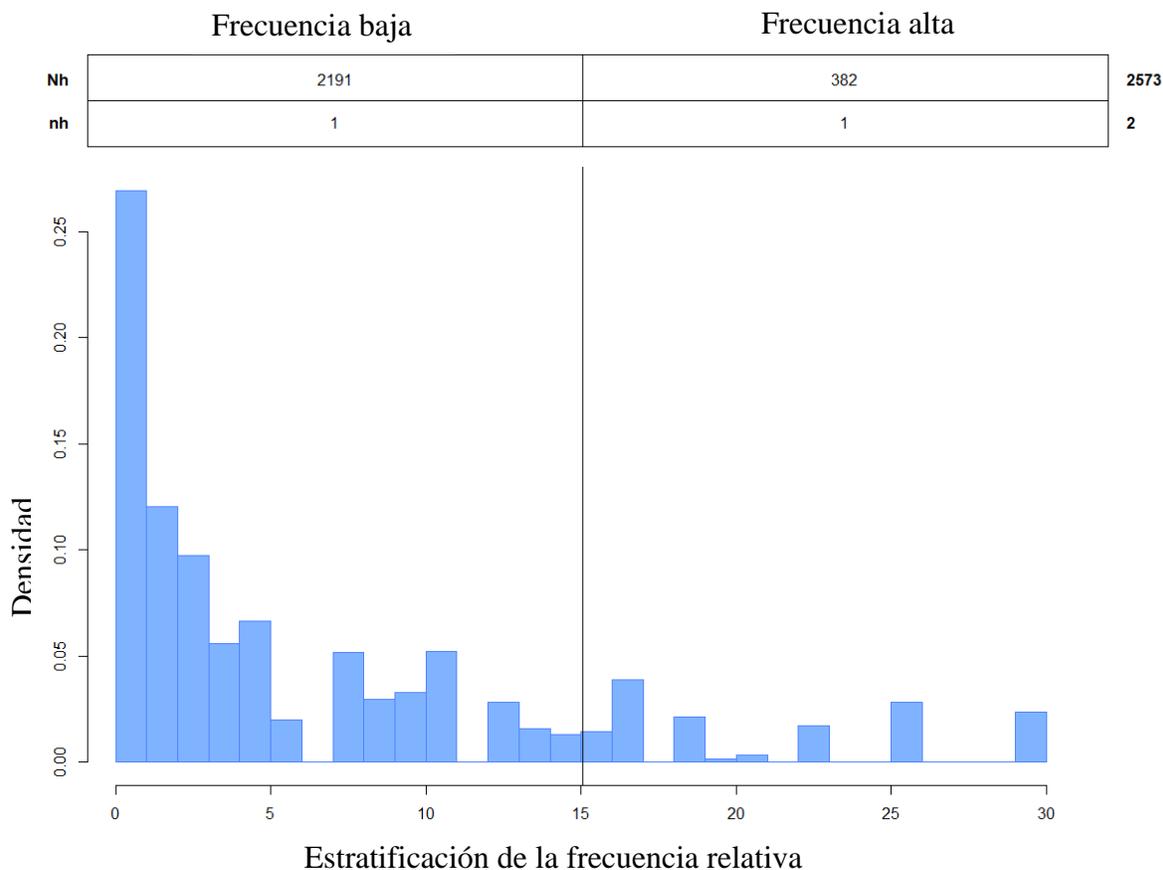
Bybee ha subrayado la importancia de la frecuencia de uso en la conformación de la estructura de las lenguas y los procesos de cambio lingüístico (2007, 2010). De manera general, ha señalado que la frecuencia puede tener un efecto conservador o innovador dependiendo del nivel y el tipo de cambio lingüístico. Los cambios fonológicos están motivados fonéticamente y afectan primero a los ítems más frecuentes (Bybee, 2007: 23-24), por lo que la frecuencia tiene un efecto innovador. Contrario a esto, el cambio analógico afecta primero los ítems menos frecuentes, lo cual le otorga un efecto conservador a la frecuencia de uso (Bybee, 2007, 2010). Dada la importancia que la frecuencia tiene en los procesos de variación y cambio, consideré pertinente evaluar el efecto que la frecuencia de uso de los ítems léxicos tiene en la marcación de número del lexicón otomí (vocabulario nativo y préstamos del español). Por ello, en esta sección evaluaré el efecto de la frecuencia de uso de las piezas léxicas en el grado de influencia del español.

Para determinar el efecto que la frecuencia del ítem léxico tiene en la marcación de número formé dos categorías, de acuerdo con la frecuencia de cada nombre en el corpus. Para ello, primero calculé la frecuencia relativa de cada palabra en el corpus, dividiendo la frecuencia absoluta de cada nombre entre la sumatoria de la frecuencia de todos los nombres en el corpus. Para un mejor manejo de estos datos, multipliqué la frecuencia relativa por 1000 de modo que el rango de esta variara entre cero y treinta.

Con esos datos construí categorías de frecuencia de los nombres a través de una estratificación de frecuencia de raíz acumulada (Dalenius & Hodges, 1959), con un coeficiente de variación máximo entre categorías igual a 0.5. Este procedimiento se llevó a

cabo con la paquetería *stratification* de R. Como resultado de lo anterior se obtuvieron dos categorías: nombres con frecuencia baja y con frecuencia alta, como se muestra enseguida.

FIGURA 4.15 *Estratificación por frecuencia*

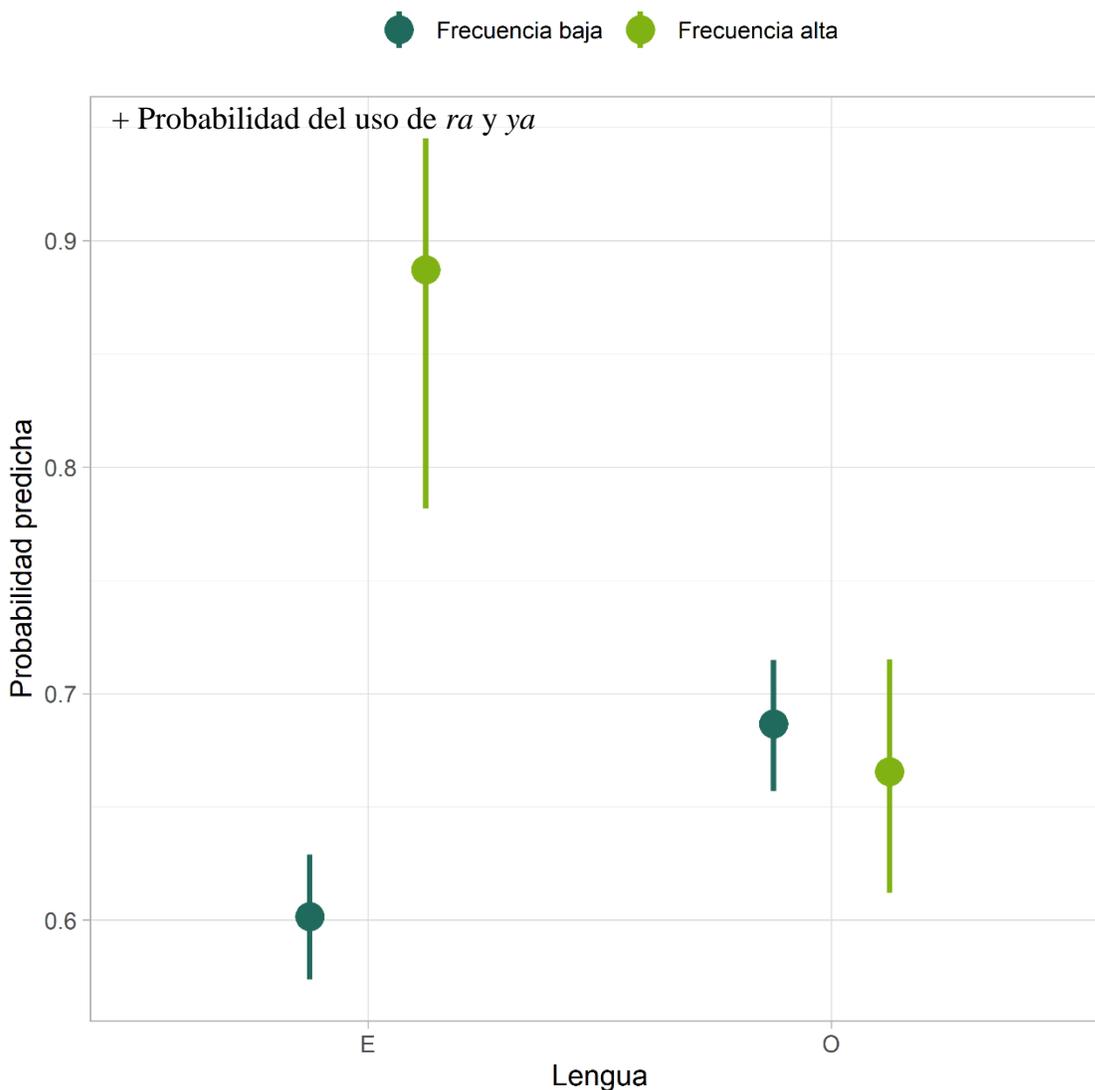


En el corpus de 2573 nombres, 2197 ítems pertenecen al grupo de frecuencia baja (85.15%), pues su índice de uso oscila entre 0 a 15 puntos. Los ítems léxicos con una frecuencia alta tienen una frecuencia relativa de entre 15 y 30 puntos y sólo comprenden 382 casos (14.85%).

En un análisis de carácter descriptivo que realicé previamente, observé que la frecuencia tenía un efecto diferente en la marcación de número de los nombres originarios del otomí y los préstamos del español. Por ello, decidí poner a prueba la hipótesis de que la marcación de número mostraba un comportamiento diferente dependiendo de la frecuencia

de uso del nombre y su origen (español u otomí). Para esto, utilicé un modelo lineal generalizado de tipo binomial en donde la variable predictora consistió en la interacción entre la categoría de frecuencia del nombre y el origen de este, mientras que la variable de respuesta fue el patrón de marcación 1 *ra* y *ya*. Para mejorar la visualización de este modelo se elaboraron gráficas de los efectos predichos por el modelo utilizando la paquetería ggpredict de R.

FIGURA 4.16 *El efecto de la frecuencia en la marcación de número de los nombres patrimoniales y los préstamos del español*



En la figura anterior se aprecia que, tal como postulé, la frecuencia de uso del ítem léxico tiene un efecto diferente en el uso de las marcas de número (Patrón 1 *ra/ya*) según el origen de los nombres. En los nombres patrimoniales no se observa una diferencia prominente en el uso de *ra* y *ya* asociado a la frecuencia de uso de los ítems léxicos. En contraste, en el grupo de préstamos del español, sí podemos ver un uso diferenciado de la morfología de número en relación con la frecuencia de uso. Los préstamos más frecuentes tienen una mayor probabilidad de presentar las marcas de número, mientras que los ítems con frecuencia baja tienen una menor probabilidad de presentar las marcas *ra* y *ya*. Aunado a esto, es importante señalar que todos los préstamos del español, independientemente de su frecuencia de uso, tienen un comportamiento similar al de los nombres patrimoniales con respecto a la marcación de número. El grupo de préstamos que se aproxima más al comportamiento de los nombres patrimoniales es el de los préstamos menos frecuentes. Los préstamos más frecuentes, por su parte, parecen ser los que lideran la generación del uso de las marcas de número, independientemente de las restricciones de la lengua otomí.

4.8 CONCLUSIONES DEL CUARTO CAPÍTULO

Los primeros resultados de este capítulo están relacionados con la composición del corpus. El 22.83% de los ítems del corpus son de origen español. Aunque este porcentaje no alcanza el 40% de elementos del español que caracteriza los procesos de relexificación, como el del náhuatl en La Malinche (Hill & Hill, 1977), sí representa una fuerte influencia del español en el lexicon del otomí de San Andrés Cuexcontitlán y Jiquipilco el Viejo. Además, este porcentaje es mayor al 9% de elementos del español que se documenta en el maya yucateco de Kopchen y Chemax (Pool Balam & le Guen, 2015).

El segundo punto a resaltar es la presencia de diferentes categorías léxicas y gramaticales del español, cuya proporción en el corpus se ordena del siguiente modo: Nombres > Conjunciones > Verbos > Pronombres > Preposiciones > Adverbios > Adjetivos > Numerales > Otros. Esta disposición de los datos concuerda, en mayor o menor medida, con las escalas de prestabilidad de Whitney (1881), Haugen (1950) y Muysken (1981) con respecto a la posición de los nombres y los verbos, pero difiere en la posición de los adjetivos, las conjunciones y las preposiciones. La posición de las conjunciones muestra similitudes con la propuesta de Matras (2007), quien pone en la primera posición de su escala a los nombres y las conjunciones, y en segundo lugar a los verbos.

El bajo porcentaje de préstamos adjetivales en el corpus es similar al documentado en otras variedades de otomí, como la de Santiago Mexquititlán. Al respecto, Hekking y Bakker (2007b) han señalado que este fenómeno se debe a que el otomí no tiene una clase léxica que funcione como modificador de una frase nominal; sin embargo, han empleado este mismo argumento para explicar la adopción masiva de conjunciones y preposiciones. Aunque este capítulo no tuvo como objetivo plantear una escala de prestabilidad, ni determinar las motivaciones del préstamo léxico, los resultados piden repensar la validez de la hipótesis del “hueco funcional”, pues parece ser insuficiente para explicar la motivación del préstamo. Asimismo, son una invitación para plantear nuevas aproximaciones en las que no sólo se consideren las restricciones estructurales, sino también las motivaciones pragmáticas y comunicativas que dan lugar al préstamo.

La influencia del español en el léxico del *ñható* es más notoria en el ámbito nominal, pues el 48.5% de los nombres del corpus son de origen español. La distribución del uso de los préstamos del español en las dos comunidades, en los diferentes grupos etarios y en los dos géneros fue la esperada. Las personas de San Andrés Cuexcontitlán, cuya comunidad se

encuentra en un proceso de desplazamiento más avanzado, son las que hacen un mayor uso de nombres del español en su otomí. Los colaboradores de la generación adulta (35-54 años) emplean un mayor número de lexemas españoles, en comparación con los hablantes de la generación mayor (55 años en adelante). En cuanto al género, los hombres usan más nombres del español que las mujeres.

Sin embargo, el uso de los préstamos del español se comportó de forma inesperada en los hablantes de los diferentes niveles de instrucción y en los diferentes tipos de bilingües. Los hablantes con mayor nivel de instrucción hacen un menor uso de los lexemas del español que los bilingües con un nivel de instrucción bajo. Por otro lado, los bilingües con mayor dominio del español son los que menos utilizan nombres del español, en comparación con los bilingües simétricos y los bilingües con mayor dominio del otomí. Esto quiere decir que la adopción de los lexemas del español no está relacionada con un mayor dominio del español y un menor conocimiento de la lengua otomí, pues son los hablantes con mayor conocimiento del *ñható* quienes más uso hacen de estos recursos.

La HIPÓTESIS GENERAL de este capítulo señala que el español ha influido de alguna manera en la marcación de número del otomí. Para poder comprobarla se planteó una serie de HIPÓTESIS ESPECÍFICAS. La primera indica que el español no ha influido de manera directa en la marcación de número del otomí, es decir, que los hablantes no han transferido la morfología de número del español a su otomí. Para poder probar esta hipótesis, en el apartado §4.4.2 se establecieron los cuatro posibles patrones de marcación de número en los nombres de las narraciones otomíes y se revisó su distribución en el corpus. Los resultados mostraron que los bilingües pueden o no usar la morfología del otomí para marcar el número en los nombres (patrimoniales y préstamos) de las narraciones otomíes, pero no emplean la morfología del español. Estos primeros resultados confirman la HIPÓTESIS ESPECÍFICA 4.1,

pues no se observa una importación directa de la morfología de número del español en el otomí.

La HIPÓTESIS ESPECÍFICA 4.2 señala que el español ha influido de manera indirecta en la marcación de número nominal del otomí. Por ello, aunque la marcación de número se expresa con los recursos morfosintácticos del otomí, la aparición o no de las marcas está relacionada con alguno(s) de los condicionante(s) lingüísticos que rigen la marcación de número en español. Para poder probar esta hipótesis, se realizó un análisis general (véase §4.4.3) para determinar qué factores lingüísticos y sociolingüísticos determinan la marcación de número en los nombres de las narraciones otomíes (patrimoniales y préstamos). Los datos indicaron que la presencia de determinantes que codifican el número, la función sintáctica de la frase nominal, la distinción entre nombre propio y común, la concreción del nombre, el nivel de instrucción, el tipo de constituyente, la referencialidad, el tipo de bilingüismo, la comunidad y la (dis)continuidad inciden en la marcación de número en los nombres de las narraciones otomíes.

Estos primeros resultados no fueron suficientes para afirmar o rechazar de forma definitiva la HIPÓTESIS ESPECÍFICA 4.2. En principio, que el factor más importante en la marcación de número del otomí fuera la presencia de determinantes que codifican el número, asemeja el otomí de SAC y JV al de otras variedades, como la del Valle del Mezquital; sin embargo, que el efecto de los determinantes fuera el opuesto podría estar relacionado con el contacto con el español o ser diferencias propias de la variación geográfica.

Además, los resultados generales no abonaron en favor de la HIPÓTESIS ESPECÍFICA 4.2, que señala que la fuerza del efecto del español depende del origen histórico de los nombres, pues la presencia o ausencia de las marcas de número *ra* y *ya* no mostró estar asociada con el origen histórico de los nombres. Para saber si había alguna diferencia en la

marcación del número de los nombres patrimoniales y de origen español, y localizar algún indicio de la influencia del español en el otomí, fue necesario llevar a cabo un análisis de la marcación de número en los nombres patrimoniales y en los préstamos por separado (véase §4.5). Por otro lado, que el factor tipo de bilingüismo resultara decisivo en la alternancia de la marcación de número fue un primer indicio en favor de la HIPÓTESIS 4.3, que sugiere que la fuerza del efecto del español está asociada con el nivel de bilingüismo de los hablantes. Esta hipótesis se evaluó en mayor detalle en la sección §4.6.

En la sección §4.5 se realizó un par de análisis para determinar qué factores lingüísticos y sociales incidían en la marcación de número de los nombres patrimoniales y los nombres procedentes del español, con la finalidad de encontrar algunas diferencias que pudieran dar indicio de alguna influencia del español en el *n̄ható*. El primer dato relevante fue encontrar una mayor presencia de las marcas de número en los nombres patrimoniales (**Patrón 1 *ra / ya*: 73.9%**; Patrón 2 N: 26.1%), en comparación con los nombres de origen español (**Patrón 1 *ra / ya*: 67.3%**; Patrón 2 N: 32.7%). En principio, este patrón sugería que la influencia del español podría estar relacionada con el desuso de la morfología de número del otomí en los préstamos del español; sin embargo, como se pudo ver en el análisis por grupo lingüístico, la respuesta no era así de simple.

Los resultados de las pruebas multivariadas indicaron que la marcación de número tiene un comportamiento lingüístico muy parecido en los nombres patrimoniales y en los préstamos del español; no obstante, este último grupo léxico mostró una mayor variación sociolingüística. Con respecto a los factores lingüísticos, el predictor de mayor importancia en los dos grupos léxicos es la ausencia o presencia de determinantes que codifican valores de número. La presencia de numerales, de la marca de indefinido o la ausencia de determinantes inhiben el uso de las marcas *ra* y *ya*. El segundo predictor de mayor

importancia en los nombres patrimoniales y los préstamos es la función sintáctica de la frase nominal. Sin embargo, en los nombres *ñható* se observó una escisión que en los préstamos no: las funciones nucleares (sujeto, objeto primario, objeto secundario) favorecen la marcación de número y las funciones periféricas (complemento nominal, predicativo, adverbial) la inhiben. El tipo de constituyente es el tercer predictor de mayor importancia en los dos grupos léxicos: los nombres que son término de preposición no favorecen el uso de las marcas *ra/ya*. La concreción del nombre es el cuarto predictor en los nombres *ñható* y el tercero en los préstamos. En ambos grupos, los nombres concretos son los que promueven el uso del Patrón 1: *ra / ya*. Los dos conjuntos nominales difieren en el quinto predictor de mayor importancia. En los nombres *ñható* es la (dis)continuidad, y los nombres discontinuos favorecen el uso de las marcas de número. En los préstamos es la referencialidad de la frase nominal, y son las frases referenciales las que promueven la aparición de *ra* y *ya*.

Por otro lado, en cuanto a los predictores sociolingüísticos, se observaron ciertas diferencias. En el grupo de nombres patrimoniales sólo es relevante la comunidad y son los hablantes de SAC quienes favorecen el empleo del Patrón 1: *ra/ya*. En contraste, en los préstamos del español, los predictores sociales más importantes son el género, el nivel de instrucción y el grado de bilingüismo. Con respecto a este último, se encontró que los bilingües con mayor dominio del otomí y los bilingües simétricos no favorecen el uso de las marcas de número en los préstamos. Este efecto se puede ver en la distribución de los datos, pues son los BO (**Patrón 1 *ra / ya*: 63.5%**; Patrón 2 N: 36.5%) y los BS (**Patrón 1 *ra / ya*: 58.3%**; Patrón 2 N: 41.7%) quienes hacen un menor uso de la morfología de número, en comparación con los bilingües con mayor dominio del español (**Patrón 1 *ra / ya*: 75%**; Patrón 2 N: 25%).

Estos resultados modificaron la idea inicial de que la influencia del español podría estar relacionada con el desuso de la morfología de número del otomí, pues son los hablantes con mayor dominio del español los que promueven el uso de las marcas de número, tanto en los nombres patrimoniales como en los préstamos del español. Además, subrayaron la pertinencia de realizar un análisis de la marcación de número en los nombres del otomí y del español según el tipo de hablante, pues era posible que el efecto y las diferencias en el uso de las marcas de número relacionadas con el contacto fueran más evidentes en alguno(s) de los grupos lingüísticos.

En la sección §4.6 se evaluó la HIPÓTESIS ESPECÍFICA 4.3 de este capítulo, la cual señala que la fuerza del efecto del español está asociada con el nivel de bilingüismo de los hablantes. Para ello, se revisó el uso de las marcas de número en los nombres patrimoniales y los préstamos en los tres grupos de bilingües, pues se partió del supuesto de que cada uno representa un grado diferente de contacto e influencia del español. La marcación de número en los nombres patrimoniales empleados por los bilingües con mayor dominio del otomí fue el punto de referencia a partir del cual se establecieron las similitudes y diferencias con el resto de los grupos lingüísticos, debido a que este conjunto representa un menor grado de contacto e influencia del español. Por ello, primero determiné cuáles eran los factores lingüísticos que incidían en la marcación de número de los nombres patrimoniales en los tres tipos de bilingües.

En primer lugar, se observó que los BO hacían un menor uso de las marcas de número en los nombres patrimoniales (**Patrón 1 ra / ya: 66.4%**; Patrón 2 N: 33.6%) y los préstamos (**Patrón 1 ra / ya: 63.5%**; Patrón 2 N: 36.5%), en comparación con los BS (**Patrimoniales - Patrón 1 ra / ya: 69.1%**; Patrón 2 N: 30.9%; **Préstamos - Patrón 1 ra / ya: 58.3%**; Patrón 2 N: 41.7%) y los BE (**Patrimoniales - Patrón 1 ra / ya: 73.2%**; Patrón 2 N: 26.8%;

Préstamo - Patrón 1 ra / ya: 75%; Patrón 2 N: 25%). Posteriormente, se realizó una serie de pruebas multivariadas (regresión logística escalonada y análisis multimodelo) para determinar qué factores incidían en el uso de las marcas de número en los nombres patrimoniales y en los préstamos de los tres tipos de bilingües, con el objetivo de establecer similitudes y diferencias entre los grupos con base en el comportamiento de los bilingües con mayor dominio del otomí.

De manera general, el sistema de marcación de número en los nombres patrimoniales de los BO está determinado por la función gramatical, el tipo de determinante, el tipo de constituyente, la concreción, la continuidad y la animacidad. El sistema de los BS está regido por la función sintáctica, la concreción, el tipo de determinante y la continuidad. El sistema de los BE está asociado con los determinantes y la concreción.

Por otro lado, el sistema de marcación de número en los préstamos de los BO está determinado por el tipo de determinante, la función gramatical, el tipo de constituyente, la concreción, la distinción entre nombre propio y común, y la referencialidad. El sistema de los BS está regido por el tipo de determinante, la función sintáctica, la referencialidad y la concreción. El sistema de los BE sólo está asociado con la referencialidad de la frase.

Aunque los resultados muestran similitudes entre los grupos lingüísticos, relacionados con la función sintáctica y la concreción en los nombres patrimoniales; y con la referencialidad en los préstamos del español, también se presentan diferencias importantes entre estos.

En primer lugar, hay un incremento en el uso de las marcas de número del otomí conforme incrementa el conocimiento del español de los bilingües y decrece el dominio de la lengua otomí. Este resultado parece contraintuitivo, y se opone a la idea inicial de que la influencia del español podría estar relacionada con el desuso de la morfología de número del

otomí, pues son los hablantes con menor dominio en esta lengua los que más emplean estas marcas. El incremento en el uso de las marcas de número en los hablantes dominantes en español, más bien, parece estar relacionado con el debilitamiento de algunas restricciones que determinan la presencia / ausencia de las marcas de número en la lengua otomí.

La primera tiene que ver con el uso de las marcas de número en frases nominales con numerales. En el otomí de los bilingües con mayor dominio del otomí, y en menor medida, en el de los bilingües simétricos, el empleo de las marcas de número con frases nominales con numerales es excepcional y los pocos casos que se documentan se restringen a nombres +humanos o +animados. En contraste, en los bilingües con mayor dominio del español, la aparición de las marcas *ra* y *ya* con numerales asciende al 50 % en los dos grupos léxicos.

Se presenta un efecto similar con el tipo de constituyente sintáctico. En el grupo de los BO se presenta una fuerte distinción en el uso de las marcas de número según sea el nombre núcleo de una FN o término de preposición. En este último caso, el uso de *ra* / *ya* no se ve favorecido y los pocos casos que se registran están asociados con el empleo de preposiciones del español. Esta restricción se debilita en el grupo de bilingües simétricos, debido a la adopción de un porcentaje importante de preposiciones del español, que, a su vez, conllevan el uso de las marcas de número en un contexto antes restringido. En el grupo de BE, este comportamiento se agudiza.

En la sección §4.7 realicé un análisis de la marcación de número en tiempo aparente, mediante el uso de un índice de integración léxica (IIL) y una clasificación del léxico, a partir de la propuesta de Ito y Mester (1995), con el objetivo de establecer un eje temporal dentro de la organización del lexicón otomí. Los resultados del análisis a partir del índice de integración señalaron que cuatro factores eran relevantes en la marcación de número: el tipo de determinante, el tipo de constituyente sintáctico, la referencialidad y el índice de

integración. Con respecto a este último, se encontró que, a mayor integración de un lexema a la lengua otomí, mayor es la posibilidad de que los bilingües utilicen las marcas de número *ra* y *ya*.

En el análisis por clasificación se consideraron tres grupos léxicos: el vocabulario nativo, los préstamos establecidos y los préstamos no establecidos. Los resultados mostraron que la marcación de número está determinada por los mismos factores lingüísticos en los tres grupos léxicos y, además, estos inciden en la misma dirección, salvo en un contexto. La principal diferencia entre los tres grupos léxicos se halla en los factores sociolingüísticos. La comunidad de origen del hablante sólo incide en el vocabulario nativo, y son los habitantes de San Andrés Cuexcontitlán quienes inhiben el uso de las marcas de número. El efecto de la variable comunidad en la marcación de número del vocabulario nativo del otomí fue constante en los análisis y concuerda con lo previsto, pues es esperable la existencia de una norma de uso distinta en diferentes comunidades de habla, como ya lo habíamos observado en la sección §4.5.1, cuando advertimos algunas diferencias en el uso de las marcas de número entre el otomí del Valle del Mezquital (Acosta Aguilera, 2020) y el otomí del centro.

En los préstamos establecidos se presenta un uso diferenciado de la morfología de número según la edad y el género de los hablantes. El efecto de la edad es interesante pues, a pesar de que gran parte de estos ítems llevan un tiempo considerable en el lexicón otomí, hay un uso diferenciado según la edad de los hablantes: los bilingües con más de 55 años inhiben el uso de las marcas de número y los más jóvenes (20-34 años) lo promueven.

Por último, en la marcación de número de los préstamos no establecidos interviene el tipo de bilingüismo de los hablantes: los bilingües simétricos y los dominantes en otomí no promueven la utilización de la morfología de número. Este hecho, concuerda con uno de los planteamientos de la sección §4.6.2.4, donde se advirtió un incremento en el uso de las

marcas de número del otomí conforme incrementaba el conocimiento del español de los bilingües y decrecía el dominio de la lengua otomí, a causa del debilitamiento de algunas restricciones que determinan la presencia / ausencia de las marcas de número en la lengua otomí. En relación con este punto, es relevante el efecto que el tipo de bilingüismo tiene en la marcación del número de los préstamos no establecidos, pues el efecto que advertíamos en la sección §4.5.2 parece no aplicar a todo el conjunto de préstamos del español, sino sólo a aquellos que se encuentran en la periferia del lexicón otomí.

Las diferencias cuantitativas y sociales que inciden en la marcación de número de los tres grupos léxicos muestran que la incidencia del español depende del grado de integración de los nombres al lexicón otomí. No obstante, la homogeneidad en el comportamiento lingüístico de la marcación de número en los tres grupos lingüísticos señala que, a pesar de los cambios cuantitativos en el uso de la morfología de número, detrás de estos actúan los mismos factores lingüísticos.

Por último, se exploró el efecto de la frecuencia del ítem léxico en la marcación de número. Los resultados señalan que la frecuencia de uso sí tiene un efecto en la marcación de número, pero este difiere según el origen de los nombres. En los nombres patrimoniales no se observa una diferencia prominente en el uso de *ra* y *ya* asociado a la frecuencia de uso de los ítems léxicos. En contraste, en el grupo de préstamos del español, sí se puede ver un uso diferenciado de la morfología de número en relación con la frecuencia de uso de los ítems léxicos. Los ítems léxicos más frecuentes tienen una mayor probabilidad de presentar las marcas de número, mientras que los ítems con frecuencia baja inhiben el uso de las marcas *ra* y *ya*. El grupo de préstamos que se aproxima más al comportamiento de los nombres patrimoniales son los menos frecuentes. Los préstamos más frecuentes, por su parte, parecen

ser los que lideran la generalización del uso de las marcas de número, independientemente de las restricciones de la lengua otomí.

Los resultados presentados a lo largo de este capítulo muestran que el español ha influido en la lengua *ñható* de diversas formas. En primer lugar, se observa una influencia directa, mediante la importación de un porcentaje importante de elementos léxicos y gramaticales del español. Si se mira el ámbito nominal, la influencia léxica del español en el otomí es aún mayor, con casi el 50% de los nombres en las narraciones.

En el caso específico de la marcación de número en la frase nominal del otomí, la influencia del español sólo se hizo evidente al realizar el análisis por grupo lingüístico. Los datos mostraron un aumento en el uso de las marcas de número del otomí, tanto en los nombres patrimoniales como en los préstamos, conforme incrementa el conocimiento del español de los bilingües y decrece su dominio de la lengua otomí, debido al debilitamiento de algunas restricciones que determinan la presencia / ausencia de las marcas de número en la lengua otomí. Con base en esto, podríamos decir que las HIPÓTESIS 4.1 Y 4.3 de este capítulo se cumplen, pues el español sí influye de manera indirecta en la marcación de número del *ñható* y esta influencia es más notoria en el grupo de bilingües con mayor dominio del español.

Para poder explicar cómo el conocimiento del español condiciona las soluciones que los bilingües emplean para marcar el número en su otomí, de nueva cuenta, resulta pertinente recurrir a la idea de Johanson (2002) en torno a las propiedades de las unidades lingüísticas (véase §1.1.3). Los bilingües, sobre todo los BE, encuentran similitudes y diferencias en las propiedades combinatorias de las marcas de número del otomí *ra / ya* y el español *-(e)s* en las FNs con numerales y en las FPs, y, a partir de eso, establecen una equivalencia entre ellas. Esta equivalencia se puede apreciar en la extensión del uso de las marcas de número en

contextos donde no solían presentarse, a causa del debilitamiento de algunas restricciones presentes en la lengua otomí, que están ausentes en el español, sobre todo las relacionadas con los numerales y las preposiciones. Con respecto a este último contexto, también resultó relevante el incremento en la adopción de preposiciones del español que, además de su forma fónica, acarrearán sus propiedades combinatorias. Esta influencia directa del español, a su vez, trae consigo cambios en las propiedades combinatorias de los nombres del otomí.

Los resultados en torno a la marcación de número según la estratificación del léxico (vocabulario nativo, préstamos establecidos y préstamos no establecidos), mostraron que la marcación de número está determinada por los mismos factores lingüísticos en los tres grupos léxicos y que la principal diferencia se halla en los factores sociolingüísticos. En el vocabulario nativo, sólo incide la comunidad de origen de los hablantes. En los préstamos establecidos se presenta un uso diferenciado de la morfología de número según la edad y el género de los hablantes. Por último, en la marcación de número de los préstamos no establecidos interviene el tipo de bilingüismo de los hablantes.

La homogeneidad en el comportamiento lingüístico de la marcación de número en los tres grupos lingüísticos señala que, a pesar de los cambios cuantitativos en el uso de la morfología de número, detrás de éstos actúan los mismos factores lingüísticos. Aunado a esto, el análisis del efecto de la frecuencia en la marcación de número indicó que los préstamos más frecuentes lideran la generalización del uso de las marcas de número, independientemente de las restricciones de la lengua otomí. Estos resultados comprueban la HIPÓTESIS 4.2, que señala que el efecto del español depende del origen histórico de los nombres, pues muestran que los préstamos van en la delantera en el proceso de expansión de las marcas de número.

CAPÍTULO 5. CONCLUSIONES GENERALES

En el primer capítulo revisé algunas propuestas teóricas en torno al estudio del contacto lingüístico y el bilingüismo. Después, presenté el marco conceptual que guiaría el análisis en esta investigación. Posteriormente, proporcioné información acerca de la marcación de número en la lengua otomí, en el nivel frástico y oracional, pues esta información sería de utilidad para formular hipótesis en torno a los fenómenos que son objeto de estudio de los capítulos tres y cuatro. El marco conceptual constaba de cuatro bloques horizontales que simulaban distintos momentos en una escala de tiempo. El primero correspondía a las causas sociales del contacto lingüístico; el segundo tenía que ver con los efectos del contacto a nivel comunidad; el tercero comprendía los efectos del contacto lingüístico en el plano individual; y el último de los bloques describía el proceso de expansión o propagación de los fenómenos de contacto. En los capítulos de esta investigación se desarrollaron algunos de estos bloques.

En el segundo capítulo presenté la metodología empleada para conformar el grupo de colaboradores de San Andrés Cuexcontitlán y de Jiquipilco el Viejo, así como las herramientas para la recolección de material lingüístico. En la segunda parte de este capítulo, describí el contexto social y económico que generó un incremento en la intensidad del contacto entre el otomí y el español en las dos comunidades. Asimismo, revisé cómo estos cambios modificaron la estructura y organización de la población de SAC y JV. Para ello, describí cinco grupos de hablantes, su caracterización social y lingüística, y el uso que hacen del otomí y el español en distintos dominios sociales y con diferentes tipos de interlocutores: bilingües con mayor dominio del otomí (BO), bilingües simétricos (BS), bilingües con mayor dominio del español (BE), monolingües con conocimiento pasivo del otomí (PO) y monolingües en español (ME).

Los resultados de esta descripción mostraron la existencia de un continuo bilingüe en las dos comunidades, que van desde el monolingüismo en una lengua indígena al monolingüismo en español y viceversa. Asimismo, sugieren la existencia de un proceso de desplazamiento del *ñható* en las dos comunidades, pues esta lengua ha perdido presencia en la mayoría de los dominios sociales, incluido el hogar; en contraste, el español se está convirtiendo en la lengua imperante en todos los dominios sociales.

El capítulo tres se centró en el español de contacto, específicamente, indagué sobre el efecto que el otomí ha tenido en la concordancia plural del español de los bilingües y monolingües de SAC y JV. Para este propósito planteé cuatro hipótesis.

La HIPÓTESIS 3.1 señala que la variación en la concordancia de número es resultado del contacto que el español mantiene con el otomí. De modo que este fenómeno responde a estrategias gramaticales relacionadas con los procesos de la lengua otomí. Por ello, aunque la concordancia plural se expresa con los recursos morfosintácticos del español normativo, la aparición o no de estas marcas estará determinada por alguno(s) de los condicionantes lingüísticos que rigen la expresión gramatical de número en otomí. Para comprobarla, en el apartado §3.4, llevé a cabo el análisis de la variación en la concordancia de número en la frase nominal del español. Los resultados permitieron afirmar esta hipótesis, pues los bilingües encuentran similitudes y diferencias en las propiedades combinatorias de las marcas de número del otomí *ra / ya* y el español *-(e)s* y, a partir de eso, establecen una equivalencia entre ellas. Esta equivalencia combinatorial se puede apreciar en el uso del patrón de marcación alterna (Presencia + Ausencia de la marca: *La-s casa*), que sigue el esquema del *ñható*, en el que la marca de número sólo se hospeda en el dependiente (Presencia + Ausencia de la marca: *Ku=ya nguu* ‘Las casas’). Asimismo, hay un efecto del factor función sintáctica, que también es relevante en la marcación de número del otomí

(véase §4.3). Aunado a esto, la variación en la concordancia plural en la frase nominal está fuertemente relacionada con la pérdida de consonantes a final de palabra, un fenómeno fónico que también tiene su origen en el contacto con el otomí.

En el apartado §3.5, analicé la variación en la concordancia de número en el nivel oracional. Los resultados también afirman la HIPÓTESIS 3.1. La variación en la concordancia plural en el nivel oracional del español hablado en San Andrés Cuexcontitlán y Jiquipilco el Viejo sí está relacionada con el contacto con la lengua otomí. Los bilingües encuentran similitudes y diferencias en las propiedades combinatorias de las marcas de número del otomí. Esto se puede apreciar en los factores que condicionan la concordancia plural en el nivel oracional de esta variedad de español: la persona gramatical, la animacidad y el tipo de relación, que son algunos de los factores que determinan la marcación de número del otomí en el nivel oracional. Además, el efecto de estos es el mismo en las dos lenguas: los sujetos de 3ª persona y no humanos promueven la discordancia de número y los sujetos de 1ª y 2ª persona, +humanos son los que favorecen la concordancia.

La HIPÓTESIS 3.2 indica que la variación en la concordancia de número en el español de contacto se comporta de manera diferente en el nivel frástico y oracional, debido a que la concordancia de número en otomí se organiza de forma distinta en los dos niveles. Los resultados obtenidos en las secciones §3.4 y §3.5 confirmaron esta hipótesis pues, aunque en ambos casos los bilingües toman como modelo las propiedades combinatorias de la concordancia plural del otomí para establecer la relación de concordancia plural en su español, cada nivel tiene especificaciones diferentes, debido a que la concordancia de número en otomí se organiza de forma distinta en los dos niveles.

La HIPÓTESIS 3.3 dice que la variación en la concordancia de número en el español de contacto tiene el mismo comportamiento lingüístico en SAC y JV, pero difiere en su

distribución social en cada comunidad. Los resultados del apartado §3.6 comprobaron esta hipótesis. La variación en la concordancia plural de las frases nominales (§véase 3.6.2) está estrechamente relacionada con la pérdida de consonantes en coda a final de palabra, tanto en SAC como en JV. Asimismo, en ambas comunidades el segundo factor de mayor importancia es el tipo de bilingüismo del emisor. La única diferencia radicó en un factor social, el género de los hablantes. Este factor sólo resultó relevante en la comunidad de San Andrés Cuexcontitlán, en donde los hombres fueron los promotores del patrón de marcación que sigue el esquema de la lengua otomí (*las casa*). Por otro lado, la concordancia plural en el nivel oracional estuvo determinada por la persona gramatical y el tipo de sujeto en las dos comunidades (véase §3.6.3). Además, en ambos casos, el tipo de bilingüismo del emisor incidió en este fenómeno variable. No obstante, de nueva cuenta, hay diferencias relacionadas con el género de los colaboradores. Este factor sólo fue relevante en JV, donde fueron los hombres quienes favorecieron la discordancia de número.

La HIPÓTESIS 3.4 señala que en SAC y JV existen diferentes sistemas de concordancia de número en español que están relacionados con el tipo de bilingüismo de los hablantes. Estos sistemas no sólo difieren cuantitativamente sino también en los factores lingüísticos que los determinan. En el caso de la concordancia en la frase nominal, está el sistema empleado por los bilingües con mayor dominio del otomí (BO), condicionado por la pérdida de consonantes en coda a final de palabra y la función sintáctica de la frase nominal, ambos factores relacionados con la lengua otomí. Por otro lado, está el sistema de los bilingües simétricos (BS) y los hablantes dominantes o monolingües en español (Esp. = BE, PO, ME), el cual se encuentra definido por la pérdida de consonantes. Con respecto a la concordancia en el nivel oracional, hay dos sistemas nítidamente diferenciados, el de los bilingües con mayor dominio del otomí (BO) y el de los hablantes cuya lengua dominante o única es el

español (Esp. = BE, PO, ME). El sistema del primer grupo está organizado en función de la persona gramatical, la animacidad y la categoría del elemento concordante. Por otro lado, el sistema de concordancia de los hablantes dominantes o monolingües en español (Esp. = BE, PO, ME) está organizado por la pérdida de consonantes, la animacidad y la especificidad. El sistema de los bilingües simétricos, por su parte, tiene rasgos similares al de los bilingües con mayor dominio del otomí y al de los hablantes dominantes en español, como una especie de grupo de transición.

El capítulo cuatro se centró en el otomí de contacto, específicamente, sobre el efecto que el español había tenido en la marcación de número en la frase nominal del otomí de los bilingües y monolingües de SAC y JV. Para este propósito planteé tres hipótesis.

La HIPÓTESIS 4.1 señala que la variación en marcación de número nominal en otomí es resultado del contacto que el español mantiene con el otomí. De modo que este fenómeno responde a estrategias gramaticales relacionadas con los procesos del español. Por ello, aunque la concordancia plural se expresa con los recursos morfosintácticos del otomí, la aparición o no de estas marcas estará determinada por alguno(s) de los condicionantes lingüísticos que rigen la expresión gramatical de número en otomí. Los resultados del apartado §4.4.2 muestran que los bilingües usan la morfología del otomí para marcar el número en los nombres (patrimoniales y préstamos) de las narraciones otomíes y no hacen uso de la morfología del español. Los hallazgos de las secciones §4.5 y 4.6 permitieron comprobar esta hipótesis, el español influye en la marcación de número del *ñható*. Los bilingües, sobre todo los BE, establecen equivalencias entre las propiedades combinatorias de la marcación de número del otomí y el español. Esta influencia se puede apreciar en la extensión del uso de las marcas de número en contextos donde no solían presentarse, a causa

del debilitamiento de algunas restricciones presentes en la lengua otomí, que están ausentes en el español.

La primera de tales restricciones tiene que ver con el uso de las marcas de número en frases nominales con numerales. En el otomí de los bilingües con mayor dominio del otomí, y en menor medida, en el de los bilingües simétricos, el empleo de las marcas de número en frases nominales con numerales era excepcional y los pocos casos que se documentaron se restringen a nombres +humanos o +animados. En contraste, en los bilingües con mayor dominio del español, la aparición de las marcas *ra* y *ya* con numerales asciende al 50 % en los dos grupos léxicos. La segunda está relacionada con el tipo de constituyente sintáctico. En el grupo de los BO se presenta una fuerte distinción en el uso de las marcas de número según sea el nombre núcleo de una FN o término de preposición. En este último caso, el uso de *ra/ya* no se ve favorecido y los pocos casos que se registran están asociados con el empleo de preposiciones del español. Esta restricción se debilita en el grupo de bilingües simétricos y en el de bilingües con mayor dominio del español, debido a la adopción de un porcentaje importante de preposiciones del español, que, a su vez, llevan al uso de las marcas de número en un contexto antes restringido.

La HIPÓTESIS 4.2 señalaba que la fuerza del efecto del español dependía del origen histórico de los nombres. Para poder comprobar esta hipótesis, en el apartado §4.5 realicé un análisis de la marcación de número en los nombres patrimoniales y en los préstamos por separado. Los resultados de las pruebas multivariadas indicaron que la marcación de número tenía un comportamiento lingüístico muy parecido en los nombres patrimoniales y en los préstamos del español; no obstante, este último grupo léxico mostró una mayor variación sociolingüística. Estas diferencias sociolingüísticas se hicieron más evidentes en el apartado §4.7.4, donde analicé la marcación de número en tiempo aparente mediante una clasificación

del léxico, basada en la propuesta de Ito y Mester (1995), cuyo objetivo era establecer un eje temporal dentro de la organización del lexicón otomí. Los resultados en torno a la marcación de número según la estratificación del léxico (vocabulario nativo, préstamos establecidos y préstamos no establecidos) mostraron que la marcación de número estaba determinada por los mismos factores lingüísticos en los tres grupos léxicos y que la principal diferencia se hallaba en los factores sociolingüísticos. En el vocabulario nativo, sólo incidía la comunidad de origen de los hablantes. En los préstamos establecidos se presentaba un uso diferenciado de la morfología de número según la edad y el género de los hablantes. Por último, en la marcación de número de los préstamos no establecidos intervenía el tipo de bilingüismo de los hablantes. La homogeneidad en el comportamiento lingüístico de la marcación de número en los tres grupos lingüísticos señala que, a pesar de los cambios cuantitativos en el uso de la morfología de número, detrás de estos actúan los mismos factores lingüísticos. Aunado a esto, el análisis del efecto de la frecuencia en la marcación de número (véase § 4.7.5) indicó que los préstamos más frecuentes lideran la generalización del uso de las marcas de número, independientemente de las restricciones de la lengua otomí. Estos resultados comprueban la HIPÓTESIS ESPECÍFICA 4.2, que señala que el efecto del español depende del origen histórico de los nombres, pues muestran que los préstamos van en la delantera en el proceso de expansión de las marcas de número.

La HIPÓTESIS 4.3 indicaba que la fuerza del efecto del español estaba asociada con el nivel de bilingüismo de los hablantes. En el caso específico de la marcación de número en la frase nominal del otomí, la influencia del español sólo se hizo evidente al realizar el análisis por grupo lingüístico. Los datos mostraron un aumento en el uso de las marcas de número del otomí, tanto en los nombres patrimoniales como en los préstamos, conforme incrementaba el conocimiento del español de los bilingües y decrecía su dominio de la lengua

otomí, debido al debilitamiento de algunas restricciones que determinaban la presencia / ausencia de las marcas de número en la lengua otomí.

Al inicio de esta investigación nos preguntamos si el contacto podía afectar un mismo dominio lingüístico en el otomí y el español. Con esa pregunta en mente nos propusimos revisar el concepto de bidireccionalidad en una situación de contacto en México, específicamente en el contacto entre el otomí del centro y el español. Para ello, analizamos los procesos de variación en la marcación de número de las frases nominales del español y del otomí en contacto.

Con base en los datos y análisis presentados en los Capítulos 3 y 4, pudimos ver que es posible que el contacto afecte un mismo dominio lingüístico, en este caso la marcación de número en el otomí y el español de forma simultánea, pero de manera distinta. El otomí influye en la marcación de número del español, mediante la importación de su patrón de marcación (P-A) al español de contacto. Por otro lado, el español influye en la marcación de número del otomí, al debilitar algunas restricciones sobre la presencia / ausencia de las marcas de número en otomí.

A lo largo de este estudio pudimos observar que uno de los retos a los que nos enfrentamos cuando queremos realizar el análisis de un fenómeno de contacto en las dos lenguas es que existen pocas descripciones o corpus diacrónicos de las lenguas originarias que nos permitan comparar distintos estados de lengua o trazar procesos de cambio a causa del contacto. Para poder evaluar el efecto que el contacto ha tenido en los sistemas de marcación de las dos lenguas, consideramos pertinente revisar el uso de las marcas de número en el español y el otomí de tres grupos de bilingües, que de forma aparente podrían representar diferentes grados de contacto.

Tanto en el capítulo 3 como en el 4, el análisis de los fenómenos variables en función de tres grupos lingüísticos fue de mucha utilidad, pues nos permitió observar la coexistencia de distintos sistemas de concordancia y marcación de número en una misma comunidad. Sin embargo, es posible que la utilidad de este tipo de análisis esté restringido a situaciones de contacto en donde la lengua originaria está en un proceso de desplazamiento acelerado, mas no así en otro tipo de situaciones de contacto más largas y estables, como las de la zona andina o la yucateca (Sánchez Paraíso, 2021), en donde el uso de fenómenos asociados al contacto ya no está relacionado con el tipo de bilingüismo, pues son rasgos que han pasado a ser parte del habla monolingüe de la región.

Aportaciones teóricas de la tesis

En esta investigación se han revisado y probado empíricamente debates, hipótesis y predicciones de gran importancia en el estudio del contacto de lenguas. Las más importantes se detallan a continuación.

1. El efecto bidireccionalidad del contacto lingüístico (Granda, 1996; Palacios, 2011). Al inicio de esta investigación, nos propusimos revisar el concepto de bidireccionalidad, a partir del análisis de dos procesos de variación en la marcación de número de las frases nominales del español y del otomí en contacto. Con base en los datos y análisis presentados en los Capítulos 3 y 4, pudimos ver que el contacto afecta un mismo dominio lingüístico, la marcación de número, en el otomí y el español, aunque de manera distinta.

El otomí influye en la concordancia plural del español porque los bilingües establecen equivalencias entre las propiedades combinatorias de las marcas de número del otomí y el español. En el nivel frástico, se aprecia en el uso del patrón de marcación alterna (Presencia

+ Ausencia de la marca: *La-s casa*), que sigue el esquema del *ñható*, en el que la marca de número sólo se hospeda en el dependiente (Presencia + Ausencia de la marca: *Ku=ya nguu* ‘Las casas’). En el nivel oracional, las equivalencias se hallan en los factores que condicionan la concordancia plural: la persona gramatical, la animacidad y el tipo de relación, que son algunos de los factores que determinan la marcación de número del otomí en el nivel oracional.

Por otro lado, el español influye en la marcación de número del otomí, al debilitar algunas restricciones sobre la presencia / ausencia de las marcas de número en otomí. Los bilingües establecen equivalencias entre las propiedades combinatorias de las marcas de número del otomí *ra / ya* y el español *y-(e)s*. Estas equivalencias se expresan en el incremento de la concordancia semántica en las FNs con numeral en otomí (*Ku^uutá ya nguu* ‘Cinco casas’), como sucede en español (*Cinco casas*); y en los términos de preposiciones del español y el otomí.

Estos resultados cuestionan la pertinencia de dicotomías como la de préstamo *versus* interferencia, que revisaremos en el segundo punto, y nos invitan a repensar la manera en que nos aproximamos al estudio del contacto de lenguas. En principio, apelan a la necesidad de estudiar un mismo dominio lingüístico en las dos lenguas en contacto en pos de entender de manera holística los fenómenos inducidos por contacto. En segundo lugar, demandan el desarrollo de una teoría y un método capaz de ascir y explicar la influencia mutua en los distintos niveles lingüísticos.

2. El debate sobre la división interferencia *versus* préstamo. Thomason y Kaufman (1988) distinguen entre fenómenos de *préstamo* e *interferencia*, a partir de la (in)existencia de un proceso de adquisición imperfecta y la duración del contacto. Los resultados de esta

investigación contradicen la existencia de esta dicotomía en la situación de contacto entre el otomí y el español en SAC y JV. En primer lugar, la idea de que los préstamos sólo ocurren en bilingües que dominan las dos lenguas y las interferencias en bilingües con una adquisición incompleta de la segunda lengua no se sostiene, pues los tres tipos de bilingües que colaboraron en esta investigación (BO, BS, BE) producen préstamos e interferencias en su español y otomí, independientemente de si su proceso de adquisición es completo o no.

La segunda razón tiene que ver con el contexto social en el que ocurren los préstamos (mantenimiento) y las interferencias (desplazamiento). Debido a que el *ñható* se encuentra en un proceso de desplazamiento en SAC y JV (véase §2.2), únicamente se esperaría documentar fenómenos de interferencia, sin embargo, también se presentan casos de préstamo. De hecho, es difícil distinguir entre ambos fenómenos en el ámbito léxico, donde el uso de una palabra del español puede estar asociado tanto a un proceso de préstamo como a uno de interferencia. Por ejemplo, en el vocabulario otomí se encuentra una gran cantidad de nombres del español pertenecientes al ámbito religioso, que ingresaron alrededor del siglo XVII, y que bien podrían ser considerados como préstamos. En contraste, también se hallan elementos léxicos modernos, cuyas características morfofonológicas podrían catalogarlos como casos de interferencia.

3. Sobre las escalas de prestabilidad y la hipótesis del “hueco funcional”. Dentro de los estudios del préstamo en situaciones de contacto, se han postulado distintas jerarquías y escalas que establecen la (im)posibilidad de que un elemento sea transferido o prestado, en función de algunas restricciones estructurales y gramaticales (Whitney, 1881; Haugen 1950; Muysken, 1981; Matras, 2008). La distribución de los préstamos del español en el corpus de esta investigación (véase Tabla 4.5) sugieren el siguiente orden: Nombres > Conjunciones >

Verbos > Pronombres > Preposiciones > Adverbios > Adjetivos > Numerales > Otros. Esta disposición de los datos concuerda, en mayor o menor medida, con las escalas de prestabilidad de Whitney (1881), Haugen (1950) y Muysken (1981) con respecto a la posición de los nombres y los verbos, pero difiere en la posición de los adjetivos, las conjunciones y las preposiciones. La posición de las conjunciones muestra similitudes con la propuesta de Matras (2007), quien pone en la primera posición de su escala a los nombres y las conjunciones, y en segundo lugar a los verbos. Estos resultados sugieren que la posición de ciertas categorías léxicas en las escalas de prestabilidad es relativa y no es inherente a la clase de palabra, sino que depende de las categorías existentes en las lenguas en contacto.

Por otro lado, en investigaciones anteriores sobre el otomí en contacto con el español (Gómez Rendón, 2008; Hekking & Bakker, 2010), se ha utilizado la hipótesis del hueco funcional para explicar el bajo porcentaje de préstamos adjetivales y la adopción masiva de conjunciones y preposiciones. En el caso de los adjetivos, se ha argumentado que la falta de una clase léxica que funcione como modificador de una frase nominal en el otomí dificulta la adopción de adjetivos. Sin embargo, este mismo argumento se ha utilizado para explicar la adopción masiva de conjunciones y preposiciones. Aunque esta investigación no tuvo como objetivo plantear una escala de prestabilidad, ni determinar las motivaciones del préstamo léxico, los resultados piden repensar la validez de la hipótesis del “hueco funcional”, pues parece ser insuficiente para explicar la motivación del préstamo. Asimismo, son una invitación para plantear nuevas aproximaciones en las que no sólo se consideren las restricciones estructurales, sino también las motivaciones pragmáticas y comunicativas que dan lugar al préstamo.

4. Sobre la jerarquía de animacidad (Corbett, 2000; de Swart & de Hoop, 2018). En el estudio de la marcación de número en el español y el otomí de contacto, tomé en cuenta la animacidad de los nombres, porque tipológicamente ha mostrado ser determinante en la organización de los sistemas de marcación y concordancia de número. Específicamente en la lengua otomí, la concordancia de número en el verbo con sujeto de tercera persona depende de que el referente sea humano (Palancar, 2013: 118) o animado en un predicado recíproco. En el análisis tomé de referencia la jerarquía de animacidad de Corbett (2000: 54-87) e incorporé esta noción a partir de tres categorías: humano > animado > inanimado.

Los resultados cuantitativos del capítulo tres no siempre siguen dicho orden. Por ejemplo, en la concordancia plural oracional del español en contacto, los nombres humanos tienen un mayor porcentaje de concordancia que los nombres inanimados y estos, a su vez, que los animados, dando lugar el siguiente orden: humano > inanimado > animado, que no concuerda con la jerarquía de Corbett (2000). En principio, esta disposición parecería ir en contra de la jerarquía de animacidad; sin embargo, este ordenamiento cobra sentido si planteamos una división binaria: humano *versus* no humano (animados e inanimados). Según de Swart y de Hoop (2018), este hecho no quiere decir que se rechace la asunción de que la jerarquía de animacidad juegue un rol en la lengua, pero varios fenómenos pueden ser estudiados de forma más parsimoniosa en términos de rasgos binarios, porque la animacidad está involucrada en fenómenos lingüísticos que son binarios en sí mismos, tal es el caso de la concordancia plural.

REFERENCIAS BIBLIOGRÁFICAS

- Acosta Aguilera, Y. (2020). *Sintaxis y semántica de las frases nominales marcadas con ra y ya en el otomí del Valle del Mezquital* [Tesis de licenciatura]. Universidad Nacional Autónoma de México.
- Allan, K. (1980). Nouns and countability. *Language*, 56, 541–567.
- Alvar Ezquerro, M. (2010). *El español en México*. Universidad de Alcalá de Henares.
- Andrews, H. (1949). Phonemes and Morphophonemes of Temoayan Otomi. *International Journal of American Linguistics*, 15(4), 213–222. <https://about.jstor.org/terms>
- Appel, R., & Muysken, P. (1996). *Bilingüismo y contacto de lenguas*. Ariel.
- Arroyo Estrada, J. D. (1985). *Monografía municipal. Temoaya*. Gobierno del Estado de México.
- Arzate Becerril, J. (1996). *Matrimonio, sexualidad y ritual cristiano en el mundo hispánico: la norma y la práctica en Temoaya durante el siglo XVIII*. Universidad Autónoma del Estado de México.
- Avelino Sierra, R. (2017). *Contacto lingüístico entre el español y el otomí en San Andrés Cuexcontitlán*. Universidad Nacional Autónoma de México.
- Bakker, D., Gómez Rendón, J., & Hekking, E. (2008). Spanish meets Guarani, Otomi and Quichua: a multilingual confrontation. In T. Stolz, D. Bakker, & R. Salas (Eds.), *Aspects of Language Contact. New Theoretical, Methodological and Empirical findings with special focus on Romancisation processes* (pp. 165–238). Mouton de Gruyter.
- Barrientos López, G. (2004). *Otomíes del Estado de México*. Comisión Nacional para el Desarrollo de los Pueblos Indígenas.

- Bartholomew, D. A. (1954). Palabras prestadas del español en el dialecto otomí. *Revista Mexicana de Estudios Antropológicos*, *14*, 169–171.
- Bartholomew, D. A. (1994). Panorama of studies in Otopamean languages. In L. Manrique, Y. Lastra, & D. Bartholomew (Eds.), *Panorama de los estudios de las lenguas indígenas en México* (pp. 335–377). Abya-Yala.
- Bialystok, E., Luk, G., Peets, K. F., & Yang, S. (2010). Receptive vocabulary differences in monolingual and bilingual children. *Bilingualism: Language and Cognition*, *13*(4), 525–531. <https://doi.org/10.1017/S1366728909990423>
- Blom, J.-P., & Gumperz, J. J. (1972). Social meaning in linguistic structures: Code-switching in Norway. In J. Gumperz & D. Hymes (Eds.), *Directions in Sociolinguistics* (pp. 407–434). Holt, Rinehart, Winston.
- Bloomfield, L. (1956). *Language*. George Allen & Unwin.
- Bosque, I. (1999). El nombre común. In I. Bosque, V. Demonte, & (dirs.) (Eds.), *Gramática descriptiva de la Lengua Española. Volumen 1: Sintaxis básica de las clases de palabras* (pp. 3–76). Espasa.
- Burnham, K. P., & Anderson, D. R. (2002). *Model Selection and Multimodel Inference. A Practical Information-Theoretic Approach* (2a ed.). Springer.
- Bybee, J. (2007). *Frequency of Use and the Organization of Language*. Oxford University Press. <https://doi.org/10.1093/acprof:oso/9780195301571.001.0001>
- Bybee, J. (2010). *Language, Usage and Cognition*. Cambridge University Press. <https://doi.org/10.1017/CBO9780511750526>
- Calcagno, V., & Mazancourt, C. de. (2010). glmulti: An R Package for Easy Automated Model Selection with (Generalized) Linear Models. *Journal of Statistical Software*, *34*(12), 1–28. <https://doi.org/10.18637/jss.v034.i12>

- Canavos, G. C. (1988). *Probabilidad y estadística. Aplicaciones y métodos*. McGraw-Hill.
- Corbett, G. G. (2000). *Number*. Cambridge University Press.
<https://doi.org/10.1017/CBO9781139164344>
- Dalenius, T., & Hodges, J. L. (1959). Minimum Variance Stratification. *Journal of the American Statistical Association*, 54(285), 88–101.
<https://doi.org/10.1080/01621459.1959.10501501>
- de Swart, P., & de Hoop, H. (2018). Shifting animacy. *Theoretical Linguistics*, 44(1–2), 1–23. <https://doi.org/10.1515/tl-2018-0001>
- Diebold, A. R. (1961). Incipient Bilingualism. *Language*, 37(1), 97–112.
http://www.jstor.orgURL:http://www.jstor.org/stable/411253http://www.jstor.org/stable/411253?seq=1&cid=pdf-reference#references_tab_contents
- Dik, S. C. (1997). *The Theory of Functional Grammar. Volume I: The structure of the clause*. (2a ed. rev.). Mouton de Gruyter.
- Dixon, R. M. W. (1994). *Ergativity*. Cambridge University Press.
- Donellan, K. (1966). Reference and definite descriptions. In D. D. Steinberg & L. A. Jakobovits (Eds.), *Semantics* (pp. 100–114). Cambridge University Press.
- Dryer, M. S. (1986). Primary Objects, Secondary Objects, and Antidative. *Language*, 62(4), 808–845.
- Egland, S. (1978). *La inteligibilidad interdialectal en México: Resultados de algunos sondeos*. Instituto Lingüístico de Verano.
- Escobar, A. M. (1988). *Hacia una tipología del bilingüismo en el Perú*. Instituto de Estudios Peruanos.
- Escobar, A. M. (2000). *Contacto social y lingüístico*. Pontificia Universidad Católica del Perú.

- Escobar, A. M. (2002). Los factores lingüísticos y sociales en el español en contacto con lenguas indígenas. In Z. Estrada & R. M. Ortiz (Eds.), *VI Encuentro Internacional de Lingüística en el Noroeste Memorias* (pp. 253–275). Universidad de Sonora.
- Fábregas, A. (2018). Determiners and quantifiers in Spanish: types, tests and theories. *Borealis – An International Journal of Hispanic Linguistics*, 7(2), 1–108. <https://doi.org/10.7557/1.7.2.4546>
- Fishman, J. A. (1972). *The sociology of language: An interdisciplinary social science approach to language in society*. Newbury House Publishers.
- Flores Farfán, J. A. (1999). *Cuatreros somos y toindioa hablamos. Contactos y conflictos entre el náhuatl y el español en el sur de México*. CIESAS.
- García Fajardo, J. (2016). *Semántica de la oración. Instrumentos para su análisis* (2a ed. rev.). El Colegio de México.
- Garza Villarreal, G. (2003). *La urbanización de México en el Siglo XX*. El Colegio de México.
- Givón, T. (1984). *Syntax. A functional-typological introduction*. John Benjamins.
- Givón, T. (2001). *Syntax*. John Benjamins Publishing Company. <https://doi.org/10.1075/z.syn1>
- Gómez Rendón, J. Arsenio. (2008). *Typological and social constraints on language contact : Amerindian languages in contact with Spanish*. LOT.
- González Villarreal, R. (2018). La Reforma educativa en México: 1870-1976. *Espacio, Tiempo y Educación*, 5(1), 95–118.
- Grosjean, F. (2004). Studying Bilinguals: Methodological and Conceptual Issues. In T. K. Bhatia & W. C. Ritchie (Eds.), *The Handbook of Bilingualism* (pp. 32–63). Blackwell Publishing Ltd. <https://doi.org/10.1002/9780470756997.ch2>

- Grosjean, F. (2012). An attempt to isolate, and differentiate, transfer and interference. *International Journal of Bilingualism*, 16(1), 11–21.
- Grosjean, F. (2014). Studying bilinguals: Methodological and conceptual issues. In T. K. Bhatia & W. C. Ritchie (Eds.), *The Handbook of Bilingualism* (pp. 32–63). Blackwell.
- Guerrero, A. (2006). Hablamo(s) así todo(s) igual(es): concordancia plural en un contexto bilingüe. In P. Martín Butragueño (Ed.), *Líderes lingüísticos* (pp. 89–110). El Colegio de México.
- Guerrero, A. (2015). Patrones tonales y acento en otomí. In E. Herrera (Ed.), *Tono, acento y estructuras métricas en lenguas mexicanas* (pp. 235–260). El Colegio de México.
- Guerrero Galván, A. (2014). Aspectos cuantitativos y cualitativos del peritaje lingüístico. In P. Martín Butragueño & L. Orozco Vaca (Eds.), *Argumentos cualitativos y argumentos cuantitativos en sociolingüística. Segundo Coloquio cambio y variación lingüística* (pp. 613–642). El Colegio de México.
- Guerrero Galván, A., & San Giacomo Trinidad, M. (2014). El llamado español indígena en el contexto del bilingüismo. In R. Barriga Villanueva & P. Martín Butragueño (Eds.), *Historia Sociolingüística de México. Volumen III: Espacio, contacto y discurso político* (pp. 1459–1526). El Colegio de México.
- Gumperz, J. J. (1962). Types of linguistic communities. *Antropological Linguistics*, 4, 28–40.
- Haspelmath, M. (2005). Argument Marking in Ditransitive Alignment Types. *Linguistic Discovery*, 3(1). <https://doi.org/10.1349/PS1.1537-0852.A.280>
- Haugen, E. (1950). The Analysis of Linguistic Borrowing. *Language*, 26(2), 210–331. <https://doi.org/10.2307/410058>

- Hein, B., & Kuteva, T. (2005). *Language Contact and Grammatical Change*. Cambridge University Press.
- Hekking, E. (1995). *El otomí de Santiago Mexquititlán: desplazamiento lingüístico, préstamos y cambios gramaticales*. IFOTT.
- Hekking, E. (2001). Cambios gramaticales por el contacto entre el otomí y el español. In T. Stolz & K. Zimmermann (Eds.), *Lo propio y lo ajeno en las lenguas austronésicas y amerindias: procesos interculturales en el contacto de las indígenas con el español en el Pacífico e Hispanoamérica* (pp. 127–151). Iberoamericana-Vervuert.
- Hekking, E., & Andrés de Jesús, S. (1984). *Gramática otomí*. Universidad Autónoma de Querétaro.
- Hekking, E., & Bakker, D. (2005). Problems with the acquisition of a second language: Otomi meets Spanish. In C. Chamoreau & Y. Lastra (Eds.), *Dinámica lingüística de las lenguas en contacto* (pp. 239–276). Universidad de Sonora.
- Hekking, E., & Bakker, D. (2007a). Cambios lingüísticos en el otomí y el español producto del contacto de lenguas en el estado de Querétaro. *UniverSOS. Revista de Lenguas Indígenas y Universos Culturales*, 4, 115–140.
- Hekking, E., & Bakker, D. (2007b). The case of Otomi: a contribution to grammatical borrowing in cross-linguistic perspective. In Y. Matras & J. Sakel (Eds.), *Grammatical Borrowing in Cross-linguistic Perspective* (pp. 435–464). Mouton de Gruyter.
- Hekking, E., & Bakker, D. (2010). Tipología de los préstamos léxicos en el otomí queretano. Una contribución para el estudio sistemático y comparativo de diversas lenguas representativas del mundo desde un enfoque interlingüístico. *CIENCIA@UAQ*, 3(1), 27–47.

- Hernández Green, N. (2015). *Morfosintaxis verbal del otomí de Acazulco* [Tesis de doctorado]. Centro de Investigaciones y Estudios Superiores en Antropología Social.
- Hernández Green, N. (2021). Concordancia de número plural con participantes de 3a persona en el verbo en ñätho (otomí de San Felipe). *Semas*, 2(4), 9–34.
- Herrera Soler, H., Martínez Arias, R., & Amengual Pizarro, M. (2011). *Estadística aplicada a la investigación lingüística*. Editorial EOS.
- Hidalgo, A., & Quilis, M. (2004). *Fonética y fonología españolas*. Tirant Humanidades.
- Hill, J. H., & Hill, K. C. (1981). Regularities in Vocabulary Replacement in Modern Nahuatl. *International Journal of American Linguistics*, 47(3), 215–226.
<https://doi.org/10.1086/465690>
- Hill, J., & Hill, K. (1977). Language Death and Relexification in Tlaxcalan Nahuatl. *International Journal of the Sociology of Language*, 1977(12), 55–70.
<https://doi.org/10.1515/ijsl.1977.12.55>
- Hualde, J. I. (2005). *The sounds of Spanish*. Cambridge University Press.
- Hulstijn, J. H. (2012). The construct of language proficiency in the study of bilingualism from a cognitive perspective. *Bilingualism: Language and Cognition*, 15(2), 422–433.
<https://doi.org/10.1017/S1366728911000678>
- Instituto Nacional de Estadística. (2010). *XIII Censo General de Población y Vivienda*. Universidad Nacional Autónoma de México.
- Instituto Nacional de Estadística, G. e I. (1990). *XI Censo General de Población y Vivienda*. Instituto Nacional de Estadística, Geografía e Informática.
- Instituto Nacional de Estadística, G. e I. (2000). *XII Censo General de Población y Vivienda*. Instituto Nacional de Estadística, Geografía e Informática.

Instituto Nacional de Estadística, G. e I. (2020). *XIV Censo General de Población y Vivienda*.

Instituto Nacional de Estadística, Geografía e Informática.

Instituto Nacional de Estadística y Geografía. (2009). *Perfil sociodemográfico de la población que habla lengua indígena*. Instituto Nacional de Estadística y Geografía.

Catálogo de las Lenguas Indígenas Nacionales: Variantes Lingüísticas de México con sus autodenominaciones y referencias geoestadísticas, Diario Oficial 31 (2008).

Ionin, T., & Matushansky, O. (2018). *Cardinals. The Syntax and Semantics of Cardinal-Cointaining Expressions*. The MIT Press.

Ito, J., & Mester, A. (1995). The Core-Periphery Structure of the Lexicon and Constraints on Reranking. In J. N. Beckman, L. Walsh, & Urbanczyk (Eds.), *Papers in Optimality Theory* (pp. 181–209). GLSA Publications.

Iwaniszewski, S. (2004). La breve historia del calendario del “Códice Telleriano - Remensis.” *Estudios de Cultura Náhuatl*, 35, 45–67.

Jackendoff, R. (1991). Parts and boundaries. *Cognition*, 41, 9–45.

Johanson, L. (2002). Contact-induced linguistic change in a code-copying framework. In *Language Change: The Interplay of Internal, External and Extra-linguistic Factors. Contributions to the Sociology of Language* (pp. 285–313). Mouton de Gruyter.

Labov, W. (1996). *Principios del cambio lingüístico. Vol. I: Factores Internos*. Gredos.

Labov, W. (2001). *Principles of Linguistic Change, Volume 2: Social Factors*. Blackwell.

Laca, B. (1999). Presencia y ausencia de determinante. In I. Bosque & V. Demonte (Eds.), *Gramática descriptiva de la lengua española. Vol. 1* (pp. 891–921). Espasa Calpe.

Larralde, A., & González, S. (2013). Conceptualización y medición de lo rural. Una propuesta para clasificar el espacio rural en México. In *La situación demográfica de México* (pp. 141–157). Consejo Nacional de Población.

- Lastra, Y. (1987). El español de una familia bilingüe otomí. In *Studia Humanitatis. Homenaje a Rubén Bonifaz Nuño* (pp. 1–419). Universidad Nacional Autónoma de México.
- Lastra, Y. (1989). *Otomí de San Andrés Cuexcontitlán, Estado de México*. El Colegio de México.
- Lastra, Y. (1992). *El otomí de Toluca*. Universidad Nacional Autónoma de México.
- Lastra, Y. (1994). Préstamos y alternancias de código en otomí y en español. In C. Mackay & V. Vázquez (Eds.), *Investigaciones lingüísticas en Mesoamérica* (pp. 185–195). Universidad Nacional Autónoma de México.
- Lastra, Y. (1997). *El otomí de Ixtenco*. Universidad Nacional Autónoma de México.
- Lastra, Y. (2001). *Unidad y diversidad de la lengua. Relatos otomíes*. Universidad Nacional Autónoma de México.
- Lastra, Y. (2010). *Los otomíes: su lengua y su historia*. Universidad Nacional Autónoma de México.
- Lastra, Y., & Horcasitas, F. (1976). El náhuatl en el Distrito Federal. *Anales de Antropología*, 13, 103–136.
- Lastra, Y., & Martín Butragueño, P. (2000). El modo de vida como factor sociolingüístico en la ciudad de México. In P. Martín Butragueño (Ed.), *Estructuras en contexto. Estudios de variación y cambio* (pp. 13–43). El Colegio de México.
- Leonetti, M. (1999). El artículo. In I. Bosque, V. Demonte, & (dirs.) (Eds.), *Gramática descriptiva de la lengua española. Volumen I: Sintaxis básica de las clases de palabras* (pp. 787–890). Espasa.
- Lizárraga Navarro, G. Z. (2014). Los pronombres clíticos del español en hablantes bilingües español-otomí. *Estudios de Lingüística Aplicada*, 59, 39–66.

- Martín Butragueño, P. (2014). *Fonología variable del español de México. Volumen I: Procesos segmentales*. El Colegio de México.
- Martín Butragueño, P., & Lastra, Y. (2011). *Corpus Sociolingüístico de la Ciudad de México. Materiales de PRESEEA - México. Volumen I: Hablantes de instrucción superior*. El Colegio de México.
- Martínez, A. (2001). *Contacto de lenguas. Discordancias gramaticales y concordancias semánticas: el número*.
- Martínez, A. (2012). El español de los Andes en la Argentina. Concordancias canónicas y concordancias alternativas de número. *Revista Neue Romania*, 4, 141–164.
- Martínez, A. (2017). Un antes y un después en la teoría del número verbal: el aporte del español de los Andes. In A. Palacios (Ed.), *Variación y cambio lingüístico en situaciones de contacto* (pp. 97–112). Iberoamericana Vervuert. <https://doi.org/10.31819/9783954876648-005>
- Martínez, J. A. (1999). La concordancia. In I. Bosque & V. Demonte (Eds.), *Gramática descriptiva de la lengua española. Volumen II: Las construcciones sintácticas fundamentales. Relaciones temporales, aspectuales y modales* (pp. 2695–2786). Espasa.
- Matras, Y. (2008). The borrowability of structural categories. In Y. Matras & J. Sakel (Eds.), *Grammatical Borrowing in Cross-Linguistic Perspective* (pp. 31–74). De Gruyter Mouton. <https://doi.org/10.1515/9783110199192.31>
- Mendenhall, W., Beaver, R. J., & Beaver, B. M. (2010). *Introducción a la probabilidad y estadística* (13a ed.). Cengage Learning. <http://latinoamerica.cengage.com>
- Millones, R., Barreno, E., Vásquez, F., & Castillo, C. (2018). *Estadística descriptiva y probabilidades. Aplicaciones en la ingeniería y los negocios*. Universidad de Lima. Fondo Editorial.

- Mithun, M. (1991). Active/agentive Case Marking and Its Motivations. *Language*, 67(3), 510–546. <https://doi.org/10.2307/415036>
- Montrul, S. (2013). *El bilingüismo el mundo hispanohablante*. Wiley-Blackwell.
- Morales Peláez, I. (2006). Modelos de regresión: lineal simple y regresión logística. In A. Guillén Serra & R. Crespo Montero (Eds.), *Métodos estadísticos para enfermería nefrológica* (pp. 195–213). SEDEN.
- Moreno Fernández, F. (1998). *Principios de sociolingüística y sociología del lenguaje*. Editorial Ariel, S.A.
- Mulík, S., Corona Dzul, B., Amengual, M., & Carrasco Ortíz, H. (2021). Perfil psicolingüístico de los bilingües otomí (hñãñho)-español, migrantes de Santiago Mexquititlán a Santiago de Querétaro, México. *Cuadernos de Lingüística de El Colegio de México*, 8, 1–50. <https://doi.org/10.24201/clecm.v8i0.154>
- Muysken, P. (1981). Halfway between Quechua and Spanish: The case of relexification. In A. Highfield & A. Valdman (Eds.), *Historicity and variation in creoles studies* (pp. 52–78). Karoma.
- Myers-Scotton, C. (1993). *Social motivations for codeswitching. Evidence from Africa*. Clarendon.
- Navarro, T. (1946). *Estudios de fonología española*. Syracuse University Press.
- Nichols, J. (1986). Head-marking and dependent-marking grammar. *Language*, 62(1), 56–119. <https://doi.org/10.1353/lan.1986.0014>
- Olate Vinet, A., Pineda Carrasco, R., Hasler Sandoval, F., & Soto Vergara, G. (2019). La concordancia de número en el castellano hablado por bilingües mapuzugun-castellano: “Y después ya cambió eso cuando aprendió a hablar wingka los mapuche”. *Círculo de*

Lingüística Aplicada a La Comunicación, 78, 211–232.
<https://doi.org/10.5209/clac.64379>

Palacios, A. (2005). Aspectos teóricos y metodológicos del contacto de lenguas: el sistema pronominal del español en áreas de contacto con lenguas amerindias. In V. Noll, K. Zimmermann, & I. Neuman-Holzschuh (Eds.), *El español en América: Aspectos teóricos, particularidades, contactos* (pp. 63–94). Iberoamericana/Vervuert.

Palacios, A. (2017). *Variación y cambio lingüístico en situaciones de contacto*. Iberoamericana/Vervuert.

Palacios Alcaine, A. (2011). Nuevas perspectivas en el estudio del cambio inducido por contacto: hacia un modelo dinámico del contacto de lenguas. *Revista Lenguas Modernas*, 38, 17–38.

Palancar, E. L. (2006). Property Concepts in Otomi: A Language with No Adjectives. *International Journal of American Linguistics*, 72(3), 325–366.

Palancar, E. L. (2009). *Gramática y Textos del hñõñhõ. Otomí de San Ildefonso Tultepec Querétaro. Volumen I: Gramática*. Plaza y Valdes Editores.

Palancar, E. L. (2013). The evolution of number in Otomi. *Studies in Language*, 37(1), 94–142. <https://doi.org/10.1075/sl.37.1.03pal>

Palma Patricio, A. (2018). *Tecnologías de la Información y Comunicación e Identidad en Jiquipilco el Viejo y Tonanitla, Estado de México* [Tesis doctoral]. Centro de Investigaciones y Estudios Superiores en Antropología Social.

Pearson, B. Z., Fernández, S. C., & Oller, D. K. (1993). Lexical Development in Bilingual Infants and Toddlers: Comparison to Monolingual Norms. *Language Learning*, 43(1), 93–120. <https://doi.org/10.1111/j.1467-1770.1993.tb00174.x>

Pearson, R. (2020). *Package ‘GoodmanKruskal.’*

- Pfeiler, B. (2009). The acquisition of numeral classifiers and optional plural marking in Yucatec Maya. In U. Stephany & M. D. Voeikova (Eds.), *Development of nominal inflection in first language acquisition* (pp. 91–110). De Gruyter. <https://doi.org/10.1515/9783110217117.91>
- Pfeiler, B. (2014). Maya and Spanish in Yucatán. In S. S. Mufwene (Ed.), *Iberian Imperialism and Language Evolution in Latin America* (pp. 205–224). University of Chicago Press. <https://doi.org/10.7208/chicago/9780226125671.003.0008>
- Pool Balam, L. I., & le Guen, O. (2015). La integración de préstamos en maya yucateco. Una perspectiva interaccional acompañada de una reflexión sobre la ideología del lenguaje. *Amerindia*, 37(2), 343–384.
- Poplack, S. (1980). Sometimes I'll start a sentence in Spanish Y TERMINO EN ESPAÑOL: toward a typology of code-switching¹. *Linguistics*, 18(7–8), 581–618. <https://doi.org/10.1515/ling.1980.18.7-8.581>
- Poplack, S., & Levey, S. (2010). Contact-induced grammatical change: A cautionary tale . In P. Auer & J. E. Schmidt (Eds.), *Language and Space. An international handbook of linguistic variation. Volume 1: Theories and methods* (pp. 391–419). Mouton de Gruyter.
- Poplack, S., & Meechan, M. (1998). Introduction: How Languages Fit Together in Codemixing. *International Journal of Bilingualism*, 2(2), 127–138. <https://doi.org/10.1177/136700699800200201>
- Poplack, S., & Sankoff, D. (1984). Borrowing: the synchrony of integration. *Linguistics*, 22(1). <https://doi.org/10.1515/ling.1984.22.1.99>
- Pozas Loyo, J. (2016). *El artículo indefinido. Origen y gramaticalización*. El Colegio de México.

- Quilis, A. (1997). *Principios de fonología y fonética españolas*. Arco/Libros.
- R CORE TEAM. (2021). *R: A language and environment for statistical computing*. R Foundation for Statistical Computing.
- RAE y ASALE. (2010). *Nueva Gramática de la Lengua Española. Manual*. Espasa.
- Ramírez Cruz, H. (2009). La inconcordancia de género y número en el contacto de lenguas. *Forma y Función*, 22(2), 163–195.
- Romaine, S. (1989). *Bilingualism*. Blackwell.
- Salazar Mondragón, G. (2019). *El Altepétl de Xiquipilco: historia de una ciudad otomí* [Tesis de licenciatura]. Universidad Autónoma del Estado de México.
- San Giacomo Trinidad, M., & Peperkamp, S. (2008). Presencia del español en náhuatl: estudio sociolingüístico de la adaptación de préstamos. In M. Westmoreland & J. Thomas (Eds.), *Selected Proceedings of the Fourth Workshop on Spanish Sociolinguistics* (pp. 149–156). Cascadilla Press.
- Sánchez Paraíso, M. (2021). La conciencia de la norma lingüística en hablantes de español andino peruano: el caso de los pronombres átonos de 3.a persona. In S. Gómez Seibane, M. Sánchez, & A. Palacios (Eds.), *Traspassando lo lingüístico: Factores esenciales en el contacto de lenguas* (pp. 49–68). Iberoamericana Vervuert. <https://doi.org/10.31819/9783968692340-004>
- Sankoff, D., & Poplack, S. (1981). A formal grammar for code-switching ¹. *Paper in Linguistics*, 14(1), 3–45. <https://doi.org/10.1080/08351818109370523>
- Silva-Corvalán, C. (2001). *Sociolingüística y pragmática del español*. Georgetown University Press.
- Silva-Corvalán, C., & Enrique-Arias, A. (2018). *Sociolingüística y pragmática del español* (2a ed.). Georgetown University Press.

- Soler Arechalde, M. Á. (2012). *La concordancia de número en español. Cuatro casos de alternancia*. Universidad Nacional Autónoma de México.
- Soustelle, J. (1993). *La familia otomi-pame del centro de México título original en francés: La Famille Otomi-Pame du Mexique Central*. Instituto Mexiquense de Cultura y Universidad Autónoma del Estado de México.
- SPSS. (2020). *Statistical Package for the Social Sciences* (13 para Windows).
- Tagliamonte, S. A., Poplack, S., & Eze, E. (1997). Plural Marking Patterns in Nigerian Pidgin English. *Journal of Pidgin and Creole Languages*, 12(1), 103–129.
<https://doi.org/10.1075/jpcl.12.1.04tag>
- Tejeda Alvarado, F. (2019). *El Lienzo de Jucutacato. Códice Postcortesiano*. Consejo Editorial H. Cámara de Diputados.
- Terborg, R. (2011). La situación del otomí de San Cristóbal Huichochitlán del Estado de México. In *Muerte y vitalidad de las lenguas indígenas y las presiones de sus hablantes* (pp. 197–220). Universidad Nacional Autónoma de México.
- Thomason, S. G. (2001). *Language contact*. Edinburgh University Press.
- Thomason, S. G., & Kaufman, T. (1988). *Language Contact, Creolization, and Genetic Linguistics*. University of California Press.
- Torres Cacoullós, R., & Travis, C. E. (2018). *Bilingualism in the Community. Code-switching and Grammars in Contact*. Cambridge University Press.
- Torres Sánchez, N. (2018). *Aquí hablamos tepehuano y allá español. Un estudio de la situación de bilingüismo incipiente entre español y tepehuano del sureste odam en Santa María de Ocotán y Durango*. El Colegio de México.
- Torres Sánchez, N. (2019). ¿Cómo medir el bilingüismo individual en una situación de contacto? *Lingüística Mexicana. Nueva Época*, 1(3), 63–89.

- Treffers-Daller, J. (2018). The Measurement of Bilingual Abilities. In A. de Houwer & L. Ortega (Eds.), *The Cambridge Handbook of Bilingualism* (pp. 289–306). Cambridge University Press. <https://doi.org/10.1017/9781316831922.016>
- Uth, M. U., & Gutiérrez Bravo, R. (2018). La marcación de número en los préstamos españoles del maya yucateco: variación y restricciones. *Cuadernos de Lingüística de El Colegio de México*, 5(1), 183–224. <https://doi.org/10.24201/clecm.v5i1.99>
- Villaseñor y Sanches, J. A. (1952). *Theatro Americano. Descripción general de los reynos y provincias de la Nueva España y sus jurisdicciones*. Editora Nacional.
- Voigtlander, K., & Echegoyen, A. (1985). *Luces contemporáneas del otomí: Gramática del otomí de La Sierra*. Instituto Lingüístico de Verano.
- Weinreich, U. (1953). *Languages in Contact. Findings and problems* (4a ed.). Mouton.
- Whitney, W. D. (1881). On mixture in language. *TAPA*, 12, 5–26.

APÉNDICE CAPÍTULO 2

APÉNDICE 2A. CUESTIONARIO DE HABILIDADES LINGÜÍSTICAS EN ESPAÑOL

Nombre: _____

Objetivo: determinar el nivel de competencia (fónica, léxica y morfosintáctica) del colaborador en español, a partir de la presencia/ausencia de transferencias del otomí en esos niveles. Este cuestionario se realizará en español y todas las respuestas del informante serán grabadas.

Nivel fonológico

Instrucciones: el entrevistador proporcionará información o una descripción de la palabra que requiere del informante.

Ejemplo: E: ¿qué es lo contrario a un hombre? I: una mujer.

1. correr (¿cómo se le llama a la acción de caminar rápido?)
2. raro (¿cómo se le dice a una persona que es extraña?)
3. amarillo (¿de qué color es el sol?)
4. lluvia (¿cómo se le dice a las gotas de agua que caen del cielo?)
5. lombriz (animal de forma redonda que vive en la tierra)
6. corazón (parte del cuerpo que bombea sangre)
7. café (bebida caliente de color oscuro que se suele beber en la mañana)
8. fotógrafo (¿cómo se le dice a la persona que saca fotos?)
9. flaco (¿cómo se le llama a una persona que es delgada?)
10. familia (grupo de personas integrado por una mamá un papá e hijos)
11. chico (¿qué es lo contrario de grande?)
12. leche (líquido que producen las vacas y que consume el hombre)

Nivel léxico

Instrucciones: el entrevistador pondrá a escuchar al colaborador una grabación con 15 palabras en otomí, previamente grabadas por un informante bilingüe, y le pedirá que diga su correspondiente en español.

- | | |
|---------------|---------------|
| 1. documento | 8. escapar |
| 2. ayate | 9. tú |
| 3. miedo | 10. él |
| 4. calzón | 11. adentro |
| 5. alegrarse | 12. afuera |
| 6. arrancar | 13. horroroso |
| 7. amanecer | 14. hermoso |
| 15. friolento | |

Nivel morfosintáctico

Parte I

Instrucciones: el entrevistador le leerá al informante dos o tres construcciones sintácticas similares y le pedirá que escoja la que generalmente usa.

- a) Mi hermano habla de quedito
b) Mi hermano habla quedito
- a) En mi pueblo la gente habla de español
b) En mi pueblo la gente habla español
c) En mi pueblo la gente habla en español
- a) Todos mis hijos estaban chico
b) Todos estaban chico mis hijos
c) Todos mis hijos estaban chicos

Parte II

Instrucciones: el entrevistador empleará una serie de imágenes para obtener construcciones sintácticas en español que presentan algún grado de dificultad para los bilingües otomí-español.

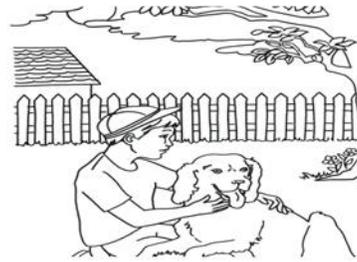
1. Concordancia del sustantivo con el verbo



2. Concordancia de número entre sustantivo y el numeral



3. Redundancia del pronombre átono de objeto directo *lo*



4. Omisión o uso alterno del pronombre reflexivo *se*



5. Omisión o uso alterno de la preposición *a* de objeto indirecto



6. Omisión o uso alterno de la preposición *a*



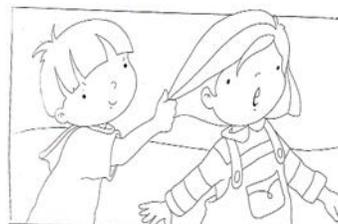
7. Omisión o uso alternativo de la preposición *en*



8. Uso de la preposición *en*



9. Omisión o uso alternativo del pronombre de objeto indirecto *le*



10. Omisión o uso alternativo de la conjunción coordinativa *y*



Parte III

Instrucciones: el entrevistador leerá algunas oraciones al colaborador y le hará preguntas sobre eso.

1. Mis tíos, mi mamá y María fueron a misa ayer.
 - a. ¿Quiénes fueron a misa mis tíos o mi mamá?
 - b. ¿María iba con mis tíos?
2. La bicicleta que traía Juan, Pedro se la regaló.
 - a. ¿Quién le regaló la bicicleta a Juan?
 - b. ¿De quién era la bicicleta?
3. Julia y Karen se peinaron
 - a. ¿Cada quién se peinó por separada?
 - b. ¿Se peinaron entre ellas?

4. Marcos le prestó una pluma a Margarita
- a. ¿A quién le prestó la pluma Marcos?
 - b. ¿Ya le prestó Marcos la pluma a Margarita?

APÉNDICE 2B CUESTIONARIO DE HABILIDADES LINGÜÍSTICAS EN OTOMÍ

Nombre: _____

Objetivo: determinar el nivel de competencia (fónica, léxica y morfosintáctica) del colaborador en otomí, a partir de la presencia/ausencia de transferencias del español en esos niveles. Este cuestionario se realizará en español y todas las respuestas del informante serán grabadas.

Nivel fonológico

Instrucciones: el entrevistador mencionará las palabras en español y el colaborador proporcionará su equivalente en otomí.

- | | |
|-----------------------------------|---------------------------|
| 1. <i>xät'ä</i> 'nopal' | 9. <i>mui</i> 'barriga' |
| 2. <i>thä</i> 'mazorca' | 10. 'ñäi' 'epazote' |
| 3. <i>tuhu</i> 'cantar' | 11. 'yo' 'perro' |
| 4. <i>gu</i> 'oreja' | 12. <i>hqi</i> 'tierra' |
| 5. <i>gin'ue</i> 'mosca' | 13. <i>xi</i> 'cáscara' |
| 6. <i>sei</i> 'pulque' | 14. <i>däza</i> 'plátano' |
| 7. <i>dęę</i> 'capulín' | 15. <i>zq</i> 'bonito' |
| 8. <i>dada, tu, mbõhõ</i> 'señor' | |

Nivel léxico

Instrucciones: el entrevistador mencionará las palabras en español y el colaborador proporcionará su equivalente en otomí

- | | |
|--------------------------------------|----------------------------|
| 1. <i>bõts'e</i> 'costilla' | 10. <i>oki</i> 'agujerado' |
| 2. <i>Ajüa/Zidada</i> 'Dios' | 11. <i>şehę</i> 'solo' |
| 3. 'bax' <i>bõts'e</i> 'chiquigüite' | 12. <i>mfege</i> 'roto' |
| 4. <i>däza</i> 'plátano' | 13. <i>nda'ä</i> 'cuál' |
| 5. <i>gä'tsi</i> 'suspirar' | 14. <i>nu'ü</i> 'aquellos' |
| 6. <i>agi</i> 'enterrar' | 15. <i>ha</i> 'en' |
| 7. 'bęts'i' 'embarrar' | |
| 8. <i>hämbi</i> 'arrebatar' | |
| 9. <i>hñü</i> 'pesado' | |

Nivel morfosintáctico

Instrucciones: el entrevistador empleará una serie de imágenes para obtener construcciones sintácticas en otomí que presentan algún grado de dificultad para los bilingües otomí-español.

1. Frase nominal (inclusión de un numeral del español)



2. Frase nominal coordinada (inserción de la conjunción y)



3. Uso de la morfología verbal



4. Oración transitiva



5. Oración transitiva



6. Oración ditransitiva



7. Oración ditransitiva



8. Oración reflexiva



9. Oración con locativo (inserción de preposiciones del español)



10. Oración locativa (inserción de preposiciones del español)



APÉNDICE 2C. CUESTIONARIO DE HABILIDADES LINGÜÍSTICAS PASIVAS EN OTOMÍ

Nombre: _____

Objetivo: determinar el nivel de conocimiento pasivo (léxico y morfosintáctico) en otomí del colaborador. Este cuestionario se realizará en español y todas las respuestas del informante serán grabadas.

Nivel léxico

Instrucciones: el entrevistador pondrá a escuchar al colaborador una grabación con 58 palabras en otomí, previamente grabadas por un informante bilingüe, y le pedirá que diga su correspondiente en español.

- | | |
|-----------------|----------------|
| 1. 'agua' | 22. 'epazote' |
| 2. 'agujereó' | 23. 'escoba' |
| 3. 'árbol' | 24. 'espuma' |
| 4. 'barriga' | 25. 'este' |
| 5. 'blanco' | 26. 'estrella' |
| 6. 'bonito' | 27. 'flor' |
| 7. 'bueno' | 28. 'gallina' |
| 8. 'caballo' | 29. 'guaje' |
| 9. 'cabeza' | 30. 'hijo' |
| 10. 'caerse' | 31. 'hoyo' |
| 11. 'caliente' | 32. 'huarache' |
| 12. 'capulín' | 33. 'lumbre' |
| 13. 'ceniza' | 34. 'mamá' |
| 14. 'chayote' | 35. 'mañana' |
| 15. 'chico' | 36. 'mañana_2' |
| 16. 'cierra' | 37. 'mazorca' |
| 17. 'come' | 38. 'metate' |
| 18. 'componer' | 39. 'México' |
| 19. 'dónde' | 40. 'negro' |
| 20. 'encontrar' | 41. 'oreja' |
| 21. 'enfermo' | 42. 'perro' |

- | | |
|----------------|----------------|
| 43. 'pie' | 51. 'sueño' |
| 44. 'piojo' | 52. 'temer' |
| 45. 'plátano' | 53. 'tierra' |
| 46. 'puerta' | 54. 'tortilla' |
| 47. 'pulque' | 55. 'bebé' |
| 48. 'sal' | 56. 'yo' |
| 49. 'señor' | 57. 'zapote' |
| 50. 'siéntate' | |

Nivel morfosintáctico

Instrucciones: el entrevistador pondrá a escuchar al colaborador una grabación con 33 oraciones en otomí, previamente grabadas por un informante bilingüe, y le pedirá que diga su correspondiente en español.

- | | |
|--|---------------------------------------|
| 1. 'A ese hombre lo mataron ayer' | 17. 'Hace calor' |
| 2. 'Ayer acabé de sembrar' | 18. 'Hace frío' |
| 3. 'Buenas tardes' | 19. 'Hasta mañana' |
| 4. 'Buenos días' | 20. 'Hoy acabo de sembrar' |
| 5. '¿Cómo es esa casa?' | 21. 'Juan sabe hablar español' |
| 6. 'Él es gordo' | 22. 'Juan se ríe de ti' |
| 7. 'Él nunca me pide tortillas' | 23. 'La mosca está en la pared' |
| 8. 'Ese árbol creció' | 24. 'La puerta se cerró' |
| 9. 'Ese árbol va a ser grande' | 25. 'Le pregunté' |
| 10. 'Ese árbol va a ser o es grande' | 26. 'Los metates de él' |
| 11. 'Ese niño se parece mucho a su papá' | 27. 'Mi papá siempre está en la casa' |
| 12. 'Ese niño se parece mucho a su papá_2' | 28. 'Nosotros somos grandes' |
| 13. 'Ese pájaro no puede volar' | 29. '¿Qué tan grande es esa casa?' |
| 14. 'Esta agujereado' | 30. 'Tengo dolor de cabeza' |
| 15. 'Esta casa es grande' | 31. '¿Tu caballo es blanco o negro?' |
| 16. 'Esta casa no es grande' | 32. 'Va a llover' |
| | 33. 'Vi animales en el camino' |

APÉNDICE 2D. CUESTIONARIO SOCIOLINGÜÍSTICO

Datos sociodemográficos

1) Nombre _____

2) Género: 1. Hombre 2. Mujer

3) Lugar de entrevista: 1. Casa 2. Calle 3. Trabajo

4) Lugar de nacimiento: _____

5) Edad: 1. 14-19 2. 20-35 3. 36-55 4. 55+

6) Domicilio actual: _____

1. San Andrés Cuexcontitlán 2. Ciudad pequeña: Toluca 3. Ciudad grande: (México)

4. Otro: _____

7) ¿Cuántos años ha vivido en su comunidad actual?

1. 1-7 años 2. 7-14 años 3. 15-22 4. 23+

8) ¿Ha salido muchas veces de su comunidad?

0. nunca 1. Pocas veces 2. Muchas 3. Periódicamente

8.1 ¿A dónde?

1. Ciudad pequeña 2. Ciudad grande 3. Norte (EE.UU.) 4. Otro

8.2 ¿Cuánto tiempo ha estado fuera de su comunidad de origen?

1. 1/2-1 año 2. 2-6 años 3. 7-15 4. 15 +

9) ¿Cuántos años ha estado en la escuela?

0. 0 años 1. 1-3 años 2. 4-6 años 3. 7-10 4. 11 +

9.1 ¿Actualmente estudia?

1. Sí 2. No

10) Estado civil:

1. Vive en pareja: (casado/unión libre) 2. Separado/Divorciado/Viudo 3. Soltero

Datos socioeconómicos

11) ¿Realiza alguna actividad económica?

1. Sí 2. No, pensionado 3. No, tareas domésticas 4. No, estudio

5. No, impedimento físico 6. No, edad 7. No, desempleado NR

12) ¿Usted trabaja para alguien o lo hace por cuenta propia?

1. Soy empleado u obrero 2. Soy peón o jornalero 3. Trabajo por cuenta propia

13) ¿Qué actividad económica realiza principalmente para subsistir?

14) ¿Realiza alguna otra actividad complementaria? ¿Cuál?

15) ¿Durante el último año usted ha migrado para trabajar?

1. Sí 2. No, no tengo como financiarme 3. No, no tengo con quien llegar ni trabajo
4. No, estudio 5. No, estoy cumpliendo algún cargo 6. No, me da miedo
7. No, no me interesa migrar NR

16) ¿Cuándo viaja para trabajar, se queda en una sola ciudad?

1. Sí, estacionaria 2. No, dos o más en temporadas diferentes, temporal
3. No, diferentes ciudades hasta volver a donde comencé a trabajar, circular NR

17) ¿Cuánto tiempo dedica a las actividades de subsistencia?

1. de 1 a 4 horas diarias 2. Más de 4 horas diarias 3. Algunos días a la semana de
1 a 4 horas cada día 5. Algunos días más de 4 horas NR

18) ¿Recibe ingresos por las actividades que realiza para subsistir?

1. Sí 2. No, actividad familiar y de autoconsumo 3. No, actividad para adquirir
experiencia 4. No, actividad familiar y los ingresos los administran los jefes de
familia NR

19) ¿Aproximadamente de cuánto es su ingreso mensual?

1. Cerca de \$1, 600 2. Entre \$1,600 y \$3,200 3. Entre \$3, 200 y \$4,000
4. Entre \$4,000 y \$8,000 6. Más de \$ 8, 000

Adquisición y competencia lingüística

20) ¿Qué lenguas habla?

1. Sólo español 2. Sólo otomí 3. Español y otomí 4. Otra NR

21) ¿Dónde aprendió el otomí? ¿Quién le enseñó a hablarlo o cómo lo aprendió?

22) ¿Dónde aprendió el español? ¿Quién le enseñó a hablarlo o cómo lo aprendió?

23) ¿A qué edad aprendió el otomí?

0. 0-5 años 1. 5-15 años 2. 16 + NR

24) ¿A qué edad aprendió el español?

0. 0-5 años 1. 5-15 años 2. 16 + NR

25) ¿Cuánto hace que no habla otomí?

0. Actualmente 1. Alguna vez lo hablo 2. Nunca lo he hablado 3. + de 1 año NR

26) ¿Cuánto hace que no habla español?

0. Actualmente 1. Alguna vez lo hablo 2. Nunca lo he hablado 3. + de 1 año NR

27) ¿Cuál lengua habla mejor?

1. Español 2. Otomí 3. Español y otomí 4. Otra NR

28) ¿Cuál lengua entiende mejor?

1. Español 2. Otomí 3. Español y otomí 4. Otra NR

29) ¿En cuál lengua piensa?

1. Español 2. Otomí 3. Español y otomí 4. Otra NR

30) ¿En cuál lengua sueña?

1. Español 2. Otomí 3. Español y otomí 4. Otra NR

31) ¿En cuál lengua hace cuentas?

1. Español 2. Otomí 3. Español y otomí 4. Otra NR

32) ¿Lee el español?

0. Sí 1. No 2. Un poco NR

33) ¿Lee el otomí?

0. Sí 1. No 2. Un poco NR

34) ¿Escribe el español?

0. Sí 1. No 2. Un poco NR

35) ¿Escribe el otomí?

0. Sí 1. No 2. Un poco NR

Uso de las lenguas

36) ¿Qué lengua habla con su familia?

1. Sólo español 2. Sólo otomí 3. Español y otomí 4. Otra NR

37) ¿Con la mamá?

1. Sólo español 2. Sólo otomí 3. Español y otomí 4. Otra NR

38) ¿Con el papá?

1. Sólo español 2. Sólo otomí 3. Español y otomí 4. Otra NR

39) ¿Con los hermanos o hermanas?

1. Sólo español 2. Sólo otomí 3. Español y otomí 4. Otra NR

40) ¿Con el abuelo?

1. Sólo español 2. Sólo otomí 3. Español y otomí 4. Otra NR

41) ¿Con la abuela?

1. Sólo español 2. Sólo otomí 3. Español y otomí 4. Otra NR

42) ¿Con sus tíos?

1. Sólo español 2. Sólo otomí 3. Español y otomí 4. Otra NR

43) ¿Con sus tías?

1. Sólo español 2. Sólo otomí 3. Español y otomí 4. Otra NR

44) ¿Con sus primos?

1. Sólo español 2. Sólo otomí 3. Español y otomí 4. Otra NR

45) ¿Con sus primas?

1. Sólo español 2. Sólo otomí 3. Español y otomí 4. Otra NR

46) ¿Con la pareja?

1. Sólo español 2. Sólo otomí 3. Español y otomí 4. Otra NR

46.1 ¿Con los hijos o hijas?

1. Sólo español 2. Sólo otomí 3. Español y otomí 4. Otra NR

46.2 ¿Con su suegro?

1. Sólo español 2. Sólo otomí 3. Español y otomí 4. Otra NR

46.3 ¿Con su suegra?

1. Sólo español 2. Sólo otomí 3. Español y otomí 4. Otra NR

46.4 ¿Con su nuera?

1. Sólo español 2. Sólo otomí 3. Español y otomí 4. Otra NR

46.5 ¿Con su yerno?

1. Sólo español 2. Sólo otomí 3. Español y otomí 4. Otra NR

46.6 ¿Con su nieto o nieta?

1. Sólo español 2. Sólo otomí 3. Español y otomí 4. Otra NR

47) ¿Con los amigos?

1. Sólo español 2. Sólo otomí 3. Español y otomí 4. Otra NR

48) ¿Con los compadres?

1. Sólo español 2. Sólo otomí 3. Español y otomí 4. Otra NR

49) ¿Con las comadres?

1. Sólo español 2. Sólo otomí 3. Español y otomí 4. Otra NR

Responda qué lengua utilizaría en las siguientes situaciones:

50) ¿En su casa con su familia?

1. + Español 2. + Otomí 3. Español y otomí 4. Otra NR

50.1 ¿En su casa con sus amigos?

1. + Español 2. + Otomí 3. Español y otomí 4. Otra NR

51) ¿En su trabajo con sus compañeros?

1. + Español 2. + Otomí 3. Español y otomí 4. Otra NR

51.1 ¿En su trabajo con su familia?

1. + Español 2. + Otomí 3. Español y otomí 4. Otra NR

51.2 ¿En su trabajo con su jefe?

1. + Español 2. + Otomí 3. Español y otomí 4. Otra NR

52) ¿En el mercado con su familia?

1. + Español 2. + Otomí 3. Español y otomí 4. Otra NR

52.1 ¿En el mercado con sus amigos?

1. + Español 2. + Otomí 3. Español y otomí 4. Otra NR

52.2 ¿En el mercado con los comerciantes?

1. + Español 2. + Otomí 3. Español y otomí 4. Otra NR

53) ¿En la ciudad con su familia?

1. + Español 2. + Otomí 3. Español y otomí 4. Otra NR

53.1 ¿En la ciudad con sus amigos? 1

1. + Español 2. + Otomí 3. Español y otomí 4. Otra NR

54) ¿Con las autoridades de San Andrés Cuexcontitlán?

1. + Español 2. + Otomí 3. Español y otomí 4. Otra NR

55) ¿En la escuela?

1. + Español 2. + Otomí 3. Español y otomí 4. Otra NR

55.1 ¿En la escuela con sus amigos?

1. + Español 2. + Otomí 3. Español y otomí 4. Otra NR

55.2 ¿En la escuela con su familia?

1. + Español 2. + Otomí 3. Español y otomí 4. Otra NR

55.3) ¿En la escuela con su profesor?

1. + Español 2. + Otomí 3. Español y otomí 4. Otra NR

56) ¿En la iglesia?

1. + Español 2. + Otomí 3. Español y otomí 4. Otra NR

56.1 ¿En la iglesia con su familia?

1. + Español 2. + Otomí 3. Español y otomí 4. Otra NR

56.2 ¿En la iglesia con sus amigos?

1. + Español 2. + Otomí 3. Español y otomí 4. Otra NR

56.3 ¿En la iglesia con el sacerdote o pastor?

1. + Español 2. + Otomí 3. Español y otomí 4. Otra NR

57) ¿En las fiestas del pueblo?

1. + Español 2. + Otomí 3. Español y otomí 4. Otra NR

57.1 ¿En las fiestas con su familia?

1. + Español 2. + Otomí 3. Español y otomí 4. Otra NR

57.2 ¿En las fiestas con sus amigos?

1. + Español 2. + Otomí 3. Español y otomí 4. Otra NR

58) ¿En las juntas o reuniones del pueblo?

1. + Español 2. + Otomí 3. Español y otomí 4. Otra NR

58.1 ¿En las juntas con su familia?

1. + Español 2. + Otomí 3. Español y otomí 4. Otra NR

58.2 ¿En las juntas con sus amigos?

1. + Español 2. + Otomí 3. Español y otomí 4. Otra NR

59) ¿En la calle de San Andrés Cuexcontitlán?

1. + Español 2. + Otomí 3. Español y otomí 4. Otra NR

59.1 ¿En la calle con su familia?

1. + Español 2. + Otomí 3. Español y otomí 4. Otra NR

59.2 ¿En la calle con sus amigos?

1. + Español 2. + Otomí 3. Español y otomí 4. Otra NR

Actitudes

60) ¿Cree que alguna lengua es más importante?

1. Español 2. Otomí 3. Ninguna o son iguales 4. Otra NR

60.1 ¿Por qué? (si elige OTRA especificar cuál)

61) ¿Cree que debería haber programas de radio y tv en otomí?

1. Sí 2. No 3. Debería haber más 4. Hay suficientes NR

62) ¿En caso de no saber, le interesaría aprender a hablar en español?

1. Sí 2. No NR

62.1 ¿Por qué?

1. Sí, me sería muy útil 2. No, porque es difícil 3. Sí, pero es muy difícil 4. No, no me sería útil

62.2 ¿Le interesaría aprender a hablar en otomí?

1. Sí 2. No NR

62.3 ¿Por qué?

1. Sí, me sería muy útil 2. No, porque es difícil 3. Sí, pero es muy difícil 4. No, no me sería útil

62.4 ¿En caso de no saber, le interesaría aprender a leer y escribir en español?

1. Sí 2. No NR

62.5 ¿Por qué?

1. Sí, me sería muy útil 2. No, porque es difícil 3. Sí, pero es muy difícil 4. No, no me sería útil

62.6 ¿Le interesaría aprender a leer y escribir en otomí?

1. Sí 2. No NR

62.7 ¿Por qué?

1. Sí, me sería muy útil 2. No, porque es difícil 3. Sí, pero es muy difícil 4. No, no me sería útil

62.8 ¿Le interesaría que hubiera libros y revistas en otomí?

1. Sí 2. No NR

63) ¿Cree que el otomí puede perderse o dejarse de hablar?

1. Sí 2. No, porque es difícil 3. No me interesa 4. Sin opinión NR

64) ¿Cree que los otomíes tienen derecho a que se les enseñe en su lengua?

1. Sí 2. No, porque es difícil 3. No me interesa 4. Sin opinión NR

65) ¿Se debería enseñar el español en las escuelas?

1. Sí 2. No NR

65.1 ¿Por qué?

1. Sí, porque es muy útil 2. No, porque es difícil 3. Sí, pero es muy difícil 4. No, porque no es útil

66) ¿Se debería enseñar el otomí en las escuelas?

1. Sí 2. No NR

66.1 ¿Por qué?

1. Sí, porque es muy útil 2. No, porque es difícil 3. Sí, pero es muy difícil 4. No, porque no es útil

67) ¿En qué lengua le gustaría que le dieran clases a sus hijos?

1. + Español 2. + Otomí 3. Español y otomí 4. Otra NR

68) ¿En caso de tener problemas con la policía, en que lengua le gustaría que le hablaran?

1. + Español 2. + Otomí 3. Español y otomí 4. Otra NR

68.1 ¿En qué lengua le gustaría que le explicaran sus derechos y obligaciones en un juzgado?

1. + Español 2. + Otomí 3. Español y otomí 4. Otra NR

68.2 ¿Cree que es necesario que exista un intérprete de otomí en el ministerio público?

1. Sí 2. No 3. Algunas veces 4. Con frecuencia NR

68.3 ¿Se ha sentido discriminado por hablar otomí?

1. Sí 2. No 3. Algunas veces 4. Con frecuencia NR

68.4 ¿Se ha sentido discriminado por no hablar otomí?

1. Sí 2. No 3. Me es igual 4. Sin opinión NR

69) ¿Sus padres le han dicho que el otomí es feo o no debe hablarlo?

1. Sí 2. No 3. Algunas veces 4. Con frecuencia NR

69.1 ¿El profesor le ha dicho que el otomí no se debe hablar?

1. Sí 2. No 3. Algunas veces 4. Con frecuencia NR

69.2 ¿Las autoridades le han dicho que el otomí no se debe hablar?

1. Sí 2. No 3. Algunas veces 4. Con frecuencia NR

69.3 ¿El sacerdote o pastor le ha dicho que el otomí no se debe hablar?

1. Sí 2. No 3. Algunas veces 4. Con frecuencia NR

69.4 ¿Personas ajenas a la comunidad le han dicho que el otomí no se debe hablar?

1. Sí 2. No 3. Algunas veces 4. Con frecuencia NR

69.5 ¿Sus compadres le han dicho que el otomí no se debe hablar?

1. Sí 2. No 3. Algunas veces 4. Con frecuencia NR

69.6 ¿Sus amigos le han dicho que el otomí no se debe hablar?

1. Sí 2. No 3. Algunas veces 4. Con frecuencia NR

69.7 ¿Usted cree que el otomí no se debe hablar?

1. Sí 2. No 3. Algunas veces 4. Con frecuencia NR

69.8 ¿Usted cree que el español no se debe hablar?

1. Sí 2. No 3. Algunas veces 4. Con frecuencia NR

70) En manos de quién cree que esta difundir lenguas como el otomí

1. Todos 2. Sus hablantes 3. La familia 4. El Estado NR

Identidad y cultura

71) ¿Cree que las que las costumbres otomíes son distintas a las personas que no son de la comunidad?

1. Sí 2. No 3. Sin opinión NR

71.1 En caso de decir SI, mencione tres aspectos en que son diferentes

1. _____

2. _____

3. _____

72) ¿Qué es lo que más le produce lealtad y sentido de pertenencia? (escoja máximo CUATRO opciones en orden de importancia)

1. La familia: ser miembro de una 2. El pueblo ñhätó: ser parte del pueblo ñhätó de San Andrés Cuexcontitlán 3. La región: ser parte de Toluca 4. El estado: ser de México

5. El país: ser mexicano 6. Otro país: norteamericano 7. El mundo: ser parte de la humanidad 8. En realidad, NO SE SIENTE PARTE de nada

73) ¿Se identifica con la cultura Otomí?

1. Sí, mucho 2. Sí, aunque no digo que soy ñhätó 3. No, porque no es distinta a la cultura mexicana 4. No, porque no es moderna NR

73.1 Puede mencionar tres aspectos que lo identifiquen más con la cultura Otomí

1. _____
2. _____
3. _____

Variación

74) Mencione tres lugares donde sabe que se habla el otomí:

1. _____
2. _____
3. _____

74.1 ¿En dónde cree que se habla mejor el otomí?

0. San Andrés Cuexcontitlán 1. San Cristóbal Huichochitlán 2. San Pablo Autopan
3. Temoaya 4. Son iguales 5. Otro lugar 6. Sin opinión

74.2 ¿Quiénes hablan mejor el otomí?

0. los niños 1. Los jóvenes 2. Los adultos 3. Los ancianos 4. Las mujeres
5. Es igual 6. Sin opinión

75) ¿Cree que el otomí que hablan los jóvenes es distinto al que hablan los ancianos?

0. No 1. Sí 2. Sí, un poco 3. Sí, es muy diferente 4. Es igual 5. Sin opinión

75.1 ¿Cree que ha cambiado el otomí últimamente?

0. No 1. Sí 2. Sí, un poco 3. Sí, es muy diferente 4. Es igual 5. Sin opinión

75.2 ¿Ese cambio ha afectado a las palabras (léxico)?

0. No 1. Sí 2. Sí, un poco 3. Sí, es muy diferente 4. Es igual 5. Sin opinión

75.3 ¿A la pronunciación?

0. No 1. Sí 2. Sí, un poco 3. Sí, es muy diferente 4. Es igual 5. Sin opinión

75.4 ¿A la entonación (tonos)?

0. No 1. Sí 2. Sí, un poco 3. Sí, es muy diferente 4. Es igual 5. Sin opinión

75.5 ¿A la forma de formar las palabras y las oraciones (gramática)?

0. No 1. Sí 2. Sí, un poco 3. Sí, es muy diferente 4. Es igual 5. Sin opinión

75.6 En caso de decir SI, mencione tras aspectos que han producido ese cambio

1 _____

2 _____

3 _____

APÉNDICE CAPÍTULO 3

APÉNDICE 3A. DISTRIBUCIÓN DE LOS PATRONES DE MARCACIÓN EN LA FN DE ACUERDO CON

LOS FACTORES LINGÜÍSTICOS Y EXTRALINGÜÍSTICOS

Factores lingüísticos								
Relación χ^2 , $p > 0.05$	Concordancia				Discordancia			
	P-P		A-P		A-A		P-A	
	F	f	F	f	F	f	F	f
N-A	215	73.9	5	1.7	5	1.7	66	22.7
D-N	1170	75.8	11	0.7	20	1.3	342	22.2
Función χ^2 , $p > 0.05$	Concordancia				Discordancia			
	P-P		A-P		A-A		P-A	
	F	f	F	f	F	f	F	f
O. indirecto	42	68.9	0	0.0	0	0.0	19	31.1
Adjunto	513	72.9	8	1.1	10	1.4	172	24.5
Sujeto	400	76.3	4	0.8	8	1.5	112	21.4
O. directo	396	77.6	4	0.8	7	1.4	103	20.2
Atributo	34	94.4	0	0.0	0	0.0	2	5.6
Categoría EC χ^2 , $p < 0.05$ V de Cramer:	Concordancia				Discordancia			
	P-P		A-P		A-A		P-A	
	F	f	F	f	F	f	F	f
Demostrativo	44	57.9	0	0.0	7	9.2	25	32.9
Numeral	187	68.5	0	0.0	0	0.0	86	31.5
Posesivo	256	72.9	2	0.6	2	0.6	91	25.9
Artículo	642	78.2	3	0.4	1	0.1	175	21.3
Adjetivo	159	86.4	5	2.7	5	2.7	15	8.2
Cuantificador	97	75.2	6	4.7	10	7.8	16	1.2
Animacidad χ^2 , $p > 0.05$	Concordancia				Discordancia			
	P-P		A-P		A-A		P-A	
	F	f	F	f	F	f	F	f
Animado	41	66.1	0	0.0	1	1.6	20	32.3
Inanimado	803	74.6	14	1.3	15	1.4	245	22.7
Humano	541	77.8	2	0.3	9	1.3	143	20.6
Especificidad χ^2 , $p > 0.05$	Concordancia				Discordancia			
	P-P		A-P		A-A		P-A	
	F	f	F	f	F	f	F	f
Específico	426	74.9	3	0.5	12	2.10	128	22.5
Inespecífico	959	75.8	13	1.0	13	1.02	280	22.1

Factores extralingüísticos

Localidad $\chi^2, p<0.05$ V de Cramer: 0.185	Concordancia				Discordancia			
	P-P		A-P		A-A		P-A	
	F	f	F	f	F	f	F	f
San Andrés	763	69.2	10	0.9	22	1.9	308	27.9
Jiquipilco	622	85.1	6	0.8	3	0.4	100	13.7

Bilingüismo $\chi^2, p<0.05$ V de Cramer:	Concordancia				Discordancia			
	P-P		A-P		A-A		P-A	
	F	f	F	f	F	f	F	f
BO	204	41.9	4	0.8	19	3.9	269	55.3
BS	448	75.4	10	1.7	5	0.8	131	22.1
BE	118	100.0	0	0.0	0	0.0	0	0.0
PO	442	98.0	2	0.4	1	0.2	6	1.3
ME	173	98.9	0	0.0	0	0.0	2	1.1

Edad $\chi^2, p<0.05$ V de Cramer: 0.226	Concordancia				Discordancia			
	P-P		A-P		A-A		P-A	
	F	f	F	f	F	f	F	f
G4	396	61.3	4	0.6	18	2.8	228	35.29
G3	374	66.5	10	1.8	6	1.1	172	30.60
G2	442	98.0	2	0.4	1	0.2	6	1.33
G1	173	98.9	0	0.0	0	0.0	2	1.15

Instrucción $\chi^2, p<0.05$ V de Cramer: 0.353	Concordancia				Discordancia			
	P-P		A-P		A-A		P-A	
	F	f	F	f	F	f	F	f
Bajo	513	54.6	13	1.4	24	2.6	390	41.5
Medio	726	97.1	3	0.4	1	0.1	18	2.4
Alto	146	100.0	0	0.0	0	0.0	0	0.0

Género $\chi^2, p<0.05$ V de Cramer: 0.090	Concordancia				Discordancia			
	P-P		A-P		A-A		P-A	
	F	f	F	f	F	f	F	f
Hombres	425	71.2	7	1.2	15	2.5	150	25.12
Mujeres	960	77.6	9	0.7	10	0.8	258	20.85

APÉNDICE 3B. RESULTADOS DE LA MEDICIÓN GOODMAN Y KRUSKAL. MODELO FRASE

NOMINAL

	Sexo	Edad	Instrucción	Bilingüismo	Localidad	Relación	Categoría	Animación	Especificidad	Función
Sexo	K = 2 0	0.01	0.01	0.11	0	0	0	0	0	0
Edad	0.02	K = 4	0.44	0.5	0.09	0.01	0	0	0	0.01
Instrucción	0.01	0.2	K = 3	0.41	0.05	0	0.01	0	0	0
Bilingüismo	0.01	0.35	0.55	K = 3	0.03	0.01	0.01	0	0	0
Localidad	0.11	0.02	0.04	0.02	K = 2	0	0	0	0	0
Relación	0	0	0	0	0	K = 2	0.14	0.02	0.06	0.01
Categoría	0.01	0	0.02	0.02	0.02	0.55	K = 6	0.14	0.56	0.04
Animación	0.01	0	0	0	0.01	0.02	0.04	K = 3	0.08	0.1
Especificidad	0	0	0	0	0	0.06	0.14	0.07	K = 2	0.03
Función	0.01	0.01	0	0.01	0	0.03	0.02	0.21	0.08	K = 5

APÉNDICE 3C. LA CONCORDANCIA DE NÚMERO EN EL NIVEL ORACIONAL DE ACUERDO CON LOS

FACTORES LINGÜÍSTICOS Y EXTRALINGÜÍSTICOS

Relación	Concordancia		Discordancia	
	F	f	F	f
FV-FN	153	73.20	56	26.79
FV-FA	96	73.84	34	26.15
FN-FV	1467	89.17	178	10.82
Categoría EC	Concordancia		Discordancia	
	F	f	F	f
Adjetivo	50	64.93	27	35.06
S. léxico	524	80.61	126	19.38
S. morfológico	1142	90.85	115	9.14
Persona	Concordancia		Discordancia	
	F	f	F	f
3ª Persona	1098	81.69	246	18.30
2ª Persona	42	89.36	5	10.63
1ª Persona	576	97.13	17	2.86
Animacidad	Concordancia		Discordancia	
	F	f	F	f
Animado	9	39.13	14	60.86
Inanimado	109	85.15	19	14.84
Humano	1598	87.17	235	12.82
Especificidad	Concordancia		No concordancia	
	F	f	F	f
Inespecífico	668	81.06	156	18.93
Específico	1048	90.34	112	9.65

Género	Concordancia		Discordancia	
	F	f	F	f
Mujeres	1131	86.07	183	13.92
Hombres	585	87.31	85	12.68
Localidad	Concordancia		Discordancia	
	F	f	F	f
San Andrés Cuexcontitlán	1066	86.10	172	13.89
Jiquipilco el Viejo	650	87.13	96	12.86
Bilingüismo	Concordancia		Discordancia	
	F	f	F	f
Bilingües dominantes en otomí	343	69.57	150	30.42
Bilingües simétricos	510	85.28	88	14.71
Bilingües dominantes en español	115	95.83	5	4.16
H. conocimiento pasivo de otomí	496	96.87	16	3.12
Monolingües den español	252	96.55	9	3.44
Edad	Concordancia		Discordancia	
	F	f	F	f
Generación 4 (+ 55 años)	496	78.11	139	21.88
Generación 3 (35-54 años)	472	81.94	104	18.05
Generación 2 (20-34 años)	496	96.87	16	3.12
Generación 1 (14-19 años)	252	96.55	9	3.44
Instrucción	Concordancia		Discordancia	
	F	f	F	f
Bajo	714	75.87	227	24.12
Medio	841	95.78	37	4.21
Alto	161	97.57	4	2.42
Género	Concordancia		Discordancia	
	F	f	F	f
Mujeres	1131	86.07	183	13.92
Hombres	585	87.31	85	12.68

APÉNDICE 3D. RESULTADOS DE LA MEDICIÓN GOODMAN Y KRUSKAL. MODELO ORACIONAL

	Sexo	Edad	Instrucción	Bilingüismo	Localidad	Relación	Persona	CEC	Animacidad	Especificidad	Sujeto.oracional
Sexo	K = 2	0.01	0.02	0.02	0.11	0	0.01	0	0	0	0
Edad	0.03	K = 4	0.46	0.54	0.1	0	0.04	0	0	0.02	0
Instrucción	0.03	0.2	K = 3	0.44	0.02	0	0.03	0	0	0.01	0
Bilingüismo	0.03	0.35	0.56	K = 3	0.03	0	0.03	0	0	0.01	0
Localidad	0.11	0.02	0.02	0.01	K = 2	0	0.01	0	0	0.02	0
Relación	0	0	0	0	0	K = 3	0.02	0.32	0.11	0	0.25
Persona	0.02	0.02	0.02	0.03	0.02	0.02	K = 3	0.05	0.03	0.34	0.06
CEC	0	0	0	0	0	0.41	0.05	K = 3	0.08	0.01	0.91
Animacidad	0	0	0	0	0.01	0.1	0.04	0.08	K = 3	0.01	0.08
Especificidad	0	0	0	0	0.02	0	0.3	0	0	K = 2	0
Sujeto.oracional	0	0	0	0	0	0.17	0.06	0.8	0.07	0	K = 2

APÉNDICE 3E. LA DISTRIBUCIÓN DE LOS TIPOS DE MARCACIÓN SEGÚN LOS FACTORES LINGÜÍSTICOS Y EXTRALINGÜÍSTICOS POR COMUNIDAD

Jiquipilco					San Andrés Cuexcontitlán				
Relación χ^2 , p>0.05	Concordancia		Discordancia		Relación χ^2 , p>0.05	Concordancia		Discordancia	
	P-P		P-A			P-P		P-A	
V de Cramer:0.001	F	f	F	f	V de Cramer:0.031	F	f	F	f
Nombre-Adj.	117	86.02	19	13.97	Nombre-Adj.	98	67.58	47	32.41
Det.-Nombre	505	86.17	81	13.82	Det.-Nombre	665	71.81	261	28.18
Función χ^2 , p<0.05	Concordancia		Discordancia		Función χ^2 , p<0.05	Concordancia		Discordancia	
	P-P		P-A			P-P		P-A	
V de Cramer:0.112	F	f	F	f	V de Cramer:0.124	F	f	F	f
Sujeto	172	81.13	40	18.87	O. indirecto	21	55.26	17	44.73
O. directo	176	86.70	27	13.30	O. directo	220	74.32	76	34.54
Adjunto	234	88.30	31	11.70	Adjunto	279	66.42	141	33.57
O. indirecto	21	91.30	2	8.70	Sujeto	228	76.00	72	24.00
Atributo	19	100.0	0	0.00	Atributo	15	88.23	2	11.76
Categoría EC χ^2 , p<0.05	Concordancia		Discordancia		Categoría EC χ^2 , p<0.05	Concordancia		Discordancia	
	P-P		P-A			P-P		P-A	
V de Cramer:0.178	F	f	F	f	V de Cramer:0.203	F	f	F	f
Posesivo	82	73.21	30	26.78	Demostrativo	24	51.06	23	48.94
Numeral	104	85.24	18	14.75	Numeral	83	54.96	68	45.04
Artículo	277	87.65	39	12.34	Posesivo	174	74.04	61	25.96
Cuantificador	50	89.28	6	10.71	Artículo	365	72.85	136	21.15
Demostrativo	20	90.90	2	9.10	Cuantificador	47	82.45	10	17.55
Adjetivo	89	94.68	5	5.31	Adjetivo	70	87.5	10	12.5
Animacidad χ^2 , p<0.05	Concordancia		Discordancia		Animacidad χ^2 , p<0.05	Concordancia		Discordancia	
	P-P		P-A			P-P		P-A	
V de Cramer:0.108	F	f	F	f	V de Cramer:0.106	F	f	F	f
Inanimado	230	81.56	52	18.44	Animado	18	36.73	31	63.27
Animado	10	83.33	2	16.66	Inanimado	421	67.90	199	32.09
Humano	382	89.25	46	10.75	Humano	311	77.36	91	22.63
Especificidad χ^2 , p<0.05	Concordancia		Discordancia		Especificidad χ^2 , p>0.05	Concordancia		Discordancia	
	P-P		P-A			P-P		P-A	
V de Cramer:0.098	F	f	F	f	V de Cramer:0.053	F	f	F	f
Inespecífico	459	88.26	61	11.74	Inespecífico	500	69.54	219	30.45
Específico	163	80.69	39	19.30	Específico	263	74.71	89	25.28

Jiquipilco el Viejo					San Andrés Cuexcontitlán						
Género		Concordancia		Discordancia		Género		Concordancia		Discordancia	
$\chi^2, p>0.05$		P-P		P-A		$\chi^2, p>0.05$		P-P		P-A	
V de Cramer: 0.065		F	f	F	f	V de Cramer: 0.020		F	f	F	f
Mujeres		532	85.25	92	14.75	Hombres		335	70.23	142	29.76
Hombres		90	91.83	8	8.16	Mujeres		428	72.06	166	27.94
Edad		Concordancia		Discordancia		Edad		Concordancia		Discordancia	
$\chi^2, p<0.05$		P-P		P-A		$\chi^2, p<0.05$		P-P		P-A	
V de Cramer: 0.238		F	f	F	f	V de Cramer: 0.474		F	f	F	f
Generación 4		227	81.07	53	18.92	Generación 4		169	49.12	175	50.87
Generación 3		155	79.08	41	20.92	Generación 3		219	62.57	131	37.42
Generación 2		240	97.56	6	2.44	Generación 2		202	100.0	0	0.00
Generación 1						Generación 1		173	98.85	2	1.15
Bilingüismo		Concordancia		Discordancia		Bilingüismo		Concordancia		Discordancia	
$\chi^2, p<0.05$		P-P		P-A		$\chi^2, p<0.05$		P-P		P-A	
V de Cramer: 0.548		F	f	F	f	V de Cramer: 0.554		F	f	F	f
BO		60	46.15	70	53.84	BO		144	41.98	199	58.01
BS		272	91.89	24	8.10	BS		176	62.19	107	37.80
BE		50	100.0	0	0.00	BE		68	100.0	0	0.00
PO		240	97.56	6	2.44	PO		202	100.0	0	0.00
ME						ME		173	98.85	2	1.15
Instrucción		Concordancia		Discordancia		Instrucción		Concordancia		Discordancia	
$\chi^2, p<0.05$		P-P		P-A		$\chi^2, p<0.05$		P-P		P-A	
V de Cramer: 0.377		F	f	F	f	V de Cramer: 0.527		F	f	F	f
Bajo		193	69.67	84	30.32	Bajo		320	51.11	306	48.88
Medio		379	95.95	16	4.05	Medio		347	99.42	2	0.58
Alto		50	100.0	0	0.00	Alto		96	100.0	0	0.00

APÉNDICE 3F. RESULTADOS DE LA MEDICIÓN GOODMAN Y KRUSKAL. MODELO FRÁSTICO PARA

JV Y SAC

	Sexo	Edad	Instrucción	Bilingüismo	Relación	Categoría	Animación	Especificidad	Función
Sexo	K = 2 0	0.03	0.1	0.01	0	0	0	0	0
Edad	0.06	K = 3 0	0.38	0.47	0	0	0	0	0
Instrucción	0.1	0.18	K = 2 0	0.26	0.01	0.01	0.01	0.01	0.01
Bilingüismo	0.04	0.39	0.57	K = 3 0	0.02	0.03	0.02	0.02	0.01
Relación	0	0	0.01	0	K = 2 0	0.15	0.02	0.05	0.01
Categoría	0.02	0.01	0.04	0.04	0.56	K = 6 0	0.11	0.56	0.05
Animación	0	0	0.01	0.01	0.02	0.03	K = 3 0	0.06	0.1
Especificidad	0	0	0.01	0.01	0.05	0.12	0.06	K = 2 0	0.03
Función	0	0.02	0.02	0.02	0.05	0.04	0.21	0.09	K = 5 0

	Sexo	Edad	Instrucción	Bilingüismo	Relación	Categoría	Animación	Especificidad	Función
Sexo	K = 2 0	0	0.01	0.01	0	0	0.01	0	0.01
Edad	0	K = 3 0	0.79	0.63	0.01	0.01	0.01	0.01	0.02
Instrucción	0.01	0.4	K = 2 0	0.56	0.01	0.01	0.01	0.01	0.01
Bilingüismo	0.02	0.6	1	K = 3 0	0.02	0.01	0.01	0.01	0.01
Relación	0	0	0.01	0.01	K = 2 0	0.13	0.01	0.06	0
Categoría	0.02	0.03	0.07	0.04	0.53	K = 6 0	0.16	0.56	0.04
Animación	0.01	0.01	0.02	0.01	0.02	0.06	K = 3 0	0.08	0.1
Especificidad	0	0	0.01	0	0.06	0.17	0.07	K = 2 0	0.03
Función	0.02	0.02	0.02	0.02	0.03	0.02	0.21	0.08	K = 5 0

APÉNDICE 3G. DISTRIBUCIÓN DE LA CONCORDANCIA PLURAL EN EL NIVEL ORACIONAL DE

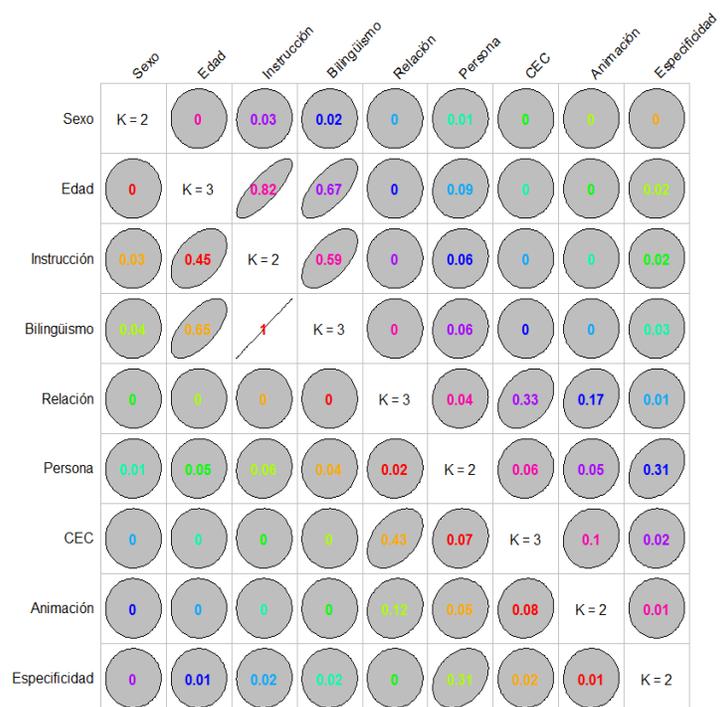
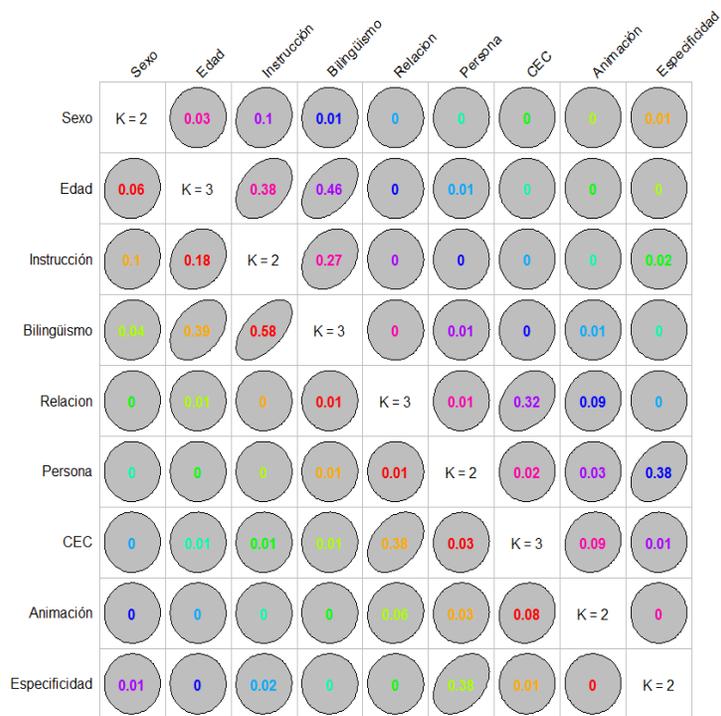
ACUERDO CON LOS FACTORES LINGÜÍSTICOS Y EXTRALINGÜÍSTICOS POR COMUNIDAD

Jiquipilco el Viejo					San Andrés Cuexcontitlán				
Relación χ^2 , p<0.05 V de Cramer: 0.106	Concordancia		Discordancia		Relación χ^2 , p<0.05 V de Cramer: 0.219	Concordancia		Discordancia	
	F	f	F	f		F	f	F	f
FV-FN	65	79.26	17	20.73	FV-FN	88	69.29	39	30.70
FV-FA	34	79.06	9	20.93	FV-FA	62	71.26	25	28.73
FN-FV	551	88.72	70	11.27	FN-FV	916	89.45	108	10.54
Persona χ^2 , p<0.05 V de Cramer: 0.194	Concordancia		Discordancia		Persona χ^2 , p<0.05 V de Cramer: 0.228	Concordancia		Discordancia	
	F	f	F	f		F	f	F	f
3ª Persona	461	83.81	89	16.18	3ª Persona	637	80.22	157	19.77
2ª Persona	2	50.00	2	50.00	2ª Persona	40	93.02	3	6.97
1ª Persona	187	97.39	5	2.60	1ª Persona	389	97.00	12	3.00
Categoría EC χ^2 , p<0.05 V de Cramer: 0.148	Concordancia		Discordancia		Categoría EC χ^2 , p<0.05 V de Cramer: 0.209	Concordancia		Discordancia	
	F	f	F	f		F	f	F	f
Adjetivo	14	66.66	7	33.33	Adjetivo	36	64.28	20	35.71
S. léxico	206	82.73	43	17.26	S. léxico	318	79.30	83	20.69
S. morfológico	430	90.33	46	9.66	S. morfológico	712	91.16	69	8.83
Animacidad χ^2 , p>0.05 V de Cramer: 0.019	Concordancia		Discordancia		Animacidad χ^2 , p<0.05 V de Cramer: 0.194	Concordancia		Discordancia	
	F	f	F	f		F	f	F	f
Animado	1	100.0	0	0.00	Animado	8	36.36	14	63.63
Inanimado	47	85.45	8	14.54	Inanimado	62	84.93	11	15.06
Humano	602	87.24	88	12.75	Humano	996	87.13	147	12.86
Especificidad χ^2 , p<0.05 V de Cramer: 0.139	Concordancia		Discordancia		Especificidad χ^2 , p<0.05 V de Cramer: 0.136	Concordancia		Discordancia	
	F	f	F	f		F	f	F	f
Inespecífico	318	82.59	67	17.40	Inespecífico	350	79.72	89	20.27
Específico	332	91.96	29	8.03	Específico	716	89.61	83	10.38

Jiquipilco el Viejo					San Andrés Cuexcontitlán				
Género $\chi^2, p>0.05$ V de Cramer: 0.025	Concordancia		Discordancia		Género $\chi^2, p>0.05$ V de Cramer: 0.043	Concordancia		Discordancia	
	F	f	F	f		F	f	F	f
Hombres	85	85.00	15	15.00	Mujeres	566	84.73	102	15.26
Mujeres	565	87.46	81	12.53	Hombres	500	87.71	70	12.28
Edad $\chi^2, p<0.05$ V de Cramer: 0.154	Concordancia		Discordancia		Edad $\chi^2, p<0.05$ V de Cramer: 0.306	Concordancia		Discordancia	
	F	f	F	f		F	f	F	f
Generación 4	248	83.78	48	16.21	Generación 4	248	73.15	91	26.84
Generación 3	166	83.00	34	17.00	Generación 3	306	81.38	70	18.61
Generación 2	236	94.40	14	5.60	Generación 2	260	99.23	2	0.76
Generación 1					Generación 1	252	96.55	9	3.44
Instrucción $\chi^2, p<0.05$ V de Cramer: 0.244	Concordancia		Discordancia		Instrucción $\chi^2, p<0.05$ V de Cramer: 0.324	Concordancia		Discordancia	
	F	f	F	f		F	f	F	f
Bajo	228	77.02	68	22.97	Bajo	486	75.34	159	24.65
Medio	375	93.75	25	6.25	Medio	466	97.48	12	2.51
Alto	47	94.00	3	6.00	Alto	114	99.13	1	0.69
Bilingüismo $\chi^2, p<0.05$ V de Cramer: 0.250	Concordancia		Discordancia		Bilingüismo $\chi^2, p<0.05$ V de Cramer: 0.356	Concordancia		Discordancia	
	F	f	F	f		F	f	F	f
BO	104	71.23	42	28.76	BO	239	68.87	108	31.12
BS	263	87.66	37	12.33	BS	247	82.88	51	17.11
BE	47	94.00	3	6.00	BE	68	97.14	2	2.85
PO	236	94.40	14	5.60	PO	260	99.23	2	0.76
ME					ME	252	96.55	9	3.44

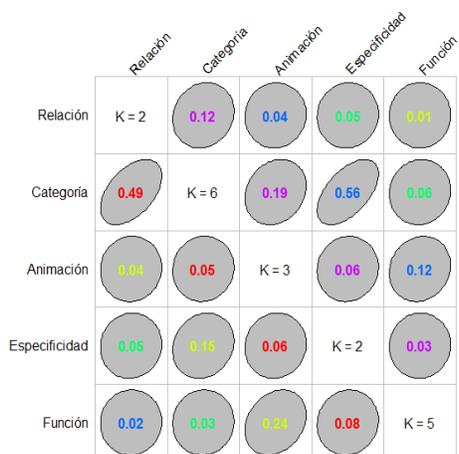
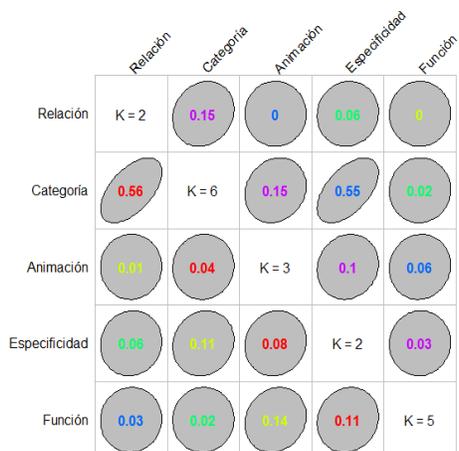
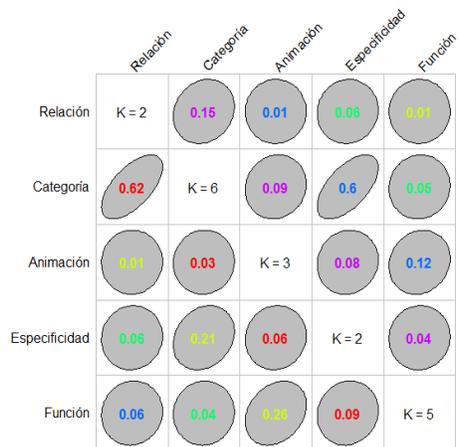
APÉNDICE 3H. RESULTADOS DE LA MEDICIÓN GOODMAN Y KRUSKAL. MODELO ORACIONAL

PARA JV Y SAC



APÉNDICE 3I. RESULTADOS DE LA MEDICIÓN GOODMAN Y KRUSKAL. MODELO FRÁSTICO POR

TIPO DE BILINGÜISMO: BO, BS, BE



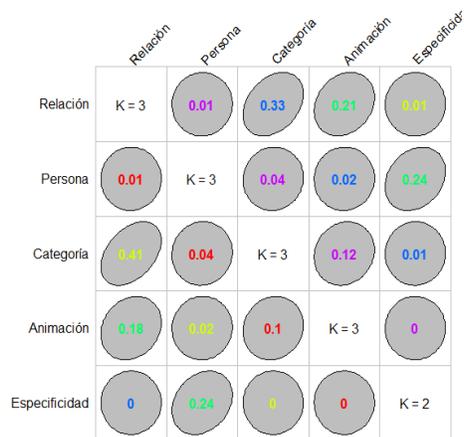
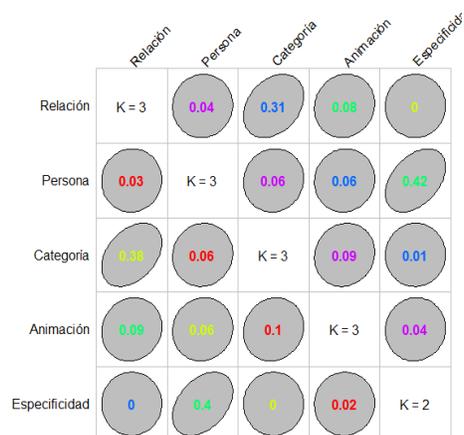
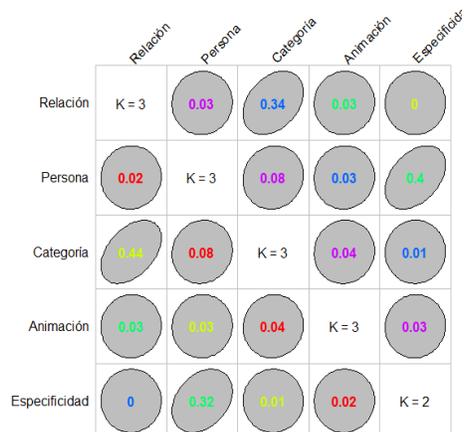
APÉNDICE 3J. RESULTADOS DE LA PRUEBA MULTIMODELO. MODELO FRÁSTICO POR TIPO DE BILINGÜISMO

Bilingües dominantes en otomí								
Factores	IPC	F	E	R	A	CEC	AIC	Var
Wi	0.986	0.550	0.322	0.278	0.125	0.119		
1	X	X					640.434	0.028
2							640.734	0.015
Bilingües simétricos								
Factores	IPC	F	E	R	A	CEC	AIC	Var
Wi	1.000	0.041	0.378	0.318	0.280	0.104		
Modelo 1	X						458.472	0.266
Modelo 2	X		X				459.821	0.267
Hablantes dominantes en español								
Factores	IPC	F	E	R	A	CEC	AIC	Var
Wi	0.996	0.086	0.269	0.276	0.401	0.044		
Modelo 1	X						78.415	0.158
Modelo 2	X				X		79.139	0.015

Tabla 3.X Modelos de la alternancia. Resultados de la selección de modelos basados en información teórica e inferencia multimodelo para la concordancia de número en los bilingües dominantes en otomí (BO), bilingües simétricos (BS) y hablantes dominantes en español (Esp). Detallamos el conjunto de modelos $\Delta AICc < 2$. Con "X" indicamos la inclusión de la variable en cada modelo. Las variables explicativas fueron: pérdida de consonantes (IPC), función sintáctica (F), especificidad (E), tipo de relación (R), animacidad (A) y categoría del elemento concordante (CEC).

APÉNDICE 3K. RESULTADOS DE LA MEDICIÓN GOODMAN Y KRUSKAL. MODELO ORACIONAL

POR TIPO DE BILINGÜISMO: BO, BS, BE.



APÉNDICE 3L. RESULTADOS DE LA PRUEBA MULTIMODELO. MODELO ORACIONAL POR TIPO DE BILINGÜISMO

Bilingües dominantes en otomí								
Modelo	PG	A	CEC	E	R	IPC	AIC	Var
Wi	1.000	0.949	0.945	0.292	0.311	0.269		
1	X	X	X				489.59	0.2153
2	X	X	X	X			491.35	0.2158
Bilingües simétricos								
Modelo	PG	A	CEC	E	R	IPC	AIC	Var
	1.000	0.352	0.880	0.299	0.651	0.990		
1	X		X		X	X	410.39	0.2111
2	X		X			X	411.73	0.2002
3	X	X	X		X	X	411.78	0.2165
4	X		X	X	X	X	411.96	0.2120
Hablantes dominantes en español								
Modelo	PG	A	CEC	E	R	IPC	AIC	Var
	0.620	0.958	0.315	0.880	0.792	0.999		
1	X	X		X	X	X	219.77	0.2323
2		X		X	X	X	220.34	0.2146

Tabla 3.X Modelos de la alternancia. Resultados de la selección de modelos basados en información teórica e inferencia multimodelo para la concordancia de número en los bilingües dominantes en otomí (BO), bilingües simétricos (BS) y hablantes dominantes en español (Esp). Detallamos el conjunto de modelos $\Delta AICc < 2$. Con "X" indicamos la inclusión de la variable en cada modelo. Las variables explicativas fueron: persona gramatical (PG), categoría del elemento concordante (CEC), animacidad (A), especificidad (E), tipo de relación (R) y pérdida de consonantes (IPC).

APÉNDICE CAPÍTULO 4

APÉNDICE 4A. DISTRIBUCIÓN DE LOS PATRONES DE MARCACIÓN EN LOS NOMBRES PATRIMONIALES SEGÚN LAS VARIABLES LINGÜÍSTICAS Y SOCIOLINGÜÍSTICAS

Determinantes $\chi^2=136.520^a$, gl = 3, p = 0.000 V de Cramer = 0.323, Sig. Aproximada: 0.000	Patrón 1 (ra / ya + N)		Patrón 2 (N)	
	F	%	F	%
FN c / Demostrativos	283	91.3	27	8.7
FN c / Indefinidos	95	80.5	23	19.5
FN s / Determinantes	506	59.3	348	40.7
FN c / Numerales	6	25.0	18	75.0
Nombre propio/común $\chi^2 = 4.177$, gl = 1, p = 0.040 V de Cramer = 0.053, Sig. Aproximada: 0.040	Patrón 1 (ra / ya + N)		Patrón 2 (N)	
	F	%	F	%
Comunes	845	68.8	383	31.2
Propios	45	57.7	33	42.3
Concreción $\chi^2=176.772^a$, gl = 1, p = 0.000 V de Cramer = 0.368, Sig. Aproximada: 0.000	Patrón 1 (ra / ya + N)		Patrón 2 (N)	
	F	f	F	f
Concreto	753	78.4	207	21.6
Abstracto	137	39.6	209	60.4
Animacidad $\chi^2=21.186^a$, gl = 2, p = 0.000 V de Cramer = 0.127, Sig. Aproximada: 0.000	Patrón 1 (ra / ya + N)		Patrón 2 (N)	
	F	f	F	f
Animado	61	92.4	5	7.6
Humano	252	70.0	108	30.0
Inanimado	577	65.6	303	34.4
(Dis)continuidad $\chi^2=41.615^a$, gl = 1, p = 0.000 V de Cramer = 0.179, Sig. Aproximada: 0.000	Patrón 1 (ra / ya + N)		Patrón 2 (N)	
	F	%	F	%
Discontinuo	672	73.6	241	26.4
Continuo	218	55.5	175	44.5
Función $\chi^2=33.927^a$, gl = 5, p = 0.000 V de Cramer = 0.528, Sig. Aproximada: 0.000	Patrón 1 (ra / ya + N)		Patrón 2 (N)	
	F	%	F	%
Objeto Secundario	20	90.9	2	9.1
Sujeto	269	89.1	33	10.9
Objeto Primario	431	83.2	87	16.8
Compl. nominal	17	50.0	17	50.0
Adverbial	131	43.4	171	56.6
Predicativo	40	17.2	106	82.8
Constituyente $\chi^2=80.671^a$, gl = 1, p = 0.000 V de Cramer = 0.249, Sig. Aproximada: 0.000	Patrón 1 (ra / ya + N)		Patrón 2 (N)	
	F	%	F	%
Núcleo nominal	830	72.5	315	27.5
Término de preposición	60	37.3	101	62.7
Referencialidad $\chi^2=78.256$, gl = 1, p = 0.000 V de Cramer = 0.243, Sig. Aproximada: 0.000	Patrón 1 (ra / ya + N)		Patrón 2 (N)	
	F	%	F	%
Frasas nominales referenciales	598	77.7	172	22.3
Frasas nominales atributivas	292	54.5	244	45.5

Localidad $\chi^2=14.727^a$, gl = 1, p = 0.000 V de Cramer = 0.106, Sig. Aproximada: 0.000	Patrón 1 (ra / ya + N)		Patrón 2 (N)	
	F	%	F	%
San Andrés Cuexcontitlán	395	63.0	232	37.0
Jiquipilco el Viejo	495	72.9	184	27.1
Edad $\chi^2=1.322^a$, gl = 1, p = 0.254 V de Cramer = 0.032, Sig. Aproximada: 0.250	Patrón 1 (ra / ya + N)		Patrón 2 (N)	
	F	%	F	%
Generación 3	383	69.9	165	30.1
Generación 4	667	66.9	251	33.1
Instrucción $\chi^2=1.849^a$, gl = 2, p = 0.397 V de Cramer = 0.038, Sig. Aproximada: 0.397	Patrón 1 (ra / ya + N)		Patrón 2 (N)	
	F	%	F	%
Alto	28	75.7	9	24.3
Bajo	781	68.3	362	31.7
Medio	81	64.3	45	35.7
Género $\chi^2=7.818^a$, gl = 1, p = 0.006 V de Cramer = 0.077, Sig. Aproximada: 0.006	Patrón 1 (ra / ya + N)		Patrón 2 (N)	
	F	%	F	%
Mujeres	673	70.3	284	29.7
Hombres	217	62.2	132	37.8
Bilingüismo $\chi^2=1.880^a$, gl = 2, p = 0.391 V de Cramer = 0.038, Sig. Aproximada: 0.391	Patrón 1 (ra / ya + N)		Patrón 2 (N)	
	F	%	F	%
Bil. +Esp	52	73.2	19	26.8
Bil. Simétricos	460	69.1	206	30.9
Bil. +Otomí	378	66.4	191	33.6

APÉNDICE 4B. RESULTADOS DE LA MEDICIÓN GOODMAN Y KRUSKAL. NOMBRES

PATRIMONIALES

	Sexo	Edad	Instrucción	Bilingüismo	Localidad	Función	Constituyente	Tipo	Propio	Animación	Contable	Determinantes	Referencialidad
Sexo	K = 3	0.01	0.02	0.05	0.3	0	0	0.01	0	0	0.01	0.01	0.01
Edad	0	K = 2	0.01	0.16	0	0	0.01	0	0	0	0	0	0
Instrucción	0.02	0.01	K = 2	0.09	0.02	0	0	0	0	0	0	0	0
Bilingüismo	0.09	0.2	0.44	K = 3	0.08	0	0	0	0.01	0	0.01	0	0
Localidad	0.29	0	0.02	0.07	K = 2	0	0	0	0	0	0.01	0.01	0.02
Función	0.01	0.01	0.01	0.01	0.02	K = 6	0.32	0.18	0.04	0.2	0.02	0.06	0.17
Constituyente	0	0.01	0	0	0	0.11	K = 2	0	0.01	0.01	0	0.02	0
Tipo	0.01	0	0	0	0	0.04	0	K = 2	0.02	0.02	0.03	0.01	0.11
Propio	0	0	0	0.01	0	0.01	0.01	0.02	K = 2	0.03	0	0	0.04
Animación	0	0	0	0	0	0.07	0.02	0.04	0.05	K = 3	0.04	0	0.01
Contable	0.01	0	0	0	0.01	0	0	0.03	0	0.02	K = 2	0.02	0.01
Determinantes	0.01	0.01	0.01	0	0.01	0.02	0.03	0.04	0	0	0.03	K = 4	0.11
Referencialidad	0.01	0	0	0	0.02	0.03	0	0.11	0.04	0.01	0.01	0.08	K = 2

APÉNDICE 4C. DISTRIBUCIÓN DE LOS PATRONES DE MARCACIÓN EN LOS PRÉSTAMOS SEGÚN LOS

FACTORES LINGÜÍSTICOS Y SOCIALES

Determinantes $\chi^2=282.802^a$, gl = 3, p = 0.000 V de Cramer = 0.472, Sig. aproximada: 0.000	Patrón 1 (ra / ya + N)		Patrón 2 (N)	
	F	%	F	%
FN c / Demostrativos	297	91.4	28	8.6
FN c / Indefinidos	120	90.2	13	9.8
FN s / Determinantes	361	46.3	418	53.7
FN c / Numerales	2	6.7	28	93.3
Propio / Común $\chi^2 = 44.477$, gl = 1, p = 0.000 V de Cramer = 0.184, Sig. aproximada: 0.000	Patrón 1 (ra / ya + N)		Patrón 2 (N)	
	F	%	F	%
Nombre común	736	64.61	403	35.38
Nombre propio	44	34.37	84	65.62
Concreción $\chi^2=34.212^a$, gl = 1, p = 0.000 V de Cramer = 0.164, Sig. aproximada: 0.000	Patrón 1 (ra / ya + N)		Patrón 2 (N)	
	F	F	F	F
Concreto	565	67.3	275	32.7
Abstracto	215	50.4	212	49.6
Animacidad $\chi^2=21.190^a$, gl = 2, p = 0.000 V de Cramer = 0.129, Sig. aproximada: 0.000	Patrón 1 (ra / ya + N)		Patrón 2 (N)	
	F	f	F	F
Animado	30	76.9	9	23.1
Humano	211	71.3	85	28.7
Inanimado	539	57.8	393	42.2
Contable $\chi^2=14.207^a$, gl = 1, p = 0.000 V de Cramer = 0.106, Sig. Aproximada: 0.000	Patrón 1 (ra / ya + N)		Patrón 2 (N)	
	F	%	F	%
Contable	573	65.0	309	35.0
No contable	207	53.8	178	46.2
Función $\chi^2=334.744^a$ gl = 5, p = 0.000 V de Cramer = 0.514, Sig. Aproximada: 0.000	Patrón 1 (ra / ya + N)		Patrón 2 (N)	
	F	%	F	%
Sujeto	219	94.4	13	5.6
Objeto Primario	343	78.9	92	21.1
Objeto Secundario	11	68.7	5	31.3
Compl. Nominal	49	45.0	60	55.0
Adverbial	129	33.9	252	66.1
Predicativo	29	30.9	65	69.1
Constituyente $\chi^2=162.505^a$ gl = 1, p = 0.000 V de Cramer = 0.358, Sig. Aproximada: 0.000	Patrón 1 (ra / ya + N)		Patrón 2 (N)	
	F	%	F	%
Núcleo nominal	712	70.3	301	29.7
Término de preposición	68	26.8	186	73.2
Referencialidad $\chi^2 = 75.74$, gl = 1, p = 0.000 V de Cramer = 0.242, Sig. Aproximada: 0.000	Patrón 1 (ra / ya + N)		Patrón 2 (N)	
	F	%	F	%
Frasas nominales referenciales	534	71.48	213	28.51
Frasas nominales atributivas	246	47.30	274	52.69

Localidad	Patrón 1 (ra / ya + N)		Patrón 2 (N)	
	F	%	F	%
$\chi^2=2.608^a$, gl = 1, p = 0.120				
V de Cramer = 0.045, Sig. Aproximada: 0.106				
San Andrés Cuexcontitlán	522	60.1	347	39.9
Jiquipilco el Viejo	258	64.8	140	35.2
Edad	Patrón 1 (ra / ya + N)		Patrón 2 (N)	
	F	%	F	%
$\chi^2=5.551^a$, gl = 1, p = 0.021				
V de Cramer = 0.066, Sig. Aproximada: 0.018				
Generación 3	399	64.9	216	35.1
Generación 4	381	58.4	271	41.6
Instrucción	Patrón 1 (ra / ya + N)		Patrón 2 (N)	
	F	%	F	%
$\chi^2=3.220^a$, gl = 2, p = 0.180				
V de Cramer = 0.050, Sig. Aproximada: 0.200				
Alto	15	78.9	4	21.1
Bajo	703	61.7	437	38.3
Medio	62	57.4	46	42.6
Género	Patrón 1 (ra / ya + N)		Patrón 2 (N)	
	F	%	F	%
$\chi^2=.927^a$, gl = 2, p = 0.347				
V de Cramer = 0.027, Sig. Aproximada: 0.336				
Hombres	324	63.2	189	36.8
Mujeres	456	60.5	298	39.5
Bilingüismo	Patrón 1 (ra / ya + N)		Patrón 2 (N)	
	F	%	F	%
$\chi^2=7.646^a$, gl = 2, p = 0.019				
V de Cramer = 0.078, Sig. Aproximada: 0.022				
Bil. +Esp	39	75.0	13	25.0
Bil. +Otomí	394	63.5	226	36.5
Bil. Simétricos	347	58.3	248	41.7

APÉNDICE 4D. RESULTADOS DE LA MEDICIÓN GOODMAN Y KRUSKAL. PRÉSTAMOS NOMINALES

	Sexo	Edad	Instrucción	Bilingüismo	Localidad	Función	Constituyente	Tipo	Propio	Animación	Contable	Determinantes	Referencialidad
Sexo	K = 3	0.02	0.08	0.01	0.14	0	0.01	0	0	0	0	0	0
Edad	0.02	K = 2	0	0.22	0	0	0	0.01	0.01	0.01	0	0	0
Instrucción	0.08	0	K = 2	0.08	0.04	0	0	0	0	0	0	0	0
Bilingüismo	0.06	0.28	0.43	K = 3	0.03	0	0.01	0.01	0	0	0	0	0
Localidad	0.14	0	0.04	0.02	K = 2	0	0.02	0	0	0	0	0	0
Función	0.01	0.01	0	0.01	0.02	K = 6	0.3	0.01	0.07	0.2	0.06	0.06	0.1
Constituyente	0.01	0	0	0.01	0.02	0.12	K = 2	0.01	0.01	0.04	0	0.02	0.01
Tipo	0	0.01	0	0.01	0	0	0.01	K = 2	0.04	0.02	0.06	0	0.08
Propio	0	0.01	0	0	0	0.01	0.01	0.04	K = 2	0.04	0.02	0.02	0.07
Animación	0	0.02	0.01	0	0	0.06	0.04	0.04	0.05	K = 3	0.09	0	0.03
Contable	0	0	0	0	0	0.01	0	0.06	0.02	0.07	K = 2	0.01	0.04
Determinantes	0	0	0	0	0.01	0.02	0.03	0.03	0.03	0	0.03	K = 4	0.08
Referencialidad	0	0	0	0	0	0.02	0.01	0.08	0.07	0.02	0.04	0.05	K = 2

APÉNDICE 4E. RESULTADOS DE LA PRUEBA MULTIMODELO. NOMBRES PATRIMONIALES Y

PRÉSTAMOS

atrimoniales															
M	D	F	TC	C/A	C/D	P/C	L	I	G	B	E	R	A	AIC	Var
Wi	1.000	1.000	1.000	0.999	0.999	0.990	0.557	0.475	0.435	0.415	0.384	0.299	0.135		
1	X	X	X	X	X	X	X	X		X				1121.229	2.660471e-02
2	X	X	X	X	X	X	X							1121.293	2.575909e-02
3	X	X	X	X	X	X	X	X	X	X				1121.439	2.394310e-02
4	X	X	X	X	X	X								1121.543	2.273065e-02
5	X	X	X	X	X	X	X				X			1121.626	2.181425e-02
Préstamos															
M	D	F	TC	C/A	C/D	P/C	L	I	G	B	E	R	A	AIC	Var
Wi	1.000	1.000	0.999	1.000	0.328	0.999	0.448	0.720	0.936	0.562	0.451	0.993	0.135	1026.843	1.069206e-01
1	X	X	X	X		X	X	X	X	X		X		1027.833	6.517166e-02
2	X	X	X	X		X		X	X	X		X		1028.313	5.125318e-02
3	X	X	X	X		X	X	X	X	X	X	X		1028.339	5.060233e-02
4	X	X	X	X	X	X	X	X	X	X		X		1028.565	4.518552e-02
5	X	X	X	X		X		X	X		X	X		1028.748	4.123928e-02

Modelos de la alternancia. Resultados de la selección de modelos basados en información teórica e inferencia multimodelo para la marcación de número en los nombres patrimoniales y préstamos del español. Se detalla el conjunto de modelos $\Delta AIC_c < 2$. Con "X" se indica la inclusión de la variable en cada modelo. Las variables explicativas fueron: presencia de determinantes que codifican el número (D), función sintáctica de la frase nominal (F), tipo de constituyente (TC), la distinción entre nombre concreto/abstracto (C/A), continuo/discontinuo (C/D), la distinción propia/común (P/C), localidad (L), instrucción (I), género (G), bilingüismo (B), edad (E), referencialidad (R), animacidad (A).

APÉNDICE 4F. RESULTADOS DE LA MEDICIÓN GOODMAN Y KRUSKAL. ÍNDICE DE INTEGRACIÓN

LÉXICA

	Sexo	Edad	Instrucción	Bilingüismo	Localidad	Función	Constituyente	Tipo	Propio	Animación	Contable	Determinantes	Referencialidad	Índice_lexico
Sexo	K = 2	0.01	0.04	0.02	0.22	0	0.01	0	0	0	0	0	0	0
Edad	0.01	K = 2	0.01	0.19	0	0	0	0	0	0	0	0	0	0
Instrucción	0.04	0.01	K = 2	0.08	0.03	0	0	0	0	0	0	0	0	0
Bilingüismo	0.07	0.23	0.44	K = 3	0.05	0	0	0	0	0	0	0	0	0
Localidad	0.22	0	0.03	0.04	K = 2	0.01	0.01	0	0	0	0.01	0	0.01	0
Función	0.01	0.01	0	0	0.02	K = 6	0.31	0.07	0.05	0.2	0.03	0.06	0.13	0.01
Constituyente	0.01	0	0	0	0.01	0.11	K = 2	0	0.01	0.02	0	0.02	0	0
Tipo	0	0	0	0	0	0.02	0	K = 2	0.03	0.02	0.05	0	0.09	0.01
Propio	0	0	0	0	0	0.01	0.01	0.03	K = 2	0.03	0.01	0	0.06	0.01
Animación	0	0.01	0	0	0	0.06	0.03	0.04	0.04	K = 3	0.06	0	0.02	0.01
Contable	0	0	0	0	0.01	0	0	0.05	0.01	0.04	K = 2	0.02	0.02	0.01
Determinantes	0	0	0	0	0.01	0.02	0.03	0.03	0.01	0	0.03	K = 4	0.1	0
Referencialidad	0	0	0	0	0.01	0.02	0	0.09	0.06	0.02	0.02	0.06	K = 2	0
Índice_lexico	0.28	0.15	0.16	0.15	0.26	0.26	0.2	0.66	0.57	0.61	0.62	0.16	0.3	K = 196

APÉNDICE 4G. RESULTADOS DE LA MEDICIÓN GOODMAN Y KRUSKAL. CLASIFICACIÓN

CATEGÓRICA DEL LÉXICO

	Sexo	Edad	Instrucción	Bilingüismo	Localidad	Función	Constituyente	Tipo	Propio	Animación	Contable	Determinantes	Referencialidad	Clasificación
Sexo	K = 2	0.01	0.04	0.02	0.22	0	0.01	0	0	0	0	0	0	0.01
Edad	0.01	K = 2	0.01	0.19	0	0	0	0	0	0	0	0	0	0
Instrucción	0.04	0.01	K = 2	0.08	0.03	0	0	0	0	0	0	0	0	0
Bilingüismo	0.07	0.23	0.44	K = 3	0.05	0	0	0	0	0	0	0	0	0
Localidad	0.22	0	0.03	0.04	K = 2	0.01	0.01	0	0	0	0.01	0	0.01	0.03
Función	0.01	0.01	0	0	0.02	K = 6	0.31	0.07	0.05	0.2	0.03	0.06	0.13	0.02
Constituyente	0.01	0	0	0	0.01	0.11	K = 2	0	0.01	0.02	0	0.02	0	0.01
Tipo	0	0	0	0	0	0.02	0	K = 2	0.03	0.02	0.05	0	0.09	0
Propio	0	0	0	0	0	0.01	0.01	0.03	K = 2	0.03	0.01	0	0.06	0
Animación	0	0.01	0	0	0	0.06	0.03	0.04	0.04	K = 3	0.06	0	0.02	0.01
Contable	0	0	0	0	0.01	0	0	0.05	0.01	0.04	K = 2	0.02	0.02	0
Determinantes	0	0	0	0	0.01	0.02	0.03	0.03	0.01	0	0.03	K = 4	0.1	0
Referencialidad	0	0	0	0	0.01	0.02	0	0.09	0.06	0.02	0.02	0.06	K = 2	0
Clasificación	0.02	0.01	0	0	0.05	0.01	0.01	0.01	0.01	0.01	0	0	0	K = 3